Contract No.:
MPR Reference No.: 8519-340

MATHEMATICA
Policy Research, Inc.

Report on Survey Methods for the Community Tracking Study's 1998-1999 Round Two Household Survey

Final Report

November 2001

Richard Strouse
Barbara Carlson
John Hall

Submitted to:

Center for Studying Health System Change
600 Maryland Ave., SW
Suite 550
Washington, DC 20024-2512

Project Officer:
Peter Cunningham, Ph.D.

Submitted by:

Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, NJ 08543-2393
(609) 799-3535

Project Director:
Richard Strouse

CONTENTS

CONTENTS *(continued)*

Chapter **Page**

CONTENTS *(continued)*

Chapter **Page**

CONTENTS *(continued)*

Chapter **Page**

TABLES

TABLES *(continued)*

Table **Page**

FIGURES

Figure **Page**

I. OVERVIEW

A. OBJECTIVES OF THE COMMUNITY TRACKING STUDY

The Community Tracking Study (CTS), which is funded by the Robert Wood Johnson Foundation (RWJF), is designed to provide a sound information base for decision making by health care leaders. It does so by collecting information on how the health system is evolving in 60 nationally representative communities across the United States and on the effects of those changes on people. The CTS, which has been under way since 1996, is a longitudinal project that relies on periodic site visits and surveys of households, physicians, and employers. Although many studies have examined markets in California and Minnesota, and many have analyzed local or selected data, no study has systematically examined change in a broad, nationally representative cross-section of U.S. markets. Moreover, none has analyzed the effects of changes on service delivery, cost, and quality. The CTS addresses two broad questions that are important to public and private health decision makers:

1. ***How is the health system changing?*** How are hospitals, health plans, physicians, safety net providers, and other provider groups restructuring, and what key forces are driving organizational change?

2. ***How do these changes affect people?*** How are insurance coverage, access to care, use of services, health care costs, and perceived quality of health care changing over time?

Focusing on markets is central to the design of the CTS. Understanding market changes requires a study of local markets, including the markets' culture, history, and public policies relating to health care. To track change across the United States, we randomly selected 60

nationally representative communities stratified by region, community size, and type (metropolitan or nonmetropolitan).[1]

The CTS examines 12 of the 60 communities in depth by conducting site visits and using survey samples large enough to draw conclusions about change in each community. The 12 communities comprise a randomly selected subset of sites that are metropolitan areas with more than 200,000 people (as of July 1992). We refer to them as *high-intensity sites*.

B. ANALYTIC COMPONENTS OF THE COMMUNITY TRACKING STUDY

The CTS has qualitative and quantitative components. The qualitative component consists of case studies in the 12 high-intensity sites. The first round of comprehensive case studies of the health system was conducted in 1996 and 1997; the second round was conducted in 1998 and 1999. Survey data from the 12 high-intensity sites and from 48 additional sites, listed in Table 1.1, complement this information.

The CTS also includes independent surveys of households, physicians, and employers in all 60 sites, thereby enabling researchers to explore relationships among purchasers, providers, and consumers of health care.[2] A Followback Survey, which is linked to the Household Survey, is also conducted. In the Followback Survey, the privately financed health insurance policies covering respondents to the survey of households are "followed back" to the organization that administers the policy. The purpose of the Followback Survey is to obtain information about the private policies that is more detailed and more accurate than Household Survey respondents are able to provide.

[1]The CTS covers the contiguous 48 states and the District of Columbia. Alaska and Hawaii were not part of the study.

[2]The RAND Corporation, in collaboration with the Center for Studying Health System Change (HSC), conducted the Employer Survey. All other surveys were conducted under HSC's direction.

TABLE I.1

SITES SELECTED FOR THE COMMUNITY TRACKING STUDY

High-Intensity Sites	Low-Intensity Sites		
Metropolitan Areas >200,000 Population[a]	Metropolitan Areas >200,000 Population[a]	Metropolitan Areas <200,000 Population[a]	Nonmetropolitan Areas
01-Boston (MA)	13-Atlanta (GA)	49-Dothan (AL)	52-West Central Alabama
02-Cleveland (OH)	14-Augusta (GA/SC)	50-Terre Haute (IN)	53-Central Arkansas
03-Greenville (SC)	15-Baltimore (MD)	51-Wilmington (NC)	54-Northern Georgia
04-Indianapolis (IN)	16-Bridgeport (CT)		55-Northeastern Illinois
05-Lansing (MI)	17-Chicago (IL)		56-Northeastern Indiana
06-Little Rock (AR)	18-Columbus (OH)		57-Eastern Maine
07-Miami (FL)	19-Denver (CO)		58-Eastern North Carolina
08-Newark (NJ)	20-Detroit (MI)		59-Northern Utah
09-Orange County (CA)	21-Greensboro (NC)		60-Northwestern Washington
10-Phoenix (AZ)	22-Houston (TX)		
11-Seattle (WA)	23-Huntington (WV/KY/OH)		
12-Syracuse (NY)	24-Killeen (TX)		
	25-Knoxville (TN)		
	26-Las Vegas (NV/AZ)		
	27-Los Angeles (CA)		
	28-Middlesex (NJ)		
	29-Milwaukee (WI)		
	30-Minneapolis (MN/WI)		
	31-Modesto (CA)		
	32-Nassau (NY)		
	33-New York City (NY)		
	34-Philadelphia (PA/NJ)		
	35-Pittsburgh (PA)		
	36-Portland (OR/WA)		
	37-Riverside (CA)		
	38-Rochester (NY)		
	39-San Antonio (TX)		
	40-San Francisco (CA)		
	41-Santa Rosa (CA)		
	42-Shreveport (LA)		
	43-St. Louis (MO/IL)		
	44-Tampa (FL)		
	45-Tulsa (OK)		
	46-Washington (DC/MD/VA)		
	47-West Palm Beach (FL)		
	48-Worcester (MA)		

NOTE: Numbers correspond to coding of the site identification variable in the survey.

[a]Based on 1995 Census estimates.

Data are collected on a two-year cycle, to enable researchers to track changes in the health care system over time. The Round One surveys and case studies of households and physicians, completed during 1996 and 1997, and the Followback Survey, completed in 1997 and 1998, are the baseline. Data collection for the Round Two surveys of households and physicians began in 1998 and was completed in 1999. Round Two Followback Survey data collection was conducted during 1999 and 2000. Round Two case studies were completed in 1998 and 1999. Documentation of CTS data collection activities is available at www.hschange.org.

C. THE HOUSEHOLD SURVEYS

The Round One and Round Two Household Surveys, funded by RWJF, were conducted under the direction of HSC. Mathematica Policy Research, Inc. (MPR) was the primary contractor for survey design, instrument development, sample design and implementation, most of the interviewing, weighting, and variance estimation. Battelle's Center for Public Health Research and Evaluation assisted with the telephone interviewing. Social and Scientific Systems, Inc. (SSS) was instrumental in converting the raw survey data into a data file suitable for analysis. MPR and SSS collaborated to prepare the documentation for the CTS Household Public Use File.

The Household Survey instruments asked about health insurance, use of health services, satisfaction with care, health status, and demographic information. A family informant provided information on insurance coverage, health resource use, the usual source of care, and the general health status of all family members. This informant also provided information on family income, as well as on employment, earnings, employer-offered insurance plans, and race/ethnicity for all adult family members. Each adult in the family (including the informant) responded through a self-response module to questions about unmet needs, patient trust, satisfaction with physician choice, detailed health questions, risk and smoking behaviors, and the last physician visit. The

self-response module included mostly subjective questions that could not be answered reliably by proxy respondents. The family informant responded on behalf of children about unmet needs and satisfaction with physician choice.[3] The adult family member who took the child to his or her last physician visit responded to questions about that visit. (This adult family member may not have been the family informant.) A Spanish version of the instrument was used when appropriate.

The survey instruments used in the two rounds were similar, but not identical. The Round Two instrument altered some of the questions about public insurance and added questions about chronic health conditions. The surveys were administered by telephone, using computer-assisted telephone interviewing (CATI) technology.

Although the majority of the respondents in each round were selected through list-assisted random-digit-dialing (RDD) sampling methodology, families without working telephones also were represented in the sample. Field staff using cellular telephones enabled these families to complete interviews.

A sample of the telephone numbers from the Round One RDD sample was included in the Round Two sample to improve precision for estimates of population change, reduce costs, and increase response rates. Although some individuals responded to both rounds, the samples were designed to allow separate cross-sectional estimates and do not allow for panel or longitudinal analyses using data from both the Round One and Round Two surveys. The design does allow for comparisons of cross-sectional estimates between the two rounds.

After we contacted the selected households, we determined the composition of each household, grouped household members into family insurance units (FIUs), and obtained

[3]In families with more than one child younger than age 18, one child was randomly selected for inclusion in the survey.

information on each adult in each FIU.[4] If an FIU contained one child, we collected information about that child. If it contained two or more children, we collected information about one randomly selected child.

Round Two interviews with 58,956 individuals from 32,047 FIUs were completed between July 1998 and November 1999.

In this report, we describe site selection, sample design, instrumentation and survey preparation, data collection methods, response rates, and sample weights.

[4]The FIU is based on groupings of people typically used by insurance carriers. It includes an adult household member, spouse, and dependent children up to age 18 (or ages 18 to 22 if the child is in school). A more detailed definition of the FIU is presented in Chapter II.

II. SITE SELECTION AND SAMPLE DESIGN

For both the first and second rounds of the Household Survey, interviews were administered to households in the 60 CTS sites and to an independent national sample of households. The survey has the following three-tier sample design, which makes it possible to develop estimates at the national and community (site) levels:

- The first tier is a sample of 12 communities from which a large number of households in each community was surveyed. The sample in each of these "high-intensity" sites was large enough to support estimates in each site.
- The second tier is a sample of 48 communities from which a smaller sample of households in each community was surveyed. This sample of "low-intensity" sites enables us to validate results from the high-intensity sites and permits findings to be generalized to the nation. The first and second tiers comprise the site sample.
- The third tier is a smaller, independent national sample. This supplemental sample augments the site sample and substantially increases the precision of national estimates with a relatively modest increase in total sample size.

In the following sections, we discuss site selection; the selection of households, FIUs, and individuals; sample size considerations; the unit chosen for longitudinal tracking; and procedures for selecting the RDD and field samples.

A. SITE SELECTION

The primary goal of the CTS is to track health system change and its effects on people at the local level. Determining which communities, or sites, to study was therefore the first step in designing the CTS sample. Site selection involved three activities: (1) defining sites, (2) determining how many would be studied, and (3) selecting the sites.

1. Definition of Sites

The sites were intended to encompass the range of existing local health care markets. Although these markets have no set boundaries, the intent was to define areas such that residents predominantly used health care providers located in the same area, and providers mostly served area residents. To this end, we generally defined sites to be Metropolitan Statistical Areas (MSAs) as defined by the Office of Management and Budget or, in the case of nonmetropolitan sites, to be Bureau of Economic Analysis economic areas (BEAEAs). For additional detail on the definition of CTS sites, refer to Metcalf et al. (1996).

2. Number of Sites

The next step in creating the site sample was to determine the number of high-intensity sites. We considered the trade-offs between data collection costs (the cost of conducting case studies and surveys) and the research benefits of a large sample of sites. The research benefits include a greater ability to empirically examine the relationship between system change and its effect on care delivery and consumers and increased "generalizability" of the study findings to the nation as a whole.

Despite the cost advantages of conducting intensive case studies in fewer sites, focusing on a small number of communities would have made it more difficult to distinguish between changes of general importance and changes or characteristics unique to a community. Solving this problem by increasing the number of case study sites increases the cost of data collection and analysis. To balance these competing concerns, we chose 12 sites for intensive study and added to this sample 48 sites that would be studied less intensively. The 60 high-intensity and low-intensity sites are primary sampling units (PSUs) and form the site sample (see Table I.1 in Chapter I).

Although we had no formal scientific basis for choosing 12 high-intensity sites, the number reflects a balance between the benefits of studying a range of different communities and the costs of that study. The addition of 48 low-intensity sites solved the problem of limited generalizability associated with only 12 sites and provided a benchmark for interpreting the representativeness of the high-intensity sites.

3. Site Selection

After the number of sites for the site sample was determined, the next step was to select the actual sites. The 60 sites were chosen for the first stage of sampling. Sites were sampled by stratifying them geographically by region and then selecting them randomly, with probability proportional to the size of their July 1992 civilian population. The CTS sites (or PSUs) were selected independently in three strata. The three strata were:

1. MSAs with 200,000 or more people (large MSAs)[1]
2. MSAs with fewer than 200,000 people (small MSAs)
3. Nonmetropolitan areas

For eight sites in the large MSA stratum, the population was sufficiently large that the site was selected with certainty. These eight sites were Boston (MA Portion); Philadelphia, PA-NJ PMSA; Washington/Hagerstown PMSAs; New York City; Detroit, MI PMSA; Chicago/Kenosha/Kankakee PMSAs; Houston-Galveston-Brazoria, TX CMSA; and Los Angeles-Long Beach, CA PMSA. A ninth site (Baltimore, MD PMSA) was selected with certainty in the sample to complete coverage of the major cities of the Northeast Corridor.

[1]Some sites were defined as primary metropolitan statistical areas (PMSAs) or consolidated metropolitan statistical areas (CMSAs).

In addition to the nine certainty selections, 39 sites were selected with probability proportional to size, using a sequential selection algorithm with selection control imposed on the basis of geographic region. This allocation ensured that (1) all MSAs had a chance to be selected, (2) larger MSAs had a greater chance than smaller MSAs of being selected, and (3) the site sample would have an approximately proportional allocation across geographic regions.

For the small MSAs, three sites were selected with probability proportional to size, again using a sequential selection algorithm with ordering by geographic region. For the nonmetropolitan areas, the first stage of selection was the state.[2] The states were selected with probability proportional to the size of their non-metropolitan population, using the sequential selection algorithm (again ordered by geographic region); nine states were selected. Based on county groups used by the BEA, one county group was selected within each state with probability proportional to the population in these county groups.

Of the 60 sites in the CTS sample, 48 were selected in large MSAs, 3 in small MSAs, and 9 in nonmetropolitan areas. The 12 high-intensity sites were selected randomly from the 48 large MSA sites.

Together, the high-intensity and low-intensity sites account for about 90 percent of all survey respondents. (The remaining 10 percent were selected from the supplemental sample, discussed below.) The site sample can be used to make national estimates and also may be used to make site-specific estimates for the high-intensity sites. Users should be aware that site-specific estimates for the low-intensity sites will be less precise because of the small sample size from these sites.

[2]Washington, DC, and New Jersey were excluded because they do not have any nonmetropolitan areas. Alaska and Hawaii were excluded by the CTS study design.

4. Additional Samples and Better National Estimates

Although the site sample alone will yield national estimates, these estimates would have been less precise than if we had sampled more communities, or if we had used a simple random sample of the entire U.S. population. We therefore added the *supplemental sample*, the third tier in the design of the Household Survey sample, to increase the precision of national estimates at a relatively small incremental increase in survey cost.

The supplemental sample is a relatively small, nationally representative sample of households randomly selected from the 48 states in the continental United States and the District of Columbia. It is stratified by region and a single stage selection of telephone numbers within strata. When it is added to the site sample to produce national estimates, the resulting sample is called the *combined sample.*

In addition to increasing the precision of national estimates based on the site sample, the supplemental sample slightly enhances site-specific estimates derived from the site sample. Because approximately half the U.S. population lives in the 60 site sample communities, approximately half the supplemental sample also falls within those communities. When making site-specific estimates, we can therefore augment observations from the individual site sample with observations from the supplemental sample. These are known as the *augmented site samples.*

B. HOUSEHOLD, FAMILY INSURANCE UNIT, AND INDIVIDUAL SELECTION

1. Households

At the beginning of the interview, a household informant was identified and queried about the composition of the household. Typically, the household informant was the person who answered the telephone, if he or she was an adult 18 years of age or older. The person who owned or rented the house was identified as the head of the household, or the householder.

People who usually lived in the household but who were temporarily living elsewhere, such as college students, were included in the household enumeration.

2. Family Insurance Units

Individuals in the household were grouped into one or more FIUs to ensure that a knowledgeable informant would be able to answer questions about each family member's health insurance coverage, use of health resources in the 12 months preceding the interview, and usual source of health care. The FIU also provided information on family income and on the employment, earnings, employer-offered health insurance plan, and race or ethnicity of each adult in the FIU. An FIU reflects family groupings typically used by insurance carriers and is similar to the filing unit used by Medicaid and state-subsidized insurance programs. The FIU includes an adult household member; his or her spouse, if any; and any dependent children up to 17 years of age or 18 to 22 years of age if a full-time student (even if living outside the household).[3]

All FIUs were selected to participate in the remainder of the interview as long as the FIU contained at least one civilian adult.[4] In each FIU, one informant was responsible for providing much of the information about the family and its members. Figure II.1 shows how one household of seven people could be divided into three FIUs. In this example, the household head's spouse is the household informant because the spouse answered the telephone and is

[3]The CTS's definition of FIU differs from the Census Bureau's definition of a family, which includes all people living in the dwelling who are related to the householder by blood or by marriage. The Census family often is larger than an FIU. Adult relatives living in one household would be included in a Census primary family but would be assigned to separate FIUs for the CTS Household Survey.

[4]Individuals who were not on active military duty at the time of the interview were considered to be civilians.

FIGURE II.1

EXAMPLE OF FAMILY INSURANCE UNITS IN A HYPOTEHTICAL HOUSEHOLD

Members of Household	FIU
Head of Household Head of Household's Spouse (Informant for HH and FIU1) Head of Household's Daughter (Selected) Head of Household's Son (Not Selected)	FIU 1
Head of Household's Father (Informant for FIU2) Head of Household's Mother	FIU 2
Unrelated Boarder	FIU 3

familiar with the composition of the household. The spouse is also familiar with the health care of the head of household and their children, so the spouse is also the informant for the first FIU (FIU1). The household head's father is the informant for the second FIU (FIU2), and the unrelated boarder responds for himself or herself (FIU3). The household head's daughter is the randomly selected child in FIU1 and the head's son is not included in the survey. The use of separate FIU informants ensures that survey respondents provide information about the health experiences of family members usually covered under the same health insurance plan. The main exception is families in which spouses are covered under separate plans. Here, we allowed the FIU informant to answer for his or her spouse's plan.

3. Individuals

The FIU informant answers questions about the FIU and about the health care situation and experiences of each adult FIU member and about one child (if the FIU included children). For FIUs containing more than one child, one was randomly selected.[5] (A "child" was defined as an unmarried individual younger than 18). Full-time college students were treated as adults in the survey; that is, they were asked all the questions asked of adults and could not be the randomly selected child.

[5]Selection in Round Two was random within an FIU if the FIU contained no children interviewed in Round One. If an FIU contained one child for whom data were collected in Round One, that child was selected for Round Two. In the rare case in which a Round Two FIU included two or more children who had been in different Round One FIUs in that household, we randomly selected one child. For example, assume a Round One household included two FIUs, each with a child. Suppose one FIU included a grandmother and grandchild (10 years old), and the second the grandmother's daughter (22 years old) and her child (4 years old). Assume that the grandmother and both grandchildren are still in the household for the Round Two interview, but that the 22-year-old daughter has left. Then, there would be one Round Two FIU with two children who had been selected in Round One. The Round Two procedure would be to randomly pick one of the children.

Each adult also was asked to answer a subset of subjective questions, including assessments of health, tobacco use, chronic conditions, and satisfaction with care and with aspects of the physician-patient interaction. These questions are described in Chapter III.

4. Individuals Excluded from the Survey

The CATI survey instrument imposed a maximum of eight people per household for inclusion in the survey. The household informant identified all members of the household; in the rare instance of households exceeding eight people, interviewers were instructed to first list all the adults in the household, and then list as many children as possible up to the maximum.

Some household members were classified as ineligible and were not included on the file. To avoid giving unmarried full-time college students multiple chances of selection, they were excluded from sampled dwellings in which their parents did not reside. Unmarried children younger than age 18 with no parent or guardian in the household also were excluded. Adults on active military duty were classified as ineligible; however, they could have acted as an FIU informant as long as there was at least one civilian adult in the family. FIUs in which all adults were active duty military personnel were considered ineligible for the survey.

Some FIUs (those listed by but not including the household informant) did not respond to the interview. Nonresponding FIUs were excluded from the file but were statistically represented by responding families. A small number of people had such high levels of nonresponse to individual survey questions that they were considered to be nonrespondents and so were excluded from the file. For Round Two, 16 people were excluded for this reason. They, too, were statistically represented by survey respondents. Adult family members who did not respond to the self-response module were included on the file as long as the core interview contained responses for them.

C. SAMPLE SIZE CONSIDERATIONS FOR THE HOUSEHOLD SURVEY

Research objectives drive sample size and design requirements for any survey. For the CTS, the objectives include describing and analyzing change at the site level, describing and analyzing subgroups of special interest, making cross-site comparisons of communities, and producing national estimates. In this section, we review the original sample size considerations related to the CTS Household Survey and the selection of FIUs and individuals. We discuss sample size requirements for (1) site-based estimates for measuring change over two interview rounds, and for making cross-site comparisons; and (2) national estimates and comparisons. Finally, we include tables showing the actual number of FIUs and people that were interviewed, by site and by sample.

1. Requirements for Site-Based Estimates

The design called for the capability to make point-in-time estimates and to measure change over time. In the 12 high-intensity sites, the design called for interviews with approximately 1,225 FIUs (combined RDD and field samples) in each site per round. In addition, the supplemental sample was expected to provide additional FIUs for each site (approximately 25 FIUs per round per high-intensity site, with the number varying by size of site).

The following design issues were considered in estimating the sample size requirement of 1,225 FIUs per round of interviewing for each of the 12 high-intensity sites:

- Minimizing design effects resulting from clustering of multiple FIUs within households and from sampling methods for coverage of nontelephone households[6]

[6]For some surveys, a simple random sample variance formula may approximate the sampling variance. However, the CTS sample design is complex, and the simple random sample variance would underestimate the sampling variance substantially. Departures from a simple random sample design result in a "design effect" that is defined as the ratio of the sampling variance, given the actual sample design, to the sampling variance of a hypothetical simple

- Allowing for analyses of subgroups of interest
- Measuring and testing hypotheses about change over two interviewing rounds
- Making cross-site comparisons

As a basis for estimating sample size within sites, we used a simple random sample of 400, which permits descriptions of binomial attributes with 95 percent confidence limits no greater than five percentage points from the estimate. If all or a portion of a sample is clustered, or if portions of the sample are over- or underrepresented, design effects resulting from clustering and weighting would decrease the effective sample size (the number of observations in a simple random sample with equivalent precision) from the nominal sample of 400 to less than 400. Therefore, we increased the nominal sample size to achieve an effective sample of 400. We projected that the effects of within-household clustering of the telephone sample would produce design effects of approximately 1.25, requiring a nominal sample size of approximately 500 to result in an effective sample size of 400.

A goal of the CTS is tracking change over time and testing hypotheses about the causes of change. Measuring change over multiple interviewing rounds requires larger samples than does conducting cross-sectional surveys. For the CTS, we developed a mixed longitudinal/cross-sectional design that assumed the second round of the Household Survey would include a mix of households interviewed for the first time and households that had previously been interviewed in Round One (Metcalf et al. 1996). We also assumed that approximately 40 to 45 percent of the households interviewed in Round Two would have been interviewed in the first round. To measure changes over time (say, five percentage points for a midrange percentage), we estimated

---

(continued)
random sample with the same number of observations. Sampling error estimation methods are discussed in Chapter V.

that an effective sample of about 975 per round would provide adequate power (70 percent power for a two-tailed test at the 95 percent confidence level). After compensating for design effects of approximately 1.25, this calculation produced a target nominal sample size of approximately 1,225 FIUs.[7]

The sample size required to describe differences in the attributes of two sites is identical to that required to compare *independent* cross-sections for a single site. For these comparisons, an effective sample size of 975 is sufficient to detect differences of five to six percentage points with 70 percent power (assuming a two-tailed test and a 95 percent confidence level). We concluded that an effective sample size of 975 per site for each interviewing round, *combined with a mixed longitudinal/cross-sectional design over time,* was an appropriate sample size for each of the 12 high-intensity sites. Again assuming a design effect of 1.25 from clustering of FIUs within households and weighting for nonresponse, an effective sample of 975 would be produced by a nominal sample size of about 1,225 FIUs.

For low-intensity sites, the sample sizes available did not allow for precise individual site-level analyses. We initially set a sample target of 375 FIUs per site but reduced the target slightly to allocate more data collection resources to obtaining higher response rates.

2. National Estimates, the Second-Tier Sample of Sites, and the Supplemental Sample

We also wanted to track changes in a way that would enable us to make statements about the nation, as well as about how individual sites compare with the nation. From this national sampling perspective, a sample of 12 metropolitan sites with populations of 200,000 or more

[7]This calculation assumed gains in precision from overlap at the FIU level. Specifically, the assumption was that the effective sample size per round required with a partial overlap would be 78 percent as large as that required with no overlap. If there is any positive correlation over time, a design with some overlap will have a greater precision for estimates of change than will a design with no overlap (independent samples). The gain in precision depends on the degree of overlap and the correlation between observations at the two points in time.

would restrict sample inferences to the population in metropolitan areas of that size and would therefore result in poor precision for national estimates.

A sample of 60 sites would increase the precision for large metropolitan areas and would expand the generalizability of the household sample to small metropolitan areas and to nonmetropolitan areas. In addition, we decided to augment the clustered site sample with an unclustered telephone sample of the entire nation. The supplemental sample would not be subject to any site-cluster-based design effects and was the most efficient method of expanding the effective size of the national sample. The initial unclustered sample size of the supplemental sample was approximately 3,500 FIUs, which we later reduced slightly; a total of 3,276 FIUs were interviewed from the first round and 3,251 from the second.

The sample design also included a field sample to increase representation of FIUs and individuals that had little or no chance of being selected as part of the RDD sample because they lacked telephone service or had frequent disconnections of their service. This population represents approximately five percent of all U.S. households. Although we concluded that a field sample was necessary, it entails much higher costs than does an RDD sample. We therefore limited the field sample to the 12 high-intensity sites, thereby representing nontelephone households in MSAs with a 1992 population of 200,000 or more. For reasons of cost, we rejected extending the field sample to represent small metropolitan areas and nonmetropolitan areas. For those areas, we developed alternative weighting procedures to represent households with intermittent telephone service (discussed in Chapter V).

3. Actual Sample Sizes

The number of FIUs and the number of individuals interviewed, by site and type of sample, are shown in Tables II.1 and II.2, respectively. A total of 32,047 FIUs and 58,596 people were interviewed in Round Two, compared with 32,732 FIUs and 60,446 people in Round One. The

TABLE II.1

NUMBER OF FAMILY INSURANCE UNITS INTERVIEWED, BY SITE AND BY SAMPLE

	Sample			
Site/Geographic Area[a]	Site	Supplemental	Augmented Site	Combined
High-Intensity Sites				
01-Boston (MA)	1,127	51	1,178	1,178
02-Cleveland (OH)	1,189	27	1,216	1,216
03-Greenville (SC)	1,343	11	1,354	1,354
04-Indianapolis (IN)	1,235	21	1,256	1,256
05-Lansing (MI)	1,222	6	1,228	1,228
06-Little Rock (AR)	1,322	5	1,327	1,327
07-Miami (FL)	1,199	24	1,223	1,223
08-Newark (NJ)	1,253	25	1,278	1,278
09-Orange County (CA)	1,134	26	1,160	1,160
10-Phoenix(AZ)	1,276	32	1,308	1,308
11-Seattle (WA)	1,031	24	1,055	1,055
12-Syracuse (NY)	1,202	5	1,207	1,207
Low-Intensity Sites				
13-Atlanta (GA)	264	36	300	300
14-Augusta (GA/SC)	286	3	289	289
15-Baltimore (MD)	292	33	325	325
16-Bridgeport(CT)	269	6	275	275
17-Chicago (IL)	299	78	377	377
18-Columbus (OH)	274	22	296	296
19-Denver (CO)	274	45	319	319
20-Detroit (MI)	279	58	337	337
21-Greensboro (NC)	250	10	260	260
22-Houston (TX)	277	50	327	327
23-Huntington (WV/KY/OH)	288	12	300	300
24-Killeen (TX)	294	3	297	297
25-Knoxville (TN)	292	10	302	302
26-Las Vegas (NV/AZ)	288	14	302	302

TABLE II.1 *continued*

Site/Geographic Area[a]	Sample			
	Site	Supplemental	Augmented Site	Combined
27-Los Angeles (CA)	295	115	410	410
28-Middlesex (NJ)	288	21	309	309
29-Milwaukee (WI)	270	24	294	294
30-Minneapolis (MN/WI)	311	38	349	349
31-Modesto (CA)	323	7	330	330
32-Nassau (NY)	318	36	354	354
33-New York City (NY)	306	50	356	356
34-Philadelphia (PA/NJ)	291	53	344	344
35-Pittsburgh (PA)	294	24	318	318
36-Portland (OR/WA)	337	20	357	357
37-Riverside (CA)	321	30	351	351
38-Rochester (NY)	357	10	367	367
39-San Antonio (TX)	280	29	309	309
40-San Francisco (CA)	256	26	282	282
41-Santa Rosa (CA)	272	2	274	274
42-Shreveport (LA)	290	9	299	299
43-St. Louis (MO/IL)	338	21	359	359
44-Tampa (FL)	252	31	283	283
45-Tulsa (OK)	332	9	341	341
46-Washington (DC/MD)	306	77	383	383
47-W Palm Beach (FL)	241	12	253	253
48-Worcester (MA)	315	3	318	318
49-Dothan (AL)	330	2	332	332
50-Terre Haute (IN)	268	1	269	269
51-Wilmington (NC)	275	4	279	279
52-W-Cen Alabama	326	2	328	328
53-Cen Arkansas	380	12	392	392
54-N Georgia	261	11	272	272
55-NE Illinois	287	3	290	290
56-NE Indiana	273	5	278	278

TABLE II.1 *continued*

Site/Geographic Area[a]	Sample			
	Site	Supplemental	Augmented Site	Combined
57-E Maine	317	10	327	327
58-E North Carolina	290	8	298	298
59-N Utah	413	3	416	416
60-NW Washington	324	2	326	326
Areas Other than CTS Sites	—	1,904	—	1,904
Total	**28,796**	**3,251**	**30,143**	**32,047**

[a]Definitions of site boundaries are included in Metcalf et al. (1996).

TABLE II.2

NUMBER OF PEOPLE INTERVIEWED, BY SITE AND BY SAMPLE

Site/Geographic Area[a]	Sample			
	Site	Supplemental	Augmented Site	Combined
High-Intensity Sites				
01-Boston (MA)	2,007	83	2,090	2,090
02-Cleveland (OH)	2,116	51	2,167	2,167
03-Greenville (SC)	2,574	23	2,597	2,597
04-Indianapolis (IN)	2,274	38	2,312	2,312
05-Lansing (MI)	2,258	14	2,272	2,272
06-Little Rock (AR)	2,465	13	2,478	2,478
07-Miami (FL)	2,065	40	2,105	2,105
08-Newark (NJ)	2,263	45	2,308	2,308
09-Orange County (CA)	2,057	45	2,102	2,102
10-Phoenix(AZ)	2,310	64	2,374	2,374
11-Seattle (WA)	1,792	40	1,832	1,832
12-Syracuse (NY)	2,184	5	2,189	2,189
Low-Intensity Sites				
13-Atlanta (GA)	488	59	547	547
14-Augusta (GA/SC)	542	5	547	547
15-Baltimore (MD)	520	61	581	581
16-Bridgeport(CT)	506	11	517	517
17-Chicago (IL)	551	140	691	691
18-Columbus (OH)	532	44	576	576
19-Denver (CO)	501	84	585	585
20-Detroit (MI)	525	98	623	623
21-Greensboro (NC)	471	20	491	491
22-Houston (TX)	520	105	625	625
23-Huntington (WV/KY/OH)	556	19	575	575
24-Killeen (TX)	561	4	565	565
25-Knoxville (TN)	545	18	562	562
26-Las Vegas (NV/AZ)	510	24	534	534

TABLE II.2 *continued*

Site/Geographic Area[a]	Sample			
	Site	Supplemental	Augmented Site	Combined
27-Los Angeles (CA)	516	203	719	719
28-Middlesex (NJ)	555	41	596	596
29-Milwaukee (WI)	487	47	534	534
30-Minneapolis (MN/WI)	607	70	677	677
31-Modesto (CA)	615	14	629	629
32-Nassau (NY)	620	69	689	689
33-New York City (NY)	491	91	582	582
34-Philadelphia (PA/NJ)	530	95	625	625
35-Pittsburgh (PA)	512	54	566	566
36-Portland (OR/WA)	619	39	658	658
37-Riverside (CA)	621	58	679	679
38-Rochester (NY)	705	18	723	723
39-San Antonio (TX)	540	49	589	589
40-San Francisco (CA)	402	41	443	443
41-Santa Rosa (CA)	512	6	518	518
42-Shreveport (LA)	557	12	569	569
43-St. Louis (MO/IL)	627	34	661	661
44-Tampa (FL)	437	58	495	495
45-Tulsa (OK)	638	16	653	653
46-Washington (DC/MD)	558	129	687	687
47-W Palm Beach (FL)	434	20	454	454
48-Worcester (MA)	583	5	588	588
49-Dothan (AL)	619	4	623	623
50-Terre Haute (IN)	493	3	496	496
51-Wilmington (NC)	498	8	506	506
52-W-Cen Alabama	593	4	597	597
53-Cen Arkansas	723	22	745	745
54-N Georgia	498	20	518	518
55-NE Illinois	545	4	549	549
56-NE Indiana	558	13	571	571

TABLE II.2 *continued*

Site/Geographic Area[a]	Sample			
	Site	Supplemental	Augmented Site	Combined
57-E Maine	605	15	620	620
58-E North Carolina	540	18	558	558
59-N Utah	853	9	862	862
60-NW Washington	590	3	593	593
Areas Other than CTS Sites	—	3,539	—	3,539
Total	**52,974**	**5,982**	**55,417**	**58,956**

[a]Definitions of site boundaries are included in Metcalf et al. (1996).

number of nominal FIUs per augmented high-intensity site varied in Round Two from 1,055 to 1,354. This variation can be explained by the fact that sites are augmented by supplemental sample cases in proportion to their population, and by differential response rates.

In Round Two, the nominal augmented sample sizes of FIUs in the low-intensity sites ranged from 260 to 416. As shown in Table II.1, supplemental samples in large metropolitan low-intensity sites significantly increased the size of these site samples.

D. SAMPLE TRACKING AND THE LONGITUDINAL COMPONENT

The Round Two sample design makes it possible to estimate changes in the population but not to track an unbiased panel of individuals (or FIUs or households). Tracking a panel of individuals has considerable analytic appeal, but we concluded that this approach to sample tracking would be very costly and subject to differentially higher nonresponse for individuals or entire households that move between surveys. Although we had fairly complete information on addresses of Round One survey respondents at the time of the second round of the Household Survey (since survey respondents are compensated by check), we did not obtain social security numbers or other information typically used to minimize panel attrition, such as addresses of friends or relatives. We attempted to obtain these data during Round One pilot testing, but the results were too incomplete to be useful. Moreover, the time required to trace movers for whom we did not have social security numbers or information on the addresses of friends and relatives would have extended the Round Two field schedule substantially. Furthermore, given changes over time in household and FIU composition, following households or FIUs would have been extremely difficult.

Our approach to measuring changes in the population was to sample Round One telephone numbers (for the RDD component) and addresses (for the field component) for the Round Two

survey. This approach is relatively simple to implement, less costly than tracking individuals, and avoids attrition resulting from inability to locate sample members.

This approach to tracking permits researchers to estimate population changes, such as changes in the percentage of adults covered by employer health insurance. For two reasons, the precision of these change estimates is greater than if estimates were made for the independent cross-sectional samples:

1. Use of the same sites in both rounds improves the precision of estimates of change at the national level.
2. Partial overlap between rounds at the household level improves the precision of site and national estimates of change.

However, the design does not permit researchers to make unbiased estimates of change for people (or FIUs or households), as individuals were not followed if they changed telephone numbers (or addresses, for the field component).

For the RDD component, we estimated that sampling 70 percent of the telephone numbers for which an interview was completed in Round One would result in completed interviews in approximately 45 percent of Round Two households that were interviewed previously. This level of overlap is the same as that specified in the original design (Metcalf et al. 1996). That is, for an estimate of population change, the gain in precision from an overlap of 45 percent would be the same whether the overlap was achieved by sampling telephone numbers or by tracking individuals.

Although following telephone numbers for the RDD longitudinal component of the design may seem novel, the idea of following units other than the unit of observation (which, in the CTS, is the individual) is not new. The Current Population Survey retains housing units in its rotation groups, rather than follows individuals or households (Robinson 1992; and U.S. Bureau

of the Census 2000). Kish (1965) described two annual surveys (1951 and 1952) that used the same dwellings in both years, with good results for change estimates. Kish mentioned cost and practicality issues when deciding which unit to sample to achieve overlap for longitudinal studies, factors that we also considered. In addition to the CTS, another large, predominantly telephone survey—the National Survey of American Families (NSAF) (Ferraro et al. 2000)—has also used this approach.

Sampling the same telephone numbers is analogous to sampling the same dwellings. Because most people retain the same telephone number and address over a two-year interval, most of the people surveyed at those numbers or addresses will be the same. Even when the telephone number has been reassigned or the dwelling is occupied by different people, there will be some overlap; unless the neighborhood has undergone major changes, new occupants or new people assigned the telephone number are likely to have demographic characteristics similar to those of former occupants or users. Therefore, some of the statistical gains in estimates from following individuals can be obtained by following telephone numbers or addresses, rather than the individuals.

E. RANDOM-DIGIT-DIALING SAMPLE SELECTION

In this section, we describe selection of the RDD samples for the Household Survey. The RDD samples for Round One and Round Two were similar in that the same sites were included, there was a supplemental sample, and strata were defined using the same criteria in both rounds. However, because the design called for a partial overlap, the second-round RDD sample had new components. In the first round, all telephone numbers were selected for the first time, whereas the RDD sample for Round Two included three groups of telephone numbers: (1) those that had been selected in Round One *(overlap sample)*, (2) those that had no chance of selection in Round One *(new sample—new working banks)*, and (3) those that had a chance of selection in Round

One but had not actually been selected *(new sample—old working banks)*. (Working banks are defined below.)

In the rest of this section, we describe the sampling frame used to select the sample in the 60 sites and in the supplemental sample. We then discuss stratification, sample allocation, and generation and release of the RDD sample.

1. Sampling Frame

We used the Genesys Sampling System to select the RDD household sample for Round One and the new and residual samples for Round Two. The Round Two overlap sample was selected from a list of the telephone numbers that had been used in Round One.

To develop a sampling frame for a county or group of counties, Genesys first assigns each area-code/exchange combination to a unique county.[8] Assignment is based on the addresses of published telephone numbers; a published number is one that appears in a regular ("White Pages") telephone company directory. An exchange is assigned to the county by the plurality of such addresses. Although this procedure can lead to occasional misassignment of numbers (assigning a telephone household to the wrong county), the misclassification rate is very low. According to an analysis of published numbers in each of the 60 sites conducted prior to Round One, fewer than one percent of numbers assigned to any of the sites represented a household located outside that site.

Within each set of area-code/exchange combinations, Genesys defines "working banks" from which to sample telephone numbers. A *working bank* is defined as a set of 100 consecutive telephone numbers (XXX-YYY-ZZ00 to XXX-YYY-ZZ99) in which one or more numbers is a

[8]In the 10-digit telephone numbering system used in the United States (XXX-YYY-ZZZZ), the first three digits (XXX) are referred to as the area code, and the next three (YYY) as the exchange.

published residential number. Limiting the sample frame to working banks excludes approximately 3.5 percent of household numbers at any point in time (see Brick et al. 1995).

Undercoverage of household members for the Household Survey may be even less than 3.5 percent, because we selected telephone numbers at two points in time. At the second point in time, a later version of Genesys was used that included some banks that had been nonworking in the earlier version. Including the newer banks enabled us to increase the overall coverage of the sample frame.

2. Stratification and Sample Allocation

In both rounds of the Household Survey, stratification was used in the supplemental sample and in the high-intensity sites to help ensure proportionate representation. We did not stratify samples in the low-intensity sites because the samples were too small. We created five strata for the supplemental sample: one stratum for nonmetropolitan areas, and four strata of metropolitan counties in each of the four Census regions. In the high-intensity sites, we stratified geographically by such characteristics as income distribution, race/ethnicity distribution, or county, depending on the composition of a site. Strata were defined at the telephone exchange level, based on data provided by Genesys.

In high-intensity sites containing more than one county, we first stratified by county, assigning the county containing the central city of the MSA in one stratum and the other county or counties in another stratum. Next, we stratified the county containing the central city by race/ethnicity or income distributions. If that county included large black and Hispanic populations, we used both variables for stratification. If the county contained a significant fraction of only one of these population groups, or if one of these groups was dominant, we stratified by the percentage belonging to that group. For example, although Miami was approximately 18 percent black, a majority of the population was Hispanic. Therefore, we

stratified on the percentage Hispanic. For sites in which neither the black nor Hispanic population was large enough to stratify on race or ethnicity, we stratified on income. Table II.3 shows the stratification variables for the high-intensity sites.

Although the same stratification criteria were used in both rounds, some exchanges could have "switched" strata or even sites after Round One. For example, the prevalence of Hispanic households in a Miami exchange may have changed between rounds. In practice, few such changes occurred. Only 3 of the 11,250 Round One exchanges in the site sample were assigned to a different site for Round Two, and only 59 were assigned to a different stratum within a site. In the supplemental sample, 62 of the 6,935 Round One exchanges were assigned to a different stratum for Round Two. Two rules were devised to deal with telephone numbers in exchanges that switched strata or sites:

1. If a telephone number was part of the overlap sample, it retained its Round One site and stratum.
2. If a telephone number was selected for the first time in Round Two but was part of an "old" working bank (one that had been working in Round One), it was assigned to the site and stratum to which its exchange belonged in Round One.

To determine the initial allocation of telephone numbers for each site or for the supplemental sample, we considered the projected household prevalence among generated telephone numbers, or "hit rate," in each site (or supplemental sample) and the expected response rate for each type of sample. Telephone numbers within sites were sampled with the goal of achieving equal probabilities of selection across strata. The initial allocation of telephone numbers was later adjusted on the basis of actual experience during the survey. Thus, if either the percentage of sampled telephone numbers that was residential or the response rate in a site

TABLE II.3

RANDOM-DIGIT-DIALING SAMPLE STRATA FOR HIGH-INTENSITY SITES

Site	Number of Strata	Stratifying Variables
Boston (MA)	3	Central city county (Suffolk) vs. remainder of site; within Suffolk, percentage black/Hispanic (0-49, 50-100)
Cleveland (OH)	3	Central city county (Cuyahoga) vs. remainder; within Cuyahoga, percentage black/Hispanic (0-49, 50-100)
Greenville (SC)	3	Central city county (Greenville) vs. remainder; within Greenville, percentage black (0-29, 30-100)
Indianapolis (IN)	3	Central city county (Marion) vs. remainder; within Marion, percentage black (0-49, 50-100)
Lansing (MI)	3	Central city county (Ingham) vs. remainder; within Ingham County, percentage with annual income $35,000 or higher (0-54, 55-100)
Little Rock (AR)	3	Central city county (Pulaski) vs. remainder; within Pulaski, percentage black (0-39, 40-100)
Miami (FL)	2	Percentage Hispanic (0-49, 50-100)
Newark (NJ)	3	Central city county (Essex) vs. remainder; in Essex, percentage black/Hispanic (0-49, 50-100)
Orange County (CA)	2	Percentage Hispanic (0-44, 45-100)
Phoenix (AZ)	3	Pinal County vs. Maricopa County; within Maricopa, percentage Hispanic (0-34, 35-100)
Seattle (WA)	3	Central city county (King) vs. remainder; within King, percentage with annual income $50,000 or higher (0-49, 50-100)
Syracuse (NY)	3	Central city county (Onondago) vs. remainder; within Onondago, percentage with annual income $35,000 or higher (0-49, 50-100)

was different than expected, the allocation of telephone numbers was adjusted to obtain the desired number of interviews.

We also varied the allocation of sample among overlap, old, and new working banks. For Round Two, all Round One telephone numbers were selected into the overlap sample at the same rate (75 percent), except for Round One *breakoffs*, which were sampled at half the rate of other overlap numbers.[9] For each low-intensity site, each stratum in a high-intensity site, and each stratum of the supplemental sample, we:

- Estimated the expected number of completed interviews (FIUs) from the overlap sample
- Estimated the sample size that should be generated from old working banks, given our goal of approximately equal probabilities of selection, and estimated the number of completed interviews from that sample
- Estimated the sample size that should be generated from new working banks to give numbers in these new banks the same probability of selection for Round Two as for cases in old working banks
- Estimated the number of completed interviews expected from the new working bank sample
- If the estimated number of completed interviews from the overlap and new working bank sample was less than the target number of interviews, calculated the additional amount of sampled telephone numbers needed from the old working bank[10]

[9]Breakoffs were households that began the Round One interview and refused before completing it. We undersampled this group because their Round Two response rates were extremely low.

[10]This process was iterative, as sampling from old working banks changes the probability of selection for Round Two for the overlap sample, which requires adjustment to the new working bank sample, and so on. Fortunately, the iterations converged satisfactorily after two or three attempts.

3. Sample Selection and Release

The initial sample was set at 60 percent of the total number of projected telephone numbers. The initial sample was released during August, September, and October of 1998. Subsequent sample releases were made for all sites and the supplemental sample to meet sample size and response rate targets. (See Table II.4 for sample releases.) Toward the end of the survey, sample selection was tailored to meet interviewing targets in specific sites or groups of sites. The steps taken in selecting and releasing the sample included:

- Generating samples of telephone numbers
- Removing known business and nonworking numbers from the sample, using Genesys identification procedures
- Checking against prior releases for duplicates
- Randomly sorting the sample
- Releasing sample to the automated call scheduler
- Using data collection reports to reestimate the size of future releases

The Genesys system uses systematic selection after a random start to select equal-probability RDD samples of telephone numbers for a sample release. Thus, if Genesys selects 1,000 numbers in the nonmetropolitan stratum of the supplemental sample, all these numbers will have the same probability of selection. This method of sample generation is described more fully in documentation available from Marketing Systems Group (1994 and 2000).

The overlap sample was selected randomly from Round One sample within each stratum (in high-intensity sites and the supplement) or low-intensity site. The new and old working bank samples were selected using the procedures presented above. Although the overlap sample had been generated in Round One using Genesys procedures, some area code designations had changed between rounds. Area codes for the overlap sample were therefore updated prior to the Genesys identification procedure.

TABLE II.4

RELEASE OF SAMPLE FOR ROUND TWO OF THE COMMUNITY TRACKING STUDY HOUSEHOLD SURVEY

Date	Total RDD	Round One Completes	Round One Breakoffs	Round One Other	New or Old Working Banks	Total Field Sample
August 1998	10,397	1,962	4	5,541	2,890	0
September 1998	17,390	5,369	484	6,513	5,024	0
October 1998	16,924	3,847	0	5,075	8,002	0
November 1998	0	0	0	0	0	5,432
December 1998	7,215	2,192	0	5,023	0	0
January 1999	0	0	0	0	0	84
February 1999	12,650	3,906	0	0	8,744	0
March 1999	5,320	0	0	0	5,320	878
May 1999	3,153	0	0	0	3,153	273
August 1999	1,299	526	0	773	0	0
TOTAL	**74,348**	**17,802**	**488**	**22,925**	**33,133**	**6,667**

The Genesys identification procedure involved two steps: (1) checking the sample against lists of published numbers, and (2) dialing numbers to determine whether they were nonworking. In the first step, all numbers were classified as published residential numbers, published business numbers, or other. The published residential numbers were retained, the business numbers eliminated, and the others prepared for dialing. Genesys used an automated dialer to check for the tone that precedes a recorded message stating the number dialed was not in service (termed an *intercept message*). If that tone was detected, the number was removed from the sample as nonworking. To minimize intrusiveness, the Genesys dialer would disconnect immediately if a ring was detected, and calls were made only between the hours of 9:00 a.m. and 5:00 p.m. local time. The remaining sample included numbers identified as published residential plus those not classified by the dialer as nonworking.[11]

Each RDD sample release in Table II.4 was randomly sorted before being released, as Genesys samples are ordered by area code and exchange. Randomizing ensured that each release was worked evenly and eliminated the need for sample replication. We also checked for duplicates against previously released sample. By checking against prior releases, rather than checking against the entire generated sample, we avoided eliminating numbers that Genesys may have eliminated during an earlier release, but that subsequently became working. The sample as then released to the CATI call scheduler; weekly survey reports on sample dispositions, by site, were used to determine the size of additional sample releases.

[11]In Round One, the Genesys procedure eliminated 14 percent of numbers generated. By calling a small sample of numbers eliminated by Genesys, we determined that, on average, 1.0 to 1.5 percent of those eliminated were residential numbers. For the second round, the Genesys procedure eliminated 16 percent. The difference is due to two factors. The use of overlap sample served to decrease the percentage eliminated, whereas the increase in telephone numbers in the frame served to increase it.

F. FIELD SAMPLE SELECTION

The Household Survey included a field sample to provide coverage of families and people who did not have telephones or who had substantial interruptions in telephone service. Several studies have indicated that omitting nontelephone households might lead to biased survey estimates (Thornberry and Massey 1988; Marcus and Crane 1986; and Corey and Freeman 1990). A "dual-frame" design similar to the one used in the CTS Household Survey was used for the RWJF Family Health Insurance Survey (Hall et al. 1994). Strouse et al. (1997) found that telephone-only estimates would bias survey estimates for several demographic variables (particularly income), health insurance coverage, and some satisfaction measures. However, biases for most of these measures are small because telephone coverage is high even across most vulnerable population groups; exceptions include Medicaid and Indian Health Service beneficiaries. Using first-round results from the CTS and the NSAF showed that a telephone-only approach could bias estimates for measures of health care utilization, insurance coverage, and economic status (Hall et al. 2000).

Restricting the field sample to the 12 high-intensity sites reduced some of the coverage bias for estimates for all large metropolitan sites as a whole and for estimates made for each of those sites that would result from using an RDD-only methodology. This option was far less expensive than collecting data through field interviewing in all 60 sites. However, limiting the field sample to the 12 high-intensity sites meant that families and people who did not have telephones and who lived in nonmetropolitan areas or in metropolitan areas with populations of fewer than 200,000 were not represented. (Weighting procedures to adjust for the absence of these households in national and other estimates are discussed in Chapter V.) The field sample was a geographically clustered sample that was initially designed to yield responses from 576 FIUs (635 FIUs were actually interviewed in the first round and 791 in the second).

Within the 12 high-intensity sites, the strategy was to sample geographic clusters with probability proportional to size; count, list, and select housing units within these clusters; and screen this sample for eligible households (defined below).[12] Respondents in eligible households were then interviewed over cellular telephones, which were provided by MPR field staff. Thus, all interviews were conducted by CATI methods, which prevented differential response resulting from different interviewing modes.

Selection of the field procedures was similar for both rounds of the Household Survey. The Round One report by Strouse et al. (1998) describes procedures for determining the Round One sample allocation among the 12 sites, identifying areas within the 12 sites for exclusion, establishing a measure of size for selecting clusters, stratifying clusters by county and by tract number within county, selecting clusters and listing areas, and listing addresses. Here, we discuss changes in Round Two procedures for defining field sample eligibility, allocating the sample among sites, and selecting addresses.

1. Defining Eligibility

In defining eligibility, the term *nontelephone household* meant that the household was always or intermittently without telephone service. The field component was designed to include these households. In contrast, in the approach used by the decennial Census and the Current Population Survey, households were classified as telephone or nontelephone on the basis of the presence or absence of a telephone at the time of interview.[13]

[12]The measure of size was the estimated number of nontelephone households.

[13]The Census estimates of prevalence of nontelephone households were based on a question on the "long form," asked of a large sample of decennial Census households. Question H12 asked, "Do you have a telephone in this house or apartment?"

We originally had planned to use the Census definition as a screening criterion, and to interview only households that did not have working telephones when first contacted by a field interviewer. However, based on experience in the RWJF Family Health Insurance Survey (Hall et al. 1994), and on research reported by Brick et al. (1995), we concluded that this static approach to defining telephone status would result in limitations for the CTS. The main limitation of the Census approach is its exclusion of households with substantial periods of interrupted telephone coverage that have telephone coverage at the time of the screening call. Although these households would have had a chance of being included in the telephone survey, we determined that they would have been underrepresented. The field sample for both rounds of the Household Survey therefore included any household with a history of significant interruption in service since the beginning of interviewing for the RDD sample (July 1, 1996, for Round One, and July 1, 1998, for Round Two).

We defined *significant interruption* to mean two weeks or more of interrupted service since the beginning of the RDD field period (or the date the household moved into the house, if the move occurred after that time) and used questions about the length of interruptions to adjust sample weights.[14] The only exception to the two-week rule was that households also were eligible for the field survey if members had moved to the listed address within the two weeks preceding the interview and had been without a telephone since moving in.

2. Allocation of the Sample Among the Sites

The field allocation model selected for the 12 high-intensity sites for Round One is shown in Table II.5. The Round One allocation was based on considerations of cost, sampling error, and potential coverage bias (Strouse et al. 1998).

[14]The use of these questions in weighting is discussed in Chapter V.

TABLE II.5

TELEPHONE PENETRATION, ESTIMATED NUMBER OF NONTELEPHONE HOUSEHOLDS, AND ROUND ONE PRELIMINARY FIELD ALLOCATIONS

Telephone Penetration	Households Without Telephone[a] (Percent)	Nontelephone Households[b] (Number)	Preliminary Field Allocation[b] (Number)
High Penetration			
Boston (MA)	1.9	30,456	21
Orange County (CA)	1.5	12,808	17
Seattle (WA)	2.0	15,298	22
Medium-High Penetration			
Cleveland (OH)	3.7	32,107	41
Lansing (MI)	3.2	5,078	36
Newark (NJ)	3.9	27,085	44
Syracuse (NY)	4.0	10,866	45
Medium-Low Penetration			
Indianapolis (IN)	5.0	26,340	56
Miami (FL)	5.0	34,652	56
Low Penetration			
Greenville (SC)	8.1	25,339	91
Little Rock (AR)	7.0	13,728	78
Phoenix (AZ)	6.2	52,656	69
Total	**—**	**—**	**576**

[a]Based on 1990 Census data, using Census definitions.

[b]Expected FIU interviews.

We reviewed our Round One experience and, for the most part, based second-round allocations on the actual number of Round One households interviewed in the site. An exception was Greenville, South Carolina, where the number of interviews during the first round was far below target. In Greenville, we set the target equal to the Round One target (expressed as household completes). As discussed below, we decided to list additional housing units to bring the number of interviews closer to the original (Round One) target.[15] The Round Two target allocations and completed interviews are shown in Table II.6.[16]

3. Selecting Sample for Round Two

For Round Two, we contacted all addresses listed during Round One. Our rationale for selecting those addresses for the field sample was the same as that for selecting telephone numbers for the RDD overlap sample. For the field component, we sampled all Round One addresses (without regard to Round One interview status) and supplemented them where we projected that we would fall short of the Round Two target. We did not include new areas in the Round Two sample. We assumed that dwellings found in areas that had no chance of selection in Round One would most likely be housing constructed since Round One, and that they would have a low likelihood of including households without telephones.

[15]We use the Census definition of a housing unit, that is, a structure that is occupied or intended for occupancy by person(s) living separately from other person(s) in the building and must meet one of the following criteria: (a) It has complete kitchen facilities for the exclusive use of that unit regardless of whether the kitchen is used or not, or (b) The housing unit has a separate entrance directly from the outside of the structure or through a common or public hall, lobby or vestibule.

[16]The Round Two targets were expressed in terms of households rather than FIUs because we had found in Round One that the number of FIUs per household varied substantially between sites.

TABLE II.6

FIELD ALLOCATIONS AND COMPLETES FOR ROUND TWO

	Household Interviews	
Site	Target	Completed
Boston (MA)	9	5
Cleveland (OH)	36	55
Greenville (SC)	65	96
Indianapolis (IN)	70	74
Lansing (MI)	22	26
Little Rock (AR)	64	69
Miami (FL)	34	19
Newark (NJ)	48	53
Orange County (CA)	7	7
Phoenix (AR)	54	108
Seattle (WA)	49	15
Syracuse (NY)	42	35
Total	**500**	**562**

We projected falling short of our Round One targets in Lansing, Miami, and Greenville. We therefore selected and listed additional listing areas in those sites but screened new housing units only in Lansing and Miami.[17]

During Round One, each sampled cluster had been divided into 10 replicates, each of which contained approximately one-tenth of the cluster's estimated nontelephone households. A replicate might contain several blocks; at the other extreme, one large block might comprise several replicates. The Round One release of blocks for listing consisted of a number of replicates chosen to minimize variation of overall probability of selection within a site, subject to listing enough housing units overall to meet projected targets.

4. Sample Release

The initial release took place in November 1998 and was based on the estimated amount of sample required to meet our targets. Subsequent releases were based on shortfalls in specific sites. The sample release of the field component is summarized in Table II.4.

[17]Based on Round One results, we estimated that we would need additional sample in Greenville. However, the number of Round Two completed interviews in Greenville was well above expectations, eliminating the need for more sample.

LEFT INTENTIONALLY BLANK

III. SURVEY DESIGN AND PREPARATION

A. OVERVIEW

The CTS Household Survey was the primary instrument for assessing the effects of health system change on individuals, including changes in health insurance coverage, access to care, use of health services, and satisfaction with health care. As described in Chapter II, the FIU was the primary interviewing unit for the survey, with selected subjective questions also asked of each adult FIU member. Within each FIU, questions were asked about all adults and about one randomly selected child. An adult familiar with the health care experiences of the individuals in the FIU served as the informant for other adults on questions about health insurance, employment, demographics, and health services use during the 12 months preceding the survey. The other adults in the FIU were asked to self-respond to questions about health status, chronic diseases, tobacco use, details about their last physician visit, level of satisfaction with that visit, and level of trust in their physician. The adult who took the randomly selected child to the last physician visit before the survey was asked questions about that visit, the child's health status, and any chronic disease the child had.

The length of the interview varied with the number of individuals in the FIU and the complexity of the individuals' experiences with health care. Including self-response modules, the Round Two interview averaged 37.2 minutes per FIU, approximately 3 minutes longer than in Round One (34.4 minutes). For Round Two FIUs that included one adult, the interview averaged 27.6 minutes; for those with two or more adults, the average length was 47.7 minutes.

The Round Two instrument included most of the Round One questions. Four major changes were made to the Round Two instrument:

1. A battery of questions on chronic diseases was added to support the Community Quality Index (CQI) Survey, which was conducted with a sample of Round Two CTS households after they completed interviews in the high-intensity sites. The CQI survey was designed and conducted by researchers at the RAND Corporation to obtain data on health outcomes; it included additional health questions and a request for medical records (project report forthcoming).
2. The CTS Household Survey health insurance questions were restructured to ask for multiple types of insurance coverage from individuals who had been asked in Round One only for their primary health insurance coverage.
3. Prompts were built into the CATI program to identify insurance plans included in a database created from the Round One Followback Survey, which was designed to link policies reported by household survey respondents to insurance products. We used the database to aid the Round Two survey respondents' recall of insurance policy names by listing the names of insurers and insurance products that *may have* been offered in the respondents' area.
4. Prompts were added to determine whether individuals currently in households linked to telephone numbers selected for both rounds were still members of the Round Two household. Questions were added to verify the name, relationship, gender, and age of persons in the Round One household.[1]

A few changes also were made to questions on resource use and the usual source of care.

We obtained endorsements or support from state and local health agencies for Round One; these endorsements were included in survey introductions and in letters mailed to sample members whose addresses were known. Efforts to develop a convincing survey introduction in Round One that included agency endorsements had little impact on response rates and were abandoned for Round Two.

For Round Two, we mailed advance letters and a brochure describing Round One results to households with published addresses. This group included nearly all the households that participated in Round One and that were reinterviewed. (All these households had been mailed checks for participating in Round One, so we had accurate addresses for more than 95 percent of

[1]Although the reinterview component is not an unbiased panel, we could use information linking persons participating in both rounds to determine whether there were gains in precision from reinterviewing at the person level.

them.) The group also included households that had not been interviewed in Round One but that had published addresses.

Altogether, 239 telephone interviewers (224 from MPR and 15 from Battelle) and 20 field interviewers were trained to conduct the survey. Telephone interviewers who had not previously conducted surveys received 12 hours of general training on interviewing methods, and all interviewers received a minimum of 12 hours of training on the survey instrument, supplemented by training on methods to avoid refusals. Field staff were trained in a two-hour session on screening households eligible for the nontelephone component of the survey. It was not necessary to train field staff on how to administer the survey instrument—after identifying eligible households, these staff used cellular telephones to call the MPR telephone center, from which an interviewer conducted the survey.

In the rest of this chapter, we discuss in detail the design of the Round Two survey instrument, preparation for the survey, interviewer selection and training, and the CATI system used to collect the data.

B. INSTRUMENTATION

The survey instrument for Round One was developed by staff at HSC and MPR, with consultation and review by several experts.[2] The questions on chronic diseases for Round Two were developed by RAND researchers responsible for the CQI Survey; other changes to the Round Two instrument were made by HSC and MPR staff.

Respondents to the survey were questioned about the following topics:

- Household composition

[2]See Chapter III in Strouse et al. (1998) for a discussion of the content of the Round One instrument design.

- Health insurance coverage
- Use of health services
- Unmet needs and expenses
- Usual source of care
- Patient trust and satisfaction
- Last visit to a physician or other healthcare provider
- Health status (and SF-12)
- Presence of chronic diseases
- Risk behaviors and smoking
- Employment, earnings, and income
- Demographic characteristics

The content of the Round Two instrument (organized by topic) is summarized in Table III.1; the sections of the interview in which these questions were asked are noted in parentheses. An English version of the Round Two instrument is included in Appendix A; English and Spanish versions of the CATI program are available on request from MPR.

Different respondents were asked different questions, and not all questions were asked of all respondents. For example, only the household informant was asked about household composition. Family informants were asked about the insurance coverage, service use, usual source of care, general health status, employment, earnings, income, and demographic characteristics of all FIU members selected for the survey. As part of the survey's self-response module, each adult, not just the family informant, provided information on unmet needs, patient trust, satisfaction, last visit, health status, chronic diseases, and risk behaviors.

To illustrate which sections of the survey were asked of various household members, we will use the hypothetical household shown in Figure II.1. In that example, the household head's spouse is the household informant because she answered the telephone and is

TABLE III.1

CONTENT OF THE HOUSEHOLD SURVEY

Health Insurance	
Private insurance coverage (Section B)	Covered by employer- or union-related private insurance Covered by other private insurance: Purchased directly Premium for directly purchased private insurance Provided by someone not in household
Public insurance coverage (Section B)	Covered by Medicare Covered by both Medicare and supplemental private insurance Premium for supplemental private insurance Covered by both Medicare and Medicaid Covered by Medicaid Covered by other public insurance (military, Indian Health Service, other state and local)
Uninsured (Section B)	Not covered by public or private insurance
Continuity of coverage/changes in coverage (Section B)	Currently insured; lost coverage during previous 12 months Currently uninsured; obtained coverage during previous 12 months Uninsured during all of previous 12 months Uninsured at some point during previous 12 months Reasons for losing health insurance coverage Any type of change in health coverage: Changed private insurance plans Reasons for changing private plans Whether previous plan was HMO/non-HMO Changed from public or private plans Obtained or lost coverage
Insurance plan attributes (Section B)	Whether plan requires signing up with primary care physician or clinic for routine care Whether plan requires approval or referral to see a specialist Whether plan requires choosing a physician or clinic from a book, directory, or list Whether plan is an HMO Whether plan will pay any costs for out-of-network care
Other insurance variables (Section B)	Ever enrolled in an HMO Number of total years enrolled in an HMO

TABLE III.1 *(continued)*

Access to Health Care	
Usual source of care (Section D)	Currently has/does not have a usual source of care Type of place of usual source of care Type of professional seen at usual source of care Reason for changing usual source of care
Travel/waiting time for physician visit (Section E)	Lag time between making appointment and seeing physician at last physician visit[a] Travel time to physician's office for last visit[a] Time spent in waiting room before seeing medical person at last physician visit[a]
Difficulty getting needed services in previous year (Section C)	Did not receive needed services[a] Delay in receiving needed services[a] Reasons for delay or for not receiving needed services[a]
Resource Use	
Use of ambulatory services in previous 12 months (Section C)	Number of physician visits Number of emergency room visits Number of visits to nonphysician providers (nurse practitioner, physician assistant, midwife) Whether had any mental health visits Number of surgical procedures
Use of inpatient services in previous 12 months (Section C)	Number of overnight hospital stays Number of overnight hospital stays excluding delivery/birth Number of inpatient surgical procedures Total number of nights spent in hospital
Preventive services use (Section C)	Whether person had flu shot during previous 12 months Whether person ever had mammogram (asked of women) If yes, time elapsed since last mammogram
Nature of last physician visit (Section E)	Reason for last visit: Illness or injury[a] Checkup, physical exam, other preventive care[a] Type of physician seen at last visit (PCP or specialist)[a] Whether last visit was to usual source of care[a] Whether last visit was to an emergency room[a] Whether last visit was with appointment or walk-in[a]
Costs (Section C)	Total family out-of-pocket expenses for health care during previous 12 months
Satisfaction and Patient Trust	
General satisfaction (Section E)	Overall satisfaction with health care received by family Satisfaction with choice of primary care physicians[a] Satisfaction with choice of specialists[a]
Satisfaction with last physician visit (Section E)	Satisfaction with thoroughness and carefulness of exam[a] Satisfaction with how well physician listened[a] Satisfaction with how well physician explained things[a]

TABLE III.1 *(continued)*

Patient's trust in physicians (Section D)	Agree/disagree that physician may not refer to specialist when needed[a] Agree/disagree that physician may perform unnecessary tests or procedures Agree/disagree that physician is influenced by health insurance company rules[a] Agree/disagree that physician puts patient's medical needs above all other considerations[a]
Employment and Earnings	
Employment status and characteristics (Section F)	Whether adult respondent has the following characteristics: Owned a business or farm Worked for pay or profit during previous week Had more than one job or business Worked for private company/government/self-employed/family business Average hours worked per week, at primary job and at other jobs Size of firm (number employees), at site where respondent works, and at all sites Type of industry
Earnings (Section F)	Earnings, from primary job and from all jobs
Health insurance options at place of employment (Sections B and F)	Whether eligible for health insurance coverage by employer Reasons for ineligibility Whether offered health insurance coverage by employer Reasons for declining coverage (if eligible but not covered) Whether offered multiple plans Whether offered HMO plan Whether offered non-HMO plan
Other Variables	
Demographics (Section A)	Age Gender Highest education level completed Whether interview was administered in Spanish CTS site State County[b]
Health status (Section E)	Overall health status (five-point scale, from excellent to poor)[a] Limited in moderate activity[a] Limited in climbing stairs[a] Accomplished less because of physical health[a] Limited in kind of work because of physical health[a] Accomplished less because of emotional health[a] Less careful in work because of emotional health[a] Pain interfered with work[a] How much time have health problems interfered with social activities[a] How much time calm and peaceful[a] How much time have energy[a] How much time downhearted/blue[a] SF-12 scores: Physical Component Summary; Mental Component Summary[a]

TABLE III.1 *(continued)*

Chronic conditions (Section E)	Presence of chronic conditions, including recent childbirth, acne, headaches, abnormal uterine bleeding, alcohol-related problems, cataracts, diabetes, arthritis, asthma, pulmonary disease, atrial fibrillation, hypertension, high cholesterol, atherosclerosis, heart diseases, cancers, hernia, ulcer, gastritis, depression, HIV, AIDS, ADHD, tuberculosis, sickle cell disease, ear infections[b]
Family income (Section G)	Family income Race
Consumer preferences (Section B)	Whether person would be willing to accept limited provider choice to save on out-of-pocket expenses[a]
Risk behaviors (Section E)	Whether person agrees that he/she is more likely to take risks than the average person[a]
Smoking cessation interventions (Section E)	Whether person has smoked at least 100 cigarettes in lifetime[a] Whether currently smoking cigarettes every day, some days, or not at all[a] Average number of cigarettes smoked per day during previous 30 days[a] How long since quit smoking[a] Whether stopped smoking one day or longer during previous 12 months, in effort to quit[a] Whether physician advised smoker to stop smoking during previous 12 months[a]

SOURCE: Table 2.2 of the Round Two Household Survey Public Use File.

[a]Information was obtained from the self-response module.

[b]Available on the Restricted Use File only.

familiar with the composition of the household. Because she is also familiar with the health care of other family members, she is also the informant for the first FIU (FIU1). The household head's father is the second FIU informant (FIU2), and the unrelated boarder responds for himself or herself (FIU3). The head of the household and her husband have two children; the daughter is the randomly selected child in FIU1 and the son is not included in the survey. Table III.2 presents the topics covered in the survey and lists the individual who, according to the hypothetical household in Figure II.1, responded to the questions under each section.

Most of the CTS Household Survey interviews were obtained from the RDD sampling frame. However, we used an area probability sample in the 12 high-intensity sites to conduct additional interviews with FIUs in households with intermittent or no telephone coverage. Sampled households were administered a screening interview to identify eligible households (see Appendix A). Field interviewers then gave cellular telephones to the family respondents so that the respondents could call the telephone center. This procedure enabled interviews with nontelephone households to be conducted using the CATI program.

We modified the CATI instrument slightly for field administration. Because of the high cost of making return visits to these households, we attempted to obtain proxy information about all household members from one family informant, rather than from a separate informant for each FIU, as was done for the RDD sample. However, the field interviewer attempted to obtain answers to self-response modules from each adult in the household.

In the rest of Section B, we describe the changes made to the instrument for Round Two (included in Appendix A). Appendix B provides additional information on changes, including the question numbers, variable names in the Public Use File, the text of the questions added to the survey for Round Two, and the text of the questions deleted after Round One.

TABLE III.2

SOURCE OF DATA FOR INDIVIDUALS, BY QUESTION TOPIC

(Illustrative Household Described in Figure III.1)

Family Insurance Unit Member	Question Topic												
	Household Composition	Insurance Coverage	Resource Use/ Expenses	Unmet Needs	Usual Source Of Care	Patient Trust	Satisfaction	Last Visit	General Health Status/ Chronic Conditions	Risk/ Smoking/ SF-12	Employment/ Earnings/ Employer Plan	Family Income	Race
First Family Insurance Unit													
Family Informant	**H**	**F1**	**F1**	**F1**	**F1**	**F1**	**F1**	**F1**	**F1**	**F1**	**F1**	**F1**	**F1**
Spouse	**H**	**F1**	**F1**	**SRM**	**F1**	**SRM**	**SRM**	**SRM**	**F1 and SRM**	**SRM**	**F1**	**F1**	**F1**
Randomly Selected Child	**H**	**F1**	**F1**	**F1**	**F1**	**Not Asked**	**F1**	**FC**	**F1 and FC**	**Not Asked**	**Not Asked**	**F1**	**Not Asked**
Other Children	**H**	Data not available—not randomly selected child											
Second Family Insurance Unit													
Family Informant	**H**	**F2**	**F2**	**F2**	**F2**	**F2**	**F2**	**F2**	**F2**	**F2**	**F2**		**F2**
Spouse	**H**	**F2**	**F2**	**SRM**	**F2**	**SRM**	**SRM**	**SRM**	**F2 and SRM**	**SRM**	**F2**		**F2**
Third Family Insurance Unit													
Unrelated Adult	**H**	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**	**F3**		**F3**

H	Data provided by the household informant (typically person who answers the telephone, if adult)
F1	Data provided by family informant for FIU1
SRM	Data provided by the individual adult family member via the self-response module questions
F1/SRM	Data on general health status provided by the family informant. Detailed health information provided by the individual family member
FC	Data provided by adult in family who took randomly selected child to last physician visit

Module A. Household Composition

- ***Identifying People Interviewed in Round One and Round Two (Del-ADD).*** Because part of the Round Two sample had participated in Round One, we verified whether individuals in Round One households were still in the households at Round Two. The interviewer verified the name, relationship, gender, and age of each person included in the Round One household; deleted movers; and added new household members.

- ***Random Selection of Child.*** Although we did not add questions to select a child to be interviewed in each FIU, we used information about Round One participation to select children for Round Two. If a Round Two FIU included one child younger than age 18 who had been selected for Round One, that child was selected for Round Two. However, if a Round Two FIU included two or more children younger than age 18 who had been selected for Round One within that household, one child was selected at random for Round Two. If none of the children in a Round Two FIU had been selected for Round One, then a new child was selected at random.

- ***Households Containing More than Eight People (emo1-3).*** The CATI program used for Round One and Round Two has a limit of eight people per household. If there were more than eight people in the household, questions were added to determine the number of people age 18 or older and the number younger than age 18 who were not included in the survey. These questions were added for sample weighting.

- ***Birth Month and Year (bmo1 and byr1).*** The Round One survey obtained data on age but did not ask for the birth date. Birth month and year were added for Round Two. We did not obtain the day of the month, to protect respondent confidentiality.

- ***Part-Year Residence (snow).*** In Round One, more telephone numbers were assigned an undetermined residential disposition in sites in Florida than in other parts of the country. We believed the difference may have been due to part-year residence by people with a permanent residence in other parts of the country. A question was added to Round Two to determine whether survey respondents had another residence in which they lived for more than half the year.

Module B. Health Insurance

- ***Changes in Questions on Health Insurance Coverage.*** The following changes were made to questions on health insurance coverage after Round One:

  - We structured Round One questions b1a-c hierarchically so that a person who was covered by private health insurance from an employer or union (b1a) was skipped over questions on individual coverage (b1b) and private coverage provided by someone outside the household (b1c). Similarly, someone who had individual coverage (b1b) was skipped over b1c. This structure was dropped for Round Two, and everyone was asked about private coverage from all three sources (b1a-c).

- In Round One, people age 65 or older living in households with no one younger than 65 were skipped over questions on private health insurance plans (b1a-c) and were asked a separate question on Medigap coverage (b59). We recognized that some people age 65 or older were not covered by Medicare, and that some Round One Medicare beneficiaries may not have included private coverage in b59, so we asked all Round Two family members questions b1a-c. The Medigap coverage question (b59) also was asked of all Medicare beneficiaries.
- In Round One, families whose members all were covered by private health insurance plans were skipped over questions on Medicaid and other state-subsidized health insurance plans. All Round Two family members were asked about coverage by Medicaid and other state-subsidized health plans.
- Between Round One and Round Two, many states started to offer subsidized health insurance plans through the State Children's Health Insurance Program (SCHIP) or other sources. The CATI program was modified to prompt respondents with the names of new plans (blex, bley), and to obtain coverage by Medicaid or non-SCHIP state-subsidized health programs (ble) or SCHIP programs (b1h).
- In Round One, families whose members all were covered by private plans, Medicaid, or Medicare were skipped over questions on coverage by military health care plans (CHAMPUS, CHAMP-VA, TRICARE, VA). All Round Two family members were asked about military coverage, regardless of coverage by other plans (b1f).
- In Round One, families whose members all were covered by private plans, Medicaid, Medicare, or military health plans were skipped over the question on coverage by the Indian Health Service (IHS). All Round Two family members were asked about coverage from the IHS (b1g).

- In Round Two, informants for FIUs reporting enrollment in private health insurance plans were prompted by the CATI program with lists of insurers and insurance products offered in their area (the state of residence, if the market was contained in one site, or nearby states, if the market covered more than one state) (see b2000 to zb2251).[3]

  - If the household had participated in Round One and had only one private plan that matched an insurance product in the database, the informant was asked whether he or she still had that plan (b204). If so, no further prompting was required.

[3]After the Household Survey was completed, we conducted a separate survey to link household-reported policies to insurers and other organizations providing insurance (such as third-party administrators, unions, or self-insured employers). This step was designed to obtain information about the characteristics of insurance plans that was more detailed than household survey respondents were able to provide.

- If the informant did not have the plan, he or she was asked whether the plan had changed or whether the name was incorrect. Then, the informant was asked the name of the current plan and was prompted to obtain documentation (insurance card, claim form, or insurance policy). The plan was then matched against the database, and the program listed insurance companies and products offered in the informant's state that closely matched the text reported by the informant. (Products for nearby states were shown for MSAs whose residents may have worked in more than one state.) The interviewer read the list of insurers and products and attempted to link the database insurer and product to the informant's policy. The description of the plan provided by the informant was retained regardless of whether a match was obtained.
- If the household had not participated in Round One or had participated but reported no private health insurance policies or more than one policy, the informant was asked the name of the current plan and was prompted to obtain documentation (insurance card, claim form, or insurance policy). The program followed the same sequence as outlined above.
- These prompts were repeated for up to three private policies.

- Most of the questions about the attributes of insurance policies were the same as in Round One. However, the following changes were made:
  - Questions were added to determine whether private plans may have been obtained from federal or state programs that subsidize insurance coverage and, if so, the name of the programs (b271-281).
  - If one or more family members was uninsured at the time of the interview, but other family members were covered by private health insurance policies, we asked whether family coverage was offered under the private health insurance plan (b79). If so, the informant was asked whether uninsured family members were not covered by the private plan because the plan was too costly or for other reasons (b791).
  - Round One questions on plan attributes were dropped for Medicaid, other state-subsidized health care plans, and military health plans because researchers did not plan to use these items. These items covered whether the plan included a network, whether it required a gatekeeper, whether it required specialist referrals, and whether it reimbursed any of the costs of out-of-network coverage. For Medicaid and state-subsidized HMO beneficiaries, we also dropped the question on the name of the HMO.
  - The Round One survey included a question to identify individuals denied health insurance because of poor health. This question was dropped for Round Two.
  - The Round One question on the cost of the Medigap premium was dropped.

Module C. Resource Use During the Last 12 Months

- We added questions to determine whether any physician visits during the 12 months preceding the interview were for routine preventive care (c3p1) or for an ongoing health problem (c3c1). A similar set of questions was added to determine whether any nurse practitioner or physician assistant visits during the last 12 months were for routine preventive care (c3p1) or for an ongoing health problem (c3c1).

- The Round One question on whether any care at home was obtained from a nurse or other professional was dropped, as was the question on whether obtaining the medical care the family needs was becoming easier, becoming harder, or remained the same.

Module D. Usual Source of Care/Patient Trust

- Two questions were added to obtain more information about the reasons the respondent changed the usual source care without affecting tracking. Question d151, on whether the change was related mainly to health insurance, to the quality of care received, or to some other factor, was retained. If the change was related mainly to health insurance, we added question d161, to determine whether the change was made because the employer changed health plans, the provider or place was not covered by the health plan, or for some other reason. If the change was related mainly to the quality of care, we added question d171, to determine whether the provider or place providing the usual source of care was no longer available, whether the respondent had to see a particular type of physician, whether the respondent had recently moved, whether the respondent felt it was more convenient to go to another physician, or whether there was some other reason.

Module E. Satisfaction/Risk Behaviors

- A series of questions on chronic diseases in the two years preceding the interview was added for each adult and child (see Section E of the instrument, cc1-e4e).

C. DEVELOPMENT OF SURVEY MESSAGES, ENDORSEMENTS, AND ADVANCE MATERIALS

Notifying potential respondents by mail before an initial call is made can reassure them about a survey's authenticity and purpose. The willingness of the general public to participate in a survey may also be increased by obtaining sponsorship or endorsement from well-known public organizations (usually a government agency), and by designing a convincing survey introduction describing the survey's purpose and value. We have used advance letters, salient

survey messages, and endorsements by senior government officials in other RDD health surveys to achieve high response rates without having to offer monetary incentives. For example, we achieved a 73 percent response rate for the RDD component of the RWJF Family Health Insurance Survey, which had been conducted in 10 states during 1993-1994 (Hall et al. 1994). The purpose of that survey was to provide data states needed to plan health care reform.

During Round One of the CTS Household Survey (1996-1997), we were concerned that health tracking would be less relevant and therefore less interesting to respondents than would health reform, which dominated public debate at the time the Family Health Insurance Survey was being conducted. We also realized that the public's limited awareness of foundations and their role in supporting health research might result in lower response rates than would surveys conducted on behalf of governors' offices or state health departments (the organizations sponsoring the Family Health Insurance Survey).

We therefore took several steps during Round One to increase the perceived impact and value of the CTS Household Survey, including requesting government endorsements and contracting with a communications consultant to develop and test messages for inclusion in survey introductions.[4] We found that neither advance letters, whether on state or RWJF letterhead, nor the brochure had an impact on initial cooperation rates (defined as the ratio of completed interviews to the sum of completed interviews and initial refusals). We also found that a brief survey introduction, which provided limited information about the study's purpose and sponsorship, was more effective at increasing response rates for the CTS Household Survey than was a long, detailed one.

[4]The results of these efforts are described in detail in the Round One Household Survey report (Strouse et al. 1998).

Three different survey introductions were developed for Round Two: (1) for households with telephone numbers that had completed Round One interviews, (2) for households with telephone numbers linked to other households with published addresses, and (3) for households with telephone numbers that were not linked to published addresses. (See Appendix C for these introductions and for all other materials cited in this section.) We expected that most or all of the individuals in households sampled from telephone numbers that completed Round One interviews would not have moved. Consequently, we mailed those households a letter and brochure describing the study. The letter referenced the household's previous survey participation and promised that participants would receive $25 for participating in Round Two; the survey introduction reinforced the information provided in the letter about past participation and compensation. The size of the payment was based on experimental results from Round One (Strouse et al. 1997) and on decisions to use the CTS sample as a frame for other studies for which individuals would be compensated $25. All families in the 12 high-intensity sites completing Round Two interviews were contacted shortly after the CTS survey and were asked to release medical records for a study on the quality of care (the CQI Study).[5] In addition, a sample of CTS respondents who completed Round One interviews was selected for participation in the Health Care for Communities (HCC) study, which focuses on mental health and substance abuse issues.[6] Most of the individuals sampled for that survey lived in low-intensity sites; however, some were in high-intensity sites. Individuals selected for the CQI or HCC surveys were offered $25 for participating in each of those surveys.

[5]The Community Quality Index (CQI) study is being conducted by RAND.

[6]The HCC study is being conducted by UCLA, with the interviewing shared by MPR and RAND. The first round of that survey was conducted between the time that Round One and Round Two of the CTS Household Survey were conducted.

We also mailed a letter and brochure to households with published addresses that were participating in the CTS for the first time. The letter provided background on the project but otherwise was similar to the reinterview letter; it informed potential respondents about selection methods, interview length, compensation, confidentiality, and contact procedures. The survey introduction read by interviewers when they called these households also differed in that there was no reference to prior participation.

MPR, HSC, and RWJF staff developed a brochure describing survey findings from Round One and answering commonly asked questions about surveys. The brochure, which accompanied all letters, was designed to thank Round One respondents for their participation, as well as to reinforce the study's message by providing background on selected findings. Although we did not find that the Round One brochure increased participation in that survey, we felt that it was important to interest participants in the study, because many would be selected for additional surveys or related studies.

No advanced materials could be sent to households for which we did not have published addresses. The survey introduction for this group was the initial contact with the household. The text provided a brief statement about the survey's purpose, indicated that we were conducting the study for a private foundation, and mentioned that participants would receive $25. Citing RWJF by name was not likely to be effective, as individual foundations have low name recognition. The survey introduction used by interviewers also provided examples of the questions we planned to ask to help interviewers respond to questions about the study. These questions were selected to demonstrate that the CTS Household Survey was a health research study, rather than a telemarketing call.

All checks to participating households were accompanied by a thank you letter. Additional materials and survey introductions were developed to support refusal conversion efforts; these efforts and materials are described in Chapter IV.

D. INTERVIEWER SELECTION AND TRAINING

1. RDD Sample

a. Recruitment

Most of the interviewing for the RDD sample was conducted by MPR, in its Columbia and Princeton telephone centers, with assistance from Battelle, in its St. Louis telephone center. Altogether, 239 telephone interviewers were trained for the household survey (224 from MPR and 15 from Battelle). Twenty training sessions were conducted (18 for MPR interviewers and 2 for Battelle interviewers). Project training was conducted by MPR staff. The Round Two *Training Manual* is included in Appendix D. Interviewing supervisors were given a more detailed manual that contained additional information enabling them to respond to interviewer's questions, and to resolve routine problems; this manual is included in Appendix E.

b. Telephone Interviewer Training Program

New interviewers were given MPR's standard general interviewer training program, which lasted 12 hours and was conducted in three 4-hour sessions. Topics included obtaining cooperation, understanding bias, using probing methods, using the CATI system, and resolving administrative issues. A variety of media and methods were used in training, including a video tape on the role of the interviewer and bias, role playing, and written exercises. Battelle conducted its own general interviewer training session.

Training on the survey instrument lasted 12 hours, with up to 8 hours of additional practice sessions, if necessary. The training session covered the following topics:

- An introduction to the project and sample design (see Chapter II of the *Training Manual*)
- A review of the CATI instrument (see Chapter III of the *Training Manual*)
- A video to demonstrate the logic of the survey instrument and to reinforce interviewing, probing, and recording techniques by presenting correct and incorrect behaviors (see Chapter IV of the *Training Manual*)
- Question-by-question review of the instrument, presented on a video screen (see Chapter V of the *Training Manual*)
- Review of contact procedures, advance materials, methods for gaining cooperation, and appropriate responses to respondents' questions (see Chapter VI of the *Training Manual*)
- Hands-on practice with scripted mock interviews (see Chapter VII of the *Training Manual*)
- Exercises to test respondents' skills in obtaining cooperation (see Chapter VIII of the *Training Manual*)
- Review of disposition coding and call scheduling (see Chapter IX of the *Training Manual*)
- Review of the interviewer bonus plan[7]

At the end of the survey, interviewers were debriefed on the effectiveness of training, interviewing problems (especially refusal conversion procedures), the interviewer bonus plan, and respondent incentives. The agenda used for the debriefing is included in Appendix F.

Because initial refusal rates for both rounds of the survey were high, considerable efforts were devoted to preparing interviewers for follow-up calls during which they attempted to convert refusals. Supervisors reinforced training techniques throughout the survey by monitoring calls and providing regular feedback (discussed in Chapter IV). In addition, we conducted refusal conversion sessions, during which trainers reviewed effective approaches and

[7]Interviewers participated in a bonus plan in which points were obtained for completing interviews based on difficulty in gaining cooperation. The points were converted to bonuses at several stages during the survey.

interviewers shared experiences about the success or failure of various techniques. The trainer's guide for refusal conversion training is included in Appendix G.

2. Field Sample

a. Recruitment

MPR interviewers were responsible for all field data collection. A total of 20 staff were trained to screen households in the 12 high-intensity sites. In addition, staff in Lansing, Miami, and Greenville listed additional housing units.

b. Training

Addresses for the field component were listed in Round One; additional segments were listed in Round Two in Lansing, Miami, and Greenville. In addition, the fact that 10 of the 20 Round Two staff had worked on Round One reduced the level of effort required to train the staff on screening methods. Training was conducted during a two-hour conference call during which the MPR trainer reviewed screening procedures with groups of four to six trainees. Screening interviewers were trained in the survey introduction; refusal avoidance; the telephone screener; operation of the cellular telephone; and follow-up interviewing methods, such as attempting contacts at varying times of the day and gaining entry to apartment buildings. After completing the training program, each trainee called the MPR telephone center and conducted a mock screening interview with a supervisor. The manual provided to field listers and screening interviewers is included in Appendix H. Because field staff called the MPR telephone center and then gave the respondent a cellular telephone to complete the interview, they did not have to be trained on how to conduct the survey.

E. CATI SYSTEM

All data collected for the CTS Household Survey were produced using computer programs made available through the Computer Assisted Survey Methods Program (CSM), University of California, Berkeley.[8] The CSM computer-assisted interviewing program, CASES, is one of the most widely used CATI systems on public policy surveys. More than 70 survey organizations, including the U.S. Bureau of the Census and Statistics Canada, are CSM members.

MPR used CASES to develop instruments and data cleaning programs for the CTS. In addition, we developed customized programs for allocating the sample and for controlling the distribution and timing of calls and developed specialized reports for monitoring the survey results.

[8]Neither the CSM staff nor the University of California bear any responsibility for the results or conclusions presented here.

LEFT INTENTIONALLY BLANK

IV. DATA COLLECTION

A. OVERVIEW

For Round Two of the CTS Household Survey, we interviewed 32,047 FIUs—31,278 from the RDD sample and 769 from the field sample. The FIUs included 48,724 eligible adults and 10,232 sampled children younger than age 18, for a total of 58,956 individuals (see Table IV.1) The unweighted Round Two household-level response rate was 66.2 percent and the unweighted FIU-level response rate was 63.2 percent. The weighted response rates for Round Two were 65.1 percent (household level) and 62.5 percent (FIU level). Response rates declined slightly between the two rounds, as the unweighted Round One household-level response rate was 66.1 percent and the FIU-level response rate was 63.8 percent. Round One weighted response rates were 67.0 percent (household) and 64.8 percent (FIU). Response rate calculations are shown in Section C.

In this chapter, we describe the organization of the Round Two RDD and field data collection efforts. These efforts include response rate calculations and response rate patterns in the two rounds; efforts to reduce nonresponse, including call scheduling procedures, use of Spanish-speaking interviewers, refusal conversions, monetary incentives, and selective use of proxy respondents; quality assurance procedures; and data editing and file preparation.

B. ORGANIZATION OF THE RDD AND FIELD SURVEYS

1. RDD Survey

Interviewing for the RDD sample was conducted primarily from MPR's Princeton, New Jersey, and Columbia, Maryland, telephone centers, with some assistance near the end of the survey from Battelle's St. Louis telephone center. A total of 239 interviewers worked on the RDD component of the survey. Overall, MPR interviewers completed 98.1 percent of the

TABLE IV.1

NUMBER OF INTERVIEWS COMPLETED WITH FIUs AND PERSONS IN ROUND ONE AND ROUND TWO OF THE CTS HOUSEHOLD SURVEY

	Round One	Round Two
Number of FIUs		
RDD	32,079	31,278
Field	635	769
Total	**32,732**	**32,047**
Number of Persons		
Adults	49,807	48,724
Children	10,639	10,232
Total	**60,446**	**58,956**

interviews and Batelle interviewers completed 1.9 percent. Both organizations used the same CASES CATI system. The initial CATI instrument and reporting programs, as well as updates to those programs, were transmitted from MPR to Battelle via dedicated data lines. Completed survey data and reports on field progress were transmitted daily. The survey reports enabled supervisors in each site and project management in MPR's Princeton office to monitor production and performance continuously. Several reports were produced, including:

- ***Status disposition reports*** showing daily and cumulative distributions of interim and final survey status codes (completions, various nonresponse and ineligibility dispositions, and current statuses for active cases), for the total sample; for each stratum; and for subgroups, including Spanish-speaking samples, primary and secondary FIUs, and refusal conversion samples
- ***Site status disposition reports*** showing cumulative distributions of interim and final survey disposition codes, by site
- ***Specialized weekly reports*** to monitor the results of experiments to test the effect of prepayment on response rates
- ***Daily interviewer performance reports*** to monitor last-day and cumulative performance statistics, including completions, separate self-response modules, first refusals, final refusals, number of calls, time per call, and time per completed interview

These reports were supplemented by regularly scheduled weekly conference calls with survey supervisors and by visits to the telephone centers by survey managers.

2. Field Survey

Listing and screening were conducted by MPR staff. A total of 20 interviewers screened addresses; additional listing was conducted in two sites (Greenville, SC, and Miami, FL). Field staff were supervised by MPR survey managers in the Princeton office. Reports were developed to monitor field costs and screening outcomes. Because interviews with eligible households were conducted via cellular telephone calls to MPR's Princeton telephone center, the CATI

reports were used to monitor interview production and sample dispositions, by site. Field listers and screening interviewers reported to the MPR supervisor on a weekly basis.

C. RESPONSE RATES

1. Calculation of Response Rates

Both unweighted and weighted response rates were calculated at the household and FIU levels for the RDD, field, and combined samples and for various subgroups, including sites and combinations of sites. The response rate is based on the standard definition the American Association for Public Opinion Research has proposed for surveys with unknown eligibility for some interviewing units (American Association for Public Opinion Research 2000):

(1) $RR3=I/[(I + P) + (R + NC + O) + e(UH + UO)]$,

where RR = response rate; I = complete interview; P = partial interview (insufficient data for analysis); R = refusal; NC = noncontact; O = other; UH= unknown whether household or occupied household; UO = unknown other; and e = estimated proportion of cases with unknown eligibility that are eligible.

The household-level response rate is the ratio of the number of households in which at least one FIU interview was completed to the estimated number of eligible households. This response rate is comparable to that used in many surveys, such as the Current Population Survey. We were not able to determine residency for all sampled telephone numbers (RDD) and addresses (field). Using methods described below, we estimated the number of telephone numbers with undetermined residency that were residential. Because the survey was designed to represent the civilian noninstitutionalized population, some residences were not eligible for the survey. We

also estimated survey eligibility for confirmed residential households for which the household screener was not completed.

The primary interviewing unit for the CTS Household Survey is the family insurance unit (the FIU), rather than the household. Consequently, we also computed an FIU-level response rate that is the product of the household-level response rate and the percentage of FIUs within completed households that responded.

The following sections describe how we calculated response rates for the RDD and field samples, as well as for the combination of the two samples.

a. Determining Residency for the RDD Sample

The first step in computing the RDD response rate was to estimate residency for sampled telephone numbers. Table IV.2 shows the disposition of the RDD household sample, by stratum; Table IV.3 shows the disposition of the sample at the FIU level. Residency was determined for 91.8 percent of the 74,348 sampled telephone numbers. Of the remainder, 4.5 percent were ring, no answers; 2.2 percent had mechanical answering devices, but with no way to determine residency; and 1.5 percent resulted in personal contact, but with no confirmation of residency. In this section, we review various approaches used to estimate residency in RDD surveys and then describe the procedures used for Round Two of the CTS Household Survey.

A key problem in defining the response rate for RDD surveys is that, because residency is not known for all dialed telephone numbers, the denominator of the response rate must be estimated (Frankel 1983; Groves and Lyberg 1988; and American Association for Public Opinion Research 2000). Some telephone numbers ring when dialed even though the number is not assigned for use because some business lines are not intercepted by a switchboard or message. The increasing prevalence of seasonal homes with telephones also has contributed to an increase in the percentage of unanswered calls. Modem lines may not be answered when

TABLE IV.2

FINAL ROUND TWO HOUSEHOLD-LEVEL SURVEY DISPOSITION (RDD SAMPLE)
(Numbers)

		Round One Complete	Round One Breakoff	Round One Other Nonresponse	Round One Undetermined Residency	Round One Nonresidency	Old Working Banks	New Working Banks	Total
Complete									
1, 2	Complete	12,382	97	814	163	2,234	7,652	2,632	25,974
Ineligible									
40	Not a residence	646	16	234	269	7,621	4,423	2,428	15,637
41	Not selected	24	1	23	3	61	86	64	262
42	Cell phone	50	2	18	15	115	103	121	424
49	Other ineligible	0	0	0	0	1	4	0	5
50	Nonworking	1,917	36	454	359	4,732	3,803	3,342	14,643
58	Computer/fax	50	1	19	166	129	299	139	803
Residential Nonresponding Household									
21	Refusal	1,827	259	1,,546	160	657	3625	877	8,951
22	Refusal with breakoff	105	16	48	4	70	277	78	598
28	Round One nonresidence; Round Two refusal	0	0	0	0	278	0	0	278
39	Privacy manager	1	1	0	0	2	3	3	10
67	Effort ended	121	3	77	27	69	223	80	600
76	Inaccessible with breakoff	12	0	4	0	2	10	1	29
77	Maximum calls with breakoff	0	0	1	0	2	1	0	4
Ring, No Answer									
65	No answer—Round Two	177	1	108	0	0	1,016	576	1,878
73	No answer—Round One and Round Two	0	0	0	983	0	0	0	983
75	No answer Round Two; Round One nonworking	0	0	0	0	490	0	0	490
Mechanical Answering Device (MAD)									
59	Voice and computer	63	0	22	0	0	177	57	319
69	MAD—voice only	211	8	116	0	0	405	195	935
72	Round One and Round Two MAD	0	0	0	0	147	0	0	147

Table IV.2 (*Continued*)

		Round One Complete	Round One Breakoff	Round One Other Nonresponse	Round One Undetermined Residency	Round One Nonresidency	Old Working Banks	New Working Banks	Total
74	Round One no answer; Round Two MAD	0	0	0	240	0	0	0	240
Contact/ Undetermined Residential									
20	HUDI	171	43	207	6	35	186	46	694
30	Language/ other barrier	35	3	124	7	20	129	29	347
66	Maximum calls—eligibility unknown	10	1	10	23	10	26	17	97
Total		**17,802**	**488**	**3,825**	**2,425**	**16,675**	**22,448**	**10,685**	**74,348**

HUDI = hung up during introduction.

TABLE IV.3

FINAL ROUND TWO FIU-LEVEL SURVEY DISPOSITION (RDD SAMPLE)[a]
(Numbers)

		Round One Complete	Round One Breakoff	Round One Other Nonresponse	Round One Undetermined Residency	Round One Nonresidency	Old Working Banks	New Working Banks	Total
Complete									
1, 2	Complete	14,978	115	977	189	2,676	9,,223	3,120	31,278
Residential Household—Survey-Ineligible Family									
41	Not selected	169	2	15	3	44	118	50	401
49	Secondary family ineligible	86	1	8	4	17	64	17	197
Residential Household—Survey-Eligible Family[a]									
22	Refusal with breakoff	75	1	25	1	27	80	22	231
76	Inaccessible with breakoff	3	0	1	0	0	2	0	6
77	Maximum calls with breakoff	7	0	0	1	0	2	1	11
Residential Household—Survey Eligibility for Family Undetermined[b]									
20	HUDI	6	0	0	0	1	11	2	20
21	Refusal	159	1	26	4	38	153	43	424
30	Language/ other barrier	14	0	7	0	2	19	2	44
39	Privacy manager	1	0	0	0	1	0	1	3
40	Not a residence	0	0	0	0	0	1	0	1
42	Cell phone—ineligible	0	0	0	0	0	1	4	5
44	Disconnected—unlisted	4	3	0	0	7	3	7	24
53	Cannot locate secondary family	55	2	11	2	28	51	30	179
54	Wrong number	25	2	7	1	19	23	22	99
58	Computer/fax	2	0	0	0	0	2	0	4
59	Voice and computer MAD	3	0	2	0	0	5	0	10
65	No answer	6	0	4	0	9	35	7	61
66	Maximum calls—eligibility unknown	17	0	3	0	4	7	1	32
67	Effort ended	18	0	12	0	10	27	3	70
69	MAD—voice only	91	0	12	6	14	83	30	236
98	Secondary—undetermined	15	1	3	0	5	14	4	42
Total		**15,734**	**128**	**1,113**	**211**	**2,902**	**9,924**	**3,366**	**33,378**

[a]These cases are limited to households in which at least one interview with an FIU was completed.

Table IV.3 (*Continued*)

---

[b]Residency was based on the disposition of the primary FIU at the time that family was interviewed; consequently, all secondary FIUs are assumed to be residential.

HUDI = hung up during introduction; MAD = mechanical answering device.

computers are turned off. It may not be clear from a mechanical answering device whether the number is residential. In other cases, although someone may answer the telephone, the call may be too brief to determine residency, or the person may speak a language that is not known by telephone center staff.

The percentage of numbers in which residency cannot be determined has been increasing over the last few years. For example, Piekarski et al. (1999) note that the number of telephone households increased by 11 percent from 1988 to 1998, but the number of telephone numbers that could be dialed in an RDD telephone survey increased by 80 percent. Although some of the increase reflects increased demand, many of the newly created numbers are not assigned to any user.

Various methods have been used to try to estimate residency for telephone numbers. The most common approach, developed for the Council of American Survey Research Organizations (CASRO) by Frankel (1983), attempts to determine eligibility for each sampled telephone number and then distributes the undetermined numbers in proportion to the distribution of the numbers that were determined. Most practitioners have concluded that this approach results in a larger percentage of the undetermined numbers being classified as residential than is reasonable (Keeter and Miller 1998).

An alternative to the CASRO approach is the "business office method." Under this approach, a subsample of the undetermined numbers is selected, and telephone business offices are contacted to determine whether the numbers are residential. The percentage of undetermined numbers estimated as residential according to the business office method is the ratio of the number of telephone numbers the business office has resolved as residential to the number of telephone numbers resolved. This approach has been used by the National Household Education Survey (Brick and Broene 1997), the National Immunization Survey (NIS) (Shapiro et al. 1995),

and the first round of the National Survey of American Families (Brick et al. 1998). Although evidence shows that the business office method overestimates the fraction of telephones that are residential (Shapiro et al. 1995), this approach often yields a slightly lower residential rate than does the CASRO method.[1]

The main drawbacks of the business office method are the high cost to obtain data from telephone business offices and cooperation rates that vary by business office. We attempted to contact business offices for Round One, but the response was too low to use for estimation purposes. For Round Two, we used the results of calls made to telephone business offices by the NIS to estimate residency for three classes of telephone numbers with undetermined residency: (1) ring, no answers; (2) mechanical answering devices; (3) and telephone numbers for which contact was made but residency was not determined. Using NIS data, we estimated that 27 percent of the ring, no answers were residences; for telephone numbers with a final disposition of mechanical answering devices (no personal contact), we estimated that 72 percent were residences, and for telephone numbers for which a contact was made but residency was not determined, we estimated that 88 percent were residential.[2]

Ring, No Answers. We made a minimum of 20 calls dispersed over various times of the day and days of the week before assigning a final disposition of "ring, no answer" (that is, there was no contact, nor was there any indication of a mechanical answering device). Of the 74,348

---

[1]A third approach, the survival function method, is currently being tested on data from the National Survey of American Families (Brick et al. in press). This method takes advantage of additional information about telephone numbers to provide better estimates, including the number of attempts to contact the telephone number, whether the number is listed in a telephone directory, and whether a mechanical answering device is ever encountered in a call attempt. We plan to test this method for estimating residency for Round Three of the CTS Household Survey.

[2]For Round Two, the unweighted RDD response rate using the business office method to estimate residency was 65.9 percent; it was 65.6 percent using the CASRO method. Therefore,

sampled telephone numbers, 4.5 percent (n = 3,351) were assigned this disposition. Nearly half (44 percent) of the ring, no answers were telephone numbers that had been sampled in Round One and had a final disposition of nonworking or ring, no answer on that survey (Table IV.2). Because these telephone numbers had been called at least 20 times during each of the two rounds, with an interval of two years between rounds, it is likely that all or nearly all are nonresidential. However, we did not have a basis for imputing a lower residency rate for these cases and used the NIS estimate, imputing 27 percent as residential.

Mechanical Answering Devices. A total of 2.2 percent (n = 1,641) of the sampled telephone numbers had a final disposition of "mechanical answering device with no personal contact and insufficient information to determine residency." A minimum of 50 calls had been made to these telephone numbers. Of the telephone numbers with this disposition, 19.4 percent were mixed-use (voice and computer signal) lines; 23.6 percent were voice-only devices that had been sampled in Round One, with a final disposition of ring, no answer or mechanical answering device on that survey; and the remaining 57.0 percent were voice-only that either were sampled for the first time or had been identified as residences in Round One. Using NIS data, we estimated that 72 percent of these telephone numbers were residential.

Contact but Undetermined Residential Status. A total of 1.5 percent of the sampled telephone numbers (n = 1,138) had at least one contact, but no designation of residency. This group includes numbers for which respondents answered but hung up during the introduction (61.0 percent), numbers for which respondents had language and disability barriers (30.5 percent), and other cases with the maximum number of calls (8.5 percent). If respondents hung up before residency could be determined, their telephone numbers were assigned a final

(continued)

the increase in the Round Two response rate resulting from the change to the business office method was only 0.3 percentage points.

disposition of "hung up during the introduction." This code usually was assigned three or more times before a case was closed. Households with language and disability barriers were assigned a final disposition when the interviewer determined that no additional information could be obtained. A total of 50 attempts were made before assigning a final disposition to a telephone number for which voice contact was made, but residency was not determined. Using NIS data, we estimated that 88 percent of these telephone numbers were residential.

b. Household Response Rate for the RDD Sample

To calculate an interview response rate at the household level, we first determined whether each telephone number was residential and then determined whether each household contained at least one eligible person, so that an FIU could be formed. We used the business office method to estimate residency for telephone numbers for which residency was not determined. We used data obtained from the household interview to determine whether there was an eligible person in the household so that the FIU could be formed. (See Section B.3 of Chapter II for the criteria used to determine survey eligibility for persons.)

We classified each telephone number according to the disposition codes in Table IV.2:

A. At least one eligible responding FIU in the household—codes 1, 2 (n = 5,975)

B. At least one eligible nonresponding FIU in the household—codes 22, 76, 77 (n = 631)

C. Nonresponding residential household, with insufficient information to determine whether there is an eligible FIU—codes 21, 28, 39, 67 (n = 9,839)

D. Residential household, where all FIUs in the household are ineligible—codes 41, 49 (n = 267)

E. Telephone number was nonresidential or nonworking—codes 40, 42, 50, 58 (n = 31,507)

F. Unable to determine whether telephone number was residential (n = 6,130)

- F1. Ring, no answer—codes 65, 73, 75 (n = 3,351)
- F2. Mechanical answering device—codes 59, 69, 72, 74 (n = 1,641)
- F3. Voice contact made—codes 20, 30, 66 (n = 1,138)

Within each site *s* (with the supplemental sample being treated as a separate site), we calculated a survey eligibility rate among residential households with known survey eligibility:

(2) $$ER_s = (A_s + B_s)/(A_s + B_s + D_s).$$

We then calculated a household response rate within each site as follows, using the business office method to estimate residency for telephone numbers with undetermined residency:

(3) $$HRR_s = \frac{A_s}{A_s + B_s + \{C_s + (F1_s \cdot .27) + (F2_s \cdot .72) + (F3_s \cdot .88)\} \cdot ER_s}.$$

To compute response rates involving more than one site, we first summed the number of cases in each of the categories above (for example, $A = \sum_s A_s$, $B = \sum_s B_s$, and so on) and then calculated the survey eligibility rate and household response rate:

(4) $$ER = (A + B)/(A + B + D).$$

(5) $$HRR = \frac{A}{A + B + \{C + (F1 \cdot .27) + (F2 \cdot .72) + (F3 \cdot .88)\} \cdot ER}.$$

Weighted response rates were calculated similarly, except that we used counts weighted by sampling weights, rather than unweighted counts. (By "sampling weight," we mean the inverse of the probability of selection.) Using the subscript *w* to indicate a weighted count, the weighted survey eligibility rate among those with known survey eligibility in site *s* is:

(6) $$ER_{ws} = (A_{ws} + B_{ws}) / (A_{ws} + B_{ws} + D_{ws}) .$$

The weighted household response rate within site *s* is:

(7) $$HRR_{ws} = \frac{A_{ws}}{A_{ws} + B_{ws} + \{C_{ws} + (F1_{ws} \cdot .27) + (F2_{ws} \cdot .72) + (F3_{ws} \cdot .88)\} \cdot ER_{ws}} .$$

Weighted survey eligibility and response rates involving more than one site are calculated as:

(8) $$ER_{w} = (A_{w} + B_{w}) / (A_{w} + B_{w} + D_{w}) ,$$

where $A_w = \sum_s A_{ws}$, $B_w = \sum_s B_{ws}$, and so on, and

(9) $$HRR_{w} = \frac{A_{w}}{A_{w} + B_{w} + \{C_{w} + (F1 \cdot .27) + (F2_{w} \cdot .72) + (F3_{w} \cdot .88)\} \cdot ER_{w}} .$$

c. Household Response Rate for the Field Sample

To calculate a household response rate for the field component, we had to determine whether (1) each address was an occupied residence, (2) the residence met the criteria for interruption in telephone service, and (3) there was at least one survey-eligible person in the household. To estimate eligibility rates for addresses with undetermined eligibility, we applied rates from those with unknown eligibility status. First, we classified each address according to the disposition codes in Table IV.4:

a. Eligible responding household—codes 1, 2 (n = 62)

b. Eligible nonresponding household—codes 21, 22, 66 (n = 82)

TABLE IV. 4

FINAL ROUND TWO CTS HOUSEHOLD SURVEY DISPOSITION (FIELD SAMPLE)

Disposition		Number
Complete		
1,2	Complete—eligible[a]	562
Eligible Nonresponse		
21	Refused main interview—eligible (telephone and survey)	68
22	Refused survey screener—eligible (telephone interruption)	4
66	Effort ended—eligible	10
Ineligible		
41	Not selected (all military or children)	3
45	Screened out (telephone service)	4,737
49	Other ineligible	1
Household with Unknown Eligibility		
20	Refused screener	278
30	Inaccessible	53
39	Other (never home)	4
67	Effort ended	214
Not a Residential Household		
40	Not a residence	26
47	Vacant unit	604
48	No housing unit	164
Unknown Whether a Household		
65	Maximum visits—no screener—eligibility unknown	2
Total	**6,730**	**6,730**[b]

[a]For the field component, the household informant responded for all the FIUs in the household. Consequently, the FIU-level response rate is the same as the household-level response rate.

[b]T6 total number of housing units attempted is greater than number of released (6,667 in Table II.4) because additional housing units were discovered during screening.

c. Household with insufficient information to determine whether eligible—codes 20, 30, 39, 67 (n = 549)

d. Ineligible household—codes 41, 45, 49 (n = 4,741)

e. Not a household—codes 40, 47, 48 (n = 794)

f. Unable to determine whether address was residential—code 65 (n = 2)

Within each high-intensity site *s*, we calculated two eligibility rates: (1) a household eligibility rate (proportion of addresses known to be occupied residences), and (2) a survey eligibility rate (proportion of residences known to be eligible for the field survey):

(10) $$HER_s = (a_s + b_s + c_s + d_s)/(a_s + b_s + c_s + d_s + e_s).$$

(11) $$SER_s = (a_s + b_s)/(a_s + b_s + d_s).$$

We then calculated a household response rate within each site as follows:

(12) $$HRR_s = \frac{a_s}{a_s + b_s + \{c_s + (f_s \cdot HER_s)\} \cdot SER_s}.$$

To compute response rates involving more than one site, we first summed the number of cases in each of the categories listed above (for example, $a = \sum_s a_s$, $b = \sum_s b_s$, and so on) and then calculated eligibility rates and the household response rates:

(13) $$HER = (a + b + c + d)/(a + b + c + d + e).$$

(14) $$SER = (a + b)/(a + b + d).$$

(15) $$HRR = \frac{a}{a + b + \{c + (f \cdot HER)\} \cdot SER}.$$

As with the RDD response rates, weighted response rates for the field component were calculated using counts weighted by sampling weights instead of unweighted counts.

d. Combinations of Household Response Rates

When calculating a response rate for combinations of various sample components (such as the site sample and supplemental sample combined, or the RDD sample and the field sample combined), we calculated the numerator and denominator of the response rates separately for each component and then calculated the response rate as follows. Here is the formula for the unweighted household response rate for site *s* combining the RDD and field components:

$$(16) \qquad HRR_s = \frac{A_s + a_s}{[A_s + B_s + \{C_s + (F1_s \cdot .27) + (F2_s \cdot .72) + (F3_s \cdot .88)\} \cdot ER_s] + [a_s + b_s + \{c_s + (f_s \cdot HER_s)\} \cdot SER_s]}.$$

e. RDD—Family Interview Response Rate

To calculate an interview response rate at the FIU level,[3] we began with all FIUs in responding households (that is, households with at least one eligible responding FIU). We classified each FIU according to the disposition codes in Table IV.3 as follows:

A. FIU is eligible for the survey and responded to interview—codes 1, 2 (n = 31,278)

B. FIU is eligible for the survey but did not respond to interview—codes 22, 76, 77 (n = 248)

C. FIU has undetermined survey eligibility (that is, an FIU informant did not complete the screening questions)—codes 20, 21, 30, 39, 40, 42, 44, 53, 54, 58, 59, 65, 66, 67, 69, 98 (n = 1,254)

[3]For the field component, the household informant responded for each FIU; consequently, the FIU response rate is equal to the household response rate.

D. FIU is ineligible for survey

- D1. No civilian adult in FIU—code 41 (n = 401)
- D2. Secondary FIU should not have been generated—code 49 (n = 197)

Within each site *s* (with the supplemental sample being treated as a separate site), we calculated a survey eligibility rate among FIUs with known survey eligibility:

(17) $ER_s = (A_s + B_s)/(A_s + B_s + D2_s)$.

For each site, we then calculated a conditional response rate for households with at least one completed FIU interview:

(18) $FRR_s = \frac{A_s}{A_s + B_s + \{C_s \cdot ER_s\}}$.

The combined response rate (which we will call the *FIU response rate*) for site *s* is simply the product of these two rates:

(19) $RR_s = HRR_s \cdot FRR_s$.

For any conditional response rates involving more than one site, we first summed the number of cases in each of the categories listed above (for example, $A = \sum_s A_s$, $B = \sum_s B_s$, and so on) and then calculated the survey eligibility rate and conditional response rate (within households including at least one completed FIU):

(20) $ER = (A + B)/(A + B + D2)$.

(21) $FRR = \frac{A}{A + B + \{C \cdot ER\}}$.

The FIU response rate is simply the product of the two rates:

(22) $RR = HRR \cdot FRR$.

Weighted response rates at the FIU level were calculated similarly, except that we used counts weighted by sampling weights, rather than unweighted counts. The weighted survey eligibility rate among those with known survey eligibility in site *s* is:

(23) $$ER_{ws} = (A_{ws} + B_{ws})/(A_{ws} + B_{ws} + D2_{ws}).$$

The weighted conditional response rate (for households including at least one completed FIU) within site *s* is:

(24) $$FRR_{ws} = \frac{A_{ws}}{A_{ws} + B_{ws} + \{C_{ws} \cdot ER_{ws}\}}.$$

The weighted FIU response rate for site *s* is simply the product of these two rates:

(25) $$RR_{ws} = HRR_{ws} \cdot FRR_{ws}.$$

Weighted survey eligibility and conditional response rates (involving more than one site) are calculated as:

(26) $$ER_{w} = (A_{w} + B_{w})/(A_{w} + B_{w} + D2_{w}),$$

where $A_w = \sum_s A_{ws}$, $B_w = \sum_s B_{ws}$, and so on, and

(27) $$FRR_{w} = \frac{A_{w}}{A_{w} + B_{w} + \{C_{w} \cdot ER_{w}\}}$$

The weighted FIU response rate is the product of the two rates:

(28) $$RR_{w} = HRR_{w} \cdot FRR_{w}.$$

Conditional FIU response rates combining the site and supplemental samples, and those involving both the RDD and in-person components, were calculated in the same way as were the household response rates.

2. Patterns in Household and FIU Response Rates, by Sample Type

a. Response Rates, by Sample Type

Table IV.5 shows the unweighted and weighted household-level response rates, for the Round Two sample, by sample type; Table IV.6 shows comparable FIU-level response rates. Household-level response rates were higher than FIU-level response rates because some households included multiple FIUs and some of these FIUs did not complete interviews. Although both household and FIU response rates are shown in the tables in this section, we will limit our discussion to FIU response rates, as the FIU was the primary interviewing unit for the Household Survey. In addition, patterns in response rates by sample type and geographic units were the same for households and FIUs. We also will limit our discussion to weighted response rates; here, too, the differences between weighted and unweighted response rates are generally very small and do not affect response rate patterns.

The weighted FIU RDD response rate was higher for the total Round One sample (65.5 percent) than for the new sample (59.1 percent). The total RDD Round One sample includes all telephone numbers selected for both rounds, including those linked to completed Round One interviews and to noninterviews. The difference was even greater for the unweighted RDD sample—66.7 percent for the Round One sample that was selected for Round Two, and 58.0 percent for the new sample. Thus, the decision to sample Round One telephone numbers mitigated the decline in the response rate between rounds.

The response rate was particularly high for Round Two FIUs selected from telephone numbers linked to Round One completed interviews (81.6 percent). The interval between rounds was only two years, so most of the families and persons interviewed for Round One were at the same telephone number for Round Two. Because nearly all the Round One sample had received

TABLE IV.5

ROUND TWO HOUSEHOLD-LEVEL RESPONSE RATE, BY SAMPLE TYPE
(Percents)

	Unweighted	Weighted
RDD		
Round One		
Completed interviews	83.22	83.85
Breakoffs	23.05	18.94
Nonresponse (other than breakoffs)	28.39	33.51
Undetermined whether residence	20.05	22.57
Nonresidential	62.49	60.64
Total Round One sample	69.56	67.94
New Sample		
Old working banks[a]	60.11	61.50
New working banks[b]	64.87	65.31
Total new sample	61.26	62.26
Total RDD	65.92	65.04
Field	79.17	73.30
Total Sample	**66.15**	**65.10**

[a]Working banks in existence at the time the Round One sample was selected.

[b]Working banks that were added between the end of Round One and the beginning of Round Two.

TABLE IV.6

ROUND TWO FIU-LEVEL RESPONSE RATE, BY SAMPLE TYPE
(Percents)

	Unweighted	Weighted
RDD		
Round One		
Completed interviews	80.54	81.56
Breakoffs	21.22	18.05
Nonresponse (other than breakoffs)	25.46	30.37
Undetermined whether residence	18.60	21.69
Nonresidential	58.88	57.42
Total Round One Sample	66.74	65.45
New Sample		
Old working banks[a]	56.92	58.33
New working banks[b]	61.37	62.35
Total New Sample	58.00	59.13
Total RDD	62.91	62.30
Field	**79.17**	**73.30**
Total Sample	**63.21**	**62.46**

[a]Working banks in existence at the time the Round One sample was selected.

[b]Working banks that were added between the end of Round One and the beginning of Round Two.

monetary incentives, most remembered the prior interview and were aware that they would be compensated for participation.

Not surprisingly, the weighted FIU response rate was poor for Round one breakoffs (18.1 percent) and for other nonresponses (30.4 percent).[4] Most of these households had been contacted many times in each round and had refused more than once. The weighted FIU response rate for Round One nonresidences (57.4 percent) was comparable to the rate for new telephone numbers. This result is not surprising, as the telephone numbers linked to these[5]households had been nonresidential at the time of the Round One survey and therefore were contacted for the first time in Round Two. The Round Two weighted FIU response rate for telephone numbers that had final dispositions as undetermined residences in Round One was only 21.7 percent. The low response rate for this group was not due to refusals; rather, it was due to the very large fraction of telephone numbers in this subsample with undetermined residency in both rounds, which resulted in a fraction having residency imputed. (Appendix I provides additional details on the computation of response rates for subgroups of the sample.)

The RDD weighted FIU response rate for telephone numbers sampled for the first time was higher for telephone numbers selected from new working banks (62.4 percent) than for working banks existing at the time the Round One sample was selected (old working banks; 58.3 percent). None of the telephone numbers in either subsample had been contacted previously for the CTS Household Survey, and the level of effort for interviewing was comparable for the two subsamples. However, new working banks are more likely to be generated in areas experiencing

[4]Other nonresponses include refusals prior to screening, disability and language barriers, and cases closed at the end of data collection (effort ended).

population growth. Therefore, the higher response rate for new working banks may be related to the location of areas experiencing growth.

Weighting had little impact on RDD response rates but had a significant impact on the Round Two field response rate (79.2 percent unweighted and 73.3 percent weighted). Field response rates varied substantially by site; the weighted response rate was depressed by low response rates in certain sites.

b. Patterns in Response Rates

Response rates declined from Round One to Round Two in all but one high-intensity site (Syracuse), in most (33 of 48) low-intensity sites, and for the supplemental sample (see the weighted FIU response rates in Table IV.7). For both rounds, response rates varied inversely with site population. For Round One, the weighted FIU RDD response rate ranged from a low of 56.6 percent in MSAs of 3 million or more people to a high of 72.6 percent in nonmetropolitan areas. The trend in Round Two response rates was similar, ranging from 55.3 percent for MSAs with 3 million or more to 71.0 percent for nonmetropolitan areas. The largest MSAs correspond to the largest media markets, whose residents may be subject to greater telemarketing and market research penetration.

Weighted field sample response rates are significantly higher than RDD response rates but declined between rounds, from 79.2 percent in Round One to 73.3 percent in Round Two. (As noted, weighting had a significant impact on the Round Two field response rate but had much less of an impact in Round One.) Response rates usually are higher for surveys involving personal contact, as interviewers are able to establish personal rapport with potential respondents and to assuage concerns about the legitimacy of these studies. The main source of the decline between rounds was the very low Round Two (weighted) response rates in two sites—Newark (46.9 percent) and Seattle (31.3 percent). These sites had gated communities to which we could

TABLE IV.7

WEIGHTED AND UNWEIGHTED RESPONSE RATES FOR ROUND ONE AND ROUND TWO
(Percents)

		Unweighted				Weighted			
		Round One		Round Two		Round One		Round Two	
		Household	FIU	Household	FIU	Household	FIU	Household	FIU
	RDD								
0	Supplemental Sample	66.46	64.26	67.18	64.66	66.89	64.73	65.61	63.06
1	Boston, MA Portion	59.94	56.72	61.51	57.47	59.94	56.72	59.88	55.97
2	Cleveland-Lorain-Elyria, OH PMSA	63.52	61.47	64.12	62.10	63.53	61.47	61.89	59.87
3	Greenville-Spartanburg-Anderson, SC MSA	70.77	68.96	69.39	67.19	70.77	68.96	67.60	65.33
4	Indianapolis, IN MSA	71.25	70.22	70.22	67.58	71.25	70.22	68.50	65.96
5	Lansing-East Lansing, MI MSA	71.97	70.12	70.64	68.34	71.97	70.11	68.32	66.10
6	Little Rock-North Little Rock, AR MSA	74.29	72.91	72.18	69.64	74.29	72.91	70.57	68.09
7	Miami, FL PMSA	53.21	49.89	54.94	50.58	53.22	49.90	54.04	49.74
8	Newark, NJ PMSA	59.36	56.32	56.76	53.03	59.36	56.32	55.48	51.75
9	Orange County, CA PMSA	55.83	52.35	55.37	50.26	55.83	52.35	53.90	48.66
10	Phoenix-Mesa, AZ MSA	68.54	66.88	65.29	62.06	68.53	66.87	62.98	59.80
11	Seattle-Bellevue-Everett, WA PMSA	65.90	62.89	62.38	59.23	65.90	62.88	60.53	57.52
12	Syracuse, NY MSA	69.00	67.12	71.98	70.46	69.00	67.12	69.78	68.35
13	Atlanta, GA MSA	65.44	62.68	71.19	67.86	65.44	62.68	69.08	65.81
14	Augusta-Aiken, GA-SC MSA	66.06	64.08	69.03	66.93	66.06	64.08	67.22	65.10
15	Baltimore, MD PMSA	65.03	63.47	67.07	64.86	65.03	63.47	66.06	63.83
16	Bridgeport-Danbury-Stamford, CT Portion	54.65	52.44	58.54	53.82	54.65	52.44	57.01	52.48
17	Chicago-Kenosha-Kankakee, IL-WI PMSA	60.99	57.83	57.40	53.48	60.99	57.83	55.78	51.87
18	Columbus, OH MSA	69.67	65.88	68.49	66.31	69.67	65.88	67.36	65.24
19	Denver-Boulder-Greeley, CO PMSA	66.08	64.10	65.89	61.42	66.08	64.10	64.11	59.28
20	Detroit, MI PMSA	66.44	62.78	60.83	59.76	66.44	62.78	59.30	58.27
21	Greensboro-Winston Salem-High Point, NC MSA	68.33	66.60	70.93	68.74	68.33	66.60	68.64	66.56
22	Houston-Galveston-Brazoria, TX PMSA	64.14	60.47	59.83	57.18	64.14	60.47	58.79	56.23
23	Huntington-Ashland, WV-KY-OH MSA	75.13	73.45	74.45	71.96	75.13	73.45	72.70	70.26
24	Killeen-Temple, TX MSA	73.17	70.80	73.34	71.64	73.17	70.80	72.08	70.29
25	Knoxville, TN MSA	72.06	70.48	68.74	67.37	72.06	70.48	66.91	65.60
26	Las Vegas, NV-AZ MSA	58.52	54.44	61.63	58.58	58.52	54.44	60.22	57.26
27	Los Angeles-Long Beach, CA PMSA	53.99	51.25	56.16	52.31	53.99	51.25	55.22	51.39
28	Middlesex-Trenton, NJ PMSA	66.07	63.82	65.44	62.83	66.07	63.82	64.05	61.48
29	Milwaukee-Racine, WI PMSA	68.77	67.68	70.95	68.43	68.77	67.68	67.83	65.42
30	Minneapolis-St Paul, MN-WI MSA	76.74	73.87	76.03	73.89	76.74	73.87	74.36	72.32
31	Modesto, CA MSA	64.76	63.11	66.97	62.90	64.76	63.11	65.09	60.77
32	Nassau-Suffolk, NY PMSA	59.63	56.95	61.15	55.27	59.63	56.95	59.14	53.33
33	New York City, NY PMSA	46.27	42.76	57.06	50.52	46.27	42.76	55.17	48.58

Table IV.7 (*Continued*)

		Unweighted				Weighted			
		Round One		Round Two		Round One		Round Two	
		Household	FIU	Household	FIU	Household	FIU	Household	FIU
34	Philadelphia, PA-NJ PMSA	64.80	62.38	59.52	56.80	64.80	62.38	58.38	55.61
35	Pittsburgh, PA MSA	64.99	62.88	67.83	66.03	64.99	62.88	65.61	63.91
36	Portland-Salem, OR-WA PMSA	68.46	64.87	72.14	69.11	68.46	64.87	70.02	67.03
37	Riverside-San Bernardino, CA PMSA	65.57	63.48	64.83	61.59	65.57	63.48	63.58	60.30
38	Rochester, NY MSA	70.47	68.54	70.77	68.28	70.47	68.54	68.90	66.41
39	San Antonio, TX MSA	63.45	61.39	67.27	63.22	63.45	61.39	65.17	60.95
40	San Francisco, CA PMSA	51.24	47.52	54.78	50.44	51.24	47.52	52.77	48.45
41	Santa Rosa, CA PMSA	60.15	55.84	62.21	59.38	60.15	55.84	60.37	57.62
42	Shreveport-Bossier City, LA MSA	71.65	68.22	71.25	68.48	71.65	68.22	69.98	67.21
43	St. Louis, MO-IL MSA	72.71	69.23	74.68	72.96	72.71	69.23	73.26	71.62
44	Tampa-St Petersburg-Clearwater, FL MSA	60.32	57.74	57.98	54.94	60.32	57.74	54.99	52.20
45	Tulsa, OK MSA	63.80	62.10	70.12	65.78	63.80	62.10	67.87	63.68
46	Washington-Hagerstown, DC-MD-VA-WV PMSA	65.59	63.94	68.23	65.05	65.59	63.94	67.06	64.02
47	West Palm Beach-Boca Raton, FL MSA	55.42	50.98	53.93	49.99	55.42	50.98	51.82	48.05
48	Worcester-Fitchburg, MA Portion	63.91	62.11	66.13	62.38	63.91	62.11	64.88	60.69
49	Dothan, AL MSA	69.81	68.45	72.75	71.67	69.81	68.45	70.88	69.84
50	Terre Haute, IN MSA	74.38	72.17	67.64	64.98	74.38	72.17	66.19	63.60
51	Wilmington, NC MSA	76.18	73.28	71.96	69.22	76.18	73.28	70.57	67.81
52	West-Central Alabama	71.54	69.43	70.48	68.62	71.54	69.43	69.14	67.20
53	Central Arkansas	75.94	75.34	74.61	73.45	75.94	75.34	72.85	71.72
54	Northern Georgia	72.87	69.81	70.48	67.88	72.87	69.81	68.70	66.24
55	Northeast Illinois	67.77	67.32	71.32	68.93	67.77	67.32	68.07	65.80
56	Northeast Indiana	72.74	70.76	71.05	69.52	72.74	70.76	69.02	66.68
57	Eastern Maine	80.32	79.08	81.82	81.57	80.32	79.08	80.30	80.06
58	Eastern North Carolina	77.24	75.26	75.04	72.08	77.24	75.26	73.99	71.08
59	Northern Utah	78.79	76.76	83.19	82.19	78.79	76.76	82.73	81.76
60	Northwest Washington	70.43	68.97	70.90	67.37	70.43	68.97	69.35	65.81
	All high-intensity sites	65.07	62.69	64.16	60.97	64.96	62.54	62.44	59.29
	All low-intensity sites	66.59	64.13	67.38	64.45	67.79	65.43	66.75	63.97
	Total RDD site sample	65.84	63.42	65.77	62.71	66.59	64.21	64.95	62.00
	Total RDD (supplemental + site)	65.90	63.50	65.92	62.91	66.64	64.37	65.04	62.30
	Field								
1	Boston, MA Portion	75.42	75.42	74.07	74.07	90.39	90.39	63.96	63.96
2	Cleveland-Lorain-Elyria, OH PMSA	73.25	73.25	84.71	84.71	72.15	72.15	78.68	78.68
3	Greenville-Spartanburg-Anderson, SC MSA	83.73	83.73	83.41	83.41	82.65	82.65	83.82	83.82
4	Indianapolis, IN MSA	89.10	89.10	82.99	82.99	89.74	89.74	83.12	83.12
5	Lansing-East Lansing, MI MSA	72.38	72.38	70.42	70.42	68.96	68.96	67.01	67.01
6	Little Rock-North Little Rock, AR MSA	90.43	90.43	80.97	80.97	91.04	91.04	83.44	83.44
7	Miami, FL PMSA	95.40	95.40	71.00	71.00	94.01	94.01	72.44	72.44
8	Newark, NJ PMSA	69.68	69.68	56.26	56.26	68.93	68.93	46.86	46.86

Table IV.7 (*Continued*)

		Unweighted				Weighted			
		Round One		Round Two		Round One		Round Two	
		Household	FIU	Household	FIU	Household	FIU	Household	FIU
9	Orange County, CA PMSA	53.04	53.04	61.52	61.52	56.61	56.61	66.75	66.75
10	Phoenix-Mesa, AZ MSA	93.14	93.14	96.28	96.28	89.26	89.26	97.24	97.24
11	Seattle-Bellevue-Everett, WA PMSA	69.44	69.44	26.53	26.53	72.28	72.28	31.33	31.33
12	Syracuse, NY MSA	93.22	93.22	67.38	67.38	92.53	92.53	78.28	78.28
	Total field sample	84.60	84.60	79.17	79.17	83.20	83.20	73.30	73.30
	RDD Site Sample + Field								
1	Boston, MA Portion	60.05	56.87	61.56	57.55	60.46	57.36	59.90	56.04
2	Cleveland-Lorain-Elyria, OH PMSA	63.80	61.81	65.02	63.07	63.97	62.09	63.06	61.22
3	Greenville-Spartanburg-Anderson, SC MSA	71.31	69.57	70.40	68.37	71.34	69.66	68.90	66.88
4	Indianapolis, IN MSA	72.18	71.21	71.03	68.59	73.47	72.61	69.70	67.52
5	Lansing-East Lansing, MI MSA	71.98	70.16	70.64	68.40	71.78	70.13	68.23	66.29
6	Little Rock-North Little Rock, AR MSA	75.00	73.70	72.67	70.30	76.06	74.89	72.32	70.36
7	Miami, FL PMSA	53.99	50.74	55.19	50.90	55.42	52.33	54.42	50.27
8	Newark, NJ PMSA	59.80	56.87	56.74	53.17	60.92	58.57	53.54	51.09
9	Orange County, CA PMSA	55.81	52.37	55.41	50.34	55.84	52.46	54.02	48.85
10	Phoenix-Mesa, AZ MSA	69.53	67.95	67.61	64.69	70.04	68.56	67.45	64.85
11	Seattle-Bellevue-Everett, WA PMSA	66.07	63.18	60.94	57.91	66.71	64.24	57.18	54.89
12	Syracuse, NY MSA	69.72	67.89	71.81	70.36	71.39	69.74	70.20	68.89
	High-intensity sites (with field)	65.61	63.30	64.59	61.54	66.21	64.07	62.85	60.13
	Sites Grouped by Population Size (RDD Site Sample)								
	3 million+	60.32	57.26	61.04	57.34	59.62	56.59	59.12	55.34
	2-3 million	62.55	59.87	61.98	58.43	62.36	59.66	60.05	56.52
	1-2 million	65.26	62.90	65.04	61.90	65.42	63.07	63.66	60.59
	<1 million large MSA	68.56	66.43	68.91	66.30	68.33	66.15	66.76	64.17
	Small MSA (<200,000)	73.41	71.27	70.88	68.77	73.75	71.50	69.68	67.48
	Non-MSA	74.19	72.56	74.42	72.51	74.28	72.64	73.01	71.03
	Total	66.08	63.70	65.95	62.96	67.07	64.81	65.03	62.27
	Total (RDD site + supplemental + field)	**66.16**	**63.80**	**66.15**	**63.21**	**67.04**	**64.82**	**65.10**	**62.46**

not gain access in Round Two. In addition, Newark had a high refusal rate, which exacerbated the decline.

The Round Two weighted FIU response rate for the 12 high-intensity sites (RDD and field combined) was 60.1 percent; it was 64.0 percent for the 48 low-intensity sites (RDD samples only). High-intensity sites were selected from mid-sized to large MSAs. Therefore, it is not surprising that they have lower response rates than do low-intensity sites, which also include small MSAs and nonmetropolitan areas; the latter areas have higher response rates. Among high-intensity sites (RDD and field combined), weighted Round Two FIU response rates ranged from a low of 50.0 percent in Miami to a high of 70.4 percent in Little Rock. Among low-intensity sites (sites 13 to 48), weighted Round Two response rates varied from a low of 48.1 percent in West Palm Beach to a high of 81.8 percent in northern Utah.

c. Relative Level of Effort to Complete an FIU in the RDD Sample Frame

Because response rates varied by site (and site groupings, such as size of MSA), the level of effort to complete an interview also varied by site. We defined the level of effort as the ratio of the total number of calls made to all sampled telephone numbers in the RDD sample divided by the number of completed FIU interviews (average number of calls per FIU). Interviewer labor varies directly with the number of calls made, and many other costs vary with interviewer labor. Therefore, this measure is a key component of the variable cost of conducting an FIU interview and is very useful in planning future sample allocations.

Table IV.8 shows the average number of calls per completed FIU, by site, size of MSA (and non-MSA sites), and sample type. For Round Two, the ratio of the total number of calls made to the RDD sample (938,602) divided by the number of completed FIUs (31,278) was 30.0. The ratio of total calls to FIUs for various sites to the overall mean is a measure of the relative level of effort for that site. For example, the average number of calls per FIU for the supplemental

TABLE IV.8

AVERAGE NUMBER OF CALLS PER COMPLETED RDD FIU, BY SITE, SIZE OF MSA, AND SAMPLE TYPE

		Responding Families (Number)	Total Calls (Number)	Calls per FIU (Average Number)	Relative to Overall Mean (Percent)
0	Supplemental Sample	3,251	95,338	29.33	97.73
01	Boston, MA Portion	1,120	36,726	32.79	109.27
02	Cleveland-Lorain-Elyria, OH PMSA	1,129	32,132	28.46	94.84
03	Greenville-Spartanburg-Anderson, SC MSA	1,220	30,197	24.75	82.48
04	Indianapolis, IN MSA	1,122	26,757	23.85	79.47
05	Lansing-East Lansing, MI MSA	1,185	29,561	24.95	83.13
06	Little Rock-North Little Rock, AR MSA	1,222	29,829	24.41	81.34
07	Miami, FL PMSA	1,173	57,112	48.69	162.25
08	Newark, NJ PMSA	1,195	50,313	42.10	140.30
09	Orange County, CA PMSA	1,125	47,171	41.93	139.73
10	Phoenix-Mesa, AZ MSA	1,109	33,933	30.60	101.96
11	Seattle-Bellevue-Everett, WA PMSA	1,014	35,497	35.01	116.66
12	Syracuse, NY MSA	1,150	24,725	21.50	71.65
13	Atlanta, GA MSA	264	7,681	29.09	96.96
14	Augusta-Aiken, GA-SC MSA	286	6,803	23.79	79.27
15	Baltimore, MD PMSA	292	8,551	29.28	97.59
16	Bridgeport-Danbury-Stamford, CT Portion	269	10,941	40.67	135.54
17	Chicago-Kenosha-Kankakee, IL-WI PMSA	299	11,297	37.78	125.91
18	Columbus, OH MSA	274	8,698	31.74	105.79
19	Denver-Boulder-Greeley, CO PMSA	274	9,625	35.13	117.06
20	Detroit, MI PMSA	279	8,826	31.63	105.42
21	Greensboro-Winston Salem-High Point, NC MSA	250	6,497	25.99	86.60
22	Houston-Galveston-Brazoria, TX PMSA	277	10,755	38.83	129.39
23	Huntington-Ashland, WV-KY-OH MSA	288	5,797	20.13	67.08
24	Killeen-Temple, TX MSA	294	7,160	24.35	81.16
25	Knoxville, TN MSA	292	8,017	27.46	91.49
26	Las Vegas, NV-AZ MSA	288	10,380	36.04	120.11
27	Los Angeles-Long Beach, CA PMSA	295	12,199	41.35	137.80
28	Middlesex-Trenton, NJ PMSA	288	9,997	34.71	115.67
29	Milwaukee-Racine, WI PMSA	270	6,253	23.16	77.18
30	Minneapolis-St Paul, MN-WI MSA	311	6,881	22.13	73.73

Table IV.8 (*Continued*)

		Responding Families (Number)	Total Calls (Number)	Calls per FIU (Average Number)	Relative to Overall Mean (Percent)
31	Modesto, CA MSA	323	8,904	27.57	91.86
32	Nassau-Suffolk, NY PMSA	318	11,331	35.63	118.74
33	New York City, NY PMSA	306	12,968	42.38	141.22
34	Philadelphia, PA-NJ PMSA	291	10,235	35.17	117.21
35	Pittsburgh, PA MSA	294	7,111	24.19	80.60
36	Portland-Salem, OR-WA PMSA	337	8,750	25.96	86.52
37	Riverside-San Bernardino, CA PMSA	321	9,583	29.85	99.48
38	Rochester, NY MSA	357	8,855	24.80	82.66
39	San Antonio, TX MSA	280	9,219	32.93	109.72
40	San Francisco, CA PMSA	256	12,100	47.27	157.51
41	Santa Rosa, CA PMSA	272	8,992	33.06	110.17
42	Shreveport-Bossier City, LA MSA	290	7,658	26.41	88.00
43	St Louis, MO-IL MSA	338	7,846	23.21	77.36
44	Tampa-St Petersburg-Clearwater, FL MSA	252	8,470	33.61	112.01
45	Tulsa, OK MSA	332	9,226	27.79	92.60
46	Washington-Hagerstown, DC-MD-VA-WV PMSA	306	10,145	33.15	110.48
47	West Palm Beach-Boca Raton, FL MSA	241	10,680	44.32	147.68
48	Worcester-Fitchburg, MA Portion	315	8,526	27.07	90.20
49	Dothan, AL MSA	330	7,547	22.87	76.21
50	Terre Haute, IN MSA	268	6,406	23.90	79.65
51	Wilmington, NC MSA	275	8,078	29.37	97.89
52	West-Central Alabama	326	7,681	23.56	78.52
53	Central Arkansas	380	7,649	20.13	67.08
54	Northern Georgia	261	5,972	22.88	76.25
55	Northeast Illinois	287	6,109	21.29	70.93
56	Northeast Indiana	273	5,873	21.51	71.69
57	Eastern Maine	317	5,674	17.90	59.65
58	Eastern North Carolina	290	6,636	22.88	76.25
59	Northern Utah	413	7,165	17.35	57.81
60	Northwest Washington	324	7,564	23.35	77.80
	All high-intensity sites	13,764	433,953	31.53	105.06
	All low-intensity sites	14,263	409,311	28.70	95.63
	Total RDD site sample	28,027	843,264	30.09	100.26

Table IV.8 (*Continued*)

	Responding Families (Number)	Total Calls (Number)	Calls per FIU (Average Number)	Relative to Overall Mean (Percent)
Sites Grouped by Population Size				
3 million+	3,437	120,832	35.16	117.15
2-3 million	7,950	275,243	34.62	115.37
1-2 million	4,629	147,439	31.85	106.14
<1 million large MSA	8,267	217,396	26.30	87.63
Small MSA (<200,000)	873	22,031	25.24	84.10
Non-MSA	2,871	60,323	21.01	70.02
RDD (site + supplemental sample) Grouped by Sample Type				
Round One Complete	14,978	238,328	15.91	53.02
Round One Refusal	115	5,768	50.16	167.14
Round One Other Household	977	67,875	69.47	231.51
Round One Undetermined Whether Household	189	54,569	288.72	962.15
Round One Not a Household	2,676	139,020	51.95	173.12
Old working banks	9,223	312,614	33.90	112.95
New working banks	3,120	120,428	38.60	128.63
Total RDD Supplemental + Site)	**31,278**	**938,602**	**30.01**	**100.00**

sample was 29.3, so the relative level of effort for the supplemental sample was 29.3/30.0 = 97.7 percent. Key differences in the level of effort, by site, size of MSA, and sample type are summarized here:

- ***Site Variation*.** The relative level of effort per FIU varies substantially by site. For high-intensity sites, it varies from a low of 79.5 percent of the mean in Indianapolis to a high of 162.3 percent of the mean in Miami. A similar pattern is repeated for low-intensity sites, with a ratio of nearly three to one between the site with the highest level of effort (West Palm Beach-Boca Raton, at 147.7 percent) and the site with the lowest (northern Utah, at 57.8 percent).
- ***MSA/Non-MSA Variation***. The level of effort also varies by size of MSA, from 117.2 percent of the mean for MSAs of 3 million or more to 70.0 percent of the mean for non-MSAs. The relative level of effort is about the same for MSAs of 3 million or more (117.2 percent) and for MSAs of 2 to 3 million (115.47 percent). It declines to 106.1 percent of the mean for MSAs of 1 to 2 million, to 87.6 percent for mid-sized MSAs, to 84.1 percent for small MSAs, and to 70.0 for non-MSAs.
- ***Sample Type*.** The level of effort to complete an interview with an FIU sampled from a Round One number linked to a completed interview was only 53.0 percent of the mean. The level of effort was above 100 percent of the mean for all other sampled groups and was more than nine times the mean for Round One telephone numbers with undetermined residency (962.2 percent).[5]

3. Response Rates for the Adult Self-Response Modules and Child's Physician Visit

Most of the FIU interview was conducted with an informant who answered for all sampled FIU members. However, each adult in the FIU was asked to self-respond to a subset of questions (the self-response module), including subjective assessments of health, tobacco use, satisfaction with care, and aspects of the physician–patient interaction. Efforts to obtain self-responses were successful in both rounds; 94.5 percent in Round One and 94.0 percent in Round Two completed these questions (Table IV.9).

[5]For Round Three, we used data on variation in the Round Two level of effort, by sample type, to develop an optimal sample allocation with respect to cost.

TABLE IV.9

ROUND ONE AND ROUND TWO RESPONSE RATES
FOR THE ADULT SELF-RESPONSE MODULE
(Percents)

	Round One	Round Two
Completed Module	94.5	94.0
Proxy Accepted		
Illness	0.5	0.5
Away and unavailable	0.6	0.7
Language barrier	0.1	0.1
Other or unspecified reason	1.1	0.3
	2.2	2.1
Refused		
Unable to Complete for Other Reasons	1.0	2.3
Total	100.0	100.0
Number of Adults	**49,807**	**48,724**

In certain circumstances, such as when an adult FIU member was too ill to respond, temporarily unavailable, or unwilling to respond after several interviewing efforts had been made, the family informant was allowed to complete the self-response module for that FIU member. Overall, 2.3 percent of the self-response modules in Round One and 1.6 percent in Round Two were completed by proxy respondents.

D. EFFORTS TO INCREASE RESPONSE TO THE SURVEY

In Chapter III, we described efforts to increase initial cooperation by developing survey messages and mailing advance materials to respondents. During data collection, we used a variety of efforts to increase response, including:

- Making numerous calls to determine residency (20 calls) and as many as 50 calls to complete interviews with residential households
- Offering Spanish-speaking interviewers to respondents who preferred to conduct the interview in that language
- Making multiple rounds of refusal conversion calls
- Offering monetary incentives
- Leaving messages on mechanical answering devices

1. Follow-Up Calls for the RDD Sample

Telephone numbers in the RDD sample were controlled by the CATI scheduler, which randomly assigned sampled telephone numbers to interviewers; nonscheduled calls were based on optimal calling patterns, dispersed over different times of the day and different days of the week. (As described in Chapter III, the survey introduction for the initial call varied according to whether the telephone number was linked to a household that had been interviewed in Round One and according to whether a letter and brochure describing the study had been mailed prior to the call.) Firm appointments were scheduled within a 20-minute window; other appointments

were scheduled within a 60-minute time period, based on information provided by the interviewer. Separate queues were set up for Spanish-speaking interviews and for refusal conversions (discussed below). A total of 20 calls were attempted to determine residency, and up to 50 calls (in total) were made to telephone numbers that were not assigned to the refusal conversion or Spanish queues, but that resulted in contact by voice or by mechanical answering device.

2. Follow-Up Calls for the Field Sample

Dwelling units selected for the field sample were screened by interviewers to identify households that had not had telephone service for a period of two weeks or more since the beginning of the RDD data collection period. Field interviewers made up to six visits to complete the household interview. Refusal rates were very low, and we did not make refusal conversion calls for the field sample. However, considerable efforts were made to obtain access to locked apartment buildings, which comprised a significant portion of sampled dwellings in some interviewing areas. These efforts included letters and calls by survey managers to supplement efforts by field interviewers.

3. Interviews Conducted in Spanish

We prepared a Spanish version of the CATI instrument and trained bilingual telephone interviewers to conduct interviews with family informants or adults for whom self-response modules were required and who preferred to conduct the interview in Spanish. In addition, two of the field interviewers (one in Miami and one in Orange County) spoke Spanish; in other sites, field interviewers attempted to use family members to translate the screener questions, when necessary. Overall, the percentage of FIUs conducted in Spanish increased from Round One (3.8 percent) to Round Two (4.3 percent; see Table IV.10). This finding is not surprising, given the

TABLE IV.10

SPANISH-SPEAKING FIU INTERVIEWS, BY SITE AND OVERALL

	Round One		Round Two	
	Spanish FIUs (Number)	FIU Interviews Completed in Spanish (Percent)	Spanish FIUs (Number)	FIU Interviews Completed in Spanish (Percent)
High-Intensity Sites				
Boston MA	33	2.93	49	4.38
Cleveland OH	7	0.60	5	0.44
Greenville SC	6	0.49	4	0.33
Indianapolis IN	1	0.08	3	0.27
Lansing MI	2	0.17	4	0.34
Little Rock AR	1	0.08	10	0.82
Miami FL	357	31.62	365	31.12
Newark NJ	62	5.07	60	5.02
Orange County CA	154	13.47	141	12.53
Phoenix AZ	81	6.92	83	7.48
Seattle WA	7	0.62	10	1.00
Syracuse NY	4	0.32	6	0.52
Low-Intensity Sites				
Atlanta GA	4	1.35	6	2.27
Augusta-Aiken GA/SC	1	0.34	1	0.35
Baltimore MD	0	0.00	3	1.03
Bridgeport CT	8	2.82	6	2.23
Chicago IL	10	3.41	24	8.03
Columbus OH	0	0.00	1	0.36
Denver CO	14	4.81	13	4.74
Detroit MI	2	0.65	2	0.72
Greensboro NC	2	0.74	2	0.80
Houston TX	31	11.07	25	9.03
Huntington-Ashland WV/KY/OH	0	0.00	0	0.00
Killeen TX	7	2.35	6	2.04
Knoxville TN	0	0.00	0	0.00
Las Vegas NV/AZ	15	5.62	16	5.56
Los Angeles CA	52	19.92	61	20.68
Middlesex NJ	7	2.25	10	3.47
Milwaukee WI	2	0.64	5	1.85
Minneapolis-St. Paul MN/WI	2	0.60	1	0.32
Modesto CA	39	12.75	38	11.76
Nassau NY	17	4.99	11	3.46
New York City NY	39	13.36	44	14.38
Philadelphia PA/NJ	11	3.56	6	2.06
Pittsburgh PA	0	0.00	0	0.00
Portland-Salem OR/WA	11	3.58	5	1.48
Riverside CA	30	9.87	48	14.95

TABLE IV.10 (*Continued*)

	Round One		Round Two	
	Spanish FIUs (Number)	FIU Interviews Completed in Spanish (Percent)	Spanish FIUs (Number)	FIU Interviews Completed in Spanish (Percent)
Low-Intensity Sites (continued)				
Rochester NY	3	0.85	2	0.56
San Antonio TX	16	5.35	19	6.79
San Francisco CA	15	5.34	21	8.20
Santa Rosa CA	14	4.91	9	3.31
Shreveport LA	0	0.00	0	0.00
St. Louis MO/IL	0	0.00	1	0.30
Tampa FL	0	0.00	10	3.97
Tulsa OK	0	0.00	1	0.30
Washington DC/MD	12	3.87	15	4.90
West Palm Beach FL	13	5.14	8	3.32
Worcester MA	3	0.97	10	3.17
Dothan AL	0	0.00	0	-
Terre Haute IN	0	0.00	0	-
Wilmington NC	0	0.00	0	-
West Central Alabama	0	0.00	0	-
Central Arkansas	1	0.26	8	2.11
Northern Georgia	6	2.20	6	2.30
Northeast Illinois	1	0.34	0	-
Northeast Indiana	0	0.00	0	-
Eastern Main	0	0.00	0	-
Eastern North Carolina	4	1.32	7	2.40
Northern Utah	8	2.12	3	0.73
Northwest Washington	4	1.21	9	2.78
Supplemental Sample	**114**	**3.48**	**127**	**3.91**
Total Overall	**1,233**	**3.80**	**1,330**	**4.25**

nationwide increase in the Hispanic population noted in reports on the 2000 Census. Spanish-speaking interviewers were critical in the Miami site, where just over 30 percent of the interviews were conducted in Spanish in each round. Spanish-speaking interviewers also were important in three other high-intensity sites: (1) Newark (about 5 percent of the interviews in each round were conducted in Spanish), (2) Orange County (about 13 percent in each round), and (3) Phoenix (about 7 percent in each round). In addition, interviews in Spanish comprised at least five percent of the completed FIU interviews in each round for 10 low-intensity sites.

4. Refusal Conversions

Based on our experience in Round One of the CTS Household Survey and in related surveys, we anticipated a high volume of refusals and trained a pool of our best interviewers to convert refusals. Refusal converters used information about the reason and intensity of the prior refusals in planning their calls. Overall, refusals occurred in 17,259 households, or 45.9 percent of the 37,582 residential households and other telephone numbers for which contact was made but residency was not determined. Secondary FIUs were less likely to refuse; only 1,045 of the 5,982 secondary FIUs (17.5 percent) refused the initial call.

Refusal conversion efforts were critical to achieving a high response rate, as at least one FIU interview was completed in 40.1 percent of the households that refused the initial call (Table IV.11). Conversions also increased response rates for secondary FIUs, as 35.1 percent eventually agreed to participate in the survey. (In households including more than one FIU, the first interview was completed with the primary FIU; other FIUs were known as secondary FIUs. See Section B of Chapter II.)

We attempted several rounds of refusal conversions, allowing a minimum of four weeks between refusal conversion attempts to minimize the likelihood of antagonizing respondents.

TABLE IV.11

ROUND TWO HOUSEHOLD (PRIMARY FIU) AND SECONDARY FIU REFUSAL CONVERSION RATES
(Percents)

Refusal Conversion Attempts	Household (Primary FIU)	Secondary FIU
Converted After One Refusal	23.2	26.7
Converted After Two Refusals	11.8	8.2
Converted After Three or More Refusals[a]	5.1	0.2
Total converted	40.1	35.1
Not Converted	59.9	64.9
Number of Refusals	**17,259**	**1,045**

[a]Interviewing supervisors could authorize additional efforts to convert households that refused more than twice; however, most refusals were assigned a final disposition after two refusals.

The refusal pool included respondents who hung up the telephone before the interviewer completed her introduction (HUDIs), those who said they preferred not to be interviewed (refusals), those who terminated the call after the screener was completed (breakoffs), and those with electronic privacy managers.[6] Of the 17,259 households that refused at least once, only 23.2 percent completed FIU interviews as a result of the first conversion effort. An additional 11.8 percent agreed to participate after a second refusal, and 5.1 percent agreed after three or more refusals. The cumulative refusal conversion rate after all efforts was 40.1 percent of households that refused the initial call.

As noted, only 17.5 percent of secondary FIUs refused the initial call. Survey participation by an informant for one FIU increased the likelihood that other FIU informants would agree to be surveyed. The refusal conversion rate among secondary FIUs that refused was 35.1 percent—26.7 percent after the first refusal, 8.2 percent after the second, and 0.2 percent after three or more.

Refusal conversion efforts were most successful for households sampled from telephone numbers linked to completed interviews in Round One (Table IV.12). The conversion rate for telephone numbers sampled from Round One completed interviews was 56.3 percent, compared with 22.3 percent for numbers sampled from Round one noninterviews, and compared with 39.9 percent for new sample. These results suggest that prior participation in the survey increases the likelihood of participation on the next round, even if a household informant initially refuses in that round. Multiple rounds of refusal conversions were required for each type of sample to achieve an acceptable overall conversion rate.

---

[6]A privacy manager is a call-screening device that works with Caller ID to intercept and identify incoming calls. The privacy manager requests the caller's name, which appears on the Caller ID box. The recipient can then choose to accept or reject the call, send the call to a mechanical answering device, or send a scripted rejection to solicitors.

TABLE IV.12

ROUND TWO HOUSEHOLD REFUSAL CONVERSION RATES,
BY SAMPLE TYPE
(Percents)

Refusal Conversion Attempts	Round One Completes	Round One Noninterviews	New Sample	Total Sample
Converted After One Refusal	32.7	14.6	22.2	23.2
Converted After Two Refusals	19.5	5.0	10.7	11.8
Converted After Three or More Refusals[a]	4.1	2.7	7.0	5.1
Total converted	56.3	22.3	39.9	40.1
Not Converted	43.7	77.7	60.1	59.9
Number of Refusals	**4,799**	**4,259**	**8,201**	**7,259**

[a]Interviewing supervisors could authorize additional efforts to convert households that refused more than twice; however, most refusals were assigned a final disposition after two refusals.

The impact of refusal conversion efforts on household and FIU response rates is shown in Table IV.13. Assuming no refusal conversion efforts were made and households and FIUs were retired after the first refusal, the household level response rate would have been 48.3 percent and the FIU response rate would have been 45.6 percent. Most of the gain from refusal conversions comes from the first refusal conversion call, with a gain of more than 10 percentage points. However, significant gains in response rate occur with each incremental refusal conversion effort—a gain of about five percentage points after the second refusal and two percentage points after three or more.

5. Monetary Incentives

As a result of experiments conducted in Round One (Strouse 1997) and use of the CTS Household Survey sample as a frame for other surveys that compensated respondents, we decided to offer $25 to nearly all Round Two FIUs.[7] A few Round One FIUs who had been offered $35 or $50 because they were part of Round One experiments were offered the same amount for participating in Round Two. For each round of the survey, $25 was offered to households eligible for the field component. Most of the Round One participants received compensation, so we had accurate names and addresses (that were two years old) for more than 90 percent of this sample. For other sample members, we could obtain addresses for most households with published telephone numbers.

The main decision for Round Two was whether to "prepay" respondents (that is, to include compensation with the letter and brochure describing the study sent to households for which we had contact information). In recent experiments with the Survey of Consumer Attitudes, a

[7]Surveys that provided compensation include the UCLA Health Care for Communities survey and the RAND Community Quality Index survey.

TABLE IV.13

ROUND TWO RDD CUMULATIVE RESPONSE RATE,
BY NUMBER OF REFUSAL COVERSIONS
(Percents)

	Household	FIU Level
Assumes No Refusal Conversion[a]	48.3	45.6
Assumes One Refusal Conversion[b]	58.5	55.7
Assumes up to Two Refusal Conversions[c]	63.7	60.8
Actual Survey Results[d]	65.9	62.9

[a]Response rate recomputed, assuming that no efforts were made to convert initial refusals.

[b]Response rate recomputed, assuming that efforts were made to convert only first refusals.

[c]Response rate recomputed, assuming that efforts were made to convert first and second refusals.

[d]Actual response rate for the Round Two RDD sample; supervisors could authorize more than two refusal conversions if they believed prior refusals had not been hostile, and that additional efforts might be effective.

monthly telephone survey, Singer (2000) demonstrated that prepayment of a small incentive ($5 in cash) accompanying an advance letter increased response rates relative to either promised payment or an advance letter without accompanying compensation.

The level of compensation is much higher for the CTS surveys ($25 per FIU), which are lengthy, require participation by each adult in the household, and are linked to other surveys. Consequently, we decided that cash incentives were not practical.[8] We also were concerned that sending a check prior to the initial call would remove the prepaid incentive as an option for families that refuse. In addition, we were concerned about the financial risk in prepaying incentives to such a large number of households. Therefore, we limited prepaid incentives to converting households that had refused or that had not responded after many calls. Prepaid incentives were sent by check.

Although $25 proved to be the most effective level of compensation for promised payments used in Round One and was used as the promised payment for Round Two, we were unsure whether offering $25 as a prepayment would be sufficient to encourage reconsideration by households that had already refused (often more than once). Consequently, we randomized households selected for prepayment in Round Two, mailing one-half a check for $25 and the other half a check for $50. Most of the prepayment sample consisted of households that had refused twice (n = 3,688); in addition, we included a small number of late first refusals (n = 630) and a small number of those that had been called many times without either refusing or completing an interview (n = 367).

[8]We are testing debit cards as an alternative for future surveys. On the one hand, respondents cannot use these cards without participating in the survey, as they can with cash or checks. On the other hand, the cards are mailed in advance, and they can be used like cash.

The results of this effort are summarized in Table IV.14. The $50 incentive resulted in a slightly higher completion rate (44.0 percent) than did the $25 incentive (38.4 percent). However, the difference was not large enough to warrant using such a large incentive in future surveys. A significantly larger fraction of both samples cashed checks without completing the interview (13 percent in the two samples combined) than responded without cashing checks (4.7 percent), somewhat increasing the cost of this effort relative to promised payments.

The completion rate for the prepayment effort was higher for households that had not refused but were called many times (45.2 percent) and for first refusals (56.4 percent) than for second refusals (38.4 percent; Table IV.15).[9] However, households that had been called many times without refusing or refusing only once were selected because they had not responded near the end of the survey; they are not necessarily representative of all households with many calls or one-refusal households.

Prepayment efforts were more successful with households that had participated in the Round One survey (Round One completes) or that were selected for the first time in Round Two than with households that had been classified as Round One noninterviews (Table IV.16). We limited this comparison to households that had refused at least once, as nonrefusals were a small and nonrandomly distributed subgroup of the prepayment sample. The completion rate for households linked to telephone numbers completing Round One interviews was 42.9 percent, compared with 38.5 percent for new sample, and 27.5 percent for Round One noninterviews.

[9]In evaluating the prepayment effort, we define a *completion rate* as the number of households with at least one completed FIU interview divided by the total number of final dispositions for cases that were mailed checks.

TABLE IV.14

ROUND TWO DISPOSITION OF PRIMARY FIUs' PREPAID INCENTIVES, BY AMOUNT AND WHETHER CASHED

	$25		$50		Total	
Round Two Disposition	Number	Percent	Number	Percent	Number	Percent
Complete						
Cashed	764	34.2	949	38.8	1,713	36.6
Not cashed	95	4.2	128	5.2	223	4.7
Total Completes	859	38.4	1,077	44.0	1,936	41.3
Refused						
Cashed	302	13.5	308	12.6	610	13.0
Not cashed	1,050	46.9	826	33.7	1,876	40.0
Total Refused	1,352	60.4	1,134	46.3	2,486	53.1
Ineligible[a]						
Cashed	10	0.4	29	1.2	39	0.8
Not cashed	9	0.4	43	1.8	52	1.1
Total Ineligible	19	0.8	72	3.0	91	1.9
Other[b]						
Cashed	5	0.2	103	4.2	108	2.3
Not cashed	2	0.1	62	2.5	67	1.4
Total Other	7	0.3	165	6.7	172	3.7
Total	**2,237**	**100.0**	**2,448**	**100.0**	**4,685**	**100.0**

[a]Nonresidential; nonworking telephone; cellular telephone; noncivilian household.

[b]Maximum calls; end of study; undetermined residual status.

TABLE IV.15

ROUND TWO DISPOSITION OF PRIMARY FIUs' PREPAID INCENTIVES, BY NUMBER OF REFUSALS
(Percents)

	Number of Refusals			
Round Two Disposition	0[a]	1	2	All
Complete	45.2	56.4	38.4	41.3
Refused	0.8	39.4	60.6	53.1
Ineligible[b]	12.8	2.5	0.8	1.9
Other[c]	41.1	1.8	0.3	3.7
Number of Families	**367**	**630**	**3,688**	**4,685**

[a]Never refused but offered prepaid incentives because they had been called many times without completing interviews or refusing.

[b]Nonresidential; nonworking telephone; cellular telephone; noncivilian household.

[c]Maximum calls; end of study; undetermined residential status.

TABLE IV.16

ROUND TWO DISPOSITION OF PRIMARY FIUS THAT HAD REFUSED AND WERE OFFERED PREPAID INCENTIVES, BY SAMPLE TYPE
(Percents)

	Sample Type			
Round Two Disposition	Round One Complete	Round One Noninterview	New Sample	Total
Complete	42.9	27.5	38.5	41.0
Refused	56.0	66.5	60.0	57.5
Ineligible[a]	0.8	3.5	0.9	1.0
Other[b]	0.3	2.4	0.6	0.5
Number of Families	**3,132**	**287**	**899**	**4,318**

[a]Nonresidential; nonworking telephone; cellular telephone; noncivilian household.

[b]Maximum calls; end of study; undetermined residential status.

6. Messages on Mechanical Answering Devices

Some residential households were difficult to contact because they used mechanical answering devices to screen calls. Interviewers left the following message on the devices to counter these chronic no-answers:

- ***Households that Had Never Participated in the CTS:*** I'm calling for the Community Tracking Study, a research project to see how managed care and other health care changes are affecting people. We're not selling anything or asking for money. We would like your household to participate in a brief interview, and we will send you $25 for helping us.[10] Please call Jackie Licodo at 1-800-298-3383. Thank you!

- ***Reinterviewed Households:*** I'm calling for the Community Tracking Study, the health care study your household participated in two years ago. We recently mailed you a letter about the study and would very much like to interview your household again. We will send you $25 for helping us.[10] Please call Jackie Licodo at 1-800-298-3383. Thank you!

The interviewer also was instructed to leave notes in the CATI system indicating that the message had been left on the answering device, and to reference the message when calling back the next time. A second message could be left after a one-week interval; the limit was two messages per month.

E. QUALITY ASSURANCE

1. RDD Sample

Production reports and regular on-line monitoring were used to evaluate interviewer performance. Daily production reports provided information on several performance indicators, including completed interviews and self-response modules, number of calls made, number of refusals, refusal conversions, time per call, time per interview, and the ratio of completed interviews to time spent charged to interviewing. Interviewer conduct during interviews was

evaluated primarily by having supervisors monitor actual calls, supplemented by review of interviewers' notes maintained in the CATI system. (All calls and notes recorded about monitored calls are maintained by the CATI system.)

The monitoring system enables supervisors to listen to interviews without either the interviewers' or respondents' knowledge. It also allows supervisors to view interviewers' screens while an interview is in progress. Interviewers are informed they will be monitored but do not know when observations will take place. Supervisors concentrate on identifying behavioral problems involving inaccurate presentation of information about the study; errors in reading questions; biased probes; inappropriate use of feedback in responding to questions; and any other unacceptable behavior, such as interrupting the respondent or offering a personal opinion about specific questions or about the survey. The supervisor reviews results with the interviewer after the interviewer completes her or his shift.

2. Field Sample

Ten percent of the screened households that were ineligible for the survey (that is, had telephone service with no interruption) were validated by telephone from MPR's Princeton office. All eligible households were interviewed by cellular telephone and were subject to standard monitoring procedures used for the RDD sample.

F. DATA EDITING, CODING, AND CLEANING

One of the most important advantages of computer-assisted surveys is that errors can be identified and corrected during the interview by building logic, range, and consistency checks into the program. The CATI program (CASES) also permits interviewers to back up and change

(continued)

[10]This message was not used for households participating in experiments in which large incentives were offered.

answers to previously answered questions without violating instrument logic.

Because of differences in design, separate instrument programs were written for the RDD primary FIU survey, RDD secondary FIU survey, and field survey. Separate Spanish versions of the programs were written as well, but their structures were the same as those of the corresponding English versions. Separate cleaning programs were written for each of the three survey instruments. The instrument cleaning programs enforce questionnaire logic strictly. An interview could not be certified as clean until all appropriate questions had either been answered or assigned an acceptable nonresponse value, and until the data record for each interview was consistent with the instrument program logic.

Survey questions were primarily closed ended. Questions on industry were open ended, and text responses were coded to the two-digit (1987) Standard Industrial Classification (SIC) coding structure.[11] A program was written to read text responses and, based on character strings in the text, to assign two-digit codes. Responses without recognizable patterns were manually coded; a sample of computer-generated codes also was reviewed by a coder.

Other open-ended items included personal contact information, insurance plan names, and employer names. Personal identifying information remained confidential and was maintained in a separate file used only to assign respondent payments and subsequent interviews. Information on insurance plan names and employer names was used to conduct a separate followback survey to link data provided by insurers with the household file. In addition, the survey included text responses to closed-ended questions, with options for answers that did not correspond to precoded categories. Files including text responses to these questions were delivered to HSC.

---

[11]The SIC has been replaced with the North American Industry Classification (NAIC) System. However, we chose to retain the industry categories used in Round One to maintain consistency in a longitudinal survey.

G. REFORMATTING DATA FILES AND FILE DELIVERY

A program was written to reformat the cleaned instrument responses into FIU- and person-level data files. Analysis files were then prepared in SAS, and additional edits performed. The additional edits included checks on the number of missing values for FIU- and person-level data, additional checks on relationship codes, deletion of FIU and person records for which inconsistencies among relationships could not be resolved, assignment of additional nonresponse values, and some constructed variables. Weights were applied to the data files (see Chapter V), and weighted data files were delivered to SSS, which was responsible for building the Public Use Files. Instrument cleaning, reformatting, and SAS programs used in the preparation of these files are maintained by MPR.

LEFT INTENTIONALLY BLANK

V. WEIGHTING AND ESTIMATION

A. OVERVIEW

The sample design was a complex one that used stratification, clustering, and oversampling. The use of unweighted data is likely to produce seriously biased estimates because the unweighted samples are distributed differently than are the populations they represent. Thus, weights were designed to restore proportionality to the sample and were adjusted to compensate for nonresponse at the household, FIU, and person levels. This difference in proportionality occurred for the following reasons:

- Design decisions, such as setting fixed sample sizes for sites, restricting the high-intensity sites to MSAs with populations of 200,000 or more and subsampling children, resulted in oversampling of some groups and undersampling of others.
- Sample frames did not cover the entire study population. The RDD frame omitted telephone banks of 100 numbers that contained no published household numbers, and the field sample, which excluded areas with high telephone penetration, was restricted in coverage to MSAs with populations of 200,000 or more.
- Some households had differing chances of selection because of the number of telephones they owned or interruptions in telephone service.
- Nonresponse to the survey differed among sites and among subgroups of the population.
- In Round Two, differential probabilities of selection were used depending on whether the telephone number or address was part of the Round One sample and, if so, on whether it was coded as a breakoff in Round One.

Although the correct use of weights in analyzing CTS Household Survey data substantially reduces the bias of estimates resulting from the sample design and survey nonresponse, the weights do not address the potential for bias resulting from item nonresponse or response errors. The procedures we used to impute missing data for individual variables are discussed in the Household Survey Public Use File (Technical Publication No. 21) (Center for Studying Health

System Change 2001). Estimates of sampling error that do not account for the use of weights and the complex nature of the sample are likely to be severely understated. Specialized software is required to properly estimate standard errors of estimates from this survey; procedures for using this software are also included in Technical Publication No. 25.

1. Weights Provided for Public Use

Ten analysis weights, summarized in Table V.1, are available for researchers' use when using the Round Two data. Weights were constructed to allow for both site-specific and national estimates for individuals and FIUs.[1] Site-specific estimates are made for a site or involve comparisons of sites. In contrast, national estimates involve inferences to a population broader than any one site or group of sampled sites. We use the term *national estimates* to include estimates for subgroups of the national population that are defined by geography or by economic or demographic classifications. The weights are computed using the features of the sampling design; therefore, all weights are design based.

Weights are provided for five classes of estimates, defined as follows:

1. ***Site Sample.*** Weights for national estimates that use data from the site sample
2. ***Supplemental Sample.*** Weights for national estimates that use the supplemental sample
3. ***Combined Sample.*** Weights used for national estimates that combine data from the site sample and supplemental sample
4. ***Augmented Site Sample (Site-Specific Estimates).*** Weights for site-specific estimates that use data from the site sample, augmented with observations from the supplemental sample that fell within the boundaries of the 60 sites

[1]Throughout this report, "national" refers to the population of the 48 contiguous states and the District of Columbia. It does not include Alaska or Hawaii.

TABLE V.1

NAMES OF ROUND TWO CTS HOUSEHOLD SURVEY WEIGHTS

	Analytic Sample and Estimate Type				
	Site-Specific Estimate	National Estimate			
Level of Analysis	Augmented Site Sample	Site Sample	Supplemental Sample	Combined Sample	Augmented site Sample
Person	*WTPER1*	*WTPER2*	*WTPER3*	*WTPER4*	*WTPER5*
FIU	*WTFAM1*	*WTFAM2*	*WTFAM3*	*WTFAM4*	*WTFAM5*

5. ***Augmented Site Sample (National Estimates*).** Weights for national estimates that use data from the site sample, augmented with observations from the supplemental sample that fell within the boundaries of the sites. (This weight is new to Round Two. Researchers linking site-specific data to their national estimates requested it, as they wanted to maximize the nominal sample size.)

For each of the five classes of estimates, two weights are provided: (1) a weight for analyzing FIU data, and (2) a weight for conducting person-level analyses.[2]

In many surveys, the nonresponse adjustments, poststratification adjustments, and other adjustments can introduce variation for the sampling weights. In some situations, the combination of these adjustments produces disproportionately large weights. Although technically appropriate, these large weights can decrease the accuracy of point estimates and decrease the precision. In this study, as in other surveys, we have chosen to trim extremely large weights. Although trimming can introduce some bias in survey estimates, we expected little or no bias in the survey and substantially better precision.

Trimming weights reduces sampling error by reducing the values of extremely large weights and distributing the excess among other weights. Although the difference between estimates using the trimmed weights or untrimmed weights is quite small (the extent of trimming was not great), the trimmed weights result in better precision.

[2]For Round One, we developed a separate set of weights accounting for self-response module nonresponse. These weights were used for analyses containing nonimputed variables from the self-response module. Details on these weights can be found in the methodology report for Round One (Strouse et al. 1998). We decided not to produce these weights for Round Two because we had substantially fewer self-response module nonrespondents in Round One than we had anticipated based on the preliminary data. The overall level of nonresponse was so small (3.2 percent for adults and 7.7 percent for children) that we did not think it would be worthwhile to produce a full set of self-response module weights in Round Two.

The combined weights include individual-level and FIU weights for national estimates designed to combine data from the site and supplemental samples. These weights are based on the relative variances of the two samples and enable researchers to more easily take advantage of the increased precision of the combined samples.

The weights for the augmented site sample, combined sample, and site sample for the high-intensity sites include cases from both the RDD and field components. The supplemental sample weights and site sample weights for the low-intensity sites include only RDD cases. We have assumed that most researchers making individual-level national estimates (including estimates for subgroups of the national population) will prefer to use the combined weights, which cover both the site sample and the supplemental sample. The precision of these estimates is substantially greater than the precision of estimates obtained for either sample alone or for the augmented sample, especially for estimates about subgroups.[3] However, any of these samples alone will produce unbiased estimates.

2. Constructing Weights

Each weight is the product of several factors:

- An initial weight, the inverse of the probability of selection, to correct for differences in probabilities of selection
- Nonresponse adjustment factors, to correct for differential nonresponse at the individual, FIU, and household levels
- Factors to adjust for interruptions in telephone service
- Poststratification adjustments to fit weighted counts to external estimates of the population

[3]The augmented sample can be used to maximize the nominal sample size when making national estimates based on site-level variables. If site-level explanatory variables are used, about half the supplemental sample cases cannot be included because their sites were not among the 60 sampled sites. These cases are assigned a weight of zero for this type of estimate.

Other adjustment factors for specific weights include:

- Factors to allow integration of the RDD and field components for the augmented site sample weights, site weights for the high-intensity sites, and combined weights
- A variance-based factor for the combined weights that allows the site and supplemental samples to be used together for national estimates

3. Changes in Weighting Methods Between Rounds One and Two

Round Two weighting procedures were the same as those used in Round One with three exceptions. First, we included a sample of telephone numbers selected in Round One as part of the Round Two sample. Consequently, the probability of selection of telephone numbers selected in Round two had to take into account Round One probabilities of selection. All Round One telephone numbers sampled for Round Two were selected with the same probability except for firm Round One refusals, which were selected at a lower rate. Second, we used propensity modeling procedures to adjust for household nonresponse among Round One completed interviews sampled for Round Two because we had more information about their characteristics than about other sample. For the rest of the Round Two sample we used a weighting cell adjustment, which was the method used for all sampled households in Round One. Third, we used external estimates to impute residency for telephone numbers where residency was not determined during the interview (see IV.C.1.a for the reasoning behind this decision). In addition, we added two new weights for round two: a weight to make national person-level estimates using the augmented site sample, and a comparable weight to make family-level estimates. After analyzing Round One results, we decided not to produce separate self-response module weights in Round two, as we did in round one. Round Two changes in weighting procedures are discussed in more detail in this chapter.

4. Sampling Error Estimation

Because sample-based estimates of population characteristics are not based on the full population, some element of uncertainty is always associated with these estimates. This element of uncertainty, known as *sampling error*, is an indicator of the precision of an estimate. Sampling error is generally measured in terms of the standard error or the sampling variance, which is the square of the standard error.[4] The standard error can be used to construct confidence intervals around estimates; for example, one can use the standard error to produce a range of numbers around an estimate with 95 percent confidence that the true value lies within that range.

The complexities of the CTS Household Survey design (which includes stratification, clustering, and oversampling) preclude the use of statistical packages for variance estimation that do not account for such a design in their algorithms. The variance estimates from these statistical packages may severely underestimate the sampling variance in the Household Survey. The CTS data therefore require the use of specialized techniques for estimating sampling variances; it is necessary to use survey data analysis software or options or specially developed programs designed to accommodate the statistic being estimated and the sampling design.

The sampling variance in the Household Survey is a function of the sampling design and the population parameter being estimated; it is called a *design-based sampling variance*. The CTS

[4]The sampling variance is a measure of the variation of an estimator attributable to having sampled a portion of the full population of interest, using a specific probability-based sampling design. The classical population variance is a measure of the variation among the members of the population, whereas a sampling variance is a measure of the variation of the *estimate* of a population parameter (for example, a population mean or proportion) over repeated samples. The population variance is different from the sampling variance in the sense that the population variance is a constant, independent of any sampling issues, whereas the sampling variance decreases as the sample size increases. The sampling variance is zero when the full population is observed, as in a census.

database contains "fully adjusted" sampling weights for site-specific estimates and national estimates of FIUs and persons, as well as the information on sample design parameters (that is, strata and clusters) necessary to estimate the sampling variance for a statistic.

Most common statistical estimates and analysis tools (such as percentages, percentiles, and linear and logistic regression) can be implemented using Taylor series approximation methods. Survey data software, such as SUDAAN (Shah et al. 1997), uses the Taylor series linearization procedure and can handle the multistage design, joint inclusion probabilities, and variance components in the Household Survey design.[5]

The remainder of this chapter discusses weighting procedures and sampling error estimation for the CTS Household Survey in more detail. Sections B and C discuss the weights for the RDD and field samples, respectively. Section D explains the procedure for integrating the RDD and field samples. Section E presents a topic that pertains to all the weights—it describes the procedures to identify and trim extremely large sampling weights. Section F discusses the weights for combining the site and supplemental samples for national estimates. Finally, Section G discusses appropriate methods for estimating sampling error for the Household Survey.

B. WEIGHTING THE RDD COMPONENT

Separate weights were constructed for the RDD sample components of the augmented site sample, site sample, and supplemental sample. In Section B.1, we present the general approach for constructing RDD weights at the household, FIU, and person levels. For each level, we describe the relevant sampling weights (defined here as the reciprocal of the probability of

[5]Other programs, such as Stata and Wesvar (available through SPSS), as well as newer versions of SAS, can also be used to estimate variances, but these programs do not fully take advantage of the sample design and will result in conservative estimates of the variance.

selection) and the nonresponse and poststratification adjustments to the weights. In Sections B.2 through B.4, we present specific issues pertaining to the construction of the three types of RDD sample weights.

1. General Weighting Approach

A general weighting approach was applied to the RDD weights. As explained in Chapter I, sampling took place in several stages. In the first stage, we selected the 60 sites and then randomly selected the high-intensity sites from among the 60. For the RDD sample, we selected telephone numbers, identified households, defined FIUs within households, and collected data on FIUs and people within FIUs (that is, collected data at the FIU and person level).[6] All these stages were considered in weighting. The initial weight of a unit (whether a telephone number, household, FIU, or person) is defined here as the reciprocal of its selection probability, incorporating the selection probability of the prior stage(s).

After constructing weights for site selection, we constructed initial weights for telephone numbers and adjusted for nonresponse. Then, we computed initial weights for households whose telephone numbers were sampled and adjusted the initial weights for nonresponse at the household level. The sum of the household weights was compared with a published national estimate of households, after which a poststratification adjustment factor was applied to the weights.

For FIUs in sampled households, we started with the final household weights and then adjusted for secondary FIU nonresponse; the result was an FIU-level weight. Finally, analogous steps were used to construct weights at the person level, after adjusting for the probability of selection of sampled children.

[6]People in FIUs include both adults and a sample of children.

a. Telephone Number Initial Weight

The telephone number was the second stage of selection for the site sample, and the first stage of selection for the supplemental sample. The telephone sampling weight accounted for the probability of selection of telephone numbers within each site, stratum, and Round One disposition. The probability of selection had to account for the fact that most of our sampled telephone numbers in Round Two also had a chance of being selected in Round One (as many were).

We calculated the probability of selection of a Round Two telephone number (in stratum *h*, with Round One disposition *d*) as follows[7]:

$$(1) \quad P_{hd} = (P1_h \cdot P2_{hd}) + (P3_h \cdot P4_{hd}),$$

where $P1_h$ is the probability of selection in Round One; $P2_{hd}$ is the probability of selection in Round Two given that the telephone number was selected for release in Round One; $P3_h$ is the probability of *not* being selected for release in Round One ($1 - P1_h$); and $P4_{hd}$ is the probability of selection of new cases in Round Two. The first step was to calculate these probabilities of selection for each telephone number.[8]

The probability of selection in Round One, in each site or stratum, is essentially the number of telephone numbers released out of the total number of telephone numbers in working banks.

[7]Throughout this chapter, we use the term *stratum h*. In the low-intensity sites, in which substratification was not used, stratum *h* refers to the entire site. For the high-intensity sites, it refers to the substrata within sites used in selecting the sample. For the supplemental sample, it refers to the five strata used in selecting the sample. Strata and substrata are defined in Chapter II.E.

[8]Because some assignments of telephone exchanges to sites and strata changed over time, we had to decide which round's geographic classifications to use when calculating the probabilities. For $P1_h$, $P2_{hd}$, and $P3_h$, we used geographic classifications as defined in Round One. For $P4_{hd}$, we used Round Two classifications.

A telephone number *bank* was defined as the first 8 digits of a 10-digit telephone number; a bank has 100 possible 10-digit telephone numbers associated with it. If at least 1 of these 100 possible telephone numbers was listed in a telephone directory as a residential number, then the bank was designated as a *working bank*. The general form of this probability is:

$$(2)\quad P1_h = P(\textit{phone p in stratum h, in Round One}) = \frac{n_{h1}}{N_{h1}} \cdot \frac{nrel_{h1}}{n_{h1} - nbad_{h1}},$$

where N_{h1} is the number of working telephone banks in stratum *h* in Round One times 100; n_{h1} is the number of telephone numbers initially selected in stratum *h* in Round One; $nbad_{h1}$ is the number of telephone numbers found to be nonworking or business numbers in stratum *h*, using GenesysID, in Round One; and $nrel_{h1}$ is the number of telephone numbers released in stratum *h* in Round One. All telephone numbers *released* in Round One or that *could* have been released in Round One were assigned this value for $P1_h$. Round Two telephone numbers that were in nonworking banks in Round One ("new" cases) were assigned a value of zero for $P1_h$.

The probability of selection of a telephone number in Round Two, given that it was released in Round One ($P2_{hd}$), depended on how the Round One case was resolved. We classified all Round Two released telephone numbers into one of the following seven categories depending on their Round One disposition (*d*):

1. Round One household complete
2. Round One household breakoff
3. Round One household, other nonresponse
4. Round One undetermined if household
5. Round One nonhousehold
6. Not released in Round One but could have been ("old working banks")
7. Could not have been released in Round One ("new working banks")

For each telephone number released in Round Two, we calculated the probability of selection within stratum *h* and Round One disposition:

$$(3)\quad P_{hd2} = P(\text{phone } p \text{ in stratum } h, \text{with disposition } d, \text{in Round Two}) = \frac{n_{hd2}}{N_{hd2}} \cdot \frac{nrel_{hd2}}{n_{hd2} - nbad_{hd2}},$$

where N_{hd2} (for d = 1,2,3,4,5) is the number of telephone numbers released in Round One in stratum *h* with disposition *d;* N_{hd2} (for d = 6,7) is the number of working telephone banks in stratum *h* with disposition *d* in Round Two times 100; n_{hd2} is the number of telephone numbers initially selected in stratum *h* with disposition *d* in Round Two; $nbad_{hd2}$ is the number of telephone numbers found to be nonworking or business numbers in stratum *h* with disposition *d*, using GenesysID, in Round Two[9]; and $nrel_{hd2}$ is the number of telephone numbers released in stratum *h* with disposition *d* in Round Two.

For cases that had been released in Round One (d = 1,2,3,4,5), within each site/stratum *and Round One status* we calculated $P2_{hd}$ using the formula given in Equation (3) for P_{hd2}. Round Two telephone numbers that were not released in Round One (d = 6), but that could have been (old working banks), were assigned a weighted average, within *h*, of the values of $P2_{hd}$ for the cases that had been released in Round One. Round Two telephone numbers that were in nonworking banks in Round One (new working banks; d = 7) were assigned a value of zero for $P2_{hd}$.

Probability $P3_h$ (the probability of not being selected in Round One) is simply calculated as $1 - P1_h$. Finally, we calculated $P4_{hd}$ (the probability of selection of a new telephone number in Round Two, within site/stratum and separately for telephone banks that were considered working

[9]For d = 6,7, this term also includes cases excluded because they were in the Round One sample.

in Round One and for telephone banks that were newly working in Round Two), using the formula given in Equation (3).[10] Round Two telephone numbers that had been released in Round One were assigned the value of $P4_{hd}$ calculated for telephone numbers in the old working bank group (those that could have been selected in Round One but had not been).

When calculating the probability of selection for the augmented site sample and for the combined sample, we must account for the fact that the telephone numbers had dual probabilities of selection. That is, they could have entered the sample through the site sample and through the supplemental sample.

For the augmented site sample weight, we assigned to each telephone number the "alternative site" and "alternative stratum" it would have received had the case been selected in the other sample. Thus, we determined the stratum into which each site sample telephone number would have fallen, had it been selected through the supplemental sample instead, and determined the site (and substrata, if applicable) into which each supplemental sample telephone number would have fallen, had it been selected through the site sample instead. We then attributed both the actual and alternative probabilities *P1* through *P4* to each case and calculated the probability of selection as follows[11]:

(4) $$P_{hd-site} = (P1_{h-site} \cdot P2_{hd-site}) + (P3_{h-site} \cdot P4_{hd-site}).$$

(5) $$P_{hd-supp} = (P1_{h-supp} \cdot P2_{hd-supp}) + (P3_{h-supp} \cdot P4_{hd-supp}).$$

(6) $$P_{hd} = P_{hd-site} + ((1 - P_{hd-site}) \cdot P_{hd-supp}).$$

[10]Note that the $P3_h$ term accounts for the overlap of residual cases between Round One and Round Two (d = 6). The product of $P3_h$ and $P4_{hd}$ is equivalent to the probability of selection conditional on not being sampled in Round One.

[11]For the site sample cases, the "alternative" site and strata would be those from the supplemental sample. For the supplemental sample cases, the "alternative" site and strata would be those from the site sample.

Finally, for all the weights calculated, the sampling weight for telephone *p* in stratum *h* (with Round One disposition *d*) was set equal to the reciprocal of this probability of selection:

$$(7) \quad SW(phone_{hpd}) = \frac{1}{p_{hd}}.$$

Note that these probability formulas differed slightly depending on the type of estimate (national or site-specific) and sample type (site, supplemental, or augmented) for which the weight was designed (described in more detail in Sections B.2 through B.5 of this chapter).

b. Adjustment to Telephone Weight for Undetermined Residency

For the telephone number weight, an adjustment was made for undetermined eligibility status (inability to determine whether a sampled telephone number was a working residential number). Different methods were used for telephone numbers selected from Round One completes and other samples.

Round One Completes. For telephone numbers selected from Round One completes, we used the Round One data to adjust for this type of nonresponse. A stepwise logistic regression model[12] using Round One data predicted the likelihood that we would not be able to determine the residency for the telephone number in Round Two. Table V.2 shows the variables found to be significant in the "screener model" for the site and supplemental samples. The weighting adjustment factor, $A'_{nr}(phone_c)$, was calculated based on the predicted probability from this model, $\hat{Y}_{screener}$, as follows:

[12]Significance for entry = .15; significance for exit = .10.

(8) $A'_{nr}(phone_c) = \frac{1 + exp(\hat{Y}_{screener})}{exp(\hat{Y}_{screener})}$.

We capped this adjustment factor at 2.5 to avoid excessive increases to the weights. For telephone numbers with unresolved Round Two residency (but associated with a Round One complete), we set this factor equal to zero. A "nonresponse"-adjusted telephone number weight was then calculated for these cases:[13]

(9) $W1(phone_{hpd} \in rd1) = SW(phone_{hpd}) \cdot A'_{nr}(phone_c)$, if eligibility of telephone number determined,

$W1(phone_{hpd} \in rd1) = 0$, otherwise.

[13]"Nonresponse" at this stage refers to an unresolved status of the telephone number.

TABLE V.2

VARIABLES IN NONRESPONSE PROPENSITY MODELS

Residency Determination ("Screener") Model	Interview Response Model
Intercept	Intercept
X1: Household is in nonmetropolitan area	X2: Household is in nonmetropolitan area
X2b: In supplemental sample, stratum 1[a]	X2b: In supplemental sample, stratum 1[a]
X3: In site 5 or site 12[b]	X4: In site 1,7,8,9,11[b]
X6: Site 4,6 stratum 3 /site 9, stratum 2[b]	X6: Site 3,5,10, stratum 2[b]
X8: Fewer than six Round One contacts	X7: One contact in Round One
X9: Household has two families	X8: Six or more Round One contacts
X12: One-person household	X10: Household has four or more families
X13: Household has more than five people	X12: Household has more than five people
X16: Household income $0-9,999	X13: Household has three or more children
X17: Household income $20,000-39,999	X21: Household has someone covered by military insurance
X20: Household has someone covered by Medicare	X23: Household has one or more months without telephone
X21: Householder is age 18-32	X24: Household has unpublished telephone number
X26: Somewhat satisfied with or neutral about health care	X2526: Householder is age 13-27 and not white
X27: Household has zero physician visits	X2529: Householder is age 78+ and not white
X29: Someone in household needs specialist	X26: Householder's age 13-27
X30: Everyone in household is in excellent health	X27: Householder's age 28-37
X33: Householder has two or more part-time jobs or is not working	X29: Householder's age 78+
X36: Householder is employed by state or local government	X31: Householder has 13 or more years of education
X43: Household has one or more months without telephone	X33: Householder is a proxy
	X35: Round One respondent's name not given
	X36: Health care satisfaction: missing, somewhat satisfied, very dissatisfied
	X37: Household has zero physician visits
	X38: Household has 11 or more physician visits
	X39: No one in household needs specialist
	X43: Householder has two or more full-time or part-time jobs or is not working
	X46: Householder is self-employed or has family business/farm
	X48: Householder's firm size is 1-999

[a]Variable used only for supplemental sample model.
[b]Variable used only for site sample model.

Other Samples. For the telephone numbers with undetermined residency that were not selected from Round One completes, we estimated the percentage with working residential telephone numbers from a study conducted for the National Immunization Survey (Shapiro et al. 1995) and discussed in Section C of Chapter IV. Based on this study, the status of these telephone numbers was determined as follows:

- ***Ring, No Answer.*** Of this group, 27 percent were estimated to be working residential numbers.
- ***Mechanical Answering Device.*** Of this group, 72 percent were estimated to be working residential numbers.
- ***Contact Made but Residency Not Determined.*** Of this group, 88 percent were estimated to be working residential numbers.

After this adjustment, telephone numbers with undetermined status and telephone numbers known to be ineligible were removed from the weighting process.[14] The remaining telephone numbers were those known to be working and associated with residences (that is, households).

c. Eligibility Nonresponse Adjustment to Household Weight

After adjusting for undetermined residency, we formed weighting cells to adjust for two kinds of household-level nonresponse: (1) not completing the survey screener to determine whether the household was eligible, and (2) eligible households not responding to the survey.[15] (The second adjustment is discussed in Section B.1.d.). An eligible household is a residence in

[14]After each weighting adjustment involving eligibility determination (discussed in Sections B.1.b, B.1.c, and B.1.f), we removed cases with undetermined eligibility status and cases known to be ineligible. After each adjustment involving nonresponse among known eligibles (discussed in Sections B.1.d, B.1.g, and B.1.i), we removed the nonrespondents from the remaining steps.

[15]See Section F.1 in Chapter II for the definition of eligible and ineligible households.

which there is at least one eligible adult. (See Section II.B.4 for a definition of individuals excluded from the survey.)

In defining the weighting cells, our goal was to group respondents who were similar with respect to the most important analytic variables, and to group those with similar likelihoods of having each type of nonresponse. The information used to form these cells had to be known for both nonrespondents and respondents. Although we could have defined the cells differently for each type of nonresponse adjustment, we decided to keep them the same. Based on generally accepted guidelines, we decided that each cell should contain at least 20 respondents, and that the adjustment factor in each cell should be less than two. Cells failing these criteria were combined with similar cells. The primary weighting cells were formed by crossing site, sampling strata, and Round One disposition; for the supplemental sample, the cells were defined by stratum and Round One disposition. The four Round One disposition categories were (1) Round One complete; (2) Round One noncomplete; (3) not sampled in Round One, but could have been; and (4) not sampled in Round One because its telephone bank was nonworking at that time (see Section B.1.a).

The first adjustments accounted for whether a residential household was eligible for the survey. We assumed that Round Two households with incomplete screening were eligible if the household associated with the telephone number had completed the Round One interview. We made this assumption because only 0.2 percent (24/12,523) of Round Two households that completed the screener failed to meet these criteria, and because they were eligible in Round One. For all other Round Two households with incomplete screening, we created the following household eligibility nonresponse adjustment factor:

$$(10)\ A''_{nr}(hhold_c) = \frac{\sum_{hh \in c} SW(phone_{hpd})}{\sum_{det\ hh \in c} SW(phone_{hpd})},$$

where the numerator is summed over all telephone numbers in cell *c* known to be working residential numbers (households) and over telephone numbers imputed to be households using the external empirical estimates described in the previous section, and the denominator is summed over households in cell *c* with a known survey eligibility status. A telephone number weight adjusted for determination of household eligibility was then calculated for these cases:

(11) $W1(phone_{hpd} \notin rd1) = SW(phone_{hpd}) \cdot A''_{nr}(phone_c)$, if eligibility of household determined

$W1(phone_{hpd} \notin rd1) = 0$, otherwise.

d. Interview Nonresponse Adjustment to Household Weight

We then adjusted these weights for survey nonresponse among eligible households. A responding household was one in which at least one eligible FIU responded to the survey. As was the case for the telephone number eligibility determination adjustment (Section B.1.a), a different method was used to make this adjustment for households linked to Round One completes and for all other telephone numbers sampled in Round Two. We used a stepwise logistic regression model[16] using Round One data to predict the likelihood of Round Two eligible household nonresponse. Table V.2 shows the variables found to be significant in the "response model" for the site and supplemental samples. The weighting adjustment factor, $A'_{nr}(hhold_c)$, is calculated based on the predicted probability from this model, $\hat{Y}_{response}$, as follows:

[16]Significance for entry = .15; significance for exit = .10.

(12) $A'_{nr}(hhold_c) = \dfrac{1 + exp(\hat{Y}_{response})}{exp(\hat{Y}_{response})}$.

We capped this adjustment factor at 2.5. We set the factor equal to zero for eligible but nonresponding Round Two households sampled from telephone numbers linked to a Round One complete. A nonresponse-adjusted household weight was then calculated for these cases:

(13) $W2(phone_{hpd} \in rd1) = W1(phone_{hpd}) \cdot A'_{nr}(hhold_c)$, if household responded

$W2(phone_{hpd} \in rd1) = 0$, otherwise.

We performed a weighting class adjustment for households that were not linked to Round One completes, using the same cells as defined for the household eligibility adjustment. We created a household survey nonresponse adjustment factor as follows:

(14) $A''_{nr}(survey_c) = \dfrac{\sum_{elig\ hh \in c} W1(phone_{hpd})}{\sum_{resp\ hh \in c} W1(phone_{hpd})}$,

where the numerator is summed over all eligible households in cell *c*, and the denominator is summed over responding eligible households in cell *c*. The following household weight adjusted for survey nonresponse was then calculated for these cases:

(15) $W2(phone_{hpd} \notin rd1) = W1(phone_{hpd}) \cdot A''_{nr}(survey_c)$, if eligibility of household determined

$W2(phone_{hpd} \notin rd1) = 0$, otherwise.

e. Poststratification and Other Adjustments to Household Weight

We then adjusted for multiple telephones in the household and for interruptions in telephone service.[17] Because some households have multiple nonbusiness telephone numbers,[18] a household multiplicity factor was used to adjust for the number of telephone numbers in the household. This factor, which is simply the inverse of the total number of these telephones in the household, was applied to the nonresponse-adjusted household weight:

(16) $W3(hhold_{hpd}) = W2(phone_{hpd})/(number\ of\ phones)$.

One of the last steps in creating the household-level weight was to poststratify the sum of the weights to estimated population totals. We used 1998 estimates from Marketing Systems Group (2000) (Genesys) of the number of households in each site and nationally (by whether or not in an MSA).[19,20] For site-specific estimates, we used 1990 Census data to estimate the proportion of telephone households in each site and stratum and adjusted the 1998 estimates of

[17]Question h30 in the Household Survey asked whether the household had any additional telephone numbers and, if so, how many; in the case of one or more numbers, question h31 asked whether the additional number(s) was(were) for home or business use. If h30 = 1, 2, 3, or 4 and h31 = 1 or 2 (home use or both), we then set the number of telephones equal to h30 plus one. For any remaining cases (h30 = 9 or h31 ≥2), we set the number equal to one.

[18]By "nonbusiness telephone number," we mean a telephone number from which the household received nonbusiness calls. Dual-use numbers would fall into this category.

[19]Marketing Systems Group uses intercensal estimates developed by Claritas.

[20]Marketing Systems Group defines a household according to the Census definition, which "... includes all the persons who occupy a housing unit," and defines a housing unit according to the Census definition as "a house, apartment, a mobile home, a group of rooms, or a single room that is occupied (or if vacant, is intended for occupancy) as separate living quarters." This definition differs slightly from our definition of an eligible household in that we excluded households containing only unmarried students younger than age 23 or people in the military. The students were eligible for the CTS Household Survey, but only through their parent's households.

the total number of households by these proportions to develop an estimate of telephone households in 1998.[21] For national estimates of telephone and nontelephone households (by metropolitan status), we used 1998 estimates from Marketing Systems Group (2000). The poststratification adjustment factor for telephone households is:

$$(17) \qquad A_{ps\text{-}tel}(stratum\ h) = \frac{TELHH_h}{\sum_{resp\ hh_i \in h} W3(hhold_{hpd})},$$

where $TELHH_h$ is the estimated number of telephone households in 1998. The household-level weight poststratified to telephone households is:

$$(18) \qquad WT_{tel}(hh_{hi}) = W3(hhold_{hpd}) \cdot A_{ps\text{-}tel}(stratum\ h).$$

For the supplemental sample and low-intensity site-specific weights, we used information on telephone service interruption to inflate the RDD sample weights for telephone households in order to account for nontelephone households. Even though all cases in the RDD telephone sample had working telephones when interviewed, they were asked whether they had had any interruption in telephone service during the year preceding the interview.[22] We used cases with interruptions in telephone service to represent nontelephone households and those with no reported interruptions to represent telephone households. In doing so, we adjusted weights to the

[21]This method assumes a steady proportion of nontelephone households since 1990. This assumption is consistent with Current Population Survey estimates of the proportion of households without telephones for the years 1990 through 1997 (U.S. Bureau of the Census 1994 and 1997).

[22]To determine telephone status, we used the responses to question h32 ("During the past 12 months, was there any time when you did not have a working telephone in your household for two weeks or more?") and question h33 ("For how many...months...?").

number of months of interrupted service. (An analogous procedure was used in creating the integrated weights discussed in Section D.) The interruption-adjusted weight is:

$$(19) \quad WT_{interruption}(hhold_{hi}) = \frac{W3(hhold_{hpd})}{proportion\ of\ year\ with\ phone}.$$

The poststratification adjustment factor for total households is:

$$(20) \quad A_{ps\text{-}all}(phone\ status\ g, stratum\ h) = \frac{TOTHH_{gh}}{\sum_{resp\ hhold_i\ \hat{I}\ h\ with\ phone\ status\ g} WT_{interruption}(hhold_{hi})},$$

where *phone status g* is equal to one (interruption in telephone service) or is equal to two (no known interruption in telephone service), $TOTHH_{1h} = TOTHH_h - TELHH_h$, and $TOTHH_{2h} = TELHH_h$.

A household-level weight poststratified to all households is:

$$(21) \quad WT_{all}(hhold_{ghi}) = WT_{interruption}(hhold_{hi}) \cdot A_{ps\text{-}all}(phone\ status\ g, stratum\ h).$$

f. Nonresponse Weight Adjustment for FIUs with Undetermined Eligibility

The probability of selection of each FIU was equal to the probability of selection for its household (that is, all FIUs in a selected household were selected for the interview). We therefore used the final household weight as the starting point for developing the FIU weight. Only households with at least one eligible responding FIU had a positive household weight. The FIU weights accounted for two types of FIU nonresponse within these households: (1) failure to determine whether the FIU was eligible for the survey, and (2) if determined to be eligible, the failure of the FIU to respond to the survey.

We started with an FIU-level file containing all FIUs enumerated within responding households and assigned to each FIU its final household weight. Using the same cells as defined for the telephone- and household-level adjustments, we calculated the following adjustment factor to account for FIUs with undetermined eligibility:

$$(22) \qquad A_{nr}(FIU_c) = \frac{\sum_{fiu \in c} W3(hhold_{hpd})}{\sum_{det\ fiu \in c} W3(hhold_{hpd})},$$

where the numerator is summed over all FIUs in cell *c* that are members of responding households, and the denominator is summed over all FIUs in cell *c* with a known survey eligibility status. An FIU weight adjusted for determination of eligibility was then calculated for these cases:

$$(23) \qquad W4(FIU_{hpd}) = W3(hhold_{hpd}) \cdot A_{nr}(FIU_c)\text{, if eligibility of FIU determined}$$
$$W4(FIU_{hpd}) = 0\text{, otherwise.}$$

g. Interview Nonresponse Adjustment to FIU Weight

We then adjusted these preliminary weights for FIU nonresponse among FIUs known to be eligible for the survey. For eligible but nonresponding FIUs, we performed another weighting class adjustment, again using the same cells as defined above. We created an FIU survey nonresponse adjustment factor as follows:

$$(24) \qquad A_{nr}(survey_c) = \frac{\sum_{elig\ fiu \in c} W4(FIU_{hpd})}{\sum_{resp\ fiu \in c} W4(FIU_{hpd})},$$

where the numerator is summed over all eligible FIUs in cell *c*, and the denominator is summed over responding eligible FIUs in cell *c*. An FIU weight adjusted for survey nonresponse was then calculated for these cases:

(25) $W5(FIU_{hpd}) = W4(FIU_{hpd}) \cdot A_{nr}(survey_c)$, if FIU responded

$W5(FIU_{hpd}) = 0$, otherwise.

h. Initial Person Weight

The probability of selection for each adult member of an eligible responding FIU was equal to the probability of selection of the FIU (that is, all adults in each responding FIU were selected for the interview). We therefore used the final FIU weight as the starting point for developing the person weight for adults. However, because only one child was selected at random per FIU, the within-FIU probability of selection for a child was equal to the inverse of the number of children in the FIU. The overall probability of selection for person *k* in FIU *j* in household *i* in stratum *h* can be expressed as:

(26) $$P(person_{hijk}) = \frac{P(FIU_{hij})}{(\delta \cdot numkids_{hij}) + (1-\delta)},$$

where $numkids_{hij}$ is the number of children in FIU_{hij}, and δ is equal to zero for adults and is equal to one for children. So, the initial person-level weight for all people was calculated as follows:

(27) $W6(person_{hpd}) = W5(FIU_{hpd}) \cdot ((\delta \cdot numkids_{hij}) + (1-\delta))$.

All eligible individuals in all responding FIUs were assigned this weight regardless of whether we had complete data on an individual.

i. Nonresponse Adjustment to Person Weight

The next adjustment to the person weight accounted for high levels of missing data among individuals selected for the survey. An editing program was used to determine whether a person record contained too many missing items to be usable. The editing rule was that all person records with 75 percent or more missing data for variables from Sections B through G of the questionnaire were considered to be non-respondents. Only 16 person records were deleted because of high levels of missing information. This step in the weighting process adjusted for this small amount of unit nonresponse at the person level, using the same weighting cells as defined for previous adjustments. We created a person-level survey nonresponse adjustment factor as follows:

$$(28) \qquad A_{nr}(missing_c) = \frac{\sum_{elig\ person \in c} W6(person_{hpd})}{\sum_{resp\ person \in c} W6(person_{hpd})},$$

where the numerator is summed over all eligible and selected individuals in cell *c*, and the denominator is summed over individuals with more complete responses. A person weight adjusted for survey nonresponse was then calculated for these cases:

$$(29) \qquad W7(person_{hpd}) = W6(person_{hpd}) \cdot A_{nr}(missing_c), \text{ if person responded}$$
$$W7(person_{hpd}) = 0, \text{otherwise.}$$

2. Using the Site Sample to Make National Estimates

The probability of selection of telephone number can be described as follows:

$$(30) \qquad P_{hd} = (P1_h \cdot P2_{hd}) + (P3_h \cdot P4_{hd}),$$

where $P1_h$ is the probability of selection in Round One; $P2_{hd}$ is the probability of selection in Round Two given that the telephone number was selected for release in Round One; $P3_h$ is the probability of *not* being selected for release in Round One ($1 - P1_h$); and $P4_{hd}$ is the probability of selection of new cases in Round Two. These probabilities have different meanings depending on whether a telephone number was selected in Round One, was not selected in Round One but could have been, or could not have been selected in Round One.

The *formulas* for the selection probabilities and weights at the household, FIU, and person levels (see equations [1] through [29]) and the formulas and methodologies for the nonresponse and poststratification adjustments are similar across four types of estimates: (1) site-specific estimates using the augmented site sample, (2) national estimates using the site sample, (3) national estimates using the supplemental sample, and (4) national estimates using the augmented site sample. However, the *values* of these weights and adjustment factors differ across the four types of RDD weights because the telephone selection probabilities differ. Furthermore, weights to be used for making national estimates using the site sample must also account for the probability of selection of the site, as well as for the distribution of cases in the high-intensity and low-intensity sites. (The selection of the 60 sites is discussed in detail in Metcalf et al. [1996]).

The sample size of telephone numbers in the RDD sample was k times larger in the high-intensity sites than in the low-intensity sites. When calculating the weights, we set k equal to four, which was in accordance with the relative sample sizes of high-intensity and low-intensity sites in the original design. To account for the probability of selection of any telephone number, when making national estimates, we had to use the expected number of selected telephone numbers in each site, $E(n_{sh})$, rather than the actual number of selected telephone numbers, n_{sh}.

For site *s* in stratum *h*, where the site is an MSA with 200,000 or more people, the expected number of selected telephone numbers is:

$$(31) \quad E(n_{sh}) = [n_{lo} \cdot k \cdot P(high\text{-}intensity)] + [n_{lo} \cdot P(low\text{-}intensity)]$$
$$= [n_{lo} \cdot k \cdot 12/48] + [n_{lo} \cdot 36/48]$$
$$= n_{lo} \cdot (k/4 + 3/4),$$

where n_{lo} is the number selected for a low-intensity site. For sites in small MSAs and for non-MSA sites, $E(n_{sh}) = n_{lo}$ because these sites had no chance of being selected as high-intensity sites.

The probability of selection of the telephone number in Round One for this sample and estimate type can then be defined as:

$$(32) \quad P1N(telephone_{ihs}) = PSUPROB_s \cdot \frac{E(n_{sh})}{n_{sh}} \cdot P1_h,$$

and the probability of selection of the telephone number in Round Two can be defined as:

$$(33) \quad P4N(telephone_{ihs}) = PSUPROB_s \cdot \frac{E(n_{sh})}{n_{sh}} \cdot P4_h,$$

where $PSUPROB_s$ is the probability of selection of site *s*[23] and n_{sh} is the actual number of telephone numbers selected in the site sample in stratum *h* in site *s* (set equal to n_{lo} for low-intensity sites and equal to $4 \times n_{lo}$ for high-intensity sites, for the actual calculation).

The cumulative probability of selection of telephone numbers for national estimates based on the site sample can then be computed as:

[23]See Metcalf et al. (1996) for a detailed discussion of this probability.

(34) $$PN(telephone_{ihs}) = P1N_{ihs} \cdot P2_{hd} + ((1 - P1N_{ihs}) \cdot P4N_{ihs}) .$$

Formulas representing subsequent stages of selection, nonresponse adjustments, and poststratification used this initial selection probability as their base.

3. Using the Supplemental Sample to Make National Estimates

When using only the supplemental sample to make national estimates, the probability of selection of the telephone number can be defined as:

(35) $$PN(telephone_{ih}) = P1_h \times P2_{hd} + ((1 - P1_h) \times P4_{hd}) .$$

4. Using the Augmented Site Sample to Make National Estimates

When combining the site sample and the part of the supplemental sample that falls within the site boundary to make *national* estimates, we had to modify the value of the expected sample size in each site as follows:

(36) $$E'(n_{sh}) = (n_{lo} \times (k/4 + 3/4)) + ((n_{lo} \times z) \times usprop),$$

where z is the size of the supplemental sample, as a multiple of the size of a low-intensity site,[24] and *usprop* is the proportion of the U.S. population (excluding Alaska and Hawaii) that fell within each site in July 1992. When calculating the weights, we set k equal to 4 and z equal to 10.

The actual number of telephone numbers selected in the augmented site sample in stratum h in site s is calculated as:

[24] This multiple is equal to 10, which is in accordance with the relative sample sizes of the supplemental sample and a low-intensity site in the sample allocation design (see Chapter II).

(37) $n_{sh} = (n_{lo}) + ((n_{lo} \cdot z) \cdot usprop)$ for low-intensity sites,

and

(38) $n_{sh} = (n_{lo} \cdot k) + ((n_{lo} \cdot z) \cdot usprop)$ for high-intensity sites.

The probability of selection of the telephone number in Round One for the site sample (national estimates) can then be defined as:

(39) $$P1N_{site}(telephone_{ihs}) = PSUPROB_s \cdot \frac{E'(n_{sh})}{n_{sh}} \cdot P1_{h-site},$$

and the probability of selection of the telephone number in Round Two can be defined as:

(40) $$P4N_{site}(telephone_{ihs}) = PSUPROB_s \cdot \frac{E'(n_{sh})}{n_{sh}} \cdot P4_{h-site},$$

where the second term in the equations is set equal to one for non-MSA sites. The cumulative probability of selection of telephone numbers for national estimates based on the site sample can then be computed as:

(41) $$PN_{site}(telephone_{ihs}) = P1N_{site} \cdot P2_{hds} + ((1 - P1N_{site}) \cdot P4N_{site}).$$

We then combined the two probabilities of selection for each number, one under the site sample scenario and one under the supplemental sample scenario ($P_{hd-site}$ and $P_{hd-supp}$, respectively), as follows:

(42) $$PN(telephone_{ihs}) = PN_{site} + ((1 - PN_{site}) \cdot PN_{supp}).$$

5. Using the Augmented Site Sample to Make Site-Specific Estimates

When combining the site sample and the supplemental sample to make site-specific estimates, we accounted for the dual probabilities of selection (that is, each telephone number could have entered the sample through the site sample or through the supplemental sample). First, we assigned alternative sites and strata for each telephone number in the augmented sample.[25] We then computed two probabilities of selection for each number, one under the site sample scenario and one under the supplemental sample scenario ($P_{hd-site}$ and $P_{hd-supp}$, respectively). The probability of selection of the telephone number could then be specified as:

$$(43) \qquad P(telephone_{ihs}) = P_{hd} = P_{hd-site} + ((1 - P_{hd-site}) \cdot P_{hd-supp}) .$$

C. WEIGHTS FOR THE FIELD SAMPLE

This section describes the procedures used to construct final design-based weights for the survey's field component, which was designed to include households that had little or no chance of being selected for the RDD surveys. The field survey was not designed for independent use because of its limited coverage and small sample size. However, when combined with the site-based RDD survey, the field sample improves population coverage among subgroups less likely to be included in RDD-only surveys.

We produced two sets of weights for the field survey data. Although neither set is intended to be used alone in policy analysis, these weights and the weights representing the RDD sample were used to create integrated weights for making inferences about the entire population. Field

[25]We assigned to each telephone number the "alternative site" and "alternative stratum" it would have had if the case had been selected in the other sample. So, for the site sample telephone numbers, we determined the strata into which each would have fallen if it had it been selected through the supplemental sample. For the supplemental sample telephone numbers, we determined the site (and substrata, if applicable) into which each would have fallen if it had been selected through the site sample.

sample weights for households, FIUs, and individuals were constructed for (1) individual sites in which the field survey was conducted, and (2) all MSAs with 1992 populations of 200,000 or more. We refer to the second set of weights as *national* weights. Each weight was the product of several factors that reflected differences in probabilities of selection and nonresponse. The weights also included poststratification adjustments so that the sample matched external estimates of the relevant population.

1. Steps in the Weighting Process

The first weighting factor for a unit (housing unit [HU], household, FIU, or individual) for any of the weights was the inverse of that unit's probability of selection.[26] This factor differed between weights used for site-specific estimates and weights used for national estimates.

The weights have two other components:

1. A nonresponse adjustment for FIUs or individuals within households for which no data were collected
2. Ratio adjustment(s) to estimated population totals (poststratification weights)

a. Initial Weights

The initial weight was the inverse of the overall probability of selection of a unit. Weights were computed for HUs, households, FIUs and individuals. For a listed housing unit LHU_{abci} in listing area LA_c in secondary sampling unit SSU_b and primary sampling unit PSU_a, the preliminary supplemental sample weight, *SWN*, is:

(44) $SWN(LHU)_{abci} = 1/P(LHU_{abci})$,

where:

[26]See Section II.F.2 for the definition of a housing unit.

(45) $$P(LHU_{ABCI}) = P(PSU_a) \cdot P(SSU_b \mid PSU_a) \cdot P(LA_c \mid SSU_b) \cdot P(HU \mid LA_c),$$

and the primary sampling units are the 12 high-intensity sites, secondary sampling units are areas within the sites selected with probability proportional to size within the sites, and listing areas were selected with equal probability within SSUs. The term $P(HU|LA_c)$ accounts for the fact that only a subsample of listed housing units was selected for interviewing in some listing areas. Note that, for a household *(hhold)*, the initial weight was the same as for a listed housing unit. Thus, for national estimates:

(46) $$P(hhold_{abci}) = P(LHU_{abci}),$$

(47) $$SWN(hhold)_{abci} = SWN(LHU)_{abci}.$$

For site-specific estimates, the same formula can be modified by omitting the term for the high-intensity-site selection probability. Thus, for site-level estimates:

(48) $$SWS(hhold)_{bci} = SWS(LHU)_{bci} = 1/P(LHU_{bci}),$$

(49) $$P(LHU_{abci}) = P(SSU_b|PSU_a) \cdot P(LA_c|SSU_b) \cdot P(HU|LA_c)$$

Probabilities of selection of FIUs and adults in FIUs are the same as for the household. Children were subsampled, so for the *k*th child in the *j*th FIU in household *i*, where the number of children in the FIU is $numkids_{abcij}$:

(50) $$P(Child_k|FIU_j|hhold_{abci}) = P(hhold_{abci}) \cdot 1|(numkids_{abcij}).$$

b. Nonresponse-Adjusted Weights

The first step in calculating nonresponse weights was to define weighting cells. Because the sample sizes were too small to justify creating cells smaller than a site, we decided that weighting cells for both national and site-based estimates should be the sites themselves.

After all listed housing units that were sampled for screening were assigned their initial probability weights, a series of adjustments were made. The first adjustment compensated for nonresponse to the screening interview among listed housing units (that is, for unknown eligibility). For simplicity, we will refer to a general set of weights *SW1* to denote the adjustment procedure for the national weight *(SWN)* and site-based weight *(SWS)*. For cell *c*, we define a nonresponse-adjustment factor $A_{nr}(HU_c)$ and the weight $W_1(hhold_{abci})$:

$$(51) \quad A_{nr}(HU_c) = \frac{\sum_{attempt\ \varepsilon\ k} SW1_{abci}}{\sum_{determ\ \varepsilon\ k} SW1_{abci}},$$

$$(52) \quad W_1(hhold_{abci}) = SW1_{abci} \cdot A_{nr}(HU_c), \text{ if eligibility of household was determined}$$
$$= 0, \text{ otherwise.}$$

As discussed in Chapter II.F, eligibility was imputed for some households in inaccessible buildings. Cases with imputed eligibility and cases with known eligibility were treated in the same way.

The next adjustment was for household nonresponse. Initially, we used the same weighting cells as for the previous adjustment:

$$(53) \quad A_{nr}(HH_c) = \frac{\sum_{elig\ \varepsilon\ c} W_1(hhold_{abci})}{\sum_{resp\ hh\ \varepsilon\ c} W_1(hhold_{abci})},$$

where the summation in the numerator is over all households found to be eligible in weighting cell *c*, and the denominator is summed over all responding households in weighting cell *c*.

Finally, households with completed interviews were assigned weights:

(54) $W_2\,(hhold_{abci}) = W_1(hhold_{abci}) \cdot A_{nr}(hhold_c)$, if the household responded
$= W_1(hhold_{abci})$, if the household or listed housing unit was ineligible
$= 0$, otherwise.

Because there was no nonresponse at the FIU level and only a few nonresponses due to missing data at the person level, these nonresponse adjustments were kept as simple as possible and were the same as those described in the sections on weighting the RDD sample data (Sections B.1.f, B.1.g, and B.1.i). The weighting adjustment was the ratio of the sum of weights for potential units (FIUs, adults, or children) for which data should have been obtained to the sum of weights for units for which data were obtained.

c. Poststratification

Poststratification weights were calculated in two stages. In the first stage, all households (whether eligible or not) were weighted up to the 1990 Census count of households for areas included in the frame. This weighting adjusted for factors unrelated to the intentional undercoverage introduced by the design. For a site *a*, where $g = 1$ for the included areas and $g = 0$ for excluded areas:

(55) $$A_{ps}1(a) = \frac{Censuscount_{ga} \cdot g}{\sum_{i=1}^{n_{bc}} W2(hhold_{abci})},$$

(56) $PSW1 = W_2(hhold_{abci}) \cdot A_{ps}1(a)$.

The second stage was a ratio adjustment of interviewed households to 1998 estimates of all nontelephone households (including areas excluded from the sampling frame), nationally and for each site[27]:

$$(57) \quad A_{ps}2(site) = \frac{census\ np\ (site)}{\sum_{i \epsilon resp}^{nsite} PSW1_i},$$

$$(58) \quad A_{ps}2(not) = \frac{census\ np\ (nat)}{\sum_{s=1}^{12} \sum_{i \epsilon resp}^{nsite} PSW1_{is}},$$

$$(59) \quad PSW2S_i = PSW1_i \cdot A_{ps2}(site),$$

$$(60) \quad PSW2N_i = PSW1_i \cdot A_{ps2}(nat).$$

A similar adjustment was made for individuals.

D. INTEGRATED WEIGHTS FOR THE HOUSEHOLD SURVEY

The integrated weights combined the field and RDD survey data from the site-based sample for use in making national and site-specific estimates. For areas represented by both the RDD and field components, the integrated weights accounted for the likelihood of being chosen in each of the two components. For areas not represented by the field component, the RDD survey data alone were weighted up to represent all households and individuals in those households,

[27]The 1998 estimates were synthesized from the 1990 Census proportion of nontelephone households and the July 1998 estimate of total households by Marketing Systems Group (2000).

including those without telephones. We used the following seven-step process to construct two sets of integrated weights (one for national estimates and one for site-specific estimates)[28]:

1. Postratify the RDD and field telephone components to our best estimates of the telephone and nontelephone populations, respectively
2. Create household telephone service interruption adjustment factors (IAFs) for both components (see Section D.1)
3. Apply IAFs to the weights for the separate household components
4. Apply the same IAFs to the FIU components
5. Apply the same IAFs to the person-level components
6. Join the RDD and field telephone components
7. Poststratify the joined RDD and field components again

For national estimates, the field component represented nontelephone households in large MSAs only. For households in small MSA or nonmetropolitan strata, the "integrated" weights were simply the weights that represented all households in the strata *(WT_{all})*, where those with any telephone service interruption had their weights inflated to account for the proportion of the year preceding the survey without service, and then poststratified to the estimated number of nontelephone households (by metropolitan status). Households in the strata with no interruption had their weights poststratified to the estimated number of telephone households.[29]

[28]For both national estimates and site-specific estimates, we included households from the site sample and households from the supplemental sample that were part of the augmented site sample.

[29]For national estimates based on the supplemental sample only, the "integrated" weights for all households were simply the weights that represented all households *(WT_{all})*, where those with any telephone interruption had their weights inflated to account for the proportion of the past year without service, then poststratified to the estimated number of nontelephone households (by metropolitan status); those with no interruption had their weights poststratified to the estimated number of telephone households.

For RDD households in the 48 large MSAs, we began with the weights that represented the telephone portion of the population *(WT_{tel})*. For the field households, we began with the weight that represented the nontelephone portion of the population. Both large MSA households in the RDD component that had intermittent telephone service and households in the field component that had any telephone service during the year preceding the survey were adjusted for dual selection probabilities (they had a chance of being selected into both the RDD and field components), while accounting for the length of interruption. (This adjustment is described in more detail below.) Table V.3 illustrates how the RDD and field components were combined for national estimates. Note that these steps were carried out for the national weights based on the site sample only and for the national weights based on the augmented site sample.

For site-specific estimates, the field component represented nontelephone households in the 12 high-intensity sites only. For households in the low-intensity sites, the "integrated" weights were simply the weights that represented all households in those strata *(WT_{all})*, where households with any telephone service interruption had their weights inflated to account for the proportion of the year preceding the survey without service, then poststratified to the estimated number of nontelephone households (by site); those with no interruption had their weights poststratified to the estimated number of telephone households in the site.

For RDD households in the 12 high-intensity sites, we began with the site-specific weights that represented the telephone portion of the population *(WT_{tel})*. For the field households (all of which were in the 12 high-intensity sites), we began with the site-specific weight that represented the nontelephone portion of the population. High-intensity-site households in the RDD component that had intermittent telephone service and households in the field component that had some telephone service during the year preceding the survey were adjusted for dual

TABLE V.3

INTEGRATION OF RDD AND FIELD COMPONENTS
FOR NATIONAL ESTIMATES BASED ON THE SITE OR AUGMENTED SITE SAMPLE

	RDD Component	Field Component
High-intensity sites	Represents households in large MSAs in contiguous United States with continuous or intermittent telephone service	Represents households in large MSAs in contiguous United States with intermittent or no telephone service
Other large-MSA sites Small-MSA sites Non-MSA sites	Represents all households in balance of contiguous United States.	

selection probabilities, while accounting for the length of interruption. Table V.4 illustrates how the RDD and field components were combined for site-specific estimates.

1. Telephone Service Interruption Adjustment Factor

A factor complicating the combination of the RDD and field samples was the inclusion in both components of households with interrupted telephone service during the year preceding the survey. The integrated weights assumed that (1) households with no interruption in service could have been sampled only for the telephone survey, (2) those with no telephone service could have been sampled only for the field survey, and (3) the remainder could have been sampled for both surveys. For the RDD site sample, 2.9 percent of households completing interviews had an interruption in telephone service of two weeks or longer during the year preceding the survey, but fewer than half the households were in areas eligible for the field component. For the field sample, 60.3 percent of households completing interviews (n = 335) had at least one month during the presurvey year in which they had telephone service and could have been sampled for the RDD survey.

Approximating probabilities of selection that accounted for multiplicity between the field and RDD sample frames was complicated by incomplete information on the addresses of some RDD households. To approximate these probabilities, we would have had to have had this information in order to link the households to the Census block groups in which they resided. In addition, the data available to match RDD households to block groups were based on the 1990 Census and therefore could not have accounted for housing construction since then. Finally, the level of effort to complete such a match would have been substantial, and we concluded it was not cost-effective, given the size of the samples eligible for inclusion in both surveys and the accuracy of the multiplicity estimates.

TABLE V.4

INTEGRATION OF RDD AND FIELD COMPONENTS FOR SITE-SPECIFIC ESTIMATES BASED ON THE SITE OR AUGMENTED SITE SAMPLE

	RDD Component	Field Component
High-intensity sites	Represents households in sites with continuous or intermittent telephone service	Represents households in site with intermittent or no telephone service
Other large-MSA sites	Represents all households in site	—
Small-MSA sites	Represents all households in site	—
Non-MSA sites	Represents all households in site	—

Instead, we constructed integrated weights that synthetically accounted for multiplicity by using a weighting adjustment that we termed the *telephone interruption adjustment factor* (the IAF). This factor accounted for both length of telephone interruption and multiplicity and was applied only to households in the "integration sites" (that is, sites represented by both the RDD and field components). For national estimates, integration sites included all large-MSA sites. For site-specific estimates, they included the 12 high-intensity sites only. For the RDD component, households with no telephone interruption would have been ineligible for the field component and so had an IAF set equal to one. For the field component, households with no telephone service would have had no chance of selection into the RDD component and also had an IAF equal to one. For households in the field component with some telephone availability and for households in the RDD component with some telephone interruption, we multiplied the value of IAF, as described below, by the households' postratified weights; the weights were postratified to the populations represented by their components (telephone or nontelephone). We calculated IAF_m as:

$$(61)\quad IAF_m = \frac{1/RelP_m}{1/MEDIAN(RelP)} . k \qquad m = (1,2,...,12),$$

where:

$$(62)\quad RelP_m = [PRatio . \frac{(12 - m)}{12}] + 1,$$

and

$$(63)\quad PRatio = \frac{hhold\ \ in\ \ RDD\ \ sample\ \ /\ \ telephone\ \ hhold\ \ in\ \ population}{hhold\ \ in\ \ field\ \ sample\ \ /\ \ nontelephone\ \ hhold\ \ in\ \ population},$$

and where m is the number of months without telephone service; k is a constant used to inflate or deflate the adjustment so that the sum of the weights across the two components for households with an interruption in telephone service remained the same; $RelP_m$ is the relative combined likelihood of selection into either component, estimated on the basis of the number of months with telephone service[30]; and *PRatio* is the probability of selection into the RDD component, relative to selection into the field component.

The IAF was then applied to the appropriate weight, depending on the sample component and length of telephone interruption, as follows:

(64) $WTINT_m = WT_{tel} \cdot IAF_m$, for RDD households in integration sites

$WTINT_m = PSW2 \cdot IAF_m$, for field households

$WTINT = WT_{all}$, for RDD households outside of integration sites,

where m is the number of months without telephone service. For RDD households with $m = 0$ and for field households with $m = 12$, $IAF_m = 1$.

2. Poststratification of Person-Level Integrated Weights

For national estimates, person-level weights were poststratified by sex and age group, then by sex and whether or not Hispanic, then by sex and race (black or nonblack), and then by level of education.[31] For high-intensity sites, site-specific weights were poststratified by age group,

[30]In equation (62), the first term (in square brackets) represents the likelihood of selection into the RDD component, and the second term (the number 1) reflects the likelihood of selection into the field component.

[31]Based on the 1998 Current Population Survey (excluding Alaska and Hawaii) (U.S. Bureau of the Census 1998).

race (whether or not Hispanic or black), and the estimated site population.[32] Weights for low-intensity site-specific estimates were poststratified to site totals only. After person-level weights were trimmed, weights were poststratified again by the same variables, as well as by pretrimming telephone/nontelephone distribution. (See the next section for discussion of the trimming of the person-level weights.) The re-postratification was done within site for site-specific weights.

E. TRIMMING PERSON WEIGHTS

In analyses of survey data, even a few extremely large weights can reduce the accuracy of point estimates and can inflate values of the sampling variance. To reduce the sampling variance, excessively large weights are trimmed, and the amount trimmed is distributed among the untrimmed weights to preserve the original sum of the weights. However, trimming of sampling weights can introduce bias into some point estimates, because the observation made with the trimmed weight is not accurately represented in the point estimate. The objective in weight trimming is to incorporate a reduction in the excessively large weights while minimizing the introduction of bias.

For site-specific and national estimates, we trimmed the person-level and family-level integrated weights and then assessed the effect of the trimming. We evaluated the extent of trimming and the inflation factor for the untrimmed weights necessary to preserve the original sum of the weights and then estimated the effect of the trimming on the sampling variance. We used a weight-trimming algorithm that compares each weight with the square root of the average value of the squared weight used to identify weights to be trimmed and the trimming value. This

[32]Age, race, ethnicity, and total population, by site, were based on Marketing Systems Group's estimates from 1998.

algorithm has been referred to as the "NAEP procedure" (Potter 1990). The trimmed excess was distributed among the weights that were not trimmed.

The statistical measure of the impact of the trimming was based on the design effect attributable to the variation among the sampling weights. Unequal weighting (a result of differential selection rates and response rates) has the potential to decrease precision because variation in the weights affects the variance of weighted estimates. Person-level weights were trimmed to reduce this design effect; however, the extent of trimming was limited to minimize the risk of introducing bias into the sample estimates.

Specifically, let WT_i denote a set of weights and let n denote the number of people. We first established trimming classes on the basis of characteristics of the sample (the site, or stratum in the supplemental sample) and the characteristics of the sample member (that is, adult or child). The weight-trimming algorithm establishes a cut-off point, T_c, in a trimming class, c, as:

$$(65)\quad T_c = (k \sum_{i \varepsilon c} WT_i^2 \;/\; n_c)^{1/2},$$

where n_c is the number of observations in the trimming class, k is an arbitrary number (generally assigned a value of 10), and the summation is over the observations in the trimming class. Any weight exceeding the cut-off point, T_c, is assigned the value of T_c, and excess is distributed among the untrimmed weights, thereby ensuring that the sum of the weights after trimming is the same as the sum of the weights before trimming.

Using these newly computed weights, the cut-off point was recomputed and each weight again compared with the cut-off point. If any weight exceeded the new cut-off point, the observation was assigned the value of the new cut-off point, and the other weights were inflated to compensate for the trimming.

The cut-off point generated by the algorithm was generally used as the value of the trimmed weight. In some trimming cells, the algorithm indicated a trimming level that was excessive, so a value larger than the computed cut-off point was used. Generally, we used a larger value when the adjustment seemed excessive for the weights that were less than the cut-off point or when a trimming class contained only a few observations.

The weights designed to produce site-specific estimates were evaluated separately for adults and children in each high-intensity site. Because only one child was randomly selected in each FIU, and the sample size of children was smaller than that of adults, weights for children had greater variation and were larger on average than were weights for adults. The weights for trimming were identified by using the NAEP procedure, as well as by visual inspection of outlier weights that the NAEP procedure might have missed. The assessment of the impact of trimming was evaluated by inspecting the trimming level, the magnitude of the adjustment to the untrimmed weights, and the anticipated design effect from unequal weights. The weights were trimmed for both site-sample and augmented site-sample estimates. They were trimmed for fewer than 0.1 percent of the adult and children observations (49 out of 55,417 individuals).

We used a similar method to trim the weights designed to produce national estimates by using the NAEP procedure and assessing the impact of the trimming on the design effect from unequal weights. For the site sample, the weight-trimming classes were defined by the three site-selection strata (large MSAs, small MSAs, and non-MSAs), geographic region (four regions), and adult versus child. For the supplemental sample, the weight-trimming classes were defined by the sample strata (metropolitan areas in each of four geographic regions and the nonmetropolitan areas of the United States) and adult versus child. Relatively few weights were trimmed—fewer than 120 of the more than 53,000 weights in the site sample (0.2 percent) and fewer than 30 of the 5,600 weights in the supplemental sample (0.4 percent).

For FIU-level weights for site-specific estimates, 15 of more than 30,000 FIU weights were trimmed (0.05 percent). For national estimates, 6 of the 3,000 weights in the national supplement sample (0.05 percent) and 51 of the nearly 29,000 weights in the site sample (0.2 percent) were trimmed.

F. WEIGHTS FOR COMBINING THE SITE AND SUPPLEMENTAL SAMPLE SURVEYS

The goal of the supplemental sample is to efficiently enhance the precision available from the site sample by using a combination of the site sample and the supplemental sample. That is, the objective in combining the samples was to use the full sample (the site and supplemental samples) to achieve variance for national estimates that was lower than available for either sample. To simplify the combined-sample analyses, we explored procedures to determine whether a single combined-sample weight (or a set of combined-sample weights) could be constructed that would achieve variance estimates near to the minimum variance. The following sections describe the procedure to achieve minimum variance estimates from the combined samples, and the results for computing the combined-sample weights.

For computing survey estimates combined across the two surveys, *Est(Y),* separate estimates can be computed for each sample component and combined using the equation:

(66) $Est\ (Y) = \lambda\ Y(Site) + (1 - \lambda)\ Y(Supp),$

where *Y(Site)* is the survey estimate from the site sample, *Y(Supp)* is the survey estimate from the supplemental sample, and λ is an arbitrary constant between zero and one. For the sampling variance, *V(Y),* the estimate is computed using the equation:

(67) $(Y) = \lambda^2\ V(Y(Site)) + (1 - \lambda)^2\ V(Y(Supp)),$

where *V(Y(Site))* is the sampling variance for the estimate from the site sample, and *V(Y(Supp))* is the sampling variance for the estimate from the supplemental sample. Any value of λ will result in an unbiased estimate of the survey estimate, but not necessarily an estimate with the minimum sampling variance. The value associated with minimum variance, λ, can be computed as:

(68) $\lambda = [1/V(Y(Site)|Design)] / [1 / V(Y(Site)|Design) + 1/ V(Y(Supp)|Design)]$

$= V(Y(Supp)|Design) / [V(Y(Site)|Design) + V(Y(Supp)|Design)]$.

In this case, the minimum variance is:

(69) $V(Y) = [V(Y(Site)) \text{ “ } V(Y(Supp))] / [V(Y(Site)) + V(Y(Supp))]$,

with the design designation omitted.

To compute the combined-sample estimate with minimum variance, a survey estimate is derived by first computing the estimate for each survey component, and then computing a value of λ using the estimated variance from each survey component. The combined-sample point estimate is computed using the point estimate from each survey component and this value of λ (as in equation (66)). The sampling variance is estimated using the sampling variance estimate from each component survey and the computed value of λ (as in equation (67)). Although this process produces the minimum variance estimates, it is computer intensive. In addition, because of differing values of λ among levels of a categorical variable, it results in some inconsistencies among estimates of percentages and proportions. For example, proportional distributions, such as the proportion of the population by insurance type, sometimes did not sum to 100 percent because the component proportions had different values of λ. In addition, this two-step process for computing estimates would likely pose analytic problems for regression analyses and more

complex analyses. We therefore explored the use of single or multiple values of λ to construct one or more weights that could be used with the combined sample for all analyses.

The concept was that a value (or values) of λ was needed that would result in the best estimate and smallest variance for a variety of analysis variables and key populations. Because any value would result in an unbiased estimate, the key statistic for the analysis was the change in the sampling variance relative to the minimum variance. We also evaluated the change in the survey estimate relative to the survey estimate with minimum variance. For that analysis, we identified 14 analysis variables (10 categorical and 4 continuous) and nine populations (the full population and eight subpopulations). For dichotomous variables (for example, a yes/no variable), the sampling variances for both response options were equal and therefore redundant. After removing redundant and unstable estimates (estimates with a relative standard error of 0.30 or higher), 210 pairs of estimates and sampling variances were used in the analysis.

The mean value of the λs was 0.841, with a median of 0.85; the distribution of the λs was slightly skewed, with fewer than 10 percent of the λs less than 0.75. The value of λ was affected by design effects in the site sample (that is, by the average number of people in a site) and by the correlation among responses within a site (that is, the intracluster correlation). As expected, because of the number of persons in each site, the mean of the λs for estimates for the full population was the lowest (mean, 0.790; median, 0.793). For three key subpopulations (children, blacks, and Hispanics), the mean value of λ was between 0.84 and 0.87, and the mean of the median values for the three subpopulations was 0.866. The mean of the median values (0.866) was used as the λ for combining the weights for three reasons. First, it was close to the median value for all λs (0.850). Second, the sample sizes of the three subpopulations were relatively small, and it was desirable to minimize the variance estimates for point estimates for these subpopulations. Third, the optimal λ for the full population would result in less-than-optimal

variances for subpopulations; however, a less-than-optimal λ for the full population would not substantially increase the variance for that group.

Using the single value of λ, the combined-sample weight was computed for individuals in the site sample as:

(70) $WT(Combined) = \lambda\ WT(trimmed\ site\ sample\ weight)$,

and for individuals in the supplemental sample as:

(71) $WT(Combined) = (1 - \lambda)\ WT(trimmed\ supplemental\ sample\ weight)$.

Using this weight, the full data file could be processed in a single program, using survey data analysis software, such as SUDAAN.

G. SAMPLING ERROR ESTIMATION

1. Background

The CTS Household Survey sample design is complex and therefore requires specialized techniques for estimation of sampling variances. Standard statistical packages, such as SAS and SPSS, compute variances using formulas under the assumption that the data are from a simple random sample from an infinite population. Although the simple random sample variance may approximate the sampling variance, in some surveys it is likely to substantially underestimate the sampling variance with a design as complex as the CTS's. Departures from a simple random sample design result in a design effect that is defined as the ratio of the sampling variance *(Var)* given the actual survey design to the sampling variance of a hypothetical simple random sample with the same number of observations. Thus:

(72) $$Deff = \frac{Var\ (actual\ design\ with\ n\ cases)}{Var\ (SRS\ with\ n\ cases)}.$$

The sampling variance is a measure of the variation of an estimator attributable to having sampled a portion of the full population of interest using a specific probability-based sampling design. The sampling variance represents the average squared differences of the observations from their expected value over all possible samples of the same size and using the same sampling design. The classical population variance is a measure of the variation among the *observations* in the population, whereas a sampling variance is a measure of the variation of the *estimate* of a population parameter (for example, a population mean or proportion) over repeated samples. The population variance is different from the sampling variance in the sense that the population variance is a constant, independent of any sampling issues, whereas the sampling variance decreases in size as the sample size increases. The sampling variance is zero when the full population is observed, as in a census.

Based on the sampling variance, a series of measures of reliability can be computed for a parameter estimate or statistic. The standard error is the square root of the sampling variance. Over repeated samples of the same size and using the same sampling design, we expect that the true value of the statistic would differ from the sample estimate by less than twice the standard error in approximately 95 percent of the samples. The degree of approximation depends on the distributional characteristics of the underlying observations. The relative standard error is the standard error divided by the sample estimate and is usually presented as a percentage. In general, an estimate of a population parameter with a relative standard error of 50 percent is considered unreliable and is not reported. Furthermore, an estimate with a relative standard error of greater than 30 percent may be reported but also may be identified as potentially unreliable.

For the CTS Household Survey, the sampling variance estimate, called the *design-based sampling variance*, is a function of the sampling design and the population parameter being estimated. The design-based variance assumes the use of "fully adjusted" sampling weights,

which are derived from the sampling design, with adjustments to compensate for nonresponse and for ratio-adjusting the sampling totals to external totals (for example, to data on population totals by age and race/ethnicity generated by the Bureau of the Census from the Current Population Survey).

For combined national estimates at the person level, the average design effect over a representative set of variables is 2.6. With 59,956 observations, the Household Survey has the equivalent precision of a simple random sample with a size of about 22,675. Note that the design effect is generally lower for subclasses of the population because there is less clustering of observations.

The data files for the CTS Household Survey contain a set of fully adjusted sampling weights and information on analysis parameters (that is, stratification and analysis clusters) necessary for the estimation of the sampling variance for a statistic. Because of the stratification and unequal sampling rates, it was necessary to account for the sampling weights and the sampling design features in order to compute unbiased estimates of population parameters and their associated sampling variances. The estimation of the sampling variance required the use of special survey data analysis software or specially developed programs designed to accommodate the population parameter being estimated and the sampling design. The CTS Household Survey Public Use File for Round Two (Technical Publication No. 25. www.hschange.org), contains tables of standard errors for various types of estimates and provides a link to ICPSR.

Survey estimators fall into two general classes: (1) linear estimators, and (2) nonlinear estimators. Linear estimators are weighted totals of the individuals with an attribute, or means and proportions, if the denominators are known (for example, when the denominator is a poststratum total or a sum of poststrata totals). Nonlinear estimators include proportions and means (when the denominators are unknown and are estimated from the survey), ratios, and

correlation and regression coefficients. In general, the variances of nonlinear statistics cannot be expressed in a closed form. Woodruff (1971) suggested a procedure in which a nonlinear estimator is linearized by a Taylor series approximation. The sampling variance equation is then used on this linear form (called a *linearized variate*) to produce a variance approximation for the original nonlinear estimator.

Most common statistical estimates and analytic tools (such as percentages, percentiles, and linear and logistic regression) can be implemented using Taylor series approximation methods. Survey data software, such as SUDAAN (Shah et al. 1997), uses the Taylor series linearization procedure and can handle the multistage CTS Household Survey design, joint inclusion probabilities, and the stratification and clustering components of variance.

Other software packages use the Taylor series approximations (for example, Stata, SAS SurveySelect, and PC-CARP), but they do not account for the survey design as completely as does SUDAAN. A major advantage of SUDAAN is that site selection for the Household Survey used a high sampling rate, with unequal selection probabilities, and without replacement sampling. The SUDAAN estimation algorithm incorporates a finite population correction factor. Failure to account for the finite population correction causes an overestimate of the variance for national estimates based on the site sample.

The alternative to using the Taylor series approximations is to use a replication technique, such as balanced repeated replications, jackknife, or boot strapping. WESVAR uses replication techniques to estimate sampling errors, but the current version does not allow for the incorporation of the finite population correction for unequal probability sampling.

2. Variance Estimation

The CTS Household Survey contains a series of weights that are designed for site-specific and national estimates. The site-specific weights are designed for estimates that include units

(either FIUs or individuals) from the site sample and units selected in the supplemental sample that were within the boundary of a site. The weights available for national estimates include the national site sample weights, the supplemental weights, the combined weights that incorporate the site and supplemental samples and the national weights based on the augmented site sample. All four national weights were poststratified to the same population totals to ensure comparability; however, the four national samples may not produce precisely the same point estimates. The following discussion provides the variance estimation protocols for each of these weights.

a. Site-Specific Estimate Weights Based on the Augmented Sample

Variance estimation for site-specific estimates treats the sites as sampling strata (with the supplemental sample cases treated as a separate file). Within each of the 12 high-intensity sites, additional stratification was defined by RDD sample strata (two or three strata, depending on the site; see Table II.3) or as field sample. For the RDD sample, FIUs and individuals were treated as being clustered within households. For the field sample cases, the cluster was defined as the listing area. The samples were assumed to be selected "with replacement" in all strata.

b. Weights for National Estimates Based on the Site Sample

As discussed previously, the 60 sites are a national probability sample. Nine of the sites were sufficiently large that they were selected with probability of 1.0 (that is, they were certainty selections). The remaining 51 sites were selected from among three strata: (1) MSAs with 200,000 or more persons in 1992, (2) MSAs with fewer than 200,000 persons in 1992, and (3) nonmetropolitan areas. The sites were selected with probability proportional to size within these strata, using a variation of the probability minimal replacement sequential selection procedure (Chromy 1979). Because the sampling rate of sites was sufficiently large and the Chromy

sampling algorithm could be assumed, we used the finite population correction to improve the estimates of the sampling variances.

The finite population correction is a factor that accounts for the reduction in the sampling variance occurring when the sample is selected without replacement and a relatively large proportion of the frame is included in the sample. In an equal probability sample selected without replacement, if 20 percent of the frame is included in the sample, then the value of the finite population correction is 0.80, and the estimated sampling variance is 80 percent of the sampling variance one would have obtained if the factor were ignored. For the Household Survey, the sampling percentage of sites was sufficiently high among the large MSAs, so we were able to use the finite population correction to obtain more accurate and smaller sampling variance estimates. We also used the finite population correction concept for the small MSAs, but not for the nonmetropolitan areas. For the nonmetropolitan areas, the sampling rate was sufficiently small that we assumed with-replacement sampling; thus, it was not necessary to use the finite population correction factor.

For the MSA sites, the samples were selected without replacement and with unequal probability. To account for the finite population correction, we computed the probability of selection of any pair of selected sites jointly into the sample. These joint inclusion probabilities and a site's probability of selection were used to compute the finite population correction factor using the Yates-Grundy-Sen variance estimation equation (Wolter 1985). The SUDAAN software package permits direct variance estimates based on this equation.

The stratification used in the variance estimation consisted of the following 20 analysis strata, also called *pseudostrata:*

- Nine analysis strata, one corresponding to each of the nine sites selected with certainty

- Nine analysis strata formed among the 39 noncertainty sites in the stratum of large MSAs (to facilitate the computation of the joint selection probabilities)
- One stratum for small MSAs
- One stratum for nonmetropolitan areas

In the nine analysis strata for the certainty selections, there was no first-stage variance component, and only a within-site variance component exists. For the noncertainty sample of MSAs, we assumed a two-stage design, with variance components at the first stage (assuming unequal probability and without replacement selection of the sites) and a variance component within the sites. For the nonmetropolitan sites, we assumed that the sites were selected with replacement; therefore, the variation among the first-stage units (the sites) accounted for the variance contribution from all stages of selection.

The within-site variance contributions were estimated for the 12 high-intensity sites using the stratification of the RDD sample and the field sample. In the low-intensity sites, the site sample was assumed to be a simple random sample with no stratification.

c. Weights for National Estimates Based on the Supplemental Sample

The supplemental sample is a national RDD sample using five strata—four geographic regions for areas within MSAs and the country as a whole for nonmetropolitan areas. Variance estimation assumed a simple stratified random sampling design, with households as the sites and no adjustment for the finite population correction.

d. Weights for National Estimates Based on the Combined Sample

The maximum precision for national survey estimates is obtained by combining the site sample and the supplemental sample. For computing survey estimates, combined across the two sample components, *Est(Y)*, separate estimates can be computed for each sample component and

combined using equation (66). The sampling variance of this estimate, *V(Y)*, is computed using equation (67). Section G of this chapter describes the value of λ we derived to simplify processing without substantial loss in precision. The combined weights incorporated this value.

The variance estimation protocol treated the site survey sample and the supplemental sample as separate strata. The combined-sample variance estimation used the full variance estimation protocols (as described) for each of the component designs.

e. Weights for National Estimates Based on the Augmented Site Sample

The variance estimation protocol for this weight is the same as that for national estimates based on the site sample, with one difference. The additional cases from the supplemental sample are assigned values according to the sites in which they are located.

THIS PAGE LEFT INTENTIONALLY BLANK

REFERENCES

American Association for Public Opinion Research. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. Ann Arbor, MI: AAPOR, 2000.

Brick, J.M., and P. Broene. "Unit and Item Response Rates, Weighting, and Imputation Procedures in the 1995 National Household Education Survey." Working Paper 97-06. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, 1997.

Brick, J.M., I. Flores-Cervantes, and D. Cantor. "1997 NSAF Response Rates and Methods Evaluation." Report No. 8. *Assessing the New Federalism*, 1998.

Brick, J.M., J. Montaquila, and F. Scheuren. "Estimating Residency Rates for Undetermined Numbers in RDD." *Public Opinion Quarterly,* forthcoming.

Brick, J. Michael, Joseph Waksberg, Dale Culp, and Amy Starer. "Bias in List-Assisted Telephone Samples." *Public Opinion Quarterly,* vol. 59, summer 1995, pp. 218-235.

Center for Studying Health System Change. "Community Tracking Study Household Survey Public Use File: User's Guide (Round Two, Release 1." CSHSC Technical Publication No. 21. Washington, DC: CSHSC, June 2001.

Chromy, J.R. "Sequential Sample Selection Methods." *In American Statistical Association Proceedings of the Section on Survey Research Methods.* Alexandria, VA: American Statistical Association, 1979, pp. 401-406.

Corey, C.R., and H.E. Freeman. "Use of Telephone Interviewing in Health Care Research." *HSR Reports,* vol. 25, no. 1, 1990, pp. 129-144.

Ferraro, David, J. Michael Brick, and Teresa Strickler. "Weighting Issues in an RDD Panel Survey." Paper presented at the Meetings of the American Statistical Association, Indianapolis, IN, August 2000.

Frankel, L.R. "The Report of the CASRO Task Force on Response Rates." In *Improving Data Quality in a Sample Survey,* edited by F. Wiseman. Cambridge, MA: Marketing Science Institute, 1983.

Groves, R.M., and L. Lyberg. "An Overview of Nonresponse Issues in Telephone Surveys." In *Telephone Survey Methodology,* edited by R. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls II, and J. Waksberg. New York: John Wiley & Sons, 1988.

Hall, John, Genevieve Kenney, Gary Shapiro, and Ismael Flores-Cervantes. "Bias from Excluding Households Without Telephones in Random Digit Dialing Surveys—Results of Two Surveys." In *American Statistical Association 1999 Proceedings of the Section on Survey Research Methods.* Alexandria, VA: American Statistical Association, 2000. pp. 382-387.

Hall, J., R. Strouse, B. Carlson, and R. Stapulonis. "Survey Design and Data Collection Methods for the Robert Wood Johnson Foundation's Family Health Insurance Survey." Princeton, NJ: Mathematica Policy Research, Inc., September, 1994.

Keeter, S., and C. Miller. "Consequences of Reducing Telephone Survey Nonresponse Bias—Or What Can You Do in Eight Weeks that You Can't Do in Five Days." Paper presented at the American Association for Public Opinion Research meetings, St. Louis, MO, 1998.

Kish, L. *Survey Sampling*. New York: John Wiley & Sons, 1965.

Marcus, A.C., and L.A. Crane. "Telephone Surveys in Public Health Research." *Medical Care,* vol. 24, no. 2, 1986, pp. 97-112.

Marketing Systems Group. *Genesys Sampling Systems Methodology*. Philadelphia, PA: MSG, 2000.

Marketing Systems Group. *Genesys Sampling Systems Methodology*. Philadelphia, PA: MSG, 1994.

Metcalf, Charles E., Peter Kemper, Linda T. Kohn, and Jeremy D. Pickreign. "Site Definition and Sample Design for the Community Tracking Study." CSHSC Technical Publication No. 1. Washington, DC: Center for Studying Health System Change, October 1, 1996.

Piekarski, L., G. Kaplan, and J. Prestegaard. "Telephony and Telephone Sampling: The Dynamics of Change." Paper presented at the American Association for Public Opinion Research meetings, St. Petersburg, FL, 1999.

Potter, F.J. "A Study of Procedures to Identify and Trim Extreme Sampling Weights." In *American Statistical Association Proceedings of the Section on Survey Research Methods*. Alexandria, VA: American Statistical Association, 1990, pp. 225-230.

Robinson, E.L. "The Current Population Survey Technical Summary of Design and Methodology." Washington, DC: U.S. Bureau of the Census, January 1992.

Shah, B.V., B.G. Barnwell, and G.S. Bieler. *SUDAAN User's Manual, Release 7.5*. Research Triangle Park, NC: Research Triangle Institute, 1997.

Shapiro, G., M. Battaglia, D. Camburn, J. Massey, and L. Tompkins. "Calling Local Telephone Business Company Offices to Determine the Residential Status of a Wide Class of Unresolved Telephone Numbers in a Random-Digit-Dialing Sample." Unpublished long version of paper published in *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 1995.

Singer, Eleanor. "Experiments with Incentives in Telephone Surveys." *Public Opinion Quarterly*, vol. 64, no. 2, 2000, pp. 171-188.

Strouse R., J. Hall, B. Carlson, and J. Cheng. "Impact of Nontelephone Sample of Family Health Insurance Estimates." Report to the Robert Wood Johnson Foundation. Princeton, NJ: Mathematica Policy Research, Inc., 1997.

Strouse, R., J.W. Hall, F. Potter, B.L. Carlson, P. Cunningham, J. Pascale, and R. Stapulonis. "Report on Survey Methods for the Community Tracking Study's 1996-1997 Round One Household Survey—Final." Report submitted to the Center for Studying Health System Change. Princeton, NJ: Mathematica Policy Research, Inc., November 1998.

Thornberry, Owen T., and J. Massey. "Trends in United States Telephone Coverage Across Time and Subgroups." In *Telephone Survey Methodology*, edited by R. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls II, and J. Waksberg. New York: John Wiley & Sons, 1988.

U.S. Bureau of the Census. Available online at: [http://www.census.gov/prod/2000pubs/tp63.pdf.] Accessed March 2000.

U.S. Bureau of the Census. *1998 Current Population Survey*. Washington, DC: Census, 1998.

U.S. Bureau of the Census. Unpublished tables of percentages of households with telephones in 1997.

U.S. Bureau of the Census. Unpublished tables of percentages of households with telephones for 1988-1993.

Wolter, Kirk M. *Introduction to Variance Estimation.* New York: Springer-Verlag, 1985.

Woodruff, R.S. "A Simple Method for Approximating the Variance of a Complicated Estimate." *Journal of the American Statistical Association,* vol. 66, no. 334, 1971, pp. 411-414.