# SCHIP Evaluation
# Public Use Files

## CONTENTS

## TABLES

## FIGURES

This document describes the State Children's Health Insurance Program (SCHIP) Evaluation Public Use Files. First, it provides an overview of the evaluation. Next, it explains the evaluation's sample design and the sampling weights. It then describes the survey instrument used to collect the data for the evaluation. Finally, it summarizes the data included in the public use files and provides instructions on how to work with these data.

This document contains four appendixes. Appendix A provides details on how the samples were selected. Appendix B presents the methods used to develop the sampling weights. Appendix C describes the SCHIP and Medicaid enrollment data collected for the evaluation and the codes used by the states to determine eligibility for these programs. Finally, Appendix D summarizes how selected analytic variables were constructed.

In addition to this document, the SCHIP Evaluation Public Use Files include the following files:

- *SCHIP Survey.PDF* is the survey instrument used to collect data for the SCHIP evaluation.
- *SCHIP.SAS7BDAT* and *SCHIP_MED.SAS7BDAT* are the SAS data sets constructed from data collected for the SCHIP evaluation.
- *SCHIP Codebook.PDF* is an electronic codebook that describes the variables in the SAS data sets.
- *Formats.SAS7BCAT* is a SAS format library to be used with the SAS data sets.

## A. OVERVIEW OF THE SCHIP EVALUATION

In the Balanced Budget Refinement Act of 1999, Congress mandated that the Secretary of the U.S. Department of Health and Human Services conduct an independent comprehensive evaluation of the State Children's Health Insurance Program (SCHIP). The primary goal of the evaluation was to address questions about (1) SCHIP design and implementation, (2) who enrolled in SCHIP, (3) how the program affected access to care, and (4) the experiences families had when

enrolling their children. Under contract to the Office of the Assistant Secretary for Planning and Evaluation (ASPE), Mathematica Policy Research, Inc. (MPR) and its partners—The Urban Institute and the MayaTech Corporation—conducted the SCHIP evaluation between 2000 and 2005. A detailed description of the evaluation and its findings can be found in the report, "Congressionally Mandated Evaluation of the State Children's Health Insurance Program: Final Report to Congress" (Wooldridge et al. 2005).

The SCHIP evaluation collected data from a variety of sources. The cornerstone of the data was the information collected from a cross-sectional survey. Respondents included the parents of SCHIP enrollees and recent disenrollees in the 10 study states, as well as of Medicaid enrollees and recent disenrollees in 2 of the 10 states. Other data collected for this evaluation included a national survey of state program administrators, SCHIP and Medicaid enrollment data from the 10 study states, and case studies of the 10 study states.

As part of its commitment to making secondary data available for analysis, ASPE contracted with MPR and the Urban Institute to create public use files from the enrollee survey data and the program enrollment data collected for the SCHIP evaluation. Three activities were involved in creating these files: (1) consolidating the data gathered for the evaluation; (2) creating, from these data, public use files that conform to federal confidentiality requirements; and (3) producing an electronic codebook so others can use these files without technical assistance. This document and the accompanying files represent the results of these activities.

## B. ENROLLEE SURVEY SAMPLE DESIGN

The data collected through the survey were from two distinct samples. The first included samples of recent and established SCHIP enrollees and recent SCHIP disenrollees in 10 states. This sample was designed to make inferences about SCHIP enrollees and disenrollees in each of the 10 states and to make comparisons across the states. The second included samples of recent

and established Medicaid enrollees and recent Medicaid disenrollees in 2 of the 10 states in which we drew our SCHIP samples. It was designed similarly to the first sample to make inferences about Medicaid enrollees and disenrollees in the two states. It was also designed to draw comparisons between SCHIP and Medicaid enrollees and disenrollees in those states.

The high costs of face-to-face interviews led to the adoption of a dual-frame sample design. The dual-frame design combined an unclustered sample that was interviewed by telephone only (when a telephone number could be found through centralized locating efforts) with a clustered sample that was interviewed by telephone but had an in-person field followup for locating of nontelephone households. With this approach, it was possible to achieve the greater precision associated with the unclustered design, while keeping the enhanced response and coverage rates of the face-to-face approach. For all sample members, the interview was conducted with the person who knew most about the health care needs and services received for the sampled child. Typically, that person was either a parent or a legal guardian of the child. For in-person interviewing, the field locator provided the person with a cellular telephone for completing the interview, thus ensuring a consistent mode of interview (telephone) for all sample members.

Here, we provide additional detail on the sample design and selection, focusing on (1) the state selection process, (2) the target population to be surveyed in the states, (3) sampling methodology, (4) steps to address sample domain inconsistencies, (5) final sample sizes, and (6) sampling weights.

## 1. State Selection

The state selection process flowed from three criteria specified in the legislation for the evaluation—that 10 states were to be selected that were to (1) include a significant portion of uninsured low-income children, (2) use diverse programmatic approaches to providing child health assistance, and (3) represent various geographic areas. In addition, priority was given to states that

were participating in a separate focus group study funded by ASPE. Guided by these selection criteria, we chose the following states to participate in the SCHIP evaluation: California, Colorado, Florida, Illinois, Louisiana, Missouri, New Jersey, New York, North Carolina, and Texas.

For the survey of Medicaid enrollees and recent disenrollees, we chose California and North Carolina. We chose those states based on three criteria: (1) the size of the low-income population covered by SCHIP and Medicaid, (2) the integration of the SCHIP and Medicaid enrollment systems, and (3) the interest of ASPE in conducting the Medicaid survey in states that had adopted a separate SCHIP program.

## 2. Target Population Within States

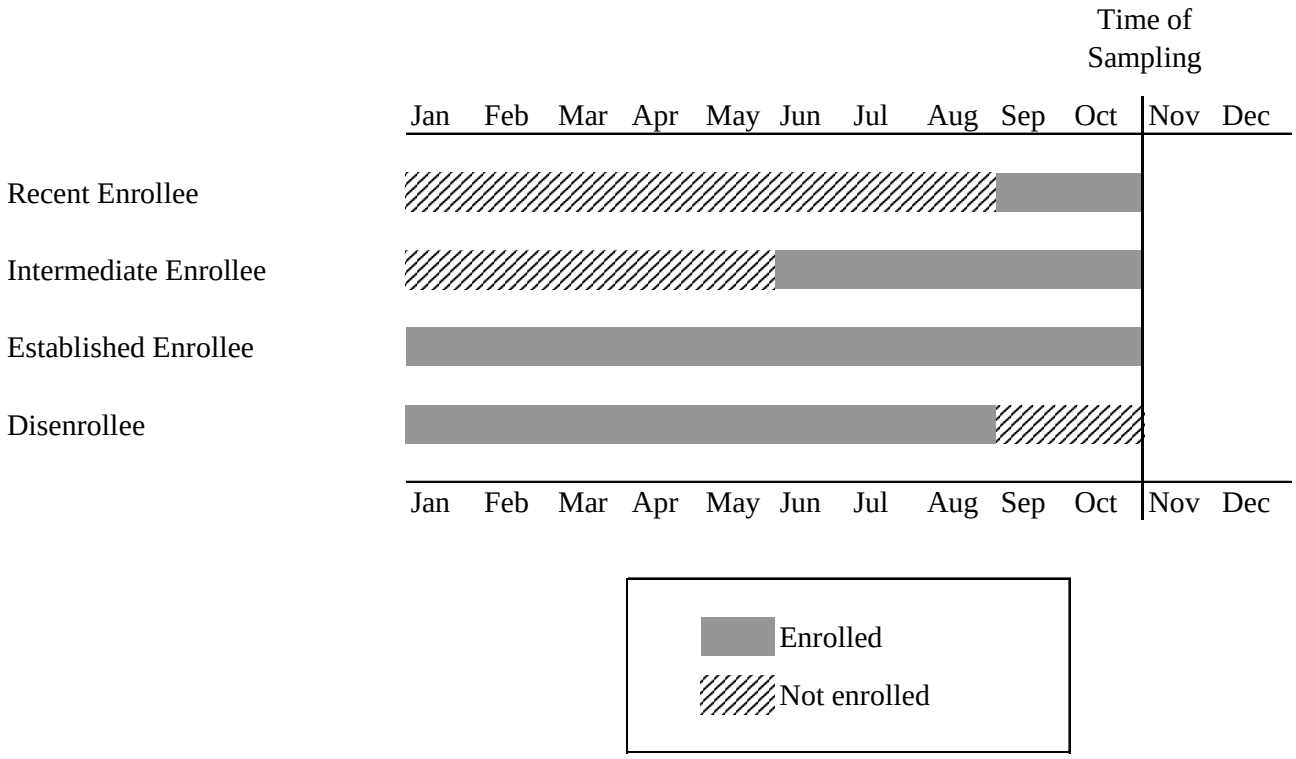
For each state, the SCHIP and Medicaid samples were drawn from a particular target population. To identify these populations, we used the following operational definitions of SCHIP and Medicaid enrollees and disenrollees:

- ***Recent Enrollees:*** Individuals enrolled in the given program (SCHIP or Medicaid) for at least one month but less than three months at the time of sampling. The enrollment spell was preceded by at least two months with no coverage in the program.
- ***Intermediate Enrollees:*** Individuals enrolled in the program for more than two months but less than six months at the time of sampling.
- ***Established Enrollees:*** Individuals enrolled for six or more months in the program at the time of sampling.
- ***Recent Disenrollees:*** Individuals disenrolled from the program at the time of sampling but enrolled in the preceding two months.

Figure 1 presents examples of individuals who would be included in each of these domains if, for example, sampling occurred in November. The target population for both the SCHIP and Medicaid samples was limited to three of those four domains: (1) recent enrollees, (2) established enrollees, and (3) recent disenrollees. Intermediate enrollees were not included in the evaluation

FIGURE 1

EXAMPLES OF SCHIP AND MEDICAID ENROLLEES AND DISENROLLEES




because they would be too far from their enrollment date to recall their preenrollment experience with a high degree of reliability but would not have been enrolled long enough to acquire experience with the program. To focus on children, the target population in both samples was further limited to individuals age 18 or younger in the two enrollee domains and to individuals age 19 or younger in the recent disenrollee domain.[1] Sampled children who had died or moved out of state were not of interest for the evaluation and were ineligible for data collection. We recorded the event leading to the ineligibility of these children to allow for complete reporting of the events leading to disenrollment.

[1] The age limit of 19 years was set for disenrollees to capture any children who had lost eligibility because of age restrictions.

In several states, the domain definitions were refined further, based on two guiding factors: (1) the enrollment process that the state used, and (2) the logistical constraints of the SCHIP enrollee databases used to select the sample. The goal of these refinements was to classify the child's enrollment status based on when the parent believed the child's health care services would be covered—a date that might differ from the date on which the state actually began paying for services. For example, some states retrospectively enrolled children as of the first day of the month in which the parent applied for SCHIP, but they might not determine the children to be eligible until one or more months after the application was received. As a result, the date that services began to be covered by the state might be month(s) earlier than the date on which the parent was notified of the child's enrollment. When defining the enrollment status, we used the child's determination/authorization date—the date on which eligibility was granted—as the start date for coverage to address this discrepancy. We did so because the determination/authorization date was likely to be the date that the parent would perceive as the start of coverage. Other states (such as New York) enrolled children at the time of application; thus, the database may contain "presumptive eligibles" who may later have been determined to be ineligible. In those states, the target population included only children for whom the determination process was completed and eligibility was confirmed. Furthermore, as in the states with retroactive enrollment, we assumed that enrollment began at the determination date.

For the Medicaid samples in California and North Carolina, several additional groups of children were excluded from the target population in order to create samples that, aside from differences in income eligibility, were equivalent and therefore comparable to the SCHIP samples in the two states. Examples of these exclusions include children who (1) resided in foster care or institutions; (2) received Social Security Income payments; (3) qualified as Medically Needy (California only); or (4) received partial benefits because of dual eligibility for Medicare,

immigrant status, or other reasons. In total, these exclusions led to the removal of about 56 and 10 percent of children from the eligibility files in California and North Carolina, respectively.

## 3. Sampling Methodology

For this study, we used data from the state SCHIP and Medicaid eligibility and enrollment files to construct the sampling frames for each state and program. Specific data elements used in the survey sampling process included:

- Application date(s) and their associated eligibility status codes
- Eligibility determination dates and their associated reason for eligibility codes
- Retroactive or presumptive eligibility status codes
- Enrollment start and end dates
- Disenrollment dates and their associated reason codes
- Individual and household identifiers
- Parent/guardian names
- Street addresses
- City, state, and zip code
- Telephone numbers
- Parent/guardian social security numbers
- Children's demographic characteristics, including age, race, and sex

Based on the sampling frame, the sample for the survey was separated into two types of households: (1) telephone households, and (2) nontelephone households. Telephone households were defined as households with telephone service for which telephone numbers could be located. Nontelephone households were defined as households without telephone service or households for which a telephone number could not be located. In each state except New Jersey, two independent

samples were selected for the SCHIP survey—one clustered and one unclustered.[2] Similarly, we drew two independent samples for the Medicaid survey in two states—one clustered and one unclustered. Telephone households were interviewed in both samples; nontelephone households were interviewed only in the clustered sample. Across both samples, telephone households were interviewed by telephone only. This restriction was necessary for the integration of the two samples. It also reduced mode effects across samples, because telephone households were always interviewed by telephone, regardless of the sample design from which they were drawn.

Each sample design was replicated with up to three different sample rounds and was fielded in each state. Each sample round was made up of sampled children from each SCHIP enrollment domain and, when applicable, from each Medicaid enrollment domain. The staged fielding of the sample was particularly important in reducing the time between sample frame construction and data collection. In addition, for states with the smallest populations of enrollees, the multiple rounds were needed to ensure that sufficient sample sizes of recent enrollees and recent disenrollees were obtained from each program. The sample for the last round for each state included a reserve sample from which additional sample cases were released for data collection if response or eligibility rates were unexpectedly low.

Because of the large population of enrollees in California and Texas, the full sample was selected from the March 2002 enrollment files. For six states (Florida, Illinois, Missouri, New Jersey, New York, and North Carolina), two sample rounds, which were based on the January and March 2002 enrollment files, were used. The samples for Colorado and Louisiana, which had the smallest enrollment populations, were selected using three sample rounds (using January, March, and May 2002 enrollment files). We avoided sampling more than one child from the same

[2] For New Jersey, we used only an unclustered design because the state is geographically small enough that the

household or sampling households in more than one sample round. Each sample draw was derived from the universe existing at the time of sampling but took into account whether a household was in the sampling frame or the sample of the prior round(s).

In each sample round, we classified children into the three domains (recent enrollees, established enrollees, and recent disenrollees), using the databases provided by the states. In states with multiple sample rounds, the populations of established enrollees overlapped extensively; however, by definition, recent enrollees and recent disenrollees were unique to a specific sample round. Enrollment status for a given child could vary from one sampling round to another. For example, established enrollees at one time could become recent disenrollees the next time.

In each round, the sample consisted of a clustered sample and an unclustered sample of children in the SCHIP domain (except for New Jersey) and the Medicaid domain (in California and North Carolina). We used sampling procedures that prevented the selection of the same child or household at subsequent rounds while preserving the probability structure of the two independent samples in each round. (Appendix A presents the complete methodology used to select the clustered and unclustered samples, including a description of the primary sampling units and strata.)

## 4. Addressing Sample Domain Inconsistencies

To illustrate the importance of addressing potential inconsistencies between the respondents' perceptions and the assigned sample domains, consider the children we selected for the recent enrollee sample. The state program files showed that almost 35 percent of the children across the 10-state sample either had spells of SCHIP coverage before enrolling (their short gaps in coverage perhaps resulting from late premium payments or renewals) or had recent spells of Medicaid

*(continued)*
use of a clustered sample was deemed unnecessary.

coverage before enrolling (often with no gaps between the two programs). In some instances, the families would not be expected to recognize their recent enrollment in SCHIP, believing instead that they had never left the program (in the case of a short gap in SCHIP coverage) or had never switched programs (in the case of a transition from Medicaid). Many of these families would therefore have reported having been covered by SCHIP for longer than indicated by the state files, often significantly so. As a result, when these families reported on key outcomes, such as prior insurance coverage or pre-SCHIP utilization of health care, they were not reporting those data for the period immediately before their current (state-determined) period of enrollment.

To address this problem and others like it, we had two options. The first was to simply drop from the survey sample any cases whose self-reported dates of entry (or exit) were inconsistent with the domains in which they had been sampled. For example, a recent enrollee who reported having been enrolled for, say, a year or more at the time of interview might be classified as ineligible for the survey and dropped from the recent enrollee sample. This approach was attractive because it was simple and would have yielded an analytic file containing reliable data for all outcomes across all sample members. However, because the approach would remove a large fraction of the children and families originally sampled for the survey, it could have led to substantial biases in our estimates of several key outcomes.

For example, suppose we had dropped from the study sample any recent enrollee who had reported being enrolled in SCHIP for an extended period, say, a year or more. This step would have eliminated the problem of interviewing recent enrollees who believed themselves enrolled for a long period of time. However, it probably also would have resulted in the removal of a disproportionate share of recent enrollees who had either transitioned from Medicaid seamlessly or who had experienced only short gaps in SCHIP coverage. In turn, any estimates of prior coverage among recent enrollees would have been biased, leading to underestimates of the share of recent

enrollees with public coverage and to overestimates of the share with private coverage or no insurance.

The second option, which we adopted, was to keep sample members who displayed inconsistency between the state enrollment data and the self-reported data and to interview families based on the self-reported information, rather than on the information from the state enrollment files. For example, if a recent enrollee informed us in the interview that he or she had been enrolled for more than a year, we interviewed that person as if he or she were an established enrollee, not a recent enrollee. This option required us to use imputation and/or nonresponse adjustment for some outcomes to account for survey data on selected sample members that were either incomplete or incorrect. For example, the person who was sampled as a recent enrollee but reported being enrolled for more than a year (and thus was interviewed as an established enrollee) was not asked questions about prior coverage because he or she might not be able to provide accurate information. Instead, we examined the program data, and if they indicated that the person had Medicaid coverage before enrolling in SCHIP, we would assign his or her prior coverage as Medicaid. This option allowed us to keep a sample that was fully representative of each study domain, making it much more likely to yield unbiased estimates of the experiences of SCHIP enrollees and disenrollees.

As Table 1 shows, the adoption of this approach led to a complex sample design. In total, the sample included 18 types of sample members across the three domains. For some sample members, survey questions were either skipped because they could not be addressed properly or were replaced by a different series of questions.[3] For example, within the recent-enrollee domain, children reported to have been enrolled at birth were not asked any questions about their pre-

[3] The survey instrument and the questions asked of sample members from different domains are discussed in greater detail in Section C of this document.

TABLE 1

SUMMARY OF THE TYPES OF SAMPLE MEMBERS

| Domain | Statuses Within Domain (Self-Reported) |
|---|---|
| Recent Enrollee | Enrolled for fewer than 12 months |
| | Born in the 6 months before SCHIP enrollment |
| | Obtained coverage at birth and is enrolled for 12 months or more |
| | Obtained coverage at birth and is enrolled for less than 12 months |
| | Enrolled for 12 months or more |
| | Disenrolled for 6 months but less than 12 months |
| | Disenrolled for 12 months or more |
| Established Enrollee | Enrolled 6 months or more |
| | Obtained coverage at birth |
| | Enrolled for less than 6 months |
| | Disenrolled 6 months but less than 12 months |
| | Disenrolled for 12 months or more |
| Disenrollees | Disenrolled for less than 12 months |
| | Currently enrolled for 6 months or more |
| | Disenrolled for 12 months or more |
| | Disenrolled for 12 months or more—recontacted and completed interview |
| Statuses That Apply to All Domains | No information on whether the sample child is enrolled |
| | Missing date(s) needed to determine duration of enrollment |

Source: 2002 congressionally mandated survey of SCHIP enrollees and disenrollees in 10 states.

SCHIP access, service use, or other experiences for obvious reasons; however, if the newborns were reported to have been enrolled for 12 months or longer at interview, we collected information about their experiences while in the program. Furthermore, we used SCHIP and Medicaid enrollment files to validate reports that children were enrolled in SCHIP at birth. We were thus able to identify children who had actually been enrolled in Medicaid at birth and had then transferred seamlessly to SCHIP. By adopting these strategies, we were able to collect as much usable information as possible on each member of the sample.

## 5. Final Sample Sizes

As described earlier, the sample design for the study allowed children to be selected for the study in either a clustered or unclustered sample. In rare instances, SCHIP children were selected for both samples, leading these children to have two records in the SCHIP analysis sample rather than one. (Children were interviewed only once, but their data were recorded in both samples.)

The resulting analysis sample for the SCHIP study, summarized in Table 2, included 16,780 records drawn from 16,680 interviews with the parents of SCHIP enrollees and disenrollees.[4] The Medicaid analysis sample, summarized in Table 3,[5] consisted of only two domains: (1) recent enrollees, and (2) established enrollees.[6] Unlike the SCHIP sample, there were no instances of dual sample selection in the Medicaid sample, so that the total sample size reported (1,833) reflects both the number of sample records and the number of completed interviews conducted with the parents of Medicaid enrollees. For both the SCHIP and Medicaid samples, the size of the unweighted sample was roughly equal across the sample domains. However, the weighted sample was much larger for the established enrollees, reflecting their larger population in relation to the other sample domains.

Within each domain, the largest subsample was the one that we would expect a respondent to self-report. For example, within the domain of recent SCHIP enrollees, the largest subsample consisted of children whose parents reported that they had been enrolled for fewer than 12 months

[4] When analyzing SCHIP enrollees and disenrollees, the full sample of 16,780 records should be used at all times.

[5] Though we combined both states of the Medicaid analysis sample in Table 3, we chose to examine each state individually throughout the analysis.

[6] Because of a combination of low response rates and the very low rates of recognition that they had actually been disenrolled, Medicaid disenrollees were not analyzed for the SCHIP evaluation and are not included in the public use files.

TABLE 2

SCHIP SURVEY: SAMPLE SIZE AND DISTRIBUTION

| Definition | Unweighted | | | | Weighted | | |
|---|---|---|---|---|---|---|---|
| | Sample Size | | | | | | |
| | Records | Interviews | Percentage of Sample Domain | Percentage of Total Sample | Sample Size | Percentage of Sample Domain | Percentage of Total Sample |
| **Recent Enrollees** | | | | | | | |
| Recent Enrollee Who Has Been Enrolled for Fewer than 12 Months | 3,330 | 3,326 | 59 | 20 | 111,658 | 61 | 6 |
| Recent Enrollee Who Was Born in the 6 Months Before SCHIP Enrollment | 67 | 67 | 1 | < 1 | 2,176 | 1 | < 1 |
| Recent Enrollee Who Obtained Coverage at Birth and Is Enrolled for 12 Months or More | 164 | 164 | 3 | 1 | 2,806 | 2 | < 1 |
| Recent Enrollee Who Obtained Coverage at Birth and Is Enrolled for Fewer than 12 Months | 37 | 37 | 1 | < 1 | 1,462 | 1 | < 1 |
| Recent Enrollee Who Has Been Enrolled for 12 Months or Longer | 1,761 | 1,756 | 31 | 10 | 55,317 | 30 | 3 |
| Recent Enrollee Who Has Been Disenrolled for 6 Months but Fewer than 12 Months | 84 | 82 | 1 | 1 | 3,160 | 2 | 0 |
| Recent Enrollee Who Has Been Disenrolled for 12 Months or Longer | 76 | 75 | 1 | < 1 | 2,294 | 1 | 0 |
| No Information on Whether Sample Child Is Enrolled | 62 | 62 | 1 | < 1 | 1,870 | 1 | 0 |
| Missing Date(s) to Determine Duration of Enrollment | 82 | 82 | 1 | < 1 | 2,361 | 1 | 0 |
| **Subtotal (Recent Enrollees)** | **5,663** | **5,651** | **100** | **34** | **183,105** | **100** | **10** |
| **Established Enrollees** | | | | | | | |
| Established Enrollee Who Has Been Enrolled 6 Months or More | 5,010 | 5,007 | 86 | 30 | 1,373,010 | 89 | 77 |
| Established Enrollee Who Obtained Coverage at Birth | 179 | 178 | 3 | 1 | 30,542 | 2 | 2 |
| Established Enrollee Enrolled for Fewer than 6 Months | 109 | 109 | 2 | 1 | 27,681 | 2 | 2 |
| Established Enrollee Who Has Been Disenrolled for 6 Months but Fewer than 12 Months | 167 | 167 | 3 | 1 | 44,873 | 3 | 3 |

TABLE 2 *(continued)*

| Definition | Unweighted Sample Size: Records | Unweighted Sample Size: Interviews | Unweighted: Percentage of Sample Domain | Unweighted: Percentage of Total Sample | Weighted: Sample Size | Weighted: Percentage of Sample Domain | Weighted: Percentage of Total Sample |
|---|---|---|---|---|---|---|---|
| Established Enrollee Who Has Been Disenrolled for 12 Months or More | 112 | 112 | 2 | 1 | 25,735 | 2 | 1 |
| No Information on Whether Sample Child Is Enrolled | 83 | 83 | 1 | < 1 | 18,398 | 1 | 1 |
| Missing Date(s) to Determine Duration of Enrollment | 177 | 137 | 2 | 1 | 26,863 | 2 | 2 |
| **Subtotal (Established Enrollees)** | **5,797** | **5,793** | **100** | **35** | **1,547,102** | **100** | **86** |
| **Disenrollees** | | | | | | | |
| Disenrollee Who Has Been Disenrolled for Less than 12 Months | 2,051 | 2,011 | 39 | 12 | 23,265 | 40 | 1 |
| Disenrollee Who Has Been Currently Enrolled for 6 Months or More | 1,762 | 1,747 | 33 | 11 | 16,980 | 29 | 1 |
| Disenrollee Who Has Been Disenrolled for 12 Months or More | 563 | 550 | 11 | 3 | 6,507 | 11 | < 1 |
| Disenrollee Who Has Been Disenrolled for 12 Months or More—Recontacted and Completed Interview | 630 | 618 | 12 | 4 | 8,352 | 14 | < 1 |
| No Information on Whether Sample Child Is Enrolled | 113 | 112 | 2 | 1 | 1,122 | 2 | < 1 |
| Missing Date(s) to Determine Duration of Enrollment | 201 | 198 | 4 | 1 | 2,177 | 4 | < 1 |
| **Subtotal (Disenrollees)** | **5,320** | **5,236** | **100** | **32** | **58,403** | **100** | **3** |
| **Total (Full Sample)** | **16,780** | **16,680** | **100** | **100** | **1,788,610** | **100** | **100** |

Source: 2002 congressionally mandated survey of SCHIP enrollees and disenrollees in 10 states.

TABLE 3

MEDICAID SURVEY: SAMPLE SIZE AND DISTRIBUTION

| Definition | Unweighted | | | Weighted | | |
|---|---|---|---|---|---|---|
| | Sample Size | Percentage of Sample Domain | Percentage of Total Sample | Sample Size | Percentage of Sample Domain | Percentage of Total Sample |
| **Recent Enrollees** | | | | | | |
| Recent Enrollee Who Has Been Enrolled for Fewer than 12 Months | 311 | 34 | 17 | 21,972 | 33 | 1 |
| Recent Enrollee Who Was Born in the 6 Months Before SCHIP Enrollment | 56 | 6 | 3 | 3,873 | 6 | < 1 |
| Recent Enrollee Who Obtained Coverage at Birth and Is Enrolled for 12 Months or More | 87 | 10 | 5 | 7,543 | 11 | < 1 |
| Recent Enrollee Who Obtained Coverage at Birth and Is Enrolled for Fewer than 12 Months | 225 | 25 | 12 | 15,581 | 23 | 1 |
| Recent Enrollee Who Has Been Enrolled for 12 Months or Longer | 186 | 20 | 10 | 13,997 | 21 | 1 |
| Recent Enrollee Who Has Been Disenrolled for 6 Months but Fewer than 12 Months | 17 | 2 | 1 | 1,581 | 2 | < 1 |
| Recent Enrollee Who Has Been Disenrolled for 12 Months or Longer | 14 | 2 | 1 | 1,225 | 2 | < 1 |
| No Information on Whether Sample Child Is Enrolled | 9 | 1 | < 1 | 1,109 | 2 | < 1 |
| Missing Date(s) to Determine Duration of Enrollment | 6 | 1 | < 1 | 497 | 1 | < 1 |
| **Subtotal (Recent Enrollees)** | **911** | **100** | **50** | **67,378** | **100** | **3** |
| **Established Enrollees** | | | | | | |
| Established Enrollee Who Has Been Enrolled 6 Months or More | 461 | 50 | 25 | 863,121 | 46 | 44 |
| Established Enrollee Who Obtained Coverage at Birth | 345 | 37 | 19 | 755,159 | 40 | 39 |
| Established Enrollee Enrolled for Fewer than 6 Months | 31 | 3 | 2 | 65,570 | 3 | 3 |
| Established Enrollee Who Has Been Disenrolled for 6 Months but Fewer than 12 Months | 25 | 3 | 1 | 55,641 | 3 | 3 |
| Established Enrollee Who Has Been Disenrolled for 12 Months or More | 28 | 3 | 2 | 69,444 | 4 | 4 |

TABLE 3 *(continued)*

| Definition | Unweighted | | | Weighted | | |
|---|---|---|---|---|---|---|
| | Sample Size | Percentage of Sample Domain | Percentage of Total Sample | Sample Size | Percentage of Sample Domain | Percentage of Total Sample |
| No Information on Whether Sample Child Is Enrolled | 16 | 2 | 1 | 38,338 | 2 | 2 |
| Missing Date(s) to Determine Duration of Enrollment | 16 | 2 | 1 | 37,777 | 2 | 2 |
| **Subtotal (Established Enrollees)** | **922** | **100** | **50** | **1,885,048** | **100** | **97** |
| **Total (Full Sample)** | **1,833** | **100** | **100** | **1,952,426** | **100** | **100** |

Source: 2002 congressionally mandated survey of Medicaid enrollees in 2 states.

(3,330 of the 5,663 records in that domain) (Table 2). Likewise, within the domain of established SCHIP enrollees, the largest subsample consisted of children whose parents reported that they had been enrolled for six months or more (5,010 of the 5,797 records in that domain). The same pattern also was true for the SCHIP-disenrollee domain, although to a lesser extent. The largest subsample reported being disenrolled for fewer than 12 months (2,051 of the 5,320 records in that domain), though a large number also reported being enrolled for six or more months (1,762).

## 6. Sampling Weights

Sampling weights were constructed for all sample members.[7] The purpose of these weights is to allow the selected sample to reflect accurately the characteristics and outcomes of the sample frame (that is, the population of SCHIP and Medicaid enrollees and disenrollees across the 10 states) from which they were drawn. For recent and established enrollees, each sample member has a single sampling weight variable, named *ENR_WGHT*. For disenrollees, each sample member has a pair of sampling weight variables. The first, named *DIS_WGHT_POP*, pertains to the entire disenrollee sample and should be used when examining remaining outcomes, including health status and sociodemographic characteristics. The second, named *DIS_WGHT*, pertains only to certain disenrollees and should be used when addressing the following key outcomes: (1) reasons for leaving SCHIP, (2) reasons for being uninsured after leaving SCHIP, (3) post-SCHIP coverage, and (4) all measures of health care access and use while enrolled in SCHIP.[8]

[7] The sampling weights take into account observations that are included in both the clustered and unclustered samples.

[8] The *DIS_WGT* variable pertains specifically to three subgroups of disenrollees: (1) disenrollees who have been disenrolled for fewer than 12 months, (2) disenrollees who have been currently enrolled for 6 months or more, and (3) disenrollees who have been disenrolled for 12 months or more and were successfully recontacted (and asked additional questions pertaining to these key outcomes). For other disenrollees, the *DIS_WGT* variable is set to zero, so there is no need to subset the disenrollee sample when analyzing outcomes relevant to this weight.

When analyzing the survey data, we recommend that the appropriate survey weight be used at all times. For a discussion of the statistical specifications used to incorporate these weights, in conjunction with the other elements of the survey design, see Section D.6 below (Computing Correct Standard Errors).

Our method for computing these weights, detailed fully in Appendix B, was as follows. First, we computed sampling weights for each design (clustered and unclustered) for each sample round and state. These within-sample round, within-design sampling weights were calculated using the product of the sampling weight of the household multiplied by the conditional sampling weight of the child, given that his or her household was selected.[9] We then combined the design-specific sample weights across rounds to create a single base sampling weight for each sampled child for each design for each state.[10] The two sets of weights (one for the unclustered sample and one for the clustered sample) were poststratified to the same average monthly enrollment population (computed from enrollment counts for data collection round enrollment files) for each domain in each state.

We subsequently conducted a nonresponse analysis to assess the response patterns for the samples. We used data available from the sampling frame (such as the age and race of the sampled child) and county-level information (such as the percentage of children living in households with family incomes under the poverty level, the percentage of households with female head of the household, and a scale denoting urbanicity). Using the results of the nonresponse analysis, we developed logistic regression models to compute response propensity scores to

[9] The sampling weight of the household is the inverse of the probability of selection of the household. The conditional sampling weight of the child is the inverse of the probability of selection of the child, given that his or her household was selected. As stated earlier, only one child per household was selected.

[10] Recall that, for California and for Texas, only one round was used, and that, for New Jersey, only the unclustered design was used.

compensate for nonresponse. The nonresponse-adjusted weight was the product of the combined-round base weight and the inverse of the response propensity score. We developed response propensity models separately for each sample (clustered and unclustered); for each domain (recent enrollees, established enrollees, and recent disenrollees); for each state; and for each study population (SCHIP and Medicaid). Finally, we used the estimated population counts in each state and each domain to poststratify within each state based on enrollment status at the time of sampling of the child. The poststratification adjustment ensured that the nonresponse-adjusted base weights summed to the estimated enrollment population for that domain in each state.[11]

## C. SURVEY INSTRUMENT

The survey instrument, *SCHIP Survey.PDF*, addressed a broad range of topics related to the ease of application and enrollment in SCHIP/Medicaid redetermination in and disenrollment from the program, health care coverage for the child, the child's health, experiences with and use of care for the child, the respondent's attitude toward health, and demographic characteristics of the household members. Whenever possible, we used survey questions that had been validated from existing surveys, including the National Health Interview Survey, the Evaluation of Five Section 1115 Medicaid Reform Demonstrations Survey, the National Survey of America's Families, the Consumer Assessment of Health Plans Survey, and the Kaiser Family Foundation National Medicaid Survey on Barriers to Medicaid for Children. Table 4 summarizes, by section, the topics in the instrument.[12] On average, the instrument took about 40 minutes to administer.

[11] Only the final analysis weights are included in the public use files. Interim weights, such as the design-specific sample weights and nonresponse adjusted weights, have not been included in the public use files because they were not analyzed for the SCHIP evaluation.

[12] A glossary of terms used throughout the survey instrument is included as an appendix to *SCHIP Survey.PDF*.

TABLE 4

SELECTED SURVEY TOPICS

| |
|---|
| **Section 1: Introduction** |
| Confirm child lives in household |
| Confirm child lives in target state |
| Confirm respondent is the person most familiar with the child's health care |
| Read confidentiality statement |
| **Section 2: Application, Enrollment, Redetermination, and Disenrollment** |
| How respondent heard about program |
| How heard about program an important part of the decision to enroll child in SCHIP/Medicaid? |
| Experiences with enrollment process |
| Experience with rejection of application |
| Number of times successfully enrolled |
| Age of child when first enrolled |
| Reason for enrollment |
| Was assistance with application process necessary? |
| Application and enrollment processes and comparisons between SCHIP and Medicaid |
| Coverage available prior to notification |
| Renewal process and experience with rejection of renewal |
| **Section 3: Health Care Coverage** |
| Current enrollment status |
| Establish end date(s) of coverage |
| Establish last or current start date |
| Establish previous end date and start date for disenrollees who enrolled again |
| Features of current, last, or previous SCHIP/Medicaid coverage |
| Premiums |
| Types of service provided |
| Co-payments |
| Prescription drug coverage |
| Period before SCHIP/Medicaid began coverage |
| If insured, features of plan |
| If uninsured, how long and why |
| Period after SCHIP/Medicaid coverage ended |
| If uninsured, how long and why |
| If insured, features of plan |
| Premiums |
| Type of service provided |

TABLE 4 *(continued)*

| |
|---|
| Co-payments |
| Prescription drug coverage |
| **Section 4: Child's Health** |
| Child's health status |
| Child's health status versus 12 months ago |
| Any impairment(s) requiring special equipment or limiting mobility |
| Existing health conditions that have been diagnosed<br>Diabetes<br>Asthma |
| Any need for doctor-prescribed medications or injections |
| Mental health or behavioral problems |
| Any need for prescription medications or injections |
| Do mental health or behavioral problems limit child's abilities at school? |
| **Section 5: Access to, Barriers to, and Satisfaction with Usual Place of Care** |
| Usual place of care child actually went to or would have gone to if sick or needed advice |
| If no usual place, why not, what type of place child would have gone to, what type of place visited |
| If usual place of care, rate features of place<br>Distance<br>Waiting time<br>Transportation<br>Particular doctor<br>How child was treated<br>Ease of care<br>Where to get advice if usual place closed<br>How long a wait for care |
| If place of care changed, main reason for change |
| Type of new place |
| Reason for visit |
| Features of this place of care |
| How well treated |
| Usual place for dental care child actually went to or would have gone to |
| If no usual place, why not |
| **Section 6: Child's Use of Health Care Services** |
| Health care services child used |
| Number of hospital visits |
| Number of nights in hospital |
| Number of emergency room visits |
| Number of times child saw a doctor, physician's assistant, nurse, or midwife |
| Use of specialists |
| Number of visits for preventive care |

TABLE 4 *(continued)*

| |
|---|
| Use of mental health professionals |
| Number of times used mental health professionals |
| Use of dentists |
| Was needed care delayed? |
| Did child take less than prescribed dose of medication? |
| Confidence that child could get needed health care |
| Satisfaction with health care received |
| How worried was respondent about meeting child's health care needs? |
| Stress about meeting child's health care needs |
| Financial problems in meeting health care needs |
| **Section 7: Parents' Characteristics and Attitudes About Health** |
| How respondent perceived own health |
| Attitude about health and health care |
| Establish household composition |
| Establish who is legal guardian of child |
| Respondent's age |
| Respondent's education level |
| Respondent's place of birth |
| Other legal guardian of child in household |
| Other legal guardian's education level |
| Other legal guardian's place of birth |
| Health insurance status of legal parents or guardians in household |
| If insured, why is child not insured by same? |
| Features of legal guardian's health insurance |
| Is legal parent/guardian married to another person who is not the legal guardian of child? |
| Can child be covered by this person's insurance? |
| Household earnings for past 12 months |
| Temporary Assistance for Needy Families recipient for past two years |
| Food Stamps recipient for past two years |
| Health care spending in past 12 months |
| Child's racial or ethnic background and language spoken in home |
| **Section 8: Telephones in Household** |
| Number of other telephone numbers used in household |
| Number working in past three months |
| Verify address |

Source: 2002 congressionally mandated survey of SCHIP enrollees and disenrollees in 10 states.

As Table 5 shows, survey respondents were asked different questions, depending on the enrollment domain in which they were sampled (recent enrollee, established enrollee, recent disenrollee) and on the information provided during the interview about when the child's coverage started and ended.[13] In addition, the wording of questions varied, depending on responses to prior questions (most notably, the dates of coverage). For example, several questions about children's service use and other topics were anchored to a specific time frame that varied both by the children's enrollment domain and the self-reported dates of enrollment. For instance, for a recent enrollee who reported a start date consistent with the sample frame drawn from state administrative data, the specified time frame was the six months before entry in SCHIP (or Medicaid, for the Medicaid sample). For an established enrollee who confirmed having been covered for at least six months, the specified time frame was the most recent six months during which the child had been covered by the program.

## D. DATA DESCRIPTION

This section describes the data in the SCHIP Evaluation Public Use Files. We begin by describing the SAS data sets that were constructed from data collected for the SCHIP evaluation. We then discuss the measures taken to maintain the confidentiality of the sample members and explain how verbatim responses were coded. Next, we define the meanings associated with missing values in the data sets and provide instructions on how to use the SAS format library. Finally, we explain how to compute the correct sampling variance when using these data.

[13] Respondents who had no information on whether the child was enrolled, who did not know the dates to determine the duration of enrollment, or who reported that the child had been disenrolled more than 12 months received a shortened survey with a set of questions solely related to demographic characteristics.

TABLE 5

SURVEY QUESTIONS ANSWERED BY RESPONDENT, BY THE SAMPLE MEMBER'S ENROLLMENT DOMAIN

| Definition | Introduction (Section 1) | Application, Enrollment Predetermination, Disenrollment (Section 2) | Child's Health Care Coverage (Section 3) | Child's Health (Section 4) | Time Frame for Sections 5-6 | Access and Barriers to Care (Section 5) | Child's Use of Health Care Services (Section 6) | Parent Characteristics (Section 7) | Telephone Coverage (Section 8) |
|---|---|---|---|---|---|---|---|---|---|
| **Statuses Within the Recent Enrollee Domain** | | | | | | | | | |
| Recent Enrollee Who Has Been Enrolled for Fewer than 12 Months | Yes | Yes | 3.1-3.9.1B, 3.20-3.44 | Yes | The 6 months before (child)'s current SCHIP coverage started | Yes | Yes | Yes | Yes |
| Recent Enrollee Who Was Born in the 6 Months Before SCHIP Enrollment | Yes | Yes | 3.1-3.9.1B, 3.20-3.44 | Yes | Before (child) was on SCHIP | Yes | Yes | Yes | Yes |
| Recent Enrollee Who Obtained Coverage at Birth and Has Been Enrolled for 12 Months or More | Yes | Yes | 3.1-3.9.1B, 3.20-3.31 | Yes | Past 6 months | Yes | Yes | Yes | Yes |
| Recent Enrollee Who Obtained Coverage at Birth and Is Enrolled for Fewer than 12 Months | Yes | Yes | 3.1-3.9.1B, 3.20-3.31 | Yes | | No | No | 7.4.a-7.4.1.9, 7.109-7.120 | 8.15 to end |
| Recent Enrollee Who Has Been Enrolled for 12 Months or Longer | Yes | Yes | 3.1-3.9.1B, 3.20-3.44 | Yes | Past 6 months | Yes | Yes | Yes | Yes |
| Recent Enrollee Who Has Been Disenrolled for 6 Months but Fewer than 12 Months | Yes | Yes | 3.1-3.9.1B, 3.20-3.44 | Yes | The 6 months before (child)'s last SCHIP coverage ended | Yes | Yes | Yes | Yes |
| Recent Enrollee Who Has Been Disenrolled for 12 Months or Longer | Yes | Yes | 3.1-3.51 | Yes | | No | No | 7.4.a-7.4.1.9, 7.109-7.120 | 8.15 to end |

TABLE 5 *(continued)*

| Definition | Introduction (Section 1) | Application, Enrollment Predetermination, Disenrollment (Section 2) | Child's Health Care Coverage (Section 3) | Child's Health (Section 4) | Time Frame for Sections 5-6 | Access and Barriers to Care (Section 5) | Child's Use of Health Care Services (Section 6) | Parent Characteristics (Section 7) | Telephone Coverage (Section 8) |
|---|---|---|---|---|---|---|---|---|---|
| **Statuses Within the Established Enrollee Domain** | | | | | | | | | |
| Established Enrollee Who Has Been Enrolled 6 Months or More | Yes | Yes | 3.1-3.9.1B, 3.20-3.44 | Yes | Past 6 months | Yes | Yes | Yes | Yes |
| Established Enrollee Who Obtained Coverage at Birth | Yes | Yes | 3.1-3.9.1B, 3.20-3.31 | Yes | Past 6 months | Yes | Yes | Yes | Yes |
| Established Enrollee Enrolled for Fewer than 6 Months | Yes | Yes | 3.1-3.9.1B, 3.20-3.44 | Yes | While the (child) was on SCHIP | Yes | Yes | Yes | Yes |
| Established Enrollee Who Has Been Disenrolled 6 Months but Fewer than 12 Months | Yes | Yes | 3.1-3.9.1B, 3.20-3.25, 3.60 to end | Yes | The 6 months before (child)'s last SCHIP coverage ended | Yes | Yes | Yes | Yes |
| Established Enrollee Who Has Been Disenrolled for 12 Months or More | Yes | Yes | 3.1-3.51 | Yes | | No | No | 7.4.a-7.4.1.9, 7.109-7.120 | 8.15 to end |
| **Statuses Within the Recent Disenrollee Domain** | | | | | | | | | |
| Disenrollee Who Has Been Disenrolled for Fewer than 12 Months | Yes | Yes | 3.1-3.9.1B, 3.20-3.25, 3.60 to end | Yes | The 6 months before (child)'s last SCHIP coverage ended | Yes | Yes | Yes | Yes |
| Disenrollee Who Has Been Currently Enrolled for 6 Months or More | Yes | Yes | 3.1-3.9.1B, 3.20-3.25, 3.60 to end | Yes | Past 6 months | Yes | Yes | Yes | Yes |
| Disenrollee Who Has Been Disenrolled for 12 Months or More | Yes | Yes | 3.1-3.51 | Yes | | No | No | 7.4.a-7.4.1.9, 7.109-7.120 | 8.15 to end |
| Disenrollee Who Has Been Disenrolled for 12 Months or More—Recontacted and Completed Interview | Yes | Yes | 3.1-3.5, 3.26, 3.60-3.65 | Yes | | No | No | 7.4.a-7.4.1.9, 7.109-7.120, 7.4.5.1-7.4.5.6, 7.90-7.101 | 8.15 to end |

TABLE 5 *(continued)*

| Definition | Introduction (Section 1) | Application, Enrollment Predetermination, Disenrollment (Section 2) | Child's Health Care Coverage (Section 3) | Child's Health (Section 4) | Time Frame for Sections 5-6 | Access and Barriers to Care (Section 5) | Child's Use of Health Care Services (Section 6) | Parent Characteristics (Section 7) | Telephone Coverage (Section 8) |
|---|---|---|---|---|---|---|---|---|---|
| **Statuses That Apply to All Domains** | | | | | | | | | |
| No Information on Whether Sample Child Is Enrolled | Yes | Yes | 3.1 | Yes | | No | No | 7.4.a-7.4.1.9, 7.109-7.120 | 8.15 to end |
| Missing Date(s) to Determine Duration of Enrollment | Yes | Yes | 3.1-3.51 | Yes | | No | No | 7.4.a-7.4.1.9, 7.109-7.120 | 8.15 to end |

Source: 2002 congressionally mandated survey of SCHIP enrollees and disenrollees in 10 states.

Note: In the table, a ""Yes" indicates that all of the survey questions were asked for that section. "No" indicates that the interviewee was skipped out of the section. When only a subset of questions was asked in a particular section, we list the question numbers that were asked in the table.

## 1. Data Sets

The SCHIP Evaluation Public Use Files include two SAS data sets. The first data set, *SCHIP.SAS7BDAT*, includes the 16,780 observations from the 10 study states used in the analysis of recent and established SCHIP enrollees and recent SCHIP disenrollees. The second data set, *SCHIP_MED.SAS7BDAT*, includes the 4,181 observations from California and North Carolina used in the analysis of recent and established Medicaid enrollees.[14] Among these observations, 1,833 are Medicaid enrollees, and 2,348 are SCHIP enrollees, also from California and North Carolina, who were included in the data set to enable comparisons between SCHIP and Medicaid enrollees.[15]

The electronic codebook, *SCHIP Codebook.PDF*, contains information about the variables in the SAS data sets. Specifically, it provides names and descriptions of the variables, along with the meaning of each value taken by the variables. In the electronic codebook, variables are categorized according to their source:

- *Sampling variables* include measures from the sampling frame, sample domain, sampling weights, and the state sampled in.
- *Survey variables* are questions taken directly from the survey instrument. These variables were named with the prefix "Q" followed by the question number, replacing the "."s in the question number with "_"s in the variable name.
- *Program variables* are the SCHIP and Medicaid enrollment data and eligibility codes provided by states. A more detailed description of these variables is also presented in Appendix C.

[14] As stated earlier, a sample of recent Medicaid disenrollees was also surveyed but was not analyzed because of low response rates and the low rates of recognition that they had actually been disenrolled. Because they were not part of the SCHIP evaluation, these data have not been included in the SAS data set.

[15] As stated earlier, we examined the two states of the Medicaid analysis separately throughout the analysis and elected never to combine these states.

- *Area Resource File variables* are based on the rural/urban continuum code from the 2001 Area Resource File. A more detailed description of these variables is also presented in Appendix D.
- *Constructed variables* are analytic measures created from survey variables, program variables, or both and used in the SCHIP evaluation. A more detailed description of the more complex constructed variables is also presented in Appendix D.

## 2. Maintaining Confidentiality

We constructed the SAS data sets in compliance with federal statutes regarding personally identifiable data, particularly the Privacy Act of 1974. This act, which regulates the collection, maintenance, and dissemination of personal information, establishes three mandates for agencies that collect personal information: (1) agencies are required to maintain the minimal amount of information about an individual that is relevant and necessary, (2) agencies must establish rules for persons working with personal information to ensure that these data remain confidential, and (3) agencies are required to develop safeguards to protect the security and confidentiality of the personal information.

Because of the sensitive nature of some information collected for the SCHIP evaluation, we implemented a masking technique on the SAS data sets to protect the privacy of the respondents and their children. The masking technique involved three basic steps. The first step was to remove all individually identifying variables from the data sets. The next step was to examine tabulations of several critical demographic characteristics of the children against other potentially identifying survey questions to determine whether any respondents or their children could be identified by any combination of such variables. The final step was to recode any potentially identifying measures.

We excluded the information used to identify and locate respondents for the survey from the SAS data sets. This information included respondent's name, address, and telephone number, and child's name. We destroyed these data after the SCHIP evaluation was completed, and we

assigned each respondent a unique, completely random, identification number. We also removed the "day" components of the date measures. Because dates, particularly dates of birth, may identify respondents or their children, we included only the month and year components of the date measures. In addition, we did not include the "other specify" responses,[16] because answers provided often contained specific information about a sample member.

Having removed all individual identifiers, we then explored the potential for sample members to be identified from a plausible combination of variables. As our first step in this process, we placed each sample member into one of 240 possible cells defined by four key, identifiable demographic characteristics of the children: their state of residence; gender (female, male); age at the time of the survey (which we classified as 4 or younger, 5 to 12, 13 or older); and race/ethnicity (which we classified as Latino, non-Latino and black, non-Latino and white, and non-Latino and neither black nor white).[17] Within each cell, we counted at least five members, suggesting that children in the study sample could not be easily identified solely from these characteristics. As our next step, we examined the distributions of responses to each potentially identifying survey question within each of the 240 cells. If this distribution led to fewer than five children being identified within a cell, the question was flagged as having potentially identifying information.

Table 6 lists the survey questions that we determined to have potentially identifying information and the approach taken to address each of them. For several questions, the responses provided were highly identifiable, and the only approach available was to entirely omit the

[16] The survey instrument contained several questions that gave respondents the opportunity to provide responses that were not among the categories listed in the survey. We have coded these responses, and the codes are included in the SAS data sets.

[17] In the Medicaid sample, sample members were grouped into 48 cells, because there were only two states.

TABLE 6

MASKING TECHNIQUES IMPLEMENTED TO MAINTAIN CONFIDENTIALITY

| Masking Technique | Questions Affected | Question Number |
|---|---|---|
| Omitted Question | Compared to 12 months ago, would you say (child)'s health is now… | 4.2 |
| | Does (child) have any impairment or health problem that requires him/her to use special equipment such as a brace, a wheelchair, or a hearing aid? Do not include ordinary eye glasses or corrective shoes. | 4.3 |
| | Does (child) have an impairment or health problem that limits his/her ability to crawl, walk, run, or play? | 4.4 |
| | Is this an impairment or health problem that has lasted or is expected to last 12 months or longer? | 4.5 |
| | Because of this impairment or health problem, does (child) need other people to help him/her with personal care needs, such as bathing, dressing, eating, or getting around? | 4.6 |
| | How old was (child) when he/she had his/her first episode of asthma? | 4.10 |
| | Has she/he taken medication or required injections for at least 3 months? | 4.12 |
| | How old was (child) when a doctor or other health professional first said that he/she had a mental health condition or behavioral problem? | 4.14 |
| | Does (child) take medication or require injections for a mental health condition or behavioral problem? | 4.15 |
| | During (timeframe1), did (child) see or talk to a mental health professional, such as a psychiatrist, psychologist, psychiatric nurse, or clinical social worker? | 6.14 |
| | How many times did (child) see or talk to a mental health professional, such as a psychiatrist, psychologist, psychiatric nurse, or clinical social worker? | 6.14.1 |
| | In what country were you born? (lper1) (Included as part of 7.4.1.9) | 7.4.1.8 |
| | In what country was he/she born? (lper2) (Included as part of 7.4.1.9) | 7.4.5.8 |
| | What is his/her relationship to (child)? (lper3) | 7.4.6.2 |
| | Is he/she (child)'s legal parent or guardian? (lper3) | 7.4.6.3 |
| | What is this person's gender? (lper3) | 7.4.6.5 |
| | What was this person's age at his/her last birthday? (lper3) | 7.4.6.6 |
| | What is the highest grade or year of schooling he/she has completed? (lper3) | 7.4.6.7 |
| | In what country was he/she born? (lper3) | 7.5 |
| | Is he/she a citizen of the United States? (lper3) | 7.6 |
| | Would you say your total household income from all sources was less than $25,000 or more than $25,000? (Included as part of 7.93) | 7.99 |
| | Would you say it was…(less than $25,000) (Included as part of 7.93) | 7.100 |
| | Would you say it was…(more than $25,000) (Included as part of 7.93) | 7.101 |
| | In the past two years, has anybody in the household received any benefits from TANF which used to be called AFDC? (Included as part of TANFORFS) | 7.102.1 |
| | Including yourself, how many people in the household received Food Stamps in the past 2 years? (Included as part of TANFORFS) | 7.102.2 |

TABLE 6 *(continued)*

| Masking Technique | Questions Affected | Question Number |
|---|---|---|
| Omitted Question *(continued)* | Would you say your household spending on health care was… (Included as part of 7.103) | 7.105 |
| | Do you consider him/her to be of Hispanic or Latino origin? (Included as part of ETH_RACE) | 7.109 |
| | What Hispanic or Latino group do you consider him/her to belong to? | 7.110 |
| | Which of the following best describes his/her racial background? (Included as part of ETH_RACE) | 7.111 |
| Set Values to "99" in Certain Cells | Has a doctor or other healthcare professional ever said that (child) had asthma? | 4.9 |
| | Does (child) take medication or require injections for his/her asthma? | 4.10.1 |
| | Does (child) take medication or require injections for any other physical condition? | 4.11 |
| | Has a doctor or other healthcare professional ever said that (child) had a mental health condition or behavioral problem? | 4.13 |
| | Has a mental health condition or behavioral problem limited (child) in his/her ability to do regular school work or to participate in the usual kind of activities done by most children his/her age? | 4.16 |
| | Are you (child)'s biological, step, adoptive parent or legal guardian? (lper1) | 7.4.1.2 |
| | What is his/her relationship to (child)? (lper2) | 7.4.5.2 |
| | Is English the main language spoken in this household? | 7.120 |
| Collapsed Response Categories | How old was (CHILD) when was enrolled in (SCHIP/MEDICAID)? | 2.12 |
| | In general, would you say (CHILD)'s health is… | 4.1 |
| | Are you (CHILD)'s biological, step, adoptive parent or legal guardian? | 7.4.1.2 |
| | What was your age at your last birthday? | 7.4.1.6 |
| | What is the highest grade or year of schooling you have completed? | 7.4.1.7 |
| | What is his/her relationship to (CHILD)? | 7.4.5.2 |
| | What was his/her age at his/her last birthday? | 7.4.5.6 |
| | What is the highest grade or year of schooling he/she has completed? | 7.4.5.7 |
| | In the past 12 months, what was the total household income from jobs and all other sources of income? | 7.93 |
| | During the past 12 months, about how much did your household spend on healthcare, that is money you or someone else in the household paid for doctors' visits, hospital stays, or prescription drugs? | 7.103 |
| | Is English the main language spoken in this household? | 7.120 |
| Combined into Constructed Variable | In the past two years, has anybody in the household received any benefits from TANF which used to be called AFDC? *and*<br>Including yourself, how many people in the household received Food Stamps in the past 2 years? (TANFORFS) | 7.102.1<br>7.102.2 |
| | Do you consider him/her to be of Hispanic or Latino origin? *and*<br>Which of the following best describes his/her racial background? (ETH_RACE) | 7.109<br>7.111 |

Source: 2002 congressionally mandated survey of SCHIP enrollees and disenrollees in 10 states.

corresponding variable from the SAS data sets.[18] For another set of questions, the responses of sample members in certain cells were identifiable due to small sample sizes; we addressed this problem by setting every sample member within the cell to a response value of "99" for the corresponding variable.[19] Alternatively, for several other questions, we were able to retain more information by collapsing responses into fewer categories rather than omitting responses of certain sample members. Finally, for a few questions, we created a constructed variable from multiple survey questions to replace the original questions that had potentially identifiable information. Specifically, in place of questions Q7_102_1 and Q7_102_2, we created the variable TANFORFS—a measure that indicates whether anybody in the household received TANF or Food Stamps in the past two years. In place of questions Q7_109 and Q7_111, we created the variable ETH_RACE—a combined race/ethnicity measure.

### 3. Back-Coded Variables

As mentioned earlier, the survey instrument contained several questions that gave respondents the opportunity to provide responses that were not among the categories listed in the survey. Though we did not include the verbatim responses in the SAS data sets for confidentiality reasons, we examined the "other specify" responses and back-coded them into either the categories listed in the survey or additional categories. Among those who provided an "other specify" response, we created a new variable to capture the categorizations of these responses. Because these responses often were very specific, newly created categories that contained fewer than 20 respondents (fewer

[18] In addition to these specific questions, we also excluded the questions from Section 1 (Introduction) and Section 8 (Questions About Telephone Coverage) from the SAS data sets. The questions in these sections were intended solely to determine eligibility and to gather identifying and locating information about respondents.

[19] If any of the response categories in one of the permutations had fewer than five respondents, we set the value of the response for all of the respondents within that permutation to 99.

than 10 respondents for the analysis of recent and established Medicaid enrollees) or less than one percent of the responses were coded as 99. The survey questions that we back-coded are:

- Q2_1_15BC—Ever heard or received information about (SCHIP/MEDICAID) at any other place or from any other person?
- Q2_10BC—What were the reasons (SCHIP/MEDICAID) rejected (CHILD)'s application?
- Q2_14BC—What was the main reason (CHILD) was enrolled in the program?
- Q2_18BC—Who gave the form to you or where did you pick it up?
- Q2_39BC—What were the reasons (CHILD)'s reapplication was rejected?
- Q3_26BC—What was the main reason this (SCHIP) coverage ended?
- Q3_34BC—What was the main reason (CHILD) was without any health insurance during this period?
- Q3_44BC—What was the main reason (CHILD)'s coverage ended?
- Q3_63_1BC—What was the main reason (CHILD) was/has been without any health insurance during this period?

**4. Missing Values**

In SAS, different missing values can be used to denote the particular reasons why a variable is not available. We used nine missing values in the SAS data sets:

1. .D, used primarily among survey variables, indicates that the respondent said, "I don't know."
2. .E, used among both survey variables and enrollment and eligibility variables, indicates that the question was erroneously skipped or that program data were missing for the sample member.
3. .L, used primarily among survey variables, indicates that the question was a logical skip.
4. .M, used exclusively among constructed variables, indicates that the variable was not constructed.
5. .N, used exclusively among enrollment and eligibility variables, indicates that program data were not provided for the given month.

6. .P, used exclusively among survey variables, indicates that the question was not asked because the respondent did not finish completing the survey.
7. .R, used primarily among survey variables, indicates that the respondent refused to provide an answer.
8. .S, used exclusively among survey variables, indicates that the question was skipped because the respondent was switched to the shortened version of the survey.
9. .U, used exclusively among back-coded variables, indicates that the response could not be back-coded.

## 5. Using the SAS Format Library

All variables in the SAS data sets included in the SCHIP Evaluation Public Use Files are numeric. Many of the survey variables are binary—that is, they have values of either 1 or 2, which indicate yes and no, respectively. Other variables in these data sets are continuous and have a wide range of values. The rest of the variables are categorical. Specifically, there are two types of categorical variables: (1) *nominal variables,* meaning that they are simply grouped into classes; and (2) *ordinal variables,* meaning that they are ordered on some continuum. An example of a nominal variable is Q2_17, the source of the SCHIP application form. This variable has values of 1 (in the mail), 2 (from someone), and 3 (from a website). An example of an ordinal variable is Q5_22, the amount of time it took to get to the usual place of care. This variable has values of 1 (less than 15 minutes), 2 (15 minutes but less than 30 minutes), 3 (30 minutes but less than 45 minutes), 4 (45 minutes but less than 1 hour), 5 (1 hour but less than 2 hours), and 6 (2 or more hours).

Given the existence of many different categorical variables in the SAS data sets, we created a SAS format library, *Formats.SAS7BCAT*, to allow the user to see the substantive meanings associated with the different values of these variables when they are used in any SAS BASE procedure. To reference the SAS format library, the following statement must be included in the SAS program: LIBNAME LIBRARY '{specific folder location of the SAS format library}.' This

library must be referenced when any of these data sets are used. A user who does not wish to see the substantive meanings associated with the different values of the categorical variables can omit the preceding LIBNAME LIBRARY statement but must include a OPTIONS NOFMTERR before any of the SAS data sets are used.

## 6. Computing Correct Standard Errors

Because of the complex survey design, special statistical software must be used to obtain the correct standard errors on statistics derived from the survey data (such as mean, frequency distributions, or ratios). The software uses three key variables that capture the survey design parameters. The first is the appropriate sampling weight variable for the sample member, discussed previously in Section B.6. The second is a stratification variable, *STRATA*, and the third is a variable to denote the first-stage sampling unit, *PSU*.

While several software packages are available to compute the standard errors with complex sample designs, we strongly recommend using the SUDAAN programming language when using the SAS data sets of the SCHIP Evaluation Public Use Files. Specifically, when using SUDAAN, three programming statements should be included in any statistical procedures:

1. On the PROC line of the given SUDAAN procedure, include the *DESIGN=WR* option, which specifies a with-replacement design.
2. Add a NEST statement to indicate the stratification variable and the first-stage sampling unit. The specific statement is *NEST STRATA PSU / MISSUNIT*.
3. Add a WEIGHT statement to indicate the sampling weight being used. For analyses involving recent and established enrollees, the statement is *WEIGHT ENR_WGHT*; for analyses involving disenrollees, the statement is either *WEIGHT DIS_WGHT_POP* or *WEIGHT DIS_WGHT*, depending on which weight variable is appropriate (see discussion in Section B.6).

The SUDAAN procedures specified above are based on classic statistical methods in which a nonlinear statistic (such as a regression coefficient) can be approximated by a Taylor series

linearization of the components within the statistic. The accuracy of the approximation depends on the sample size and the complexity of the statistic. For most commonly used nonlinear statistics (such as ratios, means, proportions, and regression coefficients), the linearized form has reliable statistical properties under large sample approximations.

## REFERENCES

Wooldridge, Judith, Genevieve Kenney, Christopher Trenholm, Lisa Dubay, Ian Hill, Myoung Kim, Lorenzo Moreno, Anna Sommers, and Stephen Zuckerman. "Congressionally Mandated Evaluation of the State Children's Health Insurance Program: Final Report to Congress." Report submitted to the U.S. Department of Health and Human Services, Assistant Secretary for Planning and Evaluation. Princeton, NJ: Mathematica Policy Research, Inc., October 2005.

The contents of this document were developed by Timothy Novak and Nora Paxton of Mathematica Policy Research, Inc. and Anna S. Sommers and Justin Yee of The Urban Institute under contract to the Office of the Assistant Secretary for Planning and Evaluation, Contract Number 100-03-0017 and Task Order Number HHSP23300002T.

# APPENDIX A

# SAMPLE SELECTION

The SCHIP evaluation used a dual-frame sample design that combined a clustered sample with an unclustered sample. The clustered sample was interviewed by telephone but had an in-person field followup, while the unclustered sample was interviewed by telephone only. This appendix explains the methods used to select these samples.

## 1. Selecting the Clustered Sample

For the clustered design, which included in-person tracking and locating, the first step in sample selection for each program was to define primary sampling units (PSUs) for each state. These PSUs were geographic areas that met a specified minimum number of total enrollees and recent disenrollees. The areas were defined based on one or more counties and, in some highly populated areas (such as Denver, Colorado, and Miami, Florida), zip code areas. The same set of PSUs was used for all sample rounds for both the Medicaid and SCHIP samples.

A composite size measure strategy was used to select sample PSUs, as well as households and children for interview.[20] As the first step, we defined a composite size measure, $S(h,i,j)$, for each household $j$ from PSU $i$ in state $h$ ($h$ = 1,2, …10) containing one or more eligible children from the three SCHIP domains and (where appropriate) the three Medicaid enrollment domains.

Let $C_d(h,i,j)$ be the total number of domain $d$ children in household $j$ from PSU $i$ of state $h$. Let $f_d(h)$ be the desired sampling rate for domain $d$ members in state $h$, or:

$$(1) \quad f_d(h) = \frac{m_d(h)}{C_d(h,+,+)} \quad ,$$

[20] See Folsom et al. (1987) for a discussion of composite size measures.

where $m_d(h)$ is the desired sample from domain $d$ ($d$ = 1, 2, …, $D$) in state $h$ and $C_d(h,+,+)$ is the total number of domain $d$ members in state $h$.[21,22] The composite size measure $S(h,i,j)$ for household $j$ from PSU $i$ of state $h$ is then defined as:

$$(2) \quad S(h,i,j) = \sum_{d=1}^{D} f_d(\mathrm{h})\, C_d(h,i,j).$$

This composite size measure was summed over all households in PSU $i$ and state $h$ to produce the size measure $S(h,i,+)$ for PSU $i$ in state $h$, which was used in selecting the first-stage sample of PSUs.[23]

In most states, 30 PSUs were selected, with probability proportional to this composite size measure and with minimal replacement, using Chromy's (1979) procedure.[24] In selecting the sample PSUs from the frame of $N_1(h)$ PSUs in state $h$, Chromy's procedure partitioned each state's $N_1(h)$ total PSUs into sampling zones of approximately equal size, based on the composite size measure $S(h,i,+)$. Exactly one PSU was selected from each zone. The zones were defined so that all pairs of PSUs had a chance of appearing together in the sample (a requirement for unbiased estimation of sampling variances).[25] Using a controlled ordering of the PSUs, this "zoned sequential selection" made possible an implicit stratification of PSUs that

[21] The domains are made up of the three SCHIP enrollment groups and, for the subset of two states, the three Medicaid enrollee groups. Thus, $D$ = 3 for eight states and $D$ = 6 for two states.

[22] The "+" sign denotes summation over all households and PSUs in state $h$.

[23] The "+" sign in $S(h,i,+)$ denotes summation over all households $j$ within PSU $i$.

[24] In California, 60 PSUs were selected; in New Jersey, no PSUs were selected.

[25] This requirement was accomplished by selecting a random starting point and treating the frame as a circular list.

ensured that sample PSUs were representative of selected variables of interest. Two of these variables were the urbanicity and the geographic location of the PSU, which ensured selection of both urban and rural PSUs and the distribution of the sample across the state.

For each domain within a state, we used a composite size measure to ensure that the desired sample sizes were achieved. The composite size measure for PSU *i* in state *h* was defined as:

$$(3) \quad S(h,i,+) = \sum_{j} S(h,i,j) = \sum_{d=1}^{D} \sum_{j} f_d(h)\, C_d(h,i,j)\,,$$

where $C_d(h,i,j)$ is the number of children in domain *d* of household *j* of PSU *i* in state *h*, and $f_d(h)$ is the desired overall sampling rate for domain *d* in state *h*. Before selection, we again used a controlled ordering procedure, this time for the households within each PSU. Some of the variables for ordering were the sampling domains and, when available, the race of the children in the household.

For each selection of the *i*th PSU from the *h*th state, $n_2(h)$ households were selected with probability proportional to their composite size.[26] When a household contained more than one enrollee type, we randomly determined the enrollee type to interview, using differential probabilities based on the desired state *h* sampling rates $f_d(h)$ for domain *d*. If more than one child was present in the sampled household for the enrollee domain selected, we randomly selected one child from the selected enrollee domain to be interviewed. Using the composite size measure for each household enabled us to oversample households with more than one eligible

[26] For some sample rounds for some states, a household was selected with certainty if the number of enrollees of a specific type (most often, recent disenrollees) was large enough to produce a composite size measure above a threshold.

child while ensuring that the selection probabilities were equal within enrollment domains, regardless of household size.

In states for which we included a second (or third) sampling round, we followed procedures designed to avoid selection of households already chosen in a previous sample round and to account for enrollees who were in the sampling frame for a prior round. By definition, recent enrollees and recent disenrollees were unique populations in each sample round. However, established enrollees could have had their status across more than one survey round (for example, in both January 2002 and March 2002). To maintain nearly equal sampling rates across the rounds, the established enrollees in round two and (as needed) in round three were divided into separate sampling strata, depending on the number of rounds for which they had that status. The sample for the later rounds was then allocated to each stratum accounting for the sampling rate in the prior round(s) of established enrollees who appeared in both the later round and an earlier round.

The composite size measure was also adjusted to ensure that households were not selected multiple times across sample rounds. We made the adjustment by creating a household-level weight for each sample round after the first round that reflected the probability of *not* being selected in the previous round. The probability was constructed as follows:

- Households that were sampled for a prior round received a score of zero.
- Households that were on the frame(s) in prior round(s) were assigned a probability equal to the likelihood of not being selected in those prior round(s).
- Households not on the frames for the prior round(s) received a probability score of 1.

The modified composite size measure defined for each household was then the product of the probability score and the round-specific composite size measure for the household. Households were then selected according to the procedures outlined above, but with this

modified composite size measure. This approach prevented the same household from being selected more than once, while ensuring nearly equal selection probabilities across sample rounds.

## 2. Selecting the Unclustered Sample

For the unclustered, telephone-only design, we first sampled households. If the household included children in two or more domains, we then selected the domain for which a child would be selected and, finally, selected the child within the domain. Among households with more than one child eligible for interview, one child was randomly selected for interview. Before sample selection, the households were sorted by the combinations of enrollment domain(s) to which their eligible children belonged (for example, recent enrollee only, recent enrollee and established enrollee, recent enrollee and recent disenrollee, established enrollee only). Then, within each combination, the households were further sorted by their race/ethnicity, metropolitan status, and geographic area. Through this process, we created an implicit stratification of the households from which to draw the sample for each domain and state.

A composite size measure was defined for each household that reflected the number of eligible children in the household (including Medicaid enrollees for the two states where they were to be sampled for the Medicaid analysis), as well as the desired, overall selection probabilities for the unclustered design. Households were selected with probability proportional to their composite size measures. For sampled households with more than one child eligible for interview, we used the desired subsampling rates for the enrollee domains in randomly sampling one child for interview. This composite size measure approach ensured that we achieved nearly equal selection probabilities within each state for each enrollee domain, regardless of the household's size. Similar to the approach used for the clustered sample, the selection process for the unclustered sample prevented selection of the same household in multiple rounds.

To account for individuals and households already selected for the clustered sample, we divided the sampling frame for the unclustered sample into two strata: (1) individuals in the geographic areas included in the sampled PSUs for the clustered sample, and (2) individuals in the rest of the state. We allocated the unclustered sample across these two strata. In the stratum of individuals in the PSUs of the clustered sample, we had to account both for households and individuals selected in any prior rounds and for the households and individuals selected in the clustered sample (for the current round and for any prior rounds). In the stratum of individuals not in the PSUs of the clustered sample, we had to account only for households and individuals selected in any prior rounds. In most states and most rounds of data collection, adequate numbers of households and individuals were available to enable us to select separate unclustered and clustered samples. In North Carolina, the number of recent disenrollees in the March extract was very small. All recent disenrollees in the North Carolina PSUs were selected for the sample. Respondents among those recent disenrollees were included as part of both the clustered sample and the unclustered sample.

## REFERENCES

Chromy, James R. "Sequential Sample Selection Methods." *1979 American Statistical Association Proceedings of the Section on Survey Research Methods*. Washington, DC: American Statistical Association, 1979, pp. 401-406.

Folsom, R.E., F. Potter, and S.R. Williams. "Notes on a Composite Measure for Self-Weighting Samples in Multiple Domains." *American Statistical Association, Proceedings of the American Statistical Association, Section on Survey Research Methods*. Washington, DC: American Statistical Association, 1987, pp. 792-796.

# APPENDIX B

# SAMPLING WEIGHTS

This appendix describes the computations of the sampling weights. The initial weights were computed in two stages: (1) the round-specific, design-specific weights; and (2) the combined-round, design-specific weights (the base weights). We then used the base weights to compute nonresponse adjustments for each design and each domain for each state. Finally, the nonresponse-adjusted base weights for each design were combined and poststratified to form the final analysis weights.

## 1. Initial (Round-Specific, Design-Specific) Weights

For California and Texas (which were sampled in a single round) and for the first sample round for the other states, initial weights for the clustered samples were computed from the inverse of the product of the selection probability for the (1) cluster, (2) household within the cluster, (3) domain type, and (4) child.

If the household included two or more children, the children could have been in (1) the same domain (for example, two children in a household both might have been recent enrollees); or (2) two or more domains (for example, one child might have been a recent enrollee and a second child might have been an established enrollee).[27] For the unclustered samples, the initial weights were computed from the inverse of the product of the selection probability for the (1) household, (2) domain type, and (3) child. For the second and third sample rounds, the initial weights also included a factor representing the probability that a household had not been selected in the prior round(s).

Because we expected variation in the eligibility and response rates in each state, we selected a reserve sample for use in ensuring an adequate number of complete interviews. The initial

[27] In California and North Carolina, some children were eligible for the samples as new enrollees in SCHIP and recent disenrollees in Medicaid. Children with this type of concurrent valid classification were accounted for in the sampling design.

weights also included a subsampling rate to reflect the proportion of the full sample (the primary and reserve samples) that was used in the survey. In some states, subsamples of nontelephone households in clustered samples were assigned to field staff for in-person locating. The initial weights accounted for this subsampling. Basically, the initial weight for each round was the inverse of the product of three to six sampling probabilities and subsampling rates. These initial weights were then poststratified by sample domain (recent enrollee, established enrollee, and recent disenrollee) to the enrollment population size in the file extract.

**2. Base (Combined-Round, Design-Specific) Weights**

For the eight states with two or three sample rounds, the initial weights summed to the enrollment population at the time of the extract. For the recent enrollees and recent disenrollees, the enrollment populations for extracts were mutually exclusive (that is, the children could not be classified as recent enrollees in both the January and March file extracts). Similarly, the same children could not be recent disenrollees in both the January and March file extracts. To compute design-specific weights for these domains that spanned all sample rounds, we combined the sample weights from the two (or three) sample rounds by multiplying the initial weight by a compositing factor based on the proportion of the sample from all sampling rounds that was used in a specific sample round. That is, if the January sample round included 180 recent enrollees and the March sample round contained 120 recent enrollees, then the weights for recent enrollees from the January sample round were multiplied by 0.60 (180/300), and the weights for recent enrollees from the March sample round were multiplied by 0.40 (120/300). After the combined-round weight was computed, we poststratified the weight to the average enrollment in that domain across the sample rounds to form the base weight.

For the established enrollees, a child in the January extract file might or might not still be an established enrollee in the March extract file. Therefore, for the six states with two sample

rounds, we had to account for the enrollment populations, which depended on the extract file in which the child was classified as an established enrollee. In particular, a child could be classified as an established enrollee (1) in January but not in March, (2) in both January and March, or (3) in March but not in January.

The round-specific weights based on the January extract provided unbiased estimates of the established enrollees who were in the January extract file but not in the March one, and of established enrollees who were in both months' extract files. The round-specific weights based on the March extract provided unbiased estimates of the established enrollees who were in both the January and March extract files, and of those who were in the March extract file but not in the January extract file.

To combine these round-specific weights, we tabulated the counts in each extract to determine the exact enrollment counts for each of the three populations (established enrollees in January only, in both January and in March, and in March only). We then poststratified the weighted counts for each sample component to the exact enrollment counts. We scaled the initial weights for the cases in both the January extract and the March extract, using the proportion of the sample in the respective January or March samples. (The initial weights for cases in only the January extract and for those in only the March extract were not changed.) These combined-round initial weights summed to the number of children who were established enrollees in either or both the January and March extract files. To compute the base weights for the established enrollees, these weights were then rescaled to the average of the enrollment in the two extracts to achieve comparability with the other states.

The base weights were computed for each design (the clustered and unclustered sample designs) for the eight states with two or three sample rounds. For Colorado and Louisiana, three sample rounds (and, therefore, three extract files) were used. A child could be an established

enrollee (1) in January, March, and May; (2) in January only; (3) in January and March but not in May; (4) in March only; (5) in March and May but not in January; or (6) in May only.[28] We used procedures analogous to those used for the states with only two sample rounds.

### 3. Nonresponse Adjustments

Nonresponse occurs in all surveys. The standard procedure to account for nonresponse is to adjust the sampling weights, thereby minimizing the potential for nonresponse bias. Weights for respondents who are similar to sample members who do not respond are adjusted to reduce the potential for nonresponse bias. We initially conducted an analysis to identify the factors that might have been related to nonresponse. Because the extract files from the states contained limited data (age and, sometimes, race) for identifying similarities between respondents and nonrespondents, we accessed county-level data from the Area Resource File (ARF) to supplement the state-provided data. The ARF contains county-level counts and other data compiled from the Census Bureau, the Bureau of Economic Analysis, the U.S. Department of Agriculture, the National Center for Health Statistics, and other sources. The data obtained from the ARF included:

- Rural/urban continuum code (10-level code)
- Population percentage for white, black/African American, Asian, American Indian/ Alaskan Native, and other
- Percentage Hispanic or Latino population
- Percentage of people 25 or older with less than nine years of school
- Percentage of people 25 or older with a high school diploma or more
- Percentage of people 25 or older with four or more years of college

[28] Children had to be enrolled for five consecutive months. Thus, by definition, a child could not be an established enrollee in January and in May but not in March.

- Median family income
- Median household income
- Percentage of families below the poverty level
- Percentage of people below the poverty level
- Percentage of families with a female head
- Percentage of people in poverty
- Percentage of people ages 0 to 17 in poverty
- Percentage of related children ages 5 to 17 in poverty

These variables were selected as measures of racial and ethnic composition and as measures related to the extent of poverty in the counties in which the sample members resided. We viewed these variables as proxy measures for unobservable factors associated with response, although the variables themselves did not imply any direct relationship with response patterns.

For the response models, we formed categories based on the characteristics of each sample to ensure that there were adequate sample counts in each category and that the categories were somewhat logical breaks in the distribution of continuous variables. We used stepwise logistic modeling to identify the variables (including both the categorized variables and the state-provided data on the child's age and race) that best explained the response pattern for each sample. Because the states and the enrollment population differed substantially, no single set of variables was consistently the best one to explain a response pattern. In general, however, response was associated with the degree of urbanicity, with lower response in some urban areas and higher response in rural areas. Other community factors that helped explain the response pattern were ethnicity and race and the percentage of children in poverty.

These response propensity models were developed separately for each domain, each sample type (clustered and unclustered), and each state. Separate models were also developed for the

Medicaid samples, again for each domain, sample type (clustered and unclustered), and state. More than 80 response propensity models were developed, with 69 developed for the SCHIP samples and 12 developed for the Medicaid samples.

## 4. Final Analysis Weights

The clustered and unclustered samples were designed so that children from telephone households would have nearly equal probabilities of selection for either design. Because of the possible similarity of responses among sample members in the same cluster (that is, the possibility of a positive intracluster correlation), the sampling variance of estimates computed using the clustered sample was expected to be somewhat larger than the sampling variance of the same estimates computed using the unclustered sample. To develop the combined-design, nonresponse-adjusted sample weight, we used the ratio of the sampling variances computed for selected outcome-related variables as a factor for computing a composite weight factor for the children in telephone households.

Specifically, to compute a survey estimate, *Est(Y),* combined across the two samples, separate estimates can be computed for each sample and combined using the equation:

(1) $Est(Y) = \lambda\ Y(Clustered) + (1 - \lambda)\ Y(Unclustered),$

where *Y(Clustered)* is the survey estimate from the clustered sample, *Y(Unclustered)* is the survey estimate from the unclustered sample, and $\lambda$ (lambda) is an arbitrary constant between 0 and 1. For the sampling variance, *V(Y),* the estimate is computed using the equation:

(2) $V(Y) = \lambda^2\ V(Y(Clustered)) + (1 - \lambda)^2\ V(Y(Unclustered)),$

where *V(Y(Clustered))* is the sampling variance for the estimate from the clustered sample and *V(Y(Unclustered))* is the sampling variance for the estimate from the unclustered sample. Any

value of lambda between 0 and 1 will result in an unbiased estimate of the survey estimate, but not necessarily in an estimate with the minimum sampling variance. A lambda value producing a sampling variance at its minimum value results in the shortest confidence interval and, by implication, the most accurate point estimate.

A value of lambda can be computed in an optimal (minimum variance) sense as:

(3) $\lambda = V(Y(Unclustered)) / [V(Y(Clustered)) + V(Y(Unclustered))]$.

In this case, the minimum variance is:

(4) $V(Y) = [V(Y(Clustered)) * V(Y(Unclustered))] / [V(Y(Clustered)) + V(Y(Unclustered))]$.

To compute a combined-sample estimate with minimum variance, survey estimates are derived by first computing the estimates for each sample component, computing a value of lambda for each pair of estimates, and then combining the point and variance estimates. Although producing the minimum variance estimates, the process is computer-intensive and results in some inconsistencies among estimates for percentages and proportions because of differing values among levels of a categorical variable.

For this study, we identified a pool of variables of interest for each domain and computed variance estimates for the clustered and unclustered samples. We used these sampling variances to compute values of lambda and used the median values of the lambdas to develop a single value for computing the combined-sample weights. The lambda values differed for each domain and state but were generally around 0.45, which indicated slightly larger sampling variances in the clustered sample (as expected). The combined weight for each sample member in the clustered sample was computed as:

(5) $WT(Combined) = \lambda\, WT(Clustered\ Nonresponse\text{-}Adjusted\ Weight)$,

and for sample members in the unclustered sample, by:

(6) $WT(Combined) = (1 - \lambda)\ WT(Unclustered\ Nonresponse\text{-}Adjusted\ Weight)$.

Children from nontelephone households were eligible for interview only when sampled for the clustered design, so their nonresponse-adjusted weight was used as their combined sample weight. This combined weight was then poststratified again to the domain-specific monthly enrollment count for each state.

# APPENDIX C

# ENROLLMENT DATA

This appendix describes the SCHIP and Medicaid enrollment data collected from the states for the SCHIP evaluation. First, it explains the data acquisition process. It then discusses the enrollment data. Finally, it describes the SCHIP and Medicaid eligibility codes that were created for the SCHIP evaluation.

## 1. Data Acquisition

Acquisition and use of these data required frequent, detailed interactions with state program staff. We first contacted senior state staff to introduce ourselves and to explain the purpose of the study, why and how the state was selected for the study, and the need for a memorandum of understanding detailing the data needs and confidentiality requirements and documents. Subsequent discussions with program staff focused on data elements that would support sampling criteria and analytic criteria, the source of program data, the format of the data available for our use, the timeliness of the data, and periodic data extracts and delivery.

Timeliness of the data was an important issue to capture the populations of recent enrollees and disenrollees. Time-related issues included (1) the time required by state and local agencies for processing initial applications and redeterminations, and (2) the use of retroactive or prospective enrollment (enrollment dates set to the application date or a date before the application date). We were concerned that delays in updating the eligibility histories could affect the timely construction of sampling frames and sampling selection. In our discussions with state program staff, we requested delivery of data by the state within two weeks of the specified data extract cutoff date. With few exceptions, the states delivered their data on time.

## 2. Enrollment Measures

We requested SCHIP enrollment histories for all children included in our survey samples of recent and established SCHIP enrollees and recent disenrollees from SCHIP. We also requested

Medicaid enrollment history data for the 10 SCHIP samples and for the samples in the two states in which we conducted a survey of Medicaid enrollees (California and North Carolina). Medicaid data were used to supplement the analysis of SCHIP enrollment and reenrollment.

The period for which we obtained enrollment records varied across states. For all 10 states, we obtained SCHIP enrollment history data from the month in which the program began in each state through December 2002. (We selected this cutoff date to coincide with the expected end of the survey field period for all states.) SCHIP enrollment histories were available for 50 to 60 months for nine states, and for 32 months for Texas. In contrast, Medicaid enrollment history data were available for only seven states. For the SCHIP samples, Medicaid data were available from the beginning of the SCHIP program in five states. The exceptions were California and Florida, for which enrollment history data began in November 2000 and in January 2001, respectively. In addition, Medicaid enrollment histories for the samples of enrollees in Medicaid were available beginning in November 2000 in California, and beginning in October 1998 in North Carolina. For all states that provided Medicaid data, these histories were available through December 2002. Medicaid enrollment histories were therefore available for 26 to 60 months for seven states.

Because enrollment files vary in their structure and content across states, we developed uniform files. The process of creating these files included data quality and consistency checks. In several instances, we contacted the states to clarify anomalies observed in specific data elements. From the state enrollment files, we created one record for each child included in the SCHIP and Medicaid survey samples and periods noted above for the 10 states. On each record, we included variables to indicate enrollment in Medicaid, a separate SCHIP program, or a Medicaid-expansion SCHIP program during each of the 60 months beginning in January 1998 (month 1) and ending in December 2002 (month 60). The Medicaid enrollment variables,

MED1—MED60, are unavailable in Colorado, New York, and Texas. The separate SCHIP enrollment variables, SCHIP1—SCHIP60, are unavailable in Louisiana and Missouri, the states with Medicaid-expansion programs. The Medicaid-expansion SCHIP enrollment variables, MSCHIP1—MSCHIP, are available only in Louisiana and Missouri and the two states with combination programs, Illinois and New Jersey.[29]

### 3. Eligibility Codes

We classified SCHIP state eligibility codes into broad categories defined by family income and, in one instance, by the age of the child (Florida). For the Medicaid codes, we classified the state eligibility codes into the four broad eligibility groups of (1) cash assistance, (2) medically needy, (3) poverty related, and (4) other. These codes correspond to the Maintenance Assistance Status codes used by the Centers for Medicare & Medicaid Services to report eligibility in the Medicaid Statistical Information System. To keep the classification manageable, we did not create subgroups defined by the Basis of Eligibility codes. The definitions of the SCHIP and Medicaid eligibility codes we used in the analysis are summarized in Tables C.1 and C.2, respectively.

[29] We also constructed variables that indicate enrollment in either a separate SCHIP program or a Medicaid-expansion program, CHIPMO1 – CHIPMO60.

TABLE C.1

CROSSWALK OF STATE ELIGIBILITY CODES INTO UNIFORM CODES, BY STATE AND PROGRAM (SCHIP)

| | State Eligibility Code Description | | | | | | |
|---|---|---|---|---|---|---|---|
| | State Eligibility Code | Program Name | Age Requirement | Income Requirement | MPR Eligibility Code | Unique MPR Eligibility Code | MPR Eligibility Code Description |
| CA[a] | | Healthy Families | 0 to 18 years | < 150% FPL | 1 | 101 | < 250% FPL |
| | | Healthy Families | 0 to 18 years | 151 to 250% FPL | 1 | 101 | < 250% FPL |
| CO[b] 01/01 to present | N | CHP+ | 0 to 18 years[c] | ≤ 40% FPL | 1 | 201 | ≤ 100% FPL |
| | A | CHP+ | 0 to 18 years[c] | 40 to 62% FPL | 1 | 201 | ≤ 100% FPL |
| | B | CHP+ | 0 to 18 years[c] | 63 to 81% FPL | 1 | 201 | ≤ 100% FPL |
| | C | CHP+ | 0 to 18 years[c] | 82 to 100% FPL | 1 | 201 | ≤ 100% FPL |
| | D | CHP+ | 6 to 18 years | 101 to 117% FPL | 2 | 202 | 101 to 150% FPL |
| | E | CHP+ | 6 to 18 years | 118 to 133% FPL | 2 | 202 | 101 to 150% FPL |
| | F- | CHP+ | 0 to 18 years | 134 to 150% FPL | 2 | 202 | 101 to 150% FPL |
| | F+ | CHP+ | 0 to 18 years | 151 to 159% FPL | 3 | 203 | 151 to 185% FPL |
| | G- | CHP+ | 0 to 18 years | 160 to 170% FPL | 3 | 203 | 151 to 185% FPL |
| | G+ | CHP+ | 0 to 18 years | 171 to 185% FPL | 3 | 203 | 151 to 185% FPL |
| 04/98 – 12/00 | N | CHP+ | 15 to 18 years | 40 to 62% FPL | 1 | 201 | ≤ 100% FPL |
| | A | CHP+ | 15 to 18 years | 63 to 81% FPL | 1 | 201 | ≤ 100% FPL |
| | B | CHP+ | 15 to 18 years | 82 to 100% FPL | 1 | 201 | ≤ 100% FPL |
| | C | CHP+ | 15 to 18 years | 101 to 117% FPL | 1 | 201 | ≤ 100% FPL |
| | D | CHP+ | 6 to 18 years | 118 to 133% FPL | 2 | 202 | 101 to 150% FPL |
| | E | CHP+ | 6 to 18 years | 134 to 150% FPL | 2 | 202 | 101 to 150% FPL |
| | F- | CHP+ | 0 to 18 years | 151 to 159% FPL | 2 | 202 | 101 to 150% FPL |
| | F+ | CHP+ | 0 to 18 years | 160 to 170% FPL | 3 | 203 | 151 to 185% FPL |
| | G- | CHP+ | 0 to 18 years | 171 to 185% FPL | 3 | 203 | 151 to 185% FPL |
| | G+ | CHP+ | 0 to 18 years | 40 to 62% FPL | 3 | 203 | 151 to 185% FPL |
| FL | MK | MediKids | 0 to 5 years | ≤ 200% FPL | 1 | 301 | MediKids |
| | HK | Healthy Kids | 5 to 18 years | ≤ 200% FPL | 2 | 302 | HealthyKids |
| | CMS | CMS | 0 to 18 years | ≤ 200% FPL | 3 | 303 | CMS |

TABLE C.1 *(continued)*

| | State Eligibility Code Description | | | | | | |
|---|---|---|---|---|---|---|---|
| | State Eligibility Code | Program Name | Age Requirement | Income Requirement | MPR Eligibility Code | Unique MPR Eligibility Code | MPR Eligibility Code Description |
| IL | K | Kidcare Assist (MSCHIP) | 0 to 18 years[d] | 47 to 100% FPL | 1 | 401 | KidCare Assist MSCHIP (< 133% FPL) |
| | L | Kidcare Assist (MSCHIP) | 0 to 18 years[d] | 47 to 100% FPL | 1 | 401 | KidCare Assist MSCHIP (< 133% FPL) |
| | H | Kidcare Assist (MSCHIP) | 5 to 18 years[e] | 101 to 133% FPL | 1 | 401 | KidCare Assist MSCHIP (< 133% FPL) |
| | I | Kidcare Assist (MSCHIP) | 5 to 18 years[e] | 101 to 133% FPL | 1 | 401 | KidCare Assist MSCHIP (< 133% FPL) |
| | N | Kidcare Assist (MSCHIP) | 0 to 18 years[d] | 101 to 133% FPL | 1 | 401 | KidCare Assist MSCHIP (< 133% FPL) |
| | O | Kidcare Assist (MSCHIP) | 0 to 18 years[d] | 101 to 133% FPL | 1 | 401 | KidCare Assist MSCHIP (< 133% FPL) |
| | 4 | Kidcare Share (SCHIP) | 1 to 18 years old | 134 to 150% FPL | 2 | 402 | KidCare Share MSCHIP (< 134 to 150% FPL) |
| | S | Kidcare Share (SCHIP) | 1 to 18 years old | 134 to 150% FPL | 2 | 402 | KidCare Share MSCHIP (< 134 to 150% FPL) |
| | Z | KidCare Premium (SCHIP) | 1 to 18 years old | 151 to 185% FPL | 3 | 403 | KidCare Premium MSCHIP (< 151 to 185% FPL) |
| LA | 007 | LACHIP | 6 to 18 years | ≤ 133% FPL | 1 | 501 | LACHIP I (< 133% FPL) |
| | 015 | LACHIP Phase II | Birth to 18 years | 133 to 150% FPL | 2 | 502 | LACHIP II (133 to 150% FPL) |
| | 055 | LACHIP Phase III | Birth to 18 years | 151 to 200% FPL | 3 | 503 | LACHIP III (151 to 200% FPL) |
| MO | C071 | MC+ for Kids | 1 to 18 years old | ≤ 185% FPL | 1 | 601 | ≤ 185% FPL |
| | C072 | MC+ for Kids | 0 to 18 years old | 186 to 225% FPL | 2 | 602 | 186 to 225% FPL |
| | C073 | MC+ for Kids | 0 to 18 years old | 126 to 300% FPL | 3 | 603 | 226 to 300% FPL |
| NJ | 484 | NJC | 0 to 18 years[d] | ≤ 100% FPL | 1 | 701 | Plan A (< 133% FPL) |
| | 485 | NJC | 6 to 18 years | 101 to 133% FPL | 1 | 701 | Plan A (< 133% FPL) |
| | 486 | KidCare | 1 to 18 years | 134 to 150% FPL | 2 | 702 | Plan B (133 to 150% FPL) |
| | 487 | KidCare | 1 to 18 years | 151 to 185% FPL | 3 | 703 | Plan C (151 to 200% FPL) |
| | 488 | KidCare | Birth to 18 years | 186 to 200% FPL | 3 | 703 | Plan C (151 to 200% FPL) |
| | 489 | KidCare Fee For Service | Birth to 3 months | 186 to 200% FPL | 3 | 703 | Plan C (151 to 200% FPL) |
| | 493 | KidCare | 0 to 18 years | 201 to 250% FPL | 4 | 704 | Plan D (201 to 350% FPL) |
| | 494 | KidCare | 0 to 18 years | 251 to 300% FPL | 4 | 704 | Plan D (201 to 350% FPL) |
| | 495 | KidCare | 0 to 18 years | 301 to 350% FPL | 4 | 704 | Plan D (201 to 350% FPL) |
| | 496 | KidCare | Birth to 3 months | 201 to 350% FPL | 4 | 704 | Plan D (201 to 350% FPL) |
| NY[f] Current | A | Child Health Plus | 6 to 18 years old | < 120% FPL | 1 | 801 | < 151% FPL |
| | B | Child Health Plus | 1 to 18 years old | 120 to 150% FPL | 1 | 801 | <151% FPL |
| | C | Child Health Plus | 1 to 18 years old | 151 to 159% FPL | 2 | 802 | 151 to 222% FPL |
| | H | Child Health Plus | 0 to 18 years old | 160 to 222% FPL | 2 | 802 | 151 to 222% FPL |
| | I | Child Health Plus | 0 to 18 years old | 160 to 222% FPL | 2 | 802 | 151 to 222% FPL |
| | L | Child Health Plus | 0 to 18 years old | 223 to 250% FPL | 3 | 803 | > 222% |
| | M | Child Health Plus | 0 to 18 years old | 223 to 250% FPL | 3 | 803 | > 222% |

TABLE C.1 *(continued)*

| | State Eligibility Code Description | | | | | | |
|---|---|---|---|---|---|---|---|
| | State Eligibility Code | Program Name | Age Requirement | Income Requirement | MPR Eligibility Code | Unique MPR Eligibility Code | MPR Eligibility Code Description |
| | S | Child Health Plus | 0 to 18 years old | > 250% FPL | 4 | 804 | Full premium |
| | [g] | Child Health Plus | | | 5 | 805 | Non-missing, unclassified |
| | P | Child Health Plus | | | 6 | 806 | Presumptive eligibility |
| Oct-98 | A | Child Health Plus | 6 to 18 years old | < 120% FPL | 1 | 801 | < 151% FPL |
| | B | Child Health Plus | 1 to 18 years old | 120 to 150% FPL | 1 | 801 | <151% FPL |
| | C | Child Health Plus | 1 to 18 years old | 151 to 159% FPL | 2 | 802 | 151 to 222% FPL |
| | H | Child Health Plus | 0 to 18 years old | 160 to 222% FPL | 2 | 802 | 151 to 222% FPL |
| | I | Child Health Plus | 0 to 18 years old | 160 to 222% FPL | 2 | 802 | 151 to 222% FPL |
| | L | Child Health Plus | 0 to 18 years old | 223 to 230% FPL | 3 | 803 | > 222% FPL |
| | M | Child Health Plus | 0 to 18 years old | 223 to 230% FPL | 3 | 803 | > 222% FPL |
| | S | Child Health Plus | 0 to 18 years old | > 230% FPL | 4 | 804 | Full premium |
| | [g] | Child Health Plus | | | 5 | 805 | Non-missing, unclassified |
| | P | Child Health Plus | | | 6 | 806 | Presumptive eligibility |
| May-98 | F | Child Health Plus | 1 to 18 years old | < 151% FPL | 1 | 801 | < 151% FPL |
| | C | Child Health Plus | 1 to 18 years old | 151 to 159% FPL | 2 | 802 | 151 to 222% FPL |
| | E | Child Health Plus | 1 to 18 years old | 151 to 159% FPL | 2 | 802 | 151 to 222% FPL |
| | K | Child Health Plus | 1 to 18 years old | 151 to 159% FPL | 2 | 802 | 151 to 222% FPL |
| | G | Child Health Plus | 0 to 18 years old | 160 to 200% FPL | 2 | 802 | 151 to 222% FPL |
| | I | Child Health Plus | 0 to 18 years old | 160 to 200% FPL | 2 | 802 | 151 to 222% FPL |
| | L | Child Health Plus | 0 to 18 years old | 160 to 200% FPL | 2 | 802 | 151 to 222% FPL |
| | H | Child Health Plus | 0 to 18 years old | 201 to 222% FPL | 2 | 802 | 151 to 222% FPL |
| | J | Child Health Plus | 0 to 18 years old | 201 to 222% FPL | 2 | 802 | 151 to 222% FPL |
| | M | Child Health Plus | 0 to 18 years old | 201 to 222% FPL | 2 | 802 | 151 to 222% FPL |
| | S | Child Health Plus | 0 to 18 years old | > 222% FPL | 4 | 804 | Full premium |
| | [g] | Child Health Plus | | | 5 | 805 | Non-missing, unclassified |
| | P | Child Health Plus | | | 6 | 806 | Presumptive Eligibility |
| Oct-97 | F | Child Health Plus | 6 to 18 years old | < 120% FPL | 1 | 801 | < 151% FPL |
| | B | Child Health Plus | 1 to 18 years old | 120 to 150% FPL | 1 | 801 | < 151% FPL |
| | D | Child Health Plus | 1 to 18 years old | 120 to 150% FPL | 1 | 801 | < 151% FPL |
| | C | Child Health Plus | 1 to 18 years old | 151 to 159% FPL | 2 | 802 | 151 to 222% FPL |
| | E | Child Health Plus | 1 to 18 years old | 151 to 159% FPL | 2 | 802 | 151 to 222% FPL |
| | G | Child Health Plus | 0 to 18 years old | 160 to 200% FPL | 2 | 802 | 151 to 222% FPL |
| | I | Child Health Plus | 0 to 18 years old | 160 to 200% FPL | 2 | 802 | 151 to 222% FPL |
| | H | Child Health Plus | 0 to 18 years old | 201 to 222% FPL | 2 | 802 | 151 to 222% FPL |
| | J | Child Health Plus | 0 to 18 years old | 201 to 222% FPL | 2 | 802 | 151 to 222% FPL |
| | S | Child Health Plus | 0 to 18 years old | > 222% FPL | 4 | 804 | Full premium |
| | [g] | Child Health Plus | | | 5 | 805 | Non-missing, unclassified |
| | P | Child Health Plus | | | 6 | 806 | Presumptive eligibility |

TABLE C.1 *(continued)*

| | State Eligibility Code Description | | | | | | |
|---|---|---|---|---|---|---|---|
| | State Eligibility Code | Program Name | Age Requirement | Income Requirement | MPR Eligibility Code | Unique MPR Eligibility Code | MPR Eligibility Code Description |
| NC | MICJN | NC Health Choice for Children | 1 to 18 years old | ≤ 150% FPL | 1 | 901 | ≤ 150% FPL |
| | MICKN | NC Health Choice for Children | 0 to 18 years old | 151 to 200% FPL | 2 | 902 | 151 to 200% FPL |
| | MICSN | NC Health Choice for Children | 0 to 18 years old | 151 to 200% FPL | 2 | 902 | 151 to 200% FPL |
| TX[h] | 0 | TexCare | < 19 years old | < 100% FPL | 1 | 991 | < 100% FPL/no co-pay |
| | 1 | TexCare | 1 to 18 years old | 100 to 150% FPL | 2 | 992 | 100 to 150% FPL |
| | 2 | TexCare | 1 to 18 years old | 151 to 185% FPL | 3 | 993 | 151 to 185% FPL |
| | 3 | TexCare | 0 to 18 years old | 186 to 200% FPL | 4 | 994 | 186 to 200% FPL |

Source: Documentation provided by the states for the enrollment history files for the samples of recent enrollees and disenrollees from the 2002 congressionally mandated survey of SCHIP enrollees and disenrollees in 10 states, supplemented with site visit report data summarized in Hill et al. (2003).

[a]California does not have SCHIP eligibility groups.

[b]Colorado does not have SCHIP eligibility groups. We used the variable "program rate," which is based on income and family size, to determine SCHIP eligibility group.

[c]Colorado does not count assets when calculating income, whereas Medicaid does. Consequently, certain children under age 18 may not qualify for Medicaid and will be covered by SCHIP. Therefore, children of any age can be found in categories N, A, B, and C (telephone conversation with Joanne Lindsay, of Colorado, on 9/19/2003).

[d]Child must be born before 10/01/1983.

[e]Child must be born after 9/30/1983.

[f]New York does not have SCHIP eligibility codes. We used the variable "payment category" to determine eligibility group.

[g]All nonmissing eligibility codes in New York that were not classified in the documentation were grouped into a separate eligibility category.

[h]Texas does not have SCHIP eligibility groups. We used the co-payment category to determine SCHIP eligibility group.

FPL= federal poverty level; MSCHIP = Medicaid-expansion SCHIP; NA = not applicable; TPL = third-party liability.

TABLE C.2

CROSSWALK OF STATE ELIGIBILITY CODES INTO UNIFORM CODES,
BY STATE AND PROGRAM (MEDICAID)

| | State Eligibility Code | Federal Eligibility Code | | | | MPR Eligibility Code | Unique MPR Eligibility Code | MPR Eligibility Code Description |
|---|---|---|---|---|---|---|---|---|
| | | MAS Code | MAS Description | BOE Code | BOE Description | | | |
| CA | 30 | 1 | Individuals receiving cash assistance | 4/5 | Child/adult | 1 | 111 | Individuals receiving cash assistance |
| | 32 | 1 | Individuals receiving cash assistance | 4/5 | Child/adult | 1 | 111 | Individuals receiving cash assistance |
| | 33 | 1 | Individuals receiving cash assistance | 4/5 | Child/adult | 1 | 111 | Individuals receiving cash assistance |
| | 35 | 1 | Individuals receiving cash assistance | 4/5 | Child/adult | 1 | 111 | Individuals receiving cash assistance |
| | 60 | 1 | Individuals receiving cash assistance | 2 | Blind/disabled | 1 | 111 | Individuals receiving cash assistance |
| | 3E | 1 | Individuals receiving cash assistance | 4/5 | Child/adult | 1 | 111 | Individuals receiving cash assistance |
| | 3L | 1 | Individuals receiving cash assistance | 4/5 | Child/adult | 1 | 111 | Individuals receiving cash assistance |
| | 3M | 1 | Individuals receiving cash assistance | 4/5 | Child/adult | 1 | 111 | Individuals receiving cash assistance |
| | 3N | 1 | Individuals receiving cash assistance | 4/5 | Child/adult | 1 | 111 | Individuals receiving cash assistance |
| | 3P | 1 | Individuals receiving cash assistance | 4/5 | Child/adult | 1 | 111 | Individuals receiving cash assistance |
| | 3R | 1 | Individuals receiving cash assistance | 4/5 | Child/adult | 1 | 111 | Individuals receiving cash assistance |
| | 3U | 1 | Individuals receiving cash assistance | 4/5 | Child/adult | 1 | 111 | Individuals receiving cash assistance |
| | 34 | 2 | Medically needy | 4/5 | Child/adult | 2 | 112 | Medically needy |
| | 37 | 2 | Medically needy | 4/5 | Child/adult | 2 | 112 | Medically needy |
| | 64 | 2 | Medically needy | 2 | Blind/disabled | 2 | 112 | Medically needy |
| | 67 | 2 | Medically needy | 2 | Blind/disabled | 2 | 112 | Medically needy |
| | 82 | 2 | Medically needy | 4 | Child | 2 | 112 | Medically needy |
| | 83 | 2 | Medically needy | 4 | Child | 2 | 112 | Medically needy |
| | 47 | 3 | Poverty related | 4 | Child | 3 | 113 | Poverty related |
| | 72 | 3 | Poverty related | 4 | Child | 3 | 113 | Poverty related |
| | 7A | 3 | Poverty related | 4 | Child | 3 | 113 | Poverty related |
| | 8P | 3 | Poverty related | 4 | Child | 3 | 113 | Poverty related |
| | 8R | 3 | Poverty related | 4 | Child | 3 | 113 | Poverty related |
| | 38 | 4 | Other | 4/5 | Child/adult | 4 | 114 | Other |
| | 39 | 4 | Other | 4/5 | Child/adult | 4 | 114 | Other |
| | 40 | 4 | Other | 8 | Foster care child | 4 | 114 | Other |
| | 42 | 4 | Other | 8 | Foster care child | 4 | 114 | Other |
| | 45 | 4 | Other | 8 | Foster care child | 4 | 114 | Other |
| | 58 | 4 | Other | 2, 1, 4/5 | Blind/disabled | 4 | 114 | Other |
| | 59 | 4 | Other | 4/5 | Child/adult | 4 | 114 | Other |
| | 74 | 4 | Other | 4 | Child | 4 | 114 | Other |
| | 3T | 4 | Other | 4/5 | Child/adult | 4 | 114 | Other |
| | 3V | 4 | Other | 4/5 | Child/adult | 4 | 114 | Other |
| | 5F | 4 | Other | 5 | Adult | 4 | 114 | Other |
| | 5K | 4 | Other | 8 | Foster care child | 4 | 114 | Other |
| | 6N | 4 | Other | 2 | Blind/disabled | 4 | 114 | Other |
| | 7C | 4 | Other | 4 | Child | 4 | 114 | Other |

TABLE C.2 *(continued)*

| | State Eligibility Code | Federal Eligibility Code: MAS Code | MAS Description | BOE Code | BOE Description | MPR Eligibility Code | Unique MPR Eligibility Code | MPR Eligibility Code Description |
|---|---|---|---|---|---|---|---|---|
| | 7J | 4 | Other | 4 | Child | 4 | 114 | Other |
| | 7K | 4 | Other | 4 | Child | 4 | 114 | Other |
| NC | MICLN[a] | 0 | Separate SCHIP | 0 | | 0 | 910 | Separate SCHIP |
| | AAFCN[b] | 1, 4 | Individuals receiving cash assistance | 4, 5, 6, 7 | | 1 | 911 | Individuals receiving cash assistance |
| | MABCY | 1 | Individuals receiving cash assistance | 2 | | 1 | 911 | Individuals receiving cash assistance |
| | MADCY | 1 | Individuals receiving cash assistance | 2 | | 1 | 911 | Individuals receiving cash assistance |
| | MAFCN | 1 | Individuals receiving cash assistance | 4, 5, 6, 7 | | 1 | 911 | Individuals receiving cash assistance |
| | MAFMN | 2 | Medically needy | 4, 5 | | 2 | 912 | Medically needy |
| | MADNN[c] | 3, 4 | Poverty related | 2 | | 3 | 913 | Poverty related |
| | MICNN | 3 | Poverty related | 4 | | 3 | 913 | Poverty related |
| | MPWFN | 3 | Poverty related | 5 | | 3 | 913 | Poverty related |
| | MPWNN | 3 | Poverty related | 3 | | 3 | 913 | Poverty related |
| | HSFNN | 4 | Other | 8 | | 4 | 914 | Other |
| | IASCN | 4 | Other | 8 | | 4 | 914 | Other |
| | MAFNN | 4 | Other | 4, 5 | | 4 | 914 | Other |

Source: Documentation provided by the states for the enrollment files for the samples of recent enrollees and disenrollees for the 2002 congressionally mandated survey of SCHIP enrollees and disenrollees in 10 states.

[a]Based on an email from Marilyn Ellwood on July 3, 2003, these children are part of the separate SCHIP program. As a result, they are given a MAS/BOE code of 00, as they are not Medicaid enrollees.

[b]Based on an email from Lorenzo Moreno, of MPR, on 6/24/2003, the MPR eligibility code for AAFCN = 1.

[c]Based on an email from Lorenzo Moreno, of MPR, on 6/24/2003, the MPR eligibility code for MADNN = 3.

BOE = basis of eligibility; MAS = maintenance assistance status.

# APPENDIX D

# CONSTRUCTED VARIABLES

This appendix explains how we constructed several types of analytic variables. The first section describes the variables taken from the Area Resource File (ARF). The next section presents the methods we used to construct variables about prior insurance coverage among recent and established enrollees. The final section describes how we constructed the variables used to analyze the experiences of disenrollees.

## 1. ARF Variables

Several variables in the SAS data sets are based on the rural/urban continuum code variable from the 2001 ARF. This code, assigned to each county in the United States by the U.S. Department of Agriculture (USDA), distinguishes metropolitan counties by size and nonmetropolitan counties by degree of urbanization and proximity to metropolitan areas. These codes were linked to each respondent in the survey based on their reported county of residence using an assigned Federal Information Processing Standards code.

METROCTY is equal to 1 (and 0 otherwise) for residents in counties assigned the following rural/urban continuum codes:

00—central counties of metropolitan areas of 1 million population or more

01—fringe counties of metropolitan areas of 1 million population or more

02—counties in metropolitan areas of 250,000 to 1,000,000 population

03—counties in metropolitan areas of less than 250,000 population

All the above codes are designated as metropolitan counties by the USDA.

NONMETRO_ADJ and NONADJCTY distinguish between nonmetropolitan counties as adjacent or not adjacent to a metropolitan area.

NONMETRO_ADJ is equal to 1 (and 0 otherwise) for residents in counties assigned the following rural/urban continuum codes:

04—urban population of 20,000 or more, adjacent to a metropolitan area

06—urban population of 2,500 to 19,999, adjacent to a metropolitan area

08—completely rural, no places with a population of 2,500 or more, adjacent to a metropolitan area

NONADJCTY is equal to 1 (and 0 otherwise) for residents in counties assigned the following rural/urban continuum codes:

05—urban population of 20,000 or more, not adjacent to a metropolitan area

07—urban population of 2,500 to 19,999, not adjacent to a metropolitan area

09—completely rural, no places with a population of 2,500 or more, not adjacent to a metropolitan area

### 2. Prior Insurance Coverage

This section discusses the methodology used for the analyses of the relationships among SCHIP, private coverage, and uninsured periods among recent and established enrollees. Methods were identical for estimates of substitution among established Medicaid enrollees, except where noted. We begin this section by describing the methodology used to assign prior coverage to the recent enrollee analytic sample. We then describe the methodology used to classify reasons reported by parents for ending private coverage and for enrolling their children in SCHIP.

#### a. Prior Coverage Among Recent Enrollees

For sample members who reported being enrolled in SCHIP for fewer than 12 months, estimates of prior coverage were taken directly from the survey data. We constructed one variable characterizing children's coverage in the month just before enrolling (the variable COV1MBEF in the SAS data sets) and another characterizing their coverage during the six months before enrolling (the variable COV6MBEF in the SAS data sets).

Only a small percentage of the sample reported two or three types of coverage "just before enrolling." For reporting purposes, we imposed a hierarchy on types of coverage to assign cases to a single type. Because our primary concern was children's access to employer coverage, we assigned a child to employer coverage if any employer coverage was reported; otherwise, we assigned the child to nongroup private, Medicaid, SCHIP, and other public coverage, in that order. We collapsed types of coverage into four categories: (1) SCHIP coverage; (2) Medicaid coverage, including Medicaid health maintenance organizations (HMOs); (3) private coverage, which included coverage from a current or past employer/union and coverage from direct purchase of insurance; and (4) other public, which included Medicare, military coverage, and coverage through the Indian Health Service. Combining information, we characterized children's coverage in the month just before enrolling as (1) uninsured, (2) private, (3) Medicaid, (4) other public, or (5) born on SCHIP.

We also characterized children's coverage during the six months before enrolling (the variable INS6MBEF in the SAS data sets) as (1) uninsured all six months, (2) private with no gap just before enrolling in SCHIP, (3) public with no gap, (4) private with gap, (5) public with gap, or (6) born on SCHIP. We did not seek to characterize the length of uninsured "gaps" but reported them as such only if the gap was less than six months and had occurred just before enrolling. In characterizing prior coverage, we incorporated only gaps in coverage that occurred immediately before joining SCHIP, even if coverage for all six months was not reported. In other words, if a parent reported his or her child as having Medicaid just before enrolling in SCHIP, with no intervening gap, but reported being covered by Medicaid for only three months, we categorized the coverage as "Medicaid with no gap."

For the sizable fraction of recent enrollees who reported coverage of more than 12 months, we did not ask any questions about the type of coverage before enrollment, as those data were

expected to be unreliable. To keep this sample in the analysis, we determined the sample members' insurance status based on data in the state enrollment files for SCHIP and Medicaid.

To assign coverage during the six months before SCHIP enrollment, we first compared the SCHIP enrollment month reported by the respondent with the enrollment month from the SCHIP enrollment file. Some respondents with long stays who were interviewed late in the survey fielding period reported lengths of coverage on SCHIP that were consistent. However, we expected some inconsistency between sources due to recall error. In the analytic phase, we therefore divided this group into two categories based on how much earlier the reported enrollment month was from the enrollment month in the state files:

1. ***Reported Enrollment Month Less than Six Months Earlier than the Enrollment Month in State Files.*** Almost one-third (32 percent) of recent enrollees reporting enrollment in SCHIP for 12 or more months fell into this category. We assumed that a discrepancy in dates of enrollment up to and including six months was due to recall error. We did not consider these discrepancies to be problematic because respondents still were referring to a time period before enrollment that overlapped with the time period about which we were asking in the survey.
2. ***Reported Enrollment Month More than Six Months Earlier than the Enrollment Month in State Files.*** Slightly more than two-thirds (68 percent) of recent enrollees reporting enrollment in SCHIP for 12 or more months fell into this category. This group presented an analytic challenge, because respondents were referring to a time period predating the six-month period before their current SCHIP enrollment spells, and they may have been reporting a coverage experience from a prior coverage spell, possibly in Medicaid.

To estimate prior coverage for these two groups, we adopted two separate imputation procedures. For the first group, which had self-reported data with few discrepancies, we relied on survey data to estimate prior coverage. For the second group, whose self-reported data were less likely to credible, we relied on information from the administrative data files.

For the first group, we used the following five-step procedure:

1. We used the six-month period before the self-reported enrollment date as the reference period to search the state administrative files.
2. From the state administrative file, we determined the number of months the child was enrolled in Medicaid during the self-reported reference period. However, we used survey data to determine whether the transition from Medicaid to SCHIP was accompanied by a gap with no coverage at all.
3. If the respondent reported being insured immediately before enrollment, we coded the child as being covered by Medicaid if we found administrative evidence of enrollment in Medicaid in the state files during the self-reported period. Otherwise, we coded children who were covered immediately before enrollment as having been covered by private insurance for all six months.
4. If the respondent reported an uninsured period of less than six months immediately before enrollment, we coded the child as moving from Medicaid to that uninsured period and then to SCHIP if we found evidence of enrollment in Medicaid. Otherwise, we coded the child as moving from private coverage to uninsured before enrolling in SCHIP.
5. If the respondent reported an uninsured period of six months or more immediately before enrollment, we coded the child as uninsured for all six months before enrollment unless we found evidence of Medicaid enrollment. In that case, we coded the child as moving uninsured to Medicaid and then directly to SCHIP. Our reasoning was that the parent may not have recognized a short spell on Medicaid before having been moved to SCHIP, but was otherwise uninsured before public coverage.

For the second group, which reported enrollment dates occurring more than six months earlier than the dates in the state files, we used the following four-step procedure:

1. We used the six-month period before the administrative enrollment date as the period of reference to search the state file.
2. Self-reported information on insurance status was overridden entirely if Medicaid or SCHIP data were found in this period, under the assumption that respondents were referring to reference periods outside our six-month period, so that their self-reports were less credible.
3. We examined the number of months the child was enrolled in Medicaid during the six-month period before the month of enrollment and whether there was a gap in enrollment in the month before SCHIP enrollment. This information was used to code the child as either being covered by Medicaid all six months or having a period of being uninsured between Medicaid and SCHIP. If we found enrollment data in

both the Medicaid and SCHIP files or in the Medicaid file alone, we coded the child as transitioning from Medicaid to SCHIP. If only SCHIP data were found, we coded the child as having a prior SCHIP episode.

4. If we found no evidence of Medicaid enrollment in the six-month period before the administrative month of enrollment, we relied on reports of uninsured periods to assign enrollees to private coverage or uninsured status. If the respondent reported some coverage, but no evidence of public coverage was found in the state files, we coded the child as having private coverage for the six months before enrollment. If the respondent reported an uninsured spell of six months or more before enrollment, and there was no evidence of Medicaid enrollment, we coded the child as being uninsured for all six months.

We also examined the enrollment records for the recent enrollees who were born on SCHIP and found evidence of Medicaid coverage before their SCHIP enrollment dates for three-quarters of them. We therefore assigned insurance coverage for these children as a seamless transition from Medicaid. Children older than age 5 and therefore born before implementation of SCHIP in January 1998, with no evidence of Medicaid or SCHIP enrollment at birth, were coded as missing prior coverage data. The remaining cases were coded as "born on SCHIP."

Colorado, New York, and Texas provided no Medicaid enrollment data from their administrative files. Therefore, we could use only state SCHIP files to determine the types of coverage for children in those states. For children reported as being insured before enrolling in SCHIP but who, according to the state files, did not have SCHIP, we could not turn to Medicaid files to determine whether the coverage was public or private. Instead, we imputed coverage status, using a regression model based on the coverage experience of two other types of recent enrollees: (1) those with complete information covered by SCHIP for more than 12 months in states with Medicaid data, and (2) recent enrollees with complete insurance information in the three states with no Medicaid data. We refer to those cases as "donor cases."

We used regression imputation to predict private or public coverage among those with coverage before SCHIP enrollment. The dependent variable was set to 1 if the donor case held

any form of private coverage during the six months before SCHIP enrollment and to 0 if the donor held only public coverage (Medicaid, SCHIP, or other public). We estimated a logistic regression because of the binary nature of the dependent variable. The model explained insurance status based on parents' work status, family structure, family income, the respondent's age and health status, the child's race/ethnicity, state of residence, and reported length of time on SCHIP. The specification for the regression achieved a high percentage of correctly predicted donor cases. We used this model for children whose prior insurance status was "insured" to assign the children a predicted probability of private coverage. Cases with a high predicted probability of private coverage were assigned private coverage.

Based on the protocol to determine prior insurance coverage within the universe of recent enrollees, we could not assign prior coverage to 350 cases and therefore had to drop those cases from the analytic sample. This group included cases coded as born on SCHIP, cases covered by SCHIP during the six months before the current enrollment, and cases missing sufficient insurance status information to classify.

**b. Reasons for Ending Private Coverage and Enrolling in SCHIP**

We analyzed reasons for ending prior coverage and enrolling in SCHIP for those with private coverage during the six months before enrollment. The reasons were used to determine whether private coverage ended voluntarily or involuntarily and to produce estimates of substitution at the time of enrollment. This section describes how we assigned reasons for transitions from private insurance in the six months before enrollment in SCHIP among recent enrollees.

Parents of recent enrollees provided information through one of three survey questions on why private coverage ended. Parents who reported their children as being privately insured just before enrolling were asked why that private coverage had ended (the variables Q3_44 and

Q3_44BC in the SAS data sets). Alternatively, parents who reported their children were uninsured at some point in the six months before enrolling were asked why their children were uninsured during that time (the variables Q3_34 and Q3_34BC in the SAS data sets). Many of the responses to that question related to private coverage that had ended. Finally, all respondents were asked why they had enrolled their children in SCHIP (the variables Q2_14 and Q2_14BC in the SAS data sets). All three questions used similar response categories, and we applied the same coding protocols to any open-ended verbatim responses that parents provided. This technique enabled us to combine responses from all three questions about why private coverage had ended.

For parents who were asked more than one of the questions, we used the responses about why private coverage had ended to assess the parents' ability to have retained private coverage for their children. For those who were asked the question but did not provide a reason, we substituted the reason why the children were uninsured. About one-fifth of cases with prior private coverage were not asked why the coverage ended or did not respond to the question about why their children were uninsured. This set of cases included primarily recent enrollees who were interviewed as established enrollees. We determined that the children had prior private coverage through our examination of administrative data, logical editing, and imputation. For these cases, we used the response to the survey question on why the parent had enrolled his or her child in SCHIP to assess why private coverage had ended. Only one case was missing responses to all the questions about reasons.

### 3. Disenrollee Experiences

This section discusses the study methodology used for the analysis of SCHIP disenrollees. We discuss the methods used to analyze the experiences of SCHIP disenrollees, focusing in particular on how we measured disenrollees' insurance coverage after leaving the program.

The most important measure in the analysis of disenrollees' experiences was the type of insurance coverage after leaving SCHIP. The other key measure we examined was the reported reason for leaving SCHIP. Development of these measures, particularly the measure of insurance coverage, was complex and required several steps.

**a. Insurance Coverage**

Our measure of insurance coverage for two groups of disenrollees—those who had exited within the past 12 months and those who had exited more than 12 months ago and were recontacted—was obtained directly from questions on the survey. The specific steps we took were as follows:

1. Based on responses to questions 3.60 and 3.63, we determined how many months the disenrollee had been uninsured after leaving SCHIP. Each of these months was coded as uninsured. If the disenrollee reported being uninsured for the "whole period" since leaving SCHIP, all months between disenrollment and the interview date (up to month 6) were coded as uninsured.
2. Based on responses to questions 3.64 and 3.64.1, we then determined how many months the disenrollee had been insured after exit (or after the spell of uninsurance, if reported above). Each of these months was then coded as insured. If the disenrollee reported being insured for the whole period, all months between disenrollment (or the end of uninsurance spell) and the interview were coded as insured.
3. For the months coded as insured, the type of insurance was coded based on responses to questions 3.65a through 3.65h. For disenrollees reported to have more than one type of coverage, we chose the first reported type of coverage as given by question 3.66.

After completing these three steps, the types of coverage were then collapsed into four categories: (1) SCHIP coverage; (2) Medicaid coverage, including Medicaid HMOs; (3) private coverage, which included coverage from a current or past employer/union and coverage from direct purchase of insurance; and (4) other/unknown coverage, which included Medicare, military coverage of any kind, coverage through the Indian Health Service, and any other type of

coverage that could not be coded. Fewer than five percent of disenrollees in each state fell into the latter category.

Because those who reported being covered by SCHIP for six or more months did not appear to recognize that they had been disenrolled from the program, the survey did not collect information about their coverage after exit. For most of these cases, the state files indicated either new spells of SCHIP coverage or Medicaid coverage in the first few months after disenrollment. This information suggests that most of the respondents did not recognize their exit either (1) because they experienced a short gap in SCHIP coverage that apparently went unnoticed, or (2) because they experienced a "seamless" transition to the Medicaid program that likewise appears to have been unrecognized. To keep these cases in the analysis, we drew on the state SCHIP and Medicaid files and followed a four-step coverage imputation procedure:

1. Using the state SCHIP files, we looked at the six months after a child's exit and identified each month that the child was shown to be covered. These months were then coded as SCHIP coverage as if the respondent had self-reported them.
2. For the seven states for which we had Medicaid enrollment data (California, Florida, Illinois, Louisiana, Missouri, New Jersey, and North Carolina), we looked at the six months after the child's exit and identified each month that the child was shown to be covered by Medicaid. If these months had not been previously imputed as SCHIP in step 1, they were coded as Medicaid as if the respondent had self-reported them.
3. For the three states for which we did not have Medicaid enrollment data (Colorado, New York, and Texas), we imputed Medicaid coverage after disenrollment, using the sample of disenrollees from three "donor states" that also had separate SCHIP programs (California, Florida, and North Carolina). The imputation was carried out as follows:
    a. We separated the disenrollees in the three donor states into groups based on their observed SCHIP coverage during the six months after exit.
    b. Within each of these groups, we identified all the possible scenarios of Medicaid coverage and calculated the frequency of each in the donor states. Each scenario was given a probability equal to this frequency.
    c. For each case subject to imputation, we determined the group to which it belonged based on the observed SCHIP coverage during the six months after exit. We then imputed the string of Medicaid coverage by selecting one of the possible scenarios identified in the previous step. The particular scenario

chosen was based on the probability assigned to it in relation to a random number between 0 and 1.

4. Any months that were not assigned SCHIP or Medicaid coverage based on the state files were imputed a value of either uninsured or private coverage. The imputation was performed as follows:

   a. If the disenrollee showed any SCHIP or Medicaid coverage during the six-month period, the undetermined months between exit and coverage (if any) were coded as uninsured. This coding was based on the assumption that few disenrollees who cycled off and back on public coverage in a short period would have obtained coverage in the intervening months.

   b. All other undetermined months were imputed through regression. Using the subsample with valid self-reported data (category 1), we first constructed a dummy variable that equaled 1 if the disenrollee was privately insured in a given month, and 0 if uninsured in the month. This dummy variable was then regressed on a series of covariates measuring key child and family demographics. Based on the coefficients from this model, we then generated the predicted probability of having private insurance in each undetermined month. This predicted value was then compared with a random digit generated between 0 and 1. If the predicted value was above the random digit, we coded the month as privately insured; if it was below the random digit, we coded the month as uninsured.

For some cases, this imputation procedure was likely to assign a coverage type that was different from what would have been reported by the respondent in the survey (had it been possible to collect this information). However, in the aggregate, we expected this procedure to yield a distribution that would be consistent with self-reported data from the survey. To investigate the degree of consistency, we studied the sample of disenrollees in the first group (those who left SCHIP within the past 12 months), whom we expected to report reliably on coverage type after exit. We compared the coverage reported in the survey for this group with the coverage derived from imputation. Results indicated similar distributions of coverage for this group of disenrollees, whether based on the reported coverage or on the imputation procedure.

Three variables in the SAS data sets indicate the type of insurance coverage disenrollees had after they left SCHIP. ENR1MAFT is the type of coverage that the disenrollee had immediately

after leaving SCHIP; ENR3MAFT is the type of coverage that the disenrollee had three months after leaving SCHIP; and ENR6MAFT is the type of insurance coverage that the disenrollee had six months after leaving SCHIP. We assigned each of these variables the following values:

0—No insurance

1—SCHIP

2—Medicaid

3—Private insurance

4—Another form of insurance

**b. Reasons for Exit and Uninsurance**

Our measures of reasons for disenrolling and for being uninsured after disenrolling are based on questions 3.26 and 3.63, respectively. Responses to these questions were open-ended; they were coded into a long list of categories by the interviewers. If responses did not fit any of the categories, the interviewers placed them in an "other specify" category and recorded them verbatim. Responses in this category were reviewed by the study team; most were then "backcoded" into existing categories. Subsequently, the response categories were reduced to a smaller number.

"Reasons for leaving SCHIP" (the variable DISREASON in the SAS data set) were grouped into six categories. Disenrollees were considered more likely to remain eligible for SCHIP if their reasons fell into one of the following three categories:

1. Failure to pay premium, which included the original categories of "could not afford premium" and "forgot to pay premium"
2. Failure to reapply, which included the original categories "did not reapply" and "too much paperwork"
3. Other reasons, which included such responses as "did not like doctors/clinic/staff where care provided," "did not like the quality of care," and "child does not get sick." This category also included a small number of miscellaneous reasons.

Disenrollees whose reasons fell into one of the following three groups were not likely to be eligible for SCHIP:

1. Child is too old, which reflected a single category
2. Eligible for other coverage, which included the original categories of "child obtained Medicaid coverage" and "child obtained other insurance"
3. Change in income or employment, which reflected a single category ("financial situation changed/not qualified")

The categories for "reasons for being uninsured" (the variable REASUNIN in the SAS data set) were also collapsed into six groups. Those whose reasons fell into any of the following three groups were again considered possibly eligible for SCHIP:

1. Unable to pay for insurance, which included the original categories of "forgot to pay premium" and "cannot afford premium"
2. Lack of access to affordable private coverage, which included the original categories of "parent(s) lost/changed job," "employer did not offer insurance," "employer stopped offering insurance," "parents got divorced/death of spouse," "benefits from former employer ran out," "no one in family employed," and "insurance costs too high"
3. Failure to reapply, which reflected a single category

The following three groups were considered not likely to be eligible for SCHIP:

1. Child is too old, which reflected a single category
2. Eligible for other coverage, which reflected a single category
3. Other reasons, which included "did not like health insurance employer offers" and "needed to be uninsured to be eligible." This category also included a few miscellaneous responses.

After reviewing the reasons for leaving SCHIP and the reasons for being uninsured, we created variables that indicate whether the disenrollee was likely to have been eligible for SCHIP immediately after leaving (SCHIPELIG_1) and six months after leaving (SCHIPELIG_6).