Community Tracking Study
Household Survey
Survey Methodology Report
(Round One)

Center for Studying Health System Change

Technical Publication No.

15

November 1998

Community Tracking Study Household Survey Survey Methodology Report (Round One)

Center for Studying Health System Change

Technical Publication No. 15

November 1998

This is one of a series of technical documents that have been done as part of the Community Tracking Study being conducted by the Center for Studying Health System Change. The study will examine changes in the local health systems and the effects of those changes on the people living in the area.

The Center welcomes your comments on this document. Write to us at 600 Maryland Avenue, SW, Suite 550, Washington, DC 20024-2512 or visit our web site at www.hschange.org.

The Center for Studying Health System Change is supported by the Robert Wood Johnson Foundation and is affiliated with Mathematica Policy Research, Inc.

CONTENTS

CONTENTS (*continued*)

CONTENTS (*continued*)

CONTENTS *(continued)*

Chapter **Page**

TABLES

TABLES *(continued)*

Table **Page**

FIGURES

EXHIBITS

I. OVERVIEW OF THE COMMUNITY TRACKING STUDY AND THE HOUSEHOLD SURVEY

The Community Tracking Study (CTS) is a national study of the rapidly changing health care market and the effects of these changes on people.[1] Funded by The Robert Wood Johnson Foundation, the study is being conducted by the Center for Studying Health System Change (HSC). Information about other aspects of the CTS is available from HSC at www.hschange.com or by e-mail (center@hschange.com).

Mathematica Policy Research, Inc. (MPR) was the primary contractor for the household survey design, instrument development, sample design and implementation, most of the interviewing, weighting, and variance estimation. Battelle, Inc. and CODA, Inc. 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 CTS Public Use File and its documentation.

A. CTS OBJECTIVES

The overall goal of the CTS--to develop an information base designed to track and analyze change in the nation's health care market and to inform public and private decision makers about these changes--has three component objectives:

1. ***Tracking Changes in Health Systems***. The study's first objective is to document changes in the health system through intensive study of selected communities. The major changes that have been reported in the health system include consolidation of the market at all levels (medical groups, hospitals, insurers, and health plans); vertical integration of providers (for example, hospitals and physicians) and of insurers and providers; increased risk sharing by providers; growth of large, national, for-profit health are enterprises; and the adoption of new techniques

[1]An overview of the Community Tracking Study is contained in Kemper et al. (1996). A description of the survey design is also included in Center for Studying Health System Change (1998).

for managing clinical care (clinical information systems, quality improvement techniques, utilization management, and so forth).

2. ***Tracking Changes in Access, Service Delivery, Cost, and Perceived Quality.*** The second objective of the study is to monitor the effects of health system change on people by tracking indicators of these effects, including favorable or unfavorable changes in access to care, service use and delivery, and quality and cost of care.

3. ***Understanding the Effects of Health System Change on People.*** The third objective of the study is to understand how differences in health systems are related to differences in access, service delivery, cost, and perceived quality. This objective will be achieved by analyzing-- qualitatively and quantitatively--the relationship between health systems and access, delivery, cost, and perceived quality.

Central to the design of the study is its community focus. This focus was established because health care delivery is primarily local and differs from one community to the next as a result of history, culture, and state and local policy. Therefore, to analyze and understand institutional changes in the delivery system and their effects on people, it is necessary to obtain information at the local level. To this end, 60 communities, listed in Table I.1, were randomly selected to form the core of the CTS and to be representative of the nation as a whole.[2] Of these communities, 12 were randomly chosen for more intensive study. They are referred to as the *high-intensity sites*.

1. Analytic Components of the Community Tracking Study

The CTS has qualitative and quantitative components. For instance, case studies are being conducted in the 12 high-intensity sites. The first round of comprehensive case studies of the health system was begun in 1996 and continued through 1997. The findings were available from HSC[3].

---

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

[3]Center for Studying Health System Change (1997a).

TABLE I.1

SITES SELECTED FOR THE COMMUNITY TRACKING STUDY

High-Intensity Sites	Low-Intensity Sites		
Metro Areas >200,000 Population	Metro Areas >200,000 Population	Metro Areas <200,000 Population	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)		
	47-West Palm Beach (FL)		
	48-Worcester (MA)		

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

This qualitative information is complemented by survey data from the 12 communities and from an additional 48 sites. In all 60 sites, HSC simultaneously conducted independent surveys of households, physicians, and employers, enabling researchers to explore relationships among purchasers, providers, and consumers of health care.[3] The Followback Survey of Health Plan Organizations is another component of the CTS. In this survey respondents to the CTS Household Survey who are covered by privately financed health insurance plans (employer, union, and privately purchased) will be "followed back" to the organization that administers the plan. This kind of survey provides information on available health plans and identifies the particular plan in which each linked policyholder is enrolled. The features of health plans measured in the followback survey include basic managed care variables, network size, and provider payment methods. Data for all survey components will be collected on a two-year cycle, allowing researchers to track changes in the health care system over time. The round one surveys and case studies, completed during 1996 and 1997, are the baseline. Data collection for round two began in August 1998.

2. The Household Survey

After selecting the sample sites, we randomly selected households within each site. We also randomly selected households for the supplemental sample, an independent national sample. We determined the composition of each household, grouped household members into family insurance units (FIUs), and obtained information on each adult in each FIU.[4] If an FIU contained one child, we collected information

---

[4]The physician survey was conducted by HSC and will be made available as a public use file. The employer survey was conducted by RAND in collaboration with HSC. Although these surveys were conducted in the same communities, they were independent of one another, and physicians or employers in the surveys cannot be linked to specific people.

[5]The FIU is based on groupings of people typically used by insurance carriers. It includes an adult (continued...)

about that child. If an FIU contained two or more children, we collected information about one randomly selected child. Figure I.1 shows an overview of survey procedures.

The Household Survey instrument covers health insurance, use of health services, satisfaction with care, health status, and demographic information. An adult in each FIU (the family informant) provided information on insurance coverage, health resource use, and usual source of care for all individuals in the FIU. This informant also provided information on family[6] income as well as on employment, earnings, employer-offered insurance plans, and race/ethnicity of each adult in the FIU. Each adult in the FIU (including the informant) responded through a self-response module to questions about unmet needs, patient trust, satisfaction with physician choice, detailed health status, 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 the randomly selected child about unmet needs and satisfaction with physician choice. 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 also fielded.

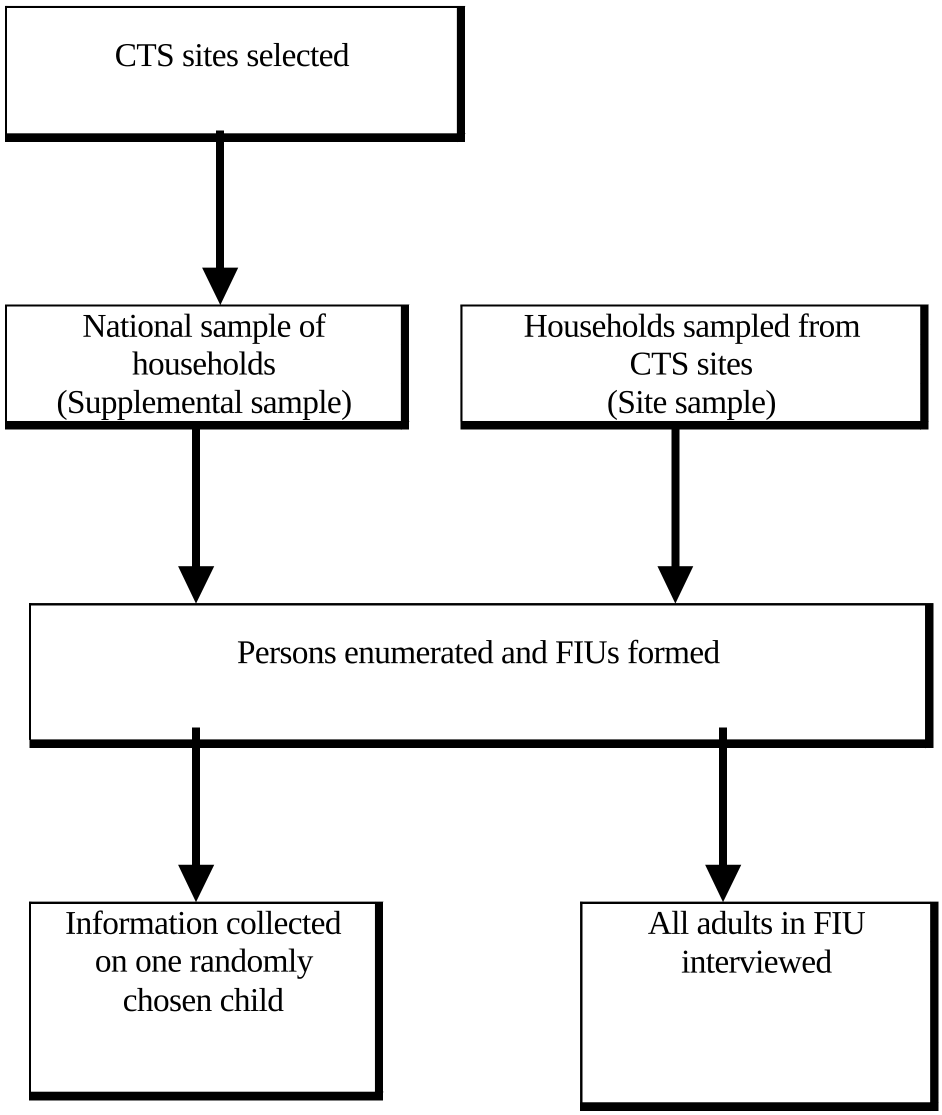
The survey was administered completely by telephone, using computer-assisted telephone interviewing technology. Although the vast majority of the respondents were selected through the use of a list-assisted random-digit-dialing sampling methodology, families without working telephones were represented in the sample as well. Field staff using cellular telephones enabled these families to complete interviews.

(...continued)
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.

[6]We use the term "family" to refer to individuals within the FIU.

FIGURE I.1

OVERVIEW OF HOUSEHOLD SURVEY PROCEDURES



Interviews with 60,446 individuals from 32,732 FIUs were completed between July 1996 and July 1997.

In the following chapters, we describe site selection and sample design, survey design and preparation, data collection procedures, and weighting and estimation.

II. SITE SELECTION AND SAMPLE DESIGN

The Household Survey was administered to households in the 60 Community Tracking Study (CTS) sites and to an independent national sample of households. The survey's three-tier sample design 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 were 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 were surveyed. This sample of "low-intensity" sites allows us to validate results from the high-intensity sites and permits findings to be generalized to the nation. The first and second tiers together are known as the *site sample*.

- The third tier is a smaller, independent national sample known as the *supplemental sample*. This sample augments the site sample and substantially increases the precision of national estimates with a relatively modest increase in total sample.

In the following sections, we discuss site selection; selection of households, family insurance units (FIUs), and individuals; sample size considerations; and procedures for selecting the random-digit-dialing (RDD) and field samples.

A. CTS SITE SELECTION

The primary goal of the CTS was 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. Three issues were central to this sample design: (1) how sites were defined, (2) how the number for study was determined, and (3) how sites were selected.

1. Definition of Sites

The first step was to define the sites. Sites were intended to encompass local health care markets. Although there are no set boundaries for these local markets, the intent was to define areas such that residents predominately used health care providers in their area and such that providers served predominately area residents. To this end, we generally defined sites to be either metropolitan statistical areas (MSAs) as defined by the Office of Management and Budget or the nonmetropolitan portions of economic areas as defined by the Bureau of Economic Analysis[1].

2. Number of Sites

The next step in creating the site sample was to determine the number of high-intensity sites. In making this decision, we considered the tradeoffs between data collection costs (case studies plus survey costs) and the research benefits of a large sample of sites. These benefits include a greater ability to empirically examine the relationship between system change and its effect on health care delivery and consumers, and to make the study findings more generalizable to the nation.

Despite the cost advantages of conducting intensive case studies in fewer sites, the smaller the number of communities, the more difficult it is to distinguish between changes of general importance and changes or characteristics unique to a community. To solve the problem by increasing the number of case study sites would make the cost of data collection and analysis prohibitively high. We therefore chose 12 sites for intensive study and added to this sample 48 sites that would be studied less intensively. These 60 high-intensity and low-intensity sites form the *site sample*.

---

[1]For more details on the definition of CTS sites, refer to Metcalf et al. (1996)

There was no formal scientific basis for settling on 12 high-intensity sites, but this number reflects a balance between the benefits of studying a range of different communities and the costs of doing so. Although this number is by no means trivial for an intensive case study design, the addition of 48 low-intensity sites solves the problem of limited generalizability associated with only 12 sites. The additional 48 sites also provide 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. Shown previously in Table I.1, the 60 sites were chosen for the first stage of sampling. Sites were sampled by stratifying them geographically by region and selecting them randomly with probability in proportion to population size. There were separate strata for (1) MSAs with a 1995 population of 200,000 or more, (2) MSAs with a 1995 population of less than 200,000, and (3) nonmetropolitan areas. This sampling approach provided maximum geographic diversity, judged critical for the 12 high-intensity sites in particular, and acceptable natural variation in city size and degree of market consolidation.[2]

The 12 high-intensity sites were selected randomly from among the 48 selected MSAs with 1995 populations of 200,000 or more. Of the 48 low-intensity sites, 36 are the remaining large metropolitan areas (also having 1995 populations of 200,000 or more), 3 are small metropolitan areas (populations of less than 200,000), and 9 are nonmetropolitan areas.

Together, the high-intensity and low-intensity sites account for about 90 percent of all survey respondents. This site sample can be used to make national estimates. The sample may also be used to

---

[2]Additional information about the number of sites and the random selection of the site sample is available in Metcalf et al. (1996).

make site-specific estimates for the high-intensity sites. However, site-specific estimates for the low-intensity sites will be less precise because of the small sample size for these sites.

4. Additional Samples and Better National Estimates

Although the site sample alone will yield national estimates, such estimates will not be as precise as they could have been had even more communities been sampled, or had the sample been a simple random sample of the entire U.S. population. The *supplemental sample*, the third tier in the design of the CTS Household Survey sample, was added to increase the precision of national estimates at a relatively small incremental increase in survey costs.

The supplemental sample is a relatively small, nationally representative sample made up of households randomly selected from the 48 states in the continental United States. It is stratified by region but essentially uses simple random sampling techniques 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 making national estimates from the site sample more precise, the supplemental sample also 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. Therefore, when a site-specific estimate is made, the individual site sample can be augmented with observations from the supplemental sample within that site. The resulting sample is known as the *augmented site sample*.

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 phone, if an adult. The person who owned or rented the house was identified as the head of the household, or the householder. Persons who usually live in the household but who were temporarily living elsewhere, including college students, were included in the household enumeration.

2. Family Insurance Units

Individuals in the household were grouped into 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 prior to the interview, and usual source of health care. As noted in Chapter I, the FIU informant also provided information on family income as well as on employment, earnings, employer offered health insurance plans, and race/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 0 to 17 years of age or 18 to 22 years of age if a full-time student (even if living outside the household).

All FIUs were selected to participate in the remainder of the interview as long as there was at least one civilian adult in the unit. (Individuals who were not on active military duty at the time of the interview were considered to be civilians.) In each FIU, one informant was responsible for providing most of the information about the family and its members. Figure II.1 show 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 he/she answered the telephone and is familiar with the

FIGURE II.1

EXAMPLE OF FIU IN A HYPOTHETICAL HOUSEHOLD

Members of Household	FIU
Head of Household	F1
Head of Household's Spouse	
Head of Household's Daughter	
Head of Household's Son	
Head of Household's Father	F2
Head of Household's Mother	
Unrelated Boarder	F3

composition of the household. Because the spouse is also familiar with the health care of the head of household and their children, the spouse is also the informant for the first FIU (F1). The household head's father is the informant for family unit two (F2), and the unrelated boarder responds for him/herself (F3). The household head's daughter is the randomly selected child in F1 and the head's son is not 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.

The CTS 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 either by blood or marriage. The Census family is often larger than an FIU. Adult relatives living within the same household would be included in a Census primary family but would be assigned to separate FIUs for the CTS survey.

3. Individuals

Each FIU informant answered question 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. (A "child" was defined as an unmarried individual younger than age 18.) Full-time college students (ages 18 to 22), even if they were living away from home at the time of the survey, were listed as household members and were included in their parents' FIU. These 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.

Each adult also was asked to answer a subset of subjective questions, including assessments of health, tobacco use, 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 computerized survey instrument imposed a maximum of eight persons per household for inclusion in the survey. The family informant identified all members of responding households; in the rare instance of households exceeding eight persons, the interviewers were instructed to list all adults in the household first, followed by as many children as possible before reaching the maximum.

Some household members were excluded from sampled households because they had multiple chances of selection or were not part of the study population. Unmarried full-time college students (ages 18 to 22) are represented in their parent's or guardian's FIU, even if they were away at college at the time of the interview. To avoid giving these individuals multiple chances of selection, they were included only in their parents' or guardians' samples and were excluded from any other sampled dwellings. Unmarried children under age 18 with no parent or guardian in a household were also excluded, because they could have been selected in their parent's household. An adult on active military duty was classified as ineligible; however, such a person could have acted as a family informant as long as there was at least one civilian adult in the FIU. Family insurance units in which all adults were active duty military personnel were considered ineligible for the survey.

C. SAMPLE SIZE CONSIDERATIONS FOR THE HOUSEHOLD SURVEY

The intended inferences that the data will be used to support ultimately 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 sample size considerations related to the 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 waves, and for making cross-site comparisons; and (2) national estimates and comparisons. Finally, we include tables showing the number of FIUs and persons that were interviewed, by site and sample.

1. Requirements for Site-Based Estimates

The design called for the capability to both make point-in-time estimates and measure change over time. In the 12 high-intensity sites, the base-year design called for interviews with approximately 1,225 FIUs (combined RDD and field samples) in 1996-1997.[3] In addition, the supplemental sample was expected to provide additional FIUs for each of the sites, depending on site size (approximately 25 FIUs per high-intensity site).

We estimated the high-intensity-site sample size requirement of 1,225 by considering the following design considerations:

- Minimizing design effects[4] resulting from clustering of multiple FIUs within households and from sampling methods for coverage of nontelephone households
- Allowing for analyses of subgroups of interest

---

[3]The site sample size for the 1998-1999 survey will be approximately the same.

[4]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 will substantially underestimate the sampling variance. 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 random sample with the same number of observations. Sampling error estimation methods are discussed in Chapter V.

- C Measuring and testing hypotheses about change over two interviewing waves
- C 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 the 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 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 related to causes of change. Measuring change over multiple interviewing waves requires larger samples. We assumed that 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 for round one (a mixed longitudinal/cross-sectional design). We also assumed that approximately 40 to 45 percent of the households interviewed for the second interviewing wave would have been interviewed in the first wave. To measure changes over time (say, five percentage points for a midrange percentage), we estimated that an effective sample of about 975 per wave 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.

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 allows for detecting 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 wave, *combined with a mixed longitudinal/cross-sectional design over time,* was an appropriate sample size for each of the 12 high-intensity sites. 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 do not allow for precise individual-site-level analyses. We initially set a sample target of 375 FIUs per site but reduced it 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

Given the scale and significance of the CTS, it is desirable to track changes in a way that permits 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 would restrict sample inferences to the population in metropolitan areas of this size and result in poor precision for national estimates.

A sample of 60 sites would increase the precision for large metropolitan areas and expand the generalizability of the household sample to small metropolitan areas and nonmetropolitan areas. In addition, we decided to augment the clustered 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 was approximately 3,500 FIUs, which we later reduced slightly; a total of 3,276 FIUs were interviewed.

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, such a sample involves much greater costs than does an RDD sample. For reasons of cost, we rejected extending the field sample to represent nonmetropolitan and small metropolitan areas. We concluded that the field sample in the 12 high-intensity sites could adequately represent nontelephone households in these metropolitan areas.[5] Thus, the design specifies that nontelephone households are not sampled in low-intensity sites or in the supplemental sample.

3. Actual Sample Sizes

The numbers of FIUs and of individuals interviewed, by site and type of sample, are shown in Tables II.1 and II.2, respectively. Altogether, 32,732 FIUs and 60,446 persons were interviewed. The number of FIUs per augmented high-intensity site varied from 1,179 to 1,419, meeting or exceeding nominal sample size requirements. The average high-intensity-site design effect attributed to clustering of FIUs within households and weighting for nonresponse was 1.13 (range, 1.07 to 1.22), which was somewhat less than the expected level of 1.25. (Estimation procedures are described in Chapter V.)

---

[5]The field sample totaled 635 FIUs in an average of six clusters of nine interviews each, in each of the 12 high-intensity sites. Although the effective sample of metropolitan nontelephone households is too small to conduct separate analyses of nontelephone families, these families represent, at most, only five percent of all metropolitan families. Thus, the design effects on estimates pertaining to *all* metropolitan families--resulting from clustering and underrepresentation of nontelephone households--is well within acceptable bounds.

TABLE II.1

NUMBER OF FAMILY INSURANCE UNITS INTERVIEWED, BY SITE AND SAMPLE

	Sample			
Site/Geographic Area[a]	Site	Supplemental	Augmented Site	Combined
Total	**29,456**	**3,276**	**30,787**	**32,732**
High-Intensity Sites				
01-Boston (MA)	1,145	34	1,179	1,179
02-Cleveland (OH)	1,211	32	1,243	1,243
03-Greenville (SC)	1,285	14	1,299	1,299
04-Indianapolis (IN)	1,316	29	1,345	1,345
05-Lansing (MI)	1,232	5	1,237	1,237
06-Little Rock (AR)	1,412	7	1,419	1,419
07-Miami (FL)	1,171	26	1,197	1,197
08-Newark (NJ)	1,282	19	1,301	1,301
09-Orange County (CA)	1,157	35	1,192	1,192
10-Phoenix(AZ)	1,250	27	1,277	1,277
11-Seattle (WA)	1,181	38	1,219	1,219
12-Syracuse (NY)	1,303	7	1,310	1,310
Low-Intensity Sites				
13-Atlanta (GA)	296	52	348	348
14-Augusta (GA/SC)	291	6	297	297
15-Baltimore (MD)	285	25	310	310
16-Bridgeport(CT)	284	6	290	290
17-Chicago (IL)	293	92	385	385
18-Columbus (OH)	296	26	322	322
19-Denver (CO)	291	40	331	331
20-Detroit (MI)	309	49	358	358
21-Greensboro (NC)	271	12	283	283
22-Houston (TX)	280	46	326	326
23-Huntington (WV/KY/OH)	307	7	314	314

	Sample			
Site/Geographic Area[a]	Site	Supplemental	Augmented Site	Combined
24-Killeen (TX)	298	1	299	299
25-Knoxville (TN)	311	15	326	326
26-Las Vegas (NV/AZ)	267	7	274	274
27-Los Angeles (CA)	261	111	372	372
28-Middlesex (NJ)	311	11	322	322
29-Milwaukee (WI)	311	26	337	337
30-Minneapolis (MN/WI)	334	37	371	371
31-Modesto (CA)	306	4	310	310
32-Nassau (NY)	341	16	357	357
33-New York City (NY)	292	59	351	351
34-Philadelphia (PA/NJ)	309	53	362	362
35-Pittsburgh (PA)	299	23	322	322
36-Portland (OR/WA)	307	19	326	326
37-Riverside (CA)	304	22	326	326
38-Rochester (NY)	355	14	369	369
39-San Antonio (TX)	299	17	316	316
40-San Francisco (CA)	281	24	305	305
41-Santa Rosa (CA)	285	5	290	290
42-Shreveport (LA)	298	7	305	305
43-St. Louis (MO/IL)	318	30	348	348
44-Tampa (FL)	268	31	299	299
45-Tulsa (OK)	292	5	297	297
46-Washington (DC/MD)	310	68	378	378
47-W Palm Beach (FL)	253	16	269	269
48-Worcester (MA)	310	11	321	321
49-Dothan (AL)	301	0	301	301
50-Terre Haute (IN)	293	0	293	293
51-Wilmington (NC)	303	6	309	309
52-W-Cen Alabama	329	3	332	332

	Sample			
Site/Geographic Area[a]	Site	Supplemental	Augmented Site	Combined
53-Cen Arkansas	379	11	390	390
54-N Georgia	273	11	284	284
55-NE Illinois	294	6	300	300
56-NE Indiana	286	4	290	290
57-E Maine	319	10	329	329
58-E North Carolina	304	10	314	314
59-N Utah	377	3	380	380
60-NW Washington	330	1	331	331
Areas other than CTS Sites	--	1,945	--	1,945

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

TABLE II.2

NUMBER OF PERSONS INTERVIEWED, BY SITE AND SAMPLE

	Sample			
Site/Geographic Area	Site	Supplemental	Augmented Site	Combined
Total	**54,371**	**6,075**	**56,798**	**60,446**
High-Intensity Sites				
01-Boston (MA)	2,024	55	2,079	2,079
02-Cleveland (OH)	2,217	59	2,276	2,276
03-Greenville (SC)	2,436	32	2,468	2,468
04-Indianapolis (IN)	2,451	56	2,507	2,507
05-Lansing (MI)	2,291	9	2,300	2,300
06-Little Rock (AR)	2,644	14	2,658	2,658
07-Miami (FL)	2,031	44	2,075	2,075
08-Newark (NJ)	2,311	33	2,344	2,344
09-Orange County (CA)	2,101	63	2,164	2,164
10-Phoenix (AZ)	2,263	47	2,310	2,310
11-Seattle (WA)	2,043	70	2,113	2,113
12-Syracuse (NY)	2,363	16	2,379	2,379
Low-Intensity Sites				
13-Atlanta (GA)	538	97	635	635
14-Augusta (GA/SC)	563	14	577	577
15-Baltimore (MD)	527	47	574	574
16-Bridgeport (CT)	548	11	559	559
17-Chicago (IL)	573	160	733	733
18-Columbus (OH)	557	48	605	605
19-Denver (CO)	558	73	631	631
20-Detroit (MI)	562	94	656	656
21-Greensboro (NC)	506	20	526	526
22-Houston (TX)	546	90	636	636

	Sample			
Site/Geographic Area	Site	Supplemental	Augmented Site	Combined
23-Huntington (WV/KY/OH)	568	8	576	576
24-Killeen (TX)	579	2	581	581
25-Knoxville (TN)	577	24	601	601
26-Las Vegas (NV/AZ)	481	11	492	492
27-Los Angeles (CA)	462	207	669	669
28-Middlesex (NJ)	572	18	590	590
29-Milwaukee (WI)	524	48	572	572
30-Minneapolis (MN/WI)	648	76	724	724
31-Modesto (CA)	606	7	613	613
32-Nassau (NY)	662	32	694	694
33-New York City (NY)	483	108	591	591
34-Philadelphia (PA/NJ)	569	95	664	664
35-Pittsburgh (PA)	544	45	589	589
36-Portland (OR/WA)	557	39	596	596
37-Riverside (CA)	574	42	616	616
38-Rochester (NY)	658	21	679	679
39-San Antonio (TX)	565	32	597	597
40-San Francisco (CA)	431	34	465	465
41-Santa Rosa (CA)	541	11	552	552
42-Shreveport (LA)	565	10	575	575
43-St. Louis (MO/IL)	590	55	645	645
44-Tampa (FL)	499	58	557	557
45-Tulsa (OK)	588	9	597	597
46-Washington (DC/MD)	551	116	667	667
47-W Palm Beach (FL)	423	22	445	445
48-Worcester (MA)	586	25	611	611
49-Dothan (AL)	558	0	558	558
50-Terre Haute (IN)	553	0	553	553
51-Wilmington (NC)	541	10	551	551

	Sample			
Site/Geographic Area	Site	Supplemental	Augmented Site	Combined
52-W-Cen Alabama	606	6	612	612
53-Cen Arkansas	770	20	790	790
54-N Georgia	511	21	532	532
55-NE Illinois	564	11	575	575
56-NE Indiana	565	8	573	573
57-E Maine	633	18	651	651
58-E North Carolina	592	17	609	609
59-N Utah	811	7	818	818
60-NW Washington	611	2	613	613
Areas other than CTS Sites	--	3,648	--	3,648

For low-intensity sites, nominal augmented sample sizes of FIUs ranged from 269 to 390. As can be seen in Table II.1, supplemental samples in large metropolitan low-intensity sites (sites 13 to 48) significantly increased the size of these site samples.

D. RANDOM-DIGIT-DIALING SAMPLE SELECTION

In this section, we describe selection of the RDD samples for the Household Survey. First, we describe the sampling frame used to select the sample in the 60 sites and the supplemental sample. We then turn to the issues of stratification and allocation, followed by a discussion of generation and release of the telephone sample.

1. Sampling Frame

We used the Genesys Sampling System to select the RDD household sample. To develop a sampling frame for a county or group of counties, Genesys first assigns each area-code/exchange combination to a unique county.[6] 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. An analysis of the published numbers in each of the 60 sites indicates that fewer than one percent of numbers assigned to one of our sites represented a household located outside that site.[7]

---

[6]In the U.S.10-digit telephone numbering system (XXX-YYY-ZZZZ), the first three digits (XXX) are referred to as the area code, and the next three (YYY) as the exchange.

[7]Data provided by Genesys showed an average of 99.1 percent of the sample in each site would reside in that site, and that the frame would cover an average of 99.2 percent of all telephone households in a 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 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). However, undercoverage is probably less than 3.5 percent because of the way that the RDD household sample for the CTS was selected. We selected telephone numbers repeatedly over a period of nine months (from August 1996 through April 1997); therefore, some banks that were nonworking early in the project could have become working as new directories were incorporated into updates of Genesys.

2. Stratification and Sample Allocation

Stratification was used for the supplemental sample and the high-intensity sites to help ensure proportionate representation. Samples within the low-intensity sites were not stratified. For the supplemental sample, we created five strata: one stratum for nonmetropolitan areas, and four strata of metropolitan counties in each of the four Census regions. Within the high-intensity sites, we stratified geographically by characteristics such as income distribution, race/ethnicity distribution, or county, depending on the composition of the site.

In sites with two or more counties, 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 the county included large black and Hispanic populations, we used both variables for stratification. However, for counties containing a significant fraction of one but not both of these population groups, or in which 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.

To determine the initial allocation of telephone numbers for each site, or 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 overall response rate. Telephone numbers within sites were sampled at equal rates across strata. This initial allocation of telephone numbers was later adjusted on the basis of actual experiences during the survey. Thus, if either the percentage of sampled telephone numbers that was residential or the response rate in a site was different than expected, the allocation of telephone numbers was adjusted to obtain the desired number of interviews. (Response rates and information on sample dispositions are discussed in Chapter IV.)

3. Sample Selection and Release

The initial sample was one-fourth of the total number of projected telephone numbers. 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

TABLE II.3

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

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

TABLE II.4

RELEASE OF SAMPLE FOR COMMUNITY TRACKING STUDY HOUSEHOLD SURVEY

Date	RDD Sample (Telephone Numbers)	Field Sample (Addresses)
17 July 1996	10,083	
24 July 1996	10,115	
22 August 1996	7,044	
20 September 1996	2,319	
24 September 1996	6,196	
28 October 1996	3,492	
29 October 1996	1,941	
31 October 1996	1,387	
1 November 1996	1,629	
6 November 1996	1,139	
8 November 1996	1,322	
11 November 1996	237	
12 November 1996	1,036	
18 November 1996	10,345	
25 November 1996	548	
18 December 1996	4,912	
2 January 1997	1,547	
4 January 1997		1,340
6 January 1997	1,548	
11 January 1997		894
18 January 1997		161
25 January 1997		341
29 January 1997	1,849	
1 February 1997		221
15 February 1997		70
22 February 1997		648
1 March 1997		784
14 March 1997	1,349	
15 March 1997		249
29 March 1997		19
12 April 1997		362
7 May 1997	599	
23 May 1997	300	
5 July 1997		57
19 July 1997		112
Total	**70,937**	**5,258**

- 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. In other words, if Genesys is set to select 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 Genesys Sampling Systems (1994).

The Genesys identification procedure involves two steps: (1) checking the sample against lists of published numbers, and (2) dialing numbers to determine whether they are nonworking. In the first step, all numbers are classified as published residential numbers, published business numbers, or other. The published residential numbers are retained, the business numbers eliminated, and the others prepared for dialing. Genesys uses an automated dialer to check for the tone that precedes a recorded message stating the number dialed is not in service (termed an "intercept message"). Numbers for which that tone is detected are removed from the sample as nonworking. To minimize intrusiveness, the Genesys dialer disconnects immediately if a ring is detected, and calls are made only between the hours of 9:00 a.m. and 5:00 p.m., local time. The remaining sample includes numbers identified as published residential plus those not classified by the dialer as nonworking.[8]

[8]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.

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 ensures that each release is worked evenly and eliminates 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 round of selection, but that subsequently became working. The sample was then released to the computer-assisted telephone interviewing (CATI) call scheduler; weekly survey reports on sample dispositions, by site, were used to determine the size of additional sample releases.

E. FIELD SAMPLE SELECTION

The CTS Household Survey included a field sample to provide coverage of families and persons who did not have telephones or who had substantial interruptions in telephone service. Several studies have indicated that omitting nontelephone households could lead to biased survey estimates (Thornberry and Massey 1988; Marcus and Crane 1986; and Corey and Freeman 1990). A similar "dual-frame" design was used for The Robert Wood Johnson Foundation (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.

Restricting the field sample to the 12 high-intensity sites reduced some of the coverage bias that would result from using an RDD-only methodology, both for estimates about all large metropolitan sites and estimates made for those sites. This option also 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 persons who do not have telephones and live in non-metropolitan areas and metropolitan areas with populations of under 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).

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). Respondents within eligible households were then interviewed over cellular telephones, which were provided by MPR field staff. Thus, all interviews were conducted by CATI, avoiding differences in response by interviewing mode.

In implementing this general strategy, we:

- Defined eligibility for the field sample
- Determined sample allocation among the 12 sites
- Identified areas within the 12 sites for exclusion
- Established a measure of size for selecting clusters
- Stratified clusters by county and by tract number within county
- Selected clusters and listing areas
- Listed addresses
- Released sample for screening

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.[9]

We originally planned to use the Census definition as a screening criterion, and interview only households that did not have working telephones when first contacted by a field interviewer. Based on experience in the RWJF Family Health Insurance Survey, and on research reported by Brick et al. (1995), we concluded that this static approach to defining telephone status produced 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. Therefore, the field sample included any households with a history of significant interruption in service since July 1, 1996, the beginning of interviewing for the RDD sample.

We defined *significant interruption* to mean two weeks or more of interrupted service since July 1, 1996 (or the date the household moved in if that occurred after July 1) and used questions about the length of interruptions to adjust sample weights.[10] 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 last two weeks prior to the interview and had been without a telephone since moving in.

[9]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?"

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

2. Determining Sample Allocation

We observed substantial differences in telephone penetration among the 12 high-intensity sites.[11] Census data show that the percentage of households without telephones in 1990 in those sites ranged from 1.5 (Orange County, California) to 8.1 (Greenville, South Carolina) (Table II.5). Because the chance that the Census would classify a household as nontelephone is proportional to the length of telephone service interruption, we assumed that estimates using Census definitions corresponded closely to our eligibility criteria. Based on this assumption, we looked at the distribution of households with and without telephone service in each site according to 1990 Census data.

The optimal allocation of the field sample among the 12 high-intensity sites would be proportional to the percentage of each site's households that was eligible for the field survey. Similarly, within a site, the optimal allocation of field interviews would be proportional to the percentage of nontelephone households within each site.[12] For example, a site in which eight percent of households was eligible for the field survey would have a greater share of its field and RDD interviews allocated to the field sample stratum than would a site in which only four percent was eligible for the field sample. These estimates led to the preliminary field allocation shown in Table II.5.

---

[11]We use the term "penetration" rather than "coverage" in referring to the percentage of the population with or without telephones because we use the term "coverage" to refer to the percentage of the study population that is covered by the sample frame.

[12]Assuming equal population variances for each stratum (telephone and nontelephone), the optimal allocation would be proportional to the relative sizes of the populations of the strata, proportional to the square root of the design effect of each method (RDD and field telephone), and inversely proportional to the square root of the cost per completed interview in each stratum.

TABLE II.5

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

Telephone Penetration	Percentage Without Telephone[1]	Nontelephone Households[1]	Preliminary Field Allocation[2]
High Penetration			
Boston (MA)	1.9	30,456	3.6
Orange County (CA)	1.5	12,808	3.0
Seattle (WA)	2.0	15,298	3.8
Medium-High Penetration			
Cleveland (OH)	3.7	32,107	7.1
Lansing (MI)	3.2	5,078	6.3
Newark (NJ)	3.9	27,085	7.6
Syracuse (NY)	4.0	10,866	7.8
Medium-Low Penetration			
Indianapolis (IN)	5.0	26,340	9.7
Miami (FL)	5.0	34,652	9.7
Low Penetration			
Greenville (SC)	8.1	25,339	15.8
Little Rock (AR)	7.0	13,728	13.5
Phoenix (AZ)	6.2	52,656	12.0

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

[2] Percentage of all expected FIU interviews.

However, this allocation did not account for between-site variation in the cost of a field telephone interview. Thus, we considered the issues of cost and coverage of the target population for the field survey. If the ratio of the cost of a field interview compared with the cost of an RDD sample interview were the same for all 12 high-intensity sites, then the optimal allocation for the field sample would vary from about 5.4 percent of all interviews in Greenville to about 0.8 percent of all interviews in Orange County. This allocation assumed that the design effect for the field sample was roughly three times that for the RDD sample, and that the cost per case in the field component was 10 times the cost per case of the RDD sample.

3. Identifying Areas for Exclusion

Cost is related to coverage. Because screening for households eligible for a field interview is very expensive, we originally proposed excluding any Census Block Group if 1990 estimates showed that fewer than five percent of households in that Block Group lacked telephones (the "five percent rule"). As Table II.6 shows, the five percent rule would have resulted in frame coverage of nontelephone households ranging from 50 percent (Orange County) to 90 percent (Greenville).[13] Because coverage was so low in many sites, we considered alternative rules based on various assumptions concerning Block Group exclusions.[14] A zero percent rule would have resulted in close to 100 percent coverage if the population had not changed since 1990 but would obviously have

---

[13]The actual coverage would be somewhat different because of (1) time elapsed since 1990 and, (2) the difference in the definition of nontelephone household used in the Census and in our eligibility criterion. Nonetheless, we expected a high correlation between the prevalence of Census-defined nontelephone households and eligibility for the CTS field component.

[14]Under the include-all rule, we would not have excluded any Block Groups, and under the zero percent rule, we would have excluded Block Groups that had no estimated nontelephone households in 1990. Under the one, two, three, and four percent rules, we would have excluded Block Groups having percentages without telephones of less than one, two, three, and four percent, respectively.

TABLE II.6

FRAME COVERAGE UNDER VARIOUS FIELD SAMPLE EXCLUSION RULES[1]
(In Percentages)

Telephone Penetration	Include All	Zero Percent Rule[2]	One Percent Rule	Two Percent Rule	Three Percent Rule	Four Percent Rule	Five Percent Rule
High Penetration							
Boston (MA)	100	98.9	97.9	89.2	78.9	71.3	61.0
Orange County (CA)	100	98.2	96.6	84.2	71.0	60.3	50.3
Seattle (WA)	100	98.6	97.3	86.3	75.2	63.8	54.9
Medium-High Penetration							
Cleveland (OH)	100	99.3	98.7	94.8	89.7	84.5	79.9
Lansing (MI)	100	98.7	97.5	90.3	83.6	76.5	68.5
Newark (NJ)	100	99.6	99.4	96.4	93.0	89.4	86.8
Syracuse (NY)	100	99.5	99.0	95.2	88.3	81.8	74.8
Medium-Low Penetration							
Indianapolis (IN)	100	99.5	99.0	95.1	91.1	85.8	80.7
Miami (FL)	100	99.1	98.2	93.8	87.7	83.3	78.7
Low Penetration							
Greenville (SC)	100	99.9	99.9	98.7	96.3	93.2	89.7
Little Rock (AR)	100	99.8	99.8	97.7	94.9	91.9	88.9
Phoenix (AZ)	100	99.7	99.6	97.5	94.5	91.4	88.1
Unweighted Average[3]	**100**	**99.2**	**98.6**	**93.3**	**87.0**	**80.7**	**75.2**
Weighted Average[3]	**100**	**99.5**	**99.1**	**9.51**	**90.8**	**80.6**	**81.2**

[1]Under the include-all rule, we would not exclude any Block Groups. Under the zero percent rule, we would exclude Block Groups that did not have any estimated nontelephone households in 1990. Under the one, two, three, and four percent rules, we would exclude Block Groups that had, respectively, less than one, two, three, and four percent without telephones.

[2]Assumed to be halfway between the include-all and the one percent rule.

[3]The unweighted average is the arithmetic mean. The weighted average is weighted by the proportion of the total 1990 nontelephone population “represented” by each site.

increased costs. Generally, by sacrificing coverage, we could have increased the efficiency of screening and reduced costs. The gains from excluding Block Groups with high telephone penetration were greatest in the sites with high telephone penetration. We derived approximate optimal allocations under each of the Block Group exclusion rules, with relatively little variation across sites.

We adopted an expanded optimal allocation strategy that weights cost, sampling error, and potential coverage bias in determining coverage rules for various sites.[15] This "mixed-rule" strategy suggested that the most efficient approach would be the five percent rule for the low- and medium-low penetration sites, a four percent or five percent rule for the medium-high penetration sites, and a three percent or four percent rule for the high-penetration sites. However, for national estimates, a two percent rule is optimal, because the sample size is larger, and the bias of undercoverage relatively more important. Site-based estimates were determined to be more important than national estimates, so we changed the exclusion rule only in cases in which it would substantially increase coverage. The field allocation model selected for the 12 high-intensity sites is shown in Table II.7.

4. Establishing a Measure of Size for Cluster Selection

We set the overall number of clusters at 72. We use the generic term "cluster," rather than primary sampling unit or secondary sampling unit, because the clusters selected within sites are secondary sample units for national estimates and primary sampling units for site-specific estimates for each of the 12 sites.

---

[15]In this case, we would try to minimize the product:

$$\mathit{Cost} \times (\mathit{Variance} + \mathit{Bias}^2).$$

TABLE II.7

FINAL DESIGN AND ALLOCATION FOR THE FIELD SAMPLE:
A MIXED-RULE STRATEGY

Telephone Penetration	Rule	Coverage (Percent)	Hit Rate (Percent)	Allocation[1] (Percent)
High Penetration				
Boston (MA)	Three percent	78.9	7.5	3.3
Orange County (CA)	Two percent	84.2	5.3	2.4
Seattle (WA)	Two percent	86.3	5.5	3.3
Medium-High Penetration				
Cleveland (OH)	Five percent	79.9	13.3	7.5
Lansing (MI)	Three percent	83.6	8.1	5.7
Newark (NJ)	Five percent	86.8	13.3	7.8
Syracuse (NY)	Three percent	88.3	11.4	7.8
Medium-Low Penetration				
Indianapolis (IN)	Four percent	85.8	11.0	9.5
Miami (FL)	Five percent	78.7	14.2	10.0
Low Penetration				
Greenville (SC)	Five percent	89.7	15.3	16.7
Little Rock (AR)	Five percent	88.9	13.0	13.8
Phoenix (AZ)	Five percent	88.1	11.5	12.0
Unweighted Average[2]	**--**	**84.9**	**10.7**	**--**
Weighted Average[2]	**--**	**88.3**	**11.9**	**--**
Total	**--**	**—**	**—**	**100**

[1]Percentage of all completed FIU interviews.

[2]The unweighted average is the arithmetic mean. The weighted average is weighted by the proportion of the total 1990 nontelephone population “represented” by each site.

The number of clusters was determined by setting a minimum of two clusters for any site; subject to this minimum, we allocated one cluster per eight field household interviews.

In sampling, we wished to avoid selecting clusters that had too few eligible households to complete an interviewing assignment; thus, clusters were defined as Block Groups or groups of Block Groups having a minimum measure of size based on estimated nontelephone households in the 1990 Census. We used 1990 Census data to select clusters for the nontelephone sample for the RWJF Family Health Insurance Survey, and found that Census estimates of prevalence were reasonable predictors of nontelephone status in 1994 (Hall et al. 1994). As discussed, we assumed that the Census definition of nontelephone coverage would correlate with our eligibility criteria. Furthermore, the 1990 Census was the only available frame. If a Block Group had less than the minimum measure of size, it was linked with one or more other Block Groups--if possible, with a Block Group in the same tract. In addition, preference was given to linking small Block Groups with large ones, although sometimes several small Block Groups had to be grouped. When there were not enough eligible Block Groups in a tract, Block Groups with less than the minimum measure of size were linked with a Block Group in a nearby Census tract.

5. Stratification of Clusters

We considered a number of explicit criteria for stratification, such as race/ethnicity and income distributions. However, none of these factors varied sufficiently within the sites to permit use of this method. In addition, the most recent data for stratification at the tract or Block Group level would have had to be based on 1990 Census data, which would have become unreliable by 1996. Therefore, we used implicit stratification, sorting clusters by county, and by tract number within county, to ensure geographic dispersion of the sample.

6. Selecting Clusters

We used probability-proportional-to-size selection methods to select clusters. To take advantage of the implicit strata, we used systematic selection after a random start. Each sampled cluster was then 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.

7. Selecting Listing Areas

An initial 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.[16] "Listing areas" were identified by selecting a random number, n_r, between 1 and 10. The n_rth replicate was chosen for listing. If the allocated number of replicates, r, was greater than 1, then we also released the $r - 1$ replicates after releasing the n_rth replicate on the list. The list was treated as circular; thus if replicate 1 was not the first selection, for any subsequent releases, it was treated as if it followed replicate 10.

---

[16]Because each cluster was selected with probability proportional to size, a cluster's probability of selection, *P(cluster)*, was proportional to its measure of size. To minimize costs, we wished to avoid (whenever possible) subsampling listed housing units; thus, we assumed that the final-stage probability would be 1.0. Given that r is the number of replicates chosen for listing, we chose r so that the project overall probability, P_{cum}, where:

$$P_{cum} = P(cluster) \cdot \frac{r}{10} \cdot 1.0,$$

would be roughly equal within each site.

8. Listing Addresses

Before listing addresses, the lister would perform a rough count of the total number of housing units in the listing area. In the majority of cases, the lister was instructed to list all housing units identified on each Census block in the assigned listing area. In some cases, before listing, we subdivided listing areas into smaller areas ("chunking"). Two types of listing areas were chunked:

1. Those containing a large block and therefore including more replicates than were allocated

2. Those in which the lister's rough count of housing units differed substantially from the expected number, because of either growth in the number of units since 1990 or discrepancies with the Census data

In the first case, we always chunked; in the second case, we chunked only if the discrepancy was large (that is, the lister reported at least twice the number of housing units shown in the Census data) and if omitted units would be expected to contain few, if any, nontelephone households. If the discrepancy was the result of new construction other than low-income housing, or the result of the exclusion of obviously high-income areas in 1990, we excluded the listing areas from the field survey; it is unlikely that these would contain eligible households. To make this determination, we considered the approximate age of construction, whether the units were designated as low- or moderate-income housing, the type of construction (that is, single-family, apartments, or mobile), and, if possible, a current estimate of average value or rent.

Chunking consisted of dividing a listing area into subareas of roughly equal size, each of which had an identifiable starting and ending address. If we chunked to reduce the listing area to *r* replicates, we would form a number of chunks equal to the total number of replicates, *R*, contained in the listing area and then randomly select *r* chunks from the total number to list. If we chunked to resolve discrepancies between the lister's count and the Census data, the lister would be instructed to form up to five chunks, from which we would randomly select a sufficient number to meet our sampling targets for that area.

As part of listing, we attempted to obtain information on listed dwelling units that could be used in sampling or that the interviewer might have found helpful. Slightly different procedures were used depending on whether housing units were listed as part of the initial effort or as part of a supplemental effort; supplemental listing was required in a few sites because the initial listing yielded too few eligible households. During the initial phase of listing, listers attempted to determine whether a listed housing unit (1) was vacant, (2) had a working telephone at the time of contact, (3) was occupied but had no telephone, or (4) had a status that could not be determined. During supplemental listing, listers classified units as vacant but used the same screener as used in interviewing to determine eligibility. In some instances, an owner or manager of a building or condominium development would not allow access to areas or buildings. We therefore gathered as much information as possible by calling the building owners or managers and by screening samples of residents, whose telephone numbers were obtained using reverse directories. The information collected was used to classify units in inaccessible buildings as (1) ineligible based on the screener, (2) eligible based on the screener, or (3) having a telephone and presumably ineligible.

If this determination could not be made, we randomly assigned eligibility status to units in proportion to the distribution of eligible units in the building.

9. Sample Release

Our procedures for releasing sample varied slightly depending on whether a unit was listed during the initial effort or during the supplemental effort. For the first release of sample (all from initial listings), we released units classified as telephone households at half the rate of other units. Housing units with working telephones were sampled because the criterion for their classification (presence or absence of a working telephone at listing) was not identical to that used in screening (presence or absence of a two-week interruption in service). Indeed, some telephone units were screened and found eligible. If subsequent releases from initial listings were needed, we released all listed units.

During supplemental listings, households that could be contacted were classified as eligible or ineligible on the basis of the screener. Thus, households classified as ineligible during supplemental listing were not released for screening. Field sample releases are shown in Table II.4.

III. SURVEY DESIGN AND PREPARATION

A. OVERVIEW

The household survey instrument was the primary data collection vehicle for assessing the effects of health system change on individuals, including assessing changes in health insurance coverage, access to care, use of health services, and satisfaction with health care. As described in Chapter II, the family insurance unit (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 one randomly selected child. An adult familiar with the health care experiences of other FIU members 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, tobacco use, details about the last physician visit, level of satisfaction with that visit, and physician trust. The adult who took the randomly selected child to the physician during the last visit before the survey was asked to answer questions about that visit. The length of the interview varied depending on the number of individuals in the FIU and the complexity of the individuals' experiences with health care. Including self-response modules, the interview required an average of 34.4 minutes to complete.

The instrument development process included a review of related instruments and methodological studies, as well as consultation with experts. We also conducted cognitive interviewing to evaluate respondent comprehension of selected questions, and a pretest to assess interview length and the computer-assisted interviewing program.

Prior to fielding the survey, we attempted to obtain endorsement or support from national and state government agencies, contracted with a communications consultant to test and develop messages about the survey's purpose and value to the public, and developed a brochure and letters to be mailed to sampled households with known addresses. We obtained statements of support (although not formal endorsements) from 26 state and local health agencies (see Appendix C), but we were not able to obtain them from national government agencies. The endorsement cited in the survey introduction stated that the survey was "supported by state health departments throughout the country, including [AGENCY NAME]." In addition, efforts to develop a convincing survey introduction built around health system tracking were less successful than we had expected. Letters were mailed to sampled households with published addresses, but cooperation rates in these households were no higher than those in households that did not receive letters. The most successful efforts to increase participation in the survey were interviewer persistence, follow-up efforts to convert refusals, and monetary incentives.

Altogether, 302 telephone interviewers and 26 field listers and screening 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 not trained to administer the survey instrument, although they were briefed during two 2-hour conference calls on the listing procedures and screening procedures that Mathematica Policy Research, Inc. (MPR) uses. Training field staff in survey instrument administration was not necessary because, after identifying eligible households, these staff used cellular telephones to call the MPR telephone center, where an interviewer conducted the survey.

In this chapter, we will discuss the design of the survey instrument; preparation for the survey, including development of the survey introduction, efforts to obtain endorsements from government organizations and design of advance materials; interviewer selection and training; and the computer-assisted telephone interviewing (CATI) system used to collect the data.

B. INSTRUMENTATION

The survey instrument was developed by staff at the Center for Studying Health System Change (HSC) and MPR, with consultation and review by several experts. In the following sections, we describe key decisions on interviewing strategy, a summary of questionnaire items, instrument design decisions, and modifications to obtain information from households in the field sample. An English version of the survey instrument is included in Appendix A (Center for Studying Health System Change 1997b). English and Spanish versions of the CATI program, as well as a modified CATI program used to conduct cellular telephone interviews with households selected from the field sample, are available on request from MPR.

1. Interviewing Strategy

We made three basic decisions concerning interviewing strategies:

1. We decided to include only questions that would be asked of all sampled individuals in the household, rather than a core set of general questions to be asked of all individuals and a longer battery of questions that would be asked only of a randomly selected individual in the FIU.

2. We opted for a mix of proxy and self-response questions to achieve a balance between data quality and interview length.

3. We anchored questions on satisfaction, access to care, and resource use to events and time periods that would allow for estimates that are representative of people's experiences.

The basis for these decisions is described in the following sections.

a. Whether to Subsample Randomly Selected Respondents

We had two reasons for deciding to include only questions that would be asked of all individuals in a household. First, any analysis that used the extended battery of questions on a randomly selected individual would require subsetting the sample to only those individuals who were asked the measures, which would result in a loss of precision. This loss of precision would have been especially problematic for health status measures, as they will be used in most analyses as independent variables. Second, many of the questions in the extended battery would have required self-response by the selected individuals, thereby significantly increasing the length of the self-response module, and resulting in additional nonresponse and higher costs.

b. Self-Response Versus Proxy Response

Most of the questions in the instrument were designed to obtain information about specific individuals in the FIU. Ideally, it would have been desirable if each adult in the FIU had responded to all person-level questions. However, that approach would have entailed gaining cooperation from the other FIU members to answer an extensive battery of questions, which would have undoubtedly resulted in additional nonresponse for these questions and higher costs. Therefore, we chose to maximize the number of items that proxy respondents could answer reliably. By reviewing the methodological literature, we determined that proxy responses on questions about demographics, employment, health insurance, usual source of care, and resource use would be of acceptable quality.

Furthermore, the use of proxy respondents in these cases may be somewhat less problematic for the Community Tracking Study (CTS) Household Survey than for other household surveys that use as the interviewing unit the Census definition of a family (all persons in a household related by blood or marriage). Because the CTS Household Survey definition of FIU is based on the insurance unit, proxy respondents were almost always spouses or parents, which helped to minimize the amount of reporting error and bias associated with proxy responses given for persons about whom the informant has little personal knowledge. The family informant answered most questions (on health status, resource care, usual source of care, and insurance coverage) about the randomly sampled child. However, if another adult had taken the child to the physician on his or her last visit before the survey was administered, that person was asked to respond to questions about the visit.

A subset of items, including subjective assessments of health, tobacco use, satisfaction with care, and aspects of the physician–patient interaction, were answered directly by each adult FIU member. The self-response questions, which averaged five to ten minutes per adult, were obtained from other FIU members after the core interview was completed or a follow-up call was scheduled, if the other adult(s) were unavailable. In a few cases (for example, when spouses were on business trips or when adult children were temporarily unavailable or away at college), we relaxed the self-response requirement and obtained proxy responses. (These cases are identified in the data files.)

c. Anchoring Questions on Satisfaction, Access, and Resource Use

For questions on access to care, satisfaction, and resource use, it was important to determine the appropriate point of reference. For example, questions on satisfaction may refer to satisfaction with overall health care (that is, satisfaction "averaged" over all visits and providers), with a particular provider (for example, the usual source of care provider), or with a particular encounter or episode of care (for example, the last visit). Because our primary analytical objective was to obtain measures enabling us to "track" change, it was important to have points of reference that would allow for estimates that are representative of people's experiences.

Some items referred to "the previous 12 months," rather than to particular physicians or medical encounters, which may not be representative of experiences with the health care system. A 12-month reference period was particularly well-suited to global satisfaction questions (for example, on overall health care), questions on barriers to care, and questions on health insurance preferences.

We felt that some questions, including satisfaction questions relating to quality of care and patient–provider interactions, were difficult to "average" over all encounters occurring during the previous 12 months. In addition, questions about the process of seeking care, such as travel time, appointment and office waiting times, and reasons for seeking care, should be anchored to either particular providers or particular encounters. It is easier for respondents to focus on a single encounter than to make summary judgments across different visits and providers (Research Triangle Institute 1995). These types of questions were anchored to the individual's last physician visit. We made this decision on the assumption that estimates based on a person's last visit and then averaged over all persons in the sample would be more representative of individual's health care system experiences than would estimates based on the usual source of care. The usual source of care may be biased because it represents the "modal" use but not necessarily the "average" use, and because the probability of having a usual source of care as well as usual site of care may be dependent on characteristics of the health system.

The use of the last visit as a point of reference has some limitations. A study conducted at the National Center for Health Statistics (NCHS) indicated that estimates based on the "last visit" tend to underrepresent "sick visits," as opposed to well or preventive visits (Makuc et al. 1994). However, time constraints and problems with respondent recall make it impractical to ask these questions about all visits occurring during the previous 12 months or about a randomly selected visit occurring during that period. Using a shorter time frame to ask about all visits, such as the previous month, would have dramatically reduced the proportion of the sample that had an encounter within that time frame. To address the issue of the underrepresentation of "sick visits," we developed a sequence of questions to obtain the dates of both the last sick visit and the last preventive visit; whichever of the two visits occurred last served as the point of reference for additional questions about satisfaction and process of care.

2. Questionnaire Items

Table III.1 summarizes the content of the survey. Questions from several survey instruments were used to develop the CTS household survey, including:

- The 1995 Current Population Survey
- The 1993-1994 Robert Wood Johnson Foundation (RWJF) Family Health Insurance Survey
- The 1987 National Medical Expenditure Survey
- The pretest version of the 1996 Medical Expenditure Panel Survey
- The 1996 National Health Insurance Survey (NHIS) Redesign
- The 1994 RWJF Access to Care Survey
- The 1996 Behavioral Risk Factor Surveillance Survey
- The National Maternal and Infant Health Care Survey
- The Agency for Health Care Policy and Research (AHCPR) Consumer Survey (draft version available in 1996)
- The SF12™ Health Survey, Standard U.S. Version 1.0, 1994

TABLE III.1

CONTENT OF THE HOUSEHOLD SURVEY

Health Insurance (Questionnaire Section B)	

Private insurance coverage	Covered by employer- or union-related private insurance
	Covered by other private insurance
	    Directly purchased
	    Premium for directly purchased private insurance
	    Provided by someone not in household

Public insurance coverage	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	Not covered by public or private insurance

Continuity of coverage/changes in coverage	Currently insured; lost coverage in past 12 months
	Currently uninsured; gained coverage in past 12 months
	Uninsured during all of past 12 months
	Uninsured at some point in past 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
	    Gained or lost coverage

Denial of coverage	Ever denied insurance coverage in past two years because of poor health
Insurance plan attributes	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	Ever enrolled in an HMO Number of total years enrolled in an HMO

TABLE III.1 *(continued)*

Access to Health Care (Questionnaire Sections C-E)	

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
	    Reasons for not having a usual source of care

Travel/waiting time for physician visit (Section E)	Lag time between making appointment and seeing physician at last physician visit* Travel time to physician's office at last visit* Time spent in waiting room before seeing medical person at last physician visit*
Difficulty getting needed services in past year (Section C)	Did not get needed services* Delayed getting needed services* Reasons for delaying or not getting needed services*
Perceived changes in access (Section C)	Getting needed medical care is easier/harder compared with three years ago
Resource Use (Questionnaire Sections C and E)	
Use of ambulatory services in past 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 there were any mental health visits Whether there were any home health visits Number of surgical procedures
Use of inpatient services in past 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 service use (Section C)	Whether person has had flu shot in past 12 months Whether person has 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*
	    Checkup, physical exam, other preventive care*
	Type of physician seen at last visit (primary care physician or specialist)*
	Whether last visit was to usual source of care*
	Whether last visit was to an emergency room*
	Whether last visit was with appointment or as walk-in*

Costs (Section C)	Total family out-of-pocket expenses for health care in past 12 months

TABLE III.1 *(continued)*

Satisfaction and Patient Trust (Questionnaire Sections D and E)	
General satisfaction (Section E)	Overall satisfaction with health care received by family Satisfaction with choice of primary care physicians* Satisfaction with choice of specialists*
Satisfaction with last doctor visit (Section E)	Satisfaction with thoroughness and carefulness of exam* Satisfaction with how well physician listened* Satisfaction with how well physician explained things*
Patient's trust in physicians (Section D)	Agree/disagree that physician may not refer to specialist when needed* Agree/disagree that physician may perform unnecessary tests or procedures* Agree/disagree that physician is influenced by health insurance company rules* Agree/disagree that physician puts patient's medical needs above all other considerations*
Employment and Earnings (Questionnaire Sections B and F)	
Employment status and characteristics (Section F)	Whether adult respondent has the following characteristics: Owned a business or farm Worked for pay or profit in past week Had more than one job or business Worked for private company/government/family/self-employed business Average hours worked per week, at primary job and at other jobs Size of firm (number employees), at site where respondent works; 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

TABLE III.1 *(continued)*

Other Variables (Questionnaire Sections A, B, E, and G)	
Demographics (Section A)	Age Gender Highest education level completed
Health status (Section E)	Overall health status (five-point scale from excellent to poor)* Limited in moderate activity* Limited in climbing stairs* Accomplished less because of physical health* Limited in kind of work because of physical health* Accomplished less because of emotional health* Less careful in work because of emotional health* Pain interfered with work* How much time health problems have interfered with social activities* How much time calm and peaceful* How much time have energy* How much time downhearted/blue* SF12 scores: Physical Component Summary; Mental Component Summary*[a]
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*
Risk behaviors (Section E)	Whether person agrees that he/she is more likely to take risks than the average person*
Smoking cessation interventions (Section E)	Whether person has smoked at least 100 cigarettes in lifetime* Whether currently smoking cigarettes every day, some days, or not at all* Average number of cigarettes smoked per day in past 30 days* How long since quit smoking* Whether stopped smoking one day or longer in past 12 months, in effort to quit* Whether physician advised smoker to stop smoking in past 12 months*

*Denotes information obtained from the Self-Response Module.

[a]For English-speaking respondents, we used the interviewer-administered version of the SF12 Health Survey, Standard US Version 1.0, 1994, The Health Institute, New England Medical Center. We also obtained an interviewer-administered version of the U.S.-Spanish SF12, which was slightly modified by project staff; modifications were reviewed and approved by New England Medical Center staff.

Many experts also were consulted about potential items that could be included in particular sections of the survey[1]. Note that the experts did not necessarily endorse the items that were eventually selected.

3. Instrument Design Decisions

The major analytical objectives and design decisions for key sections of the instrument are summarized in this section.

a. Health Insurance

Measures of health insurance coverage will be used to track aspects of health insurance (for example, the percentage uninsured in each site and percentage in managed care plans in each site) and as independent variables in analyses of satisfaction, access to care, and resource use. These measures include rates of current coverage from employment-related private insurance, other private insurance, Medicaid, Medicare, and other public sources, as well as the percentage uninsured.

Attributes of health insurance coverage, including coverage from health maintenance organizations (both public and private), whether the plan requires primary care gatekeepers, and some information on restrictions in choice of physicians, were obtained. Many individuals are unaware of or do not understand these aspects of their coverage; therefore, we were careful to select attributes for which empirical evidence shows that individuals can answer questions with a reasonable degree of accuracy. Selected questions had been pretested for the 1996 Medical

---

[1]We wish to express our appreciation to the following researchers who provided guidance on item selection: Jessica Banthin (AHCPR), David Blumenthal (Massachusetts General Hospital, Paul Cleary (Harvard University), Gary Collins (NCHS), Anne Hardy (NCHS), Doris Lefkowitz (AHCPR), Jeanne McGee (McGee and Evers Consulting), Diane McKuc (NCHS), David Mechanic (Rutgers University), Alan Monheit (AHCPR), P. Ellen Parsons (NCHS), Barbara Schone (AHCPR), Barbara Starfield (Johns Hopkins University), Anne Stratton (NCHS), and John Ware (New England Medical Center).

Expenditure Panel Survey (formerly, National Medical Expenditure Survey) conducted by the AHCPR.[2]

Many of the questions on satisfaction, access to care, and resource use have reference periods that covered the 12 months preceding the survey. Thus, it was necessary to determine whether there were any changes in individuals' health insurance coverage and time spent uninsured during that period, as at least some interactions with the health care system likely occurred prior to the period of current coverage. Although the survey provides less detail on previously held coverage than on current coverage, having some information on all health care coverage during the 12 months preceding the survey will permit researchers to address the problem analytically (for example, by controlling for the proportion of the year during which an individual had current coverage).

It will also be important to assess whether differences in health care utilization that vary by type of health insurance plan are a result of actual plan differences or of individuals with different health care needs self-selecting into the different types of plans (that is, selectivity bias). Measures that allow for the modeling of choice of health insurance plans are typically used to control for selectivity bias. For this reason, we obtained information on health insurance options at the individual's place of employment, as well as on general preferences for types of health insurance.

b. Access to Care

Access-to-care measures in population-based surveys typically include insurance coverage, health care utilization, and usual source of care. However, none of these measures ascertain whether individuals actually confront obstacles in obtaining health care. We considered several different types of measures that

[2]We recently completed a "followback" survey to insurers, third-party administrators, and self-insured employers to (1) obtain other details about survey respondents' coverage, including risk-sharing with providers, characteristics of the provider network, and other topics; and (2) validate attributes reported by the family informant.

would allow us to assess this aspect of access. Although some satisfaction questions are related to access (for example, satisfaction with travel times and appointment waiting times), these questions are subjective, and dissatisfaction with care does not necessarily equate to lack of access. We also were concerned about possible differences in interpretation and response to these items according to socioeconomic status.

For these reasons, we decided to adopt the approach used by NCHS in its redesign of the NHIS. This approach determines whether individuals delayed care for financial or nonfinancial reasons (for example, could not get an appointment soon enough, availability of providers, convenience of office hours). We found this method appealing because it determines whether individuals perceived obstacles and barriers to obtaining health care.

c. Usual Source of Care

Whether an individual has a usual source of care--as well as the type of place or provider that individuals report as their usual source of care--is a traditional measure of access to care. However, studies show that the usual source of care is limited as a measure of access because most individuals who do not have a usual source of care indicate that they “do not need” or “do not use medical care” (Hayward et al. 1991).

Although limited as an access measure, usual source of care is still useful for tracking because it allows for some understanding of the care arrangements individuals select when they need health care. Through the series of questions on usual source of care, we can determine whether individuals identify a hospital emergency department as their source of care, whether they usually see physicians or nonphysicians, whether they have a particular physician, and whether they changed their usual- source-of-care place/provider during the year preceding the survey.

d. Satisfaction with Health Care

Many consumer surveys contain extensive sets of questions on satisfaction with various aspects of health care. However, many of these items overlap with the "objective" measures that are included in the sequence of questions on access to care and last visit. In choosing between an "objective" and a "satisfaction"-type question (for example, actual travel time versus satisfaction with travel time), we selected objective measures. Differences in the degree of satisfaction or dissatisfaction with various aspects of health care may reflect different expectations or preferences, rather than characteristics of the health care system. In addition, most questions on satisfaction require self-response, which would have increased the burden on additional household members and, potentially, the level of nonresponse for these items.

Nevertheless, we wanted to include the most salient items on satisfaction, as well as items that had no "objective" counterparts. These items included satisfaction with overall health care, choice of providers, thoroughness of the examination received at the last visit, and provider–patient communication.

e. Resource Use and Process of Care

Tracking changes in the use of health services is an important goal of the study. However, the amount of detail that can be obtained on the type and quantity of services used is limited by the ability of respondents to accurately recall this information. Individuals at NCHS who were involved in the redesign of the NHIS were consulted about methodological issues, and we reviewed the results of several studies that examined the effect of length of recall on the amount of reporting error for questions about specific types of services. These studies indicated that the accuracy of recall of the characteristics of health care encounters decays continuously after one month (Jay et al. 1995).

For these reasons, we followed the approach used by NCHS in its redesign of the NHIS. We limited items based on 12-month recall to general measures of health services use (physician visits, hospital stays, surgical procedures), as well as to certain specific preventive measures, such as mammography screening for women and flu shots for elderly people.

Given the time constraints and recall problems we have discussed, it was not feasible to obtain detailed information on all medical encounters (for example, the reasons people seek health care, the types of providers they see, the "mode of entry" into the system, and the travel time to providers). These details were obtained only for the last physician visit during the 12 months prior to the survey. Aggregating the "last visit" measures across all individuals in the community will allow for a representative analysis of how health care utilization patterns are changing at the community level.

f. Patients' Trust in Their Physicians

Questions on patients' trust in their physicians was another area that required subjective evaluation by respondents. This topic is an important tracking issue because changes in the health care system have the potential to disrupt the traditional patient–provider relationship and patients' perceptions of their physicians, both of which could affect the quality of care. We consulted Dr. David Blumenthal and Dr. Paul Cleary, of Harvard University, and obtained questions developed by Dr. Cleary to measure public opinions on these issues. These opinions included perceptions of whether the respondent's physician refers the patient to specialists when needed, whether he or she performs unnecessary tests and procedures, whether he or she is influenced by health insurance company rules, and whether he or she places the patient's medical needs above all other considerations.

g. Health Status and Tobacco Use

Health status measures will be used to meet three analytical goals: (1) as independent variables in modeling choice of health plans (important for analyses of risk selection); (2) as independent variables in analyses of access to care, resource use, and satisfaction; and (3) to identify individuals with special needs who might be sampled for various studies.[3] These considerations drove the decision to use the SF12™, developed by John Ware and colleagues at the New England Medical Center Health Institute. The SF12 has wide acceptance as a measure of health status, and considerable empirical work supports the validity, reliability, and precision of this instrument. Furthermore, guidelines have been developed for interpreting the meaning of differences and changes in health status scores based on the SF12. The SF12 is also strongly associated with health services use, which is important for our second analytical objective.

We did not include an extensive list of risk behaviors, as many are not directly influenced by the health care system. We included questions on smoking behavior because this behavior is considered a serious health risk that can be influenced by health care providers.

4. Cognitive Interviewing

Cognitive interviewing was used to assess respondents' understanding of several questions.[4] These questions, with assessments of understanding shown in parentheses, included the following:

- Whether questions on managed care, designed for private insurance plans, made sense for persons covered by managed care plans under Medicare, Medicaid, and military plans (required some modifications in question wording)

---

[3]For example, scores on the mental health component of the SF12 are being used to sample people for a collateral study that researchers at UCLA and RAND are conducting on persons at risk for alcohol abuse, drug abuse, and mental health conditions.

[4]A full report on the results of cognitive interviewing is available from MPR.

- For those uninsured at the time of the interview, whether respondents could recall the type of insurance coverage they had had prior to losing their insurance coverage (yes, if bounded by 12-month recall)

- Whether respondents understood questions and answer choices for questions on unmet need (yes)

- Whether respondents could respond to subjective questions for others in the FIU on changes in access to health care and overall satisfaction with health care (yes)

- Whether respondents understood the distinction between well-patient and sick-patient visits (required some changes in question wording)

- Whether family informants could answer questions about spouses' insurance plans[5] (spouses in dual-policy FIUs were able to respond to questions about each other's insurance plans)

5. Instrument Modifications to Obtain Information from Nontelephone Households

Most of the CTS Household Survey interviews were obtained from the random-digit-dialing (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 B). Field interviewers then gave cellular telephones to the family respondents, to be used to call the MPR telephone center. Thus, interviews could 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.

[5]Other relatives and nonrelatives were assigned to separate FIUs.

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 a well-known public organization (usually a government agency), and by designing a convincing survey introduction describing the survey's purpose and value. In other RDD health surveys, we had successfully used advance letters, salient survey messages, and endorsements by senior government officials 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. The purpose of that survey was to provide data needed by states planning health care reform.

For the CTS Household Survey, we were concerned that health tracking might be less salient to respondents than health care reform, which dominated public debate when the Family Health Insurance Surveys were 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 (these organizations were sponsors for the Family Health Insurance Survey). Consequently, we took several steps to increase the survey's perceived impact and value, including requesting government endorsements and contracting with a communications consultant to develop and test messages for inclusion in survey introductions.

1. Endorsements

Our initial efforts focused on obtaining national endorsements from the National Governors' Association and the U.S. Public Health Service. Unfortunately, these organizations indicated that it was against their policies to endorse a survey they were not sponsoring. Next, we contacted the Association of State and Territorial Health Organizations (ASTHO), which sent a letter to state health directors describing the survey; however, ASTHO could not endorse it, either. We followed up by mailing background material about the study to state health directors in all states containing sampled communities. Officials in 11 states and two local communities provided letters of endorsement, and 26 health officials were willing to allow us to reference their support for the survey in survey introductions and advance letters on RWJF letterhead. The list of officials endorsing the survey is included in Appendix C.

2. Advance Letters and Survey Brochures

Working with a communications consultant and staff at RWJF, we developed a standard format for letters mailed under state or county health department or RWJF letterhead (Appendix D). Prior to the interviewer's telephone call, the letters were mailed to households with published addresses. We used state or county letterhead for sampled sites and areas of the national sample for which letters of endorsement were obtained; for other areas, we used the letterhead of the president of RWJF. The initial endorsement letter was modified as a result of feedback provided during the first few weeks of data collection. The revised letters, included in Appendix D, made reference to various monetary incentives, which were offered as part of an experiment to increase response rates. (The experiment is discussed in Chapter IV.)

In addition, we developed a four-panel brochure to accompany the letters, which was also translated into Spanish (Appendix E). The brochure described the study's goals, discussed RWJF's role in helping communities, presented a personal statement from the president of RWJF, and listed questions and answers about survey participation. The brochure included a customized insert, listing specific RWJF health projects completed in each state that contained sampled communities; this insert was excluded from areas in the national supplement that were not in one of these states.

We expected that survey cooperation rates would be higher for households receiving advance letters and brochures than in households that did not receive them, and that letters from state health officials would have a greater impact than those from RWJF, which has relatively less name recognition. Although we did not randomize assignment of advance materials to households with published addresses, we were able to compare cooperation rates between households with published addresses, which received brochures, and those with unpublished addresses, which did not. Even if cooperation rates were relatively higher in households with published addresses, we would not have been able to separate the effect of the letter and brochure from factors related to the decision to publish one's telephone number. However, if we had achieved no better cooperation rates from households receiving the letter and brochure, then it would have been reasonable to infer that the advance materials had no impact.[6]

We observed that neither the advance letters, whether from the state or RWJF, nor the brochures had any impact on initial cooperation rates, which were defined as the ratio of completed interviews to the sum of completed interviews and initial refusals. The initial cooperation rates for households included in the experimental sample are as follows:

- Households with published addresses receiving RWJF letters and brochures (55.6 percent)

[6]Note that these households also were part of an experiment to test the impact of monetary incentives on cooperation rates; however, the incentive treatments were randomized across the sample and did not interact with the use of letters.

- Households with published addresses receiving letters from state health officials and brochures (54.9 percent)

- Households with unpublished addresses (and therefore not receiving letters or brochures) in sites in which households with published addresses received RWJF letters and brochures (54.9 percent)

- Households with unpublished addresses (and therefore not receiving letters or brochures) in sites in which households with published addresses received letters from state health officials and brochures (53.3 percent)

As a result, we discontinued advance mailings to initially sampled households with published addresses. However, we used a revised letter, accompanied by an incentive check, for refusal conversion efforts conducted during the second half of the field period (see Appendix D).

3. Survey Messages

In addition to obtaining endorsements and developing advance materials to be mailed to sampled households with published addresses, we worked with the communications consultant to develop an effective survey introduction in the survey instrument. First, we reviewed survey introductions used in related health surveys and the results of opinion surveys on perceived health concerns. Then, we developed several messages built around the importance to the general public of health tracking and tested the messages through cognitive interviews and pretesting. A key issue was the broad focus of the study, which was difficult to capture in a brief statement that resonated with the general public. For example, messages that emphasized access issues, such as affordable health care or meeting the needs of the uninsured, failed to elicit responses from households that were basically satisfied with their coverage. Messages emphasizing satisfaction with health care and health plans were not particularly effective because the survey was not sponsored by the respondent’s insurer or by government agencies overseeing health plans.

Although we tested many survey introductions, we used three during the survey (see Exhibit III.1). The first was developed prior to the survey and was successfully pretested. However, the pretest sample was small and from two communities that were not included in the CTS sample. In addition, the pretest used experienced interviewers. The introduction, which emphasized specific goals of the survey, later proved to be ineffective with most households and was rejected shortly after the survey began. Review of the text on reasons for refusals and interviewer debriefings indicated that the introduction was too long and placed too much emphasis on the study's research objectives.

Refusal rates declined with the second approach, which reduced the emphasis on research objectives and increased the emphasis on state endorsements and the value of information on public health needs to communities. The second approach also included text on experimentally varied respondent incentives designed to test the impact of incentives on survey participation (discussed in Chapter IV). As a result of the experiment, which showed that monetary incentives significantly increased survey participation, we opted for an abbreviated introduction, emphasizing the incentive and state endorsements. We identified the sponsor as a "nonprofit foundation" rather than specifying it by name and further reduced discussion of study goals. We provided additional text on sponsorship, survey goals, interview content, and confidentiality on CATI screens and background material provided to interviewers. Interviewers used this information to answer questions from respondents who had indicated interest in the survey, but wanted additional information.

Our experience demonstrated the importance of developing and testing alternative formulations of survey introductions and advance materials, even for surveys conducted a few years apart about similar topics (access to health care), using the same sample frame (RDD), and funded by the same organization. We had been very successful in using a brief, effective survey introduction for the

EXHIBIT III.1

SURVEY INTRODUCTIONS USED ON THE CTS HOUSEHOLD SURVEY AND THE FAMILY HEALTH INSURANCE SURVEY

CTS Household Survey (1996-1997) Introductions

1. First Introduction (July-August 1996)

Hello, my name is ______. I'm part of a national research project on health care sponsored by the Robert Wood Johnson Foundation. [IF PUBLISHED TELEPHONE NUMBER: You may have recently received a letter and brochure describing our project.] We are interviewing over 36,000 households throughout the country about their health and health needs. We are not trying to sell anything or ask for money. Our goal is to get an accurate picture of how changes in health insurance and health care are affecting us, and to use what we learn to encourage better health care. We will share our findings with people responsible for health care and health insurance. Stories about our project also will appear in newspapers, and on television and radio. Because the survey concerns health issues, I would like to speak with the adult most familiar with the health care of the people who live in your household.

IF MORE NEEDED: We are doing this study because so much has changed in recent years, such as the growth of managed care and how we choose doctors and hospitals. Many people today are concerned that they may lose their health insurance, might have to pay more than they can afford, or they won't be able to get the care they or their families need. Our goal is to get accurate information on people's health needs and problems to decision makers in government and industry. You may be interested to know that [NAME AND TITLE OF HEALTH OFFICIAL ENDORSING STUDY IN RESPONDENTS COMMUNITY] urges residents of [STATE] to take part in the survey.

2. Second Introduction (September 1996-March 1997)

Hello my name is_______. I'm calling to ask you to take part in a major health study. This study is supported by [NAME OF HEALTH DEPARTMENT/IF STATE NOT LISTED: state health departments throughout the country], and is funded by The Robert Wood Johnson Foundation. [INCENTIVE: Because your participation is very important to our study, we will send you or your family [AMOUNT] for helping us with the survey]. We are not trying to sell anything or ask for money, and we are not associated with any political party [PHRASE DROPPED AFTER ELECTION]. We simply want to know your concerns and options about health care so communities in [STATE] and other state will have accurate information about people's health needs. Since the survey is about health issues, I'd like to speak to the person who is most familiar with the health needs of the people who live here.

IF MORE INFORMATION NEEDED: We are doing this study because health care has changed so much in recent years and we don't really know how people are being affected by these changes. This study will help [NAME OF STATE HEALTH DEPARTMENT/state health departments] and others responsible for health care answer important questions. For example, the study will help us understand what types of health care plans cover different families' needs, how satisfied people are with their insurance plans and medical providers, whether people can afford the health care they need, and how we can help people who don't have health insurance or may lose it. We are not proposing particular solutions to these problems. Our goal is to get accurate information bout people's health concerns and views and to use this information to improve health care in communities throughout the country.

3. Third and Final Introduction (March-July 1997)

Hello, my name is ______. I'm calling to offer you [$xx] to help us with a major health study. It's sponsored by a nonprofit foundation and is supported by state health departments throughout the country, [including [fill in state health department]].

We're not selling anything or asking for donations; we just want to hear about your opinions and concerns on health issues, and as I mentioned, we will pay your family [$xx} for helping us with the survey. Since the survey is about health issues, I would like to speak with an adult who lives here and is familiar with the health care of family members. Let's begin...

ADDITIONAL INFORMATION INCLUDED ON CATI SCREEN TO BE USED AS NEEDED

- We're doing the study because health care has changed so much in recent years and we don't know how these changes are affecting people like you.
- The questions are very basic --things like "Are you satisfied with your health care? How long does it take you to get to the doctor? Have you had a flu shot in the last 12 months?"
- The interview is strictly confidential and you don't have to answer any questions you don't want to.
- The study is funded by The Robert Wood Johnson Foundation, a nonprofit organization whose sole mission is to improve health care. The Foundation is not associated with any political party or private company. Since 1972, the Foundation has given more than $2 billion in grants to train doctors and nurses, to make sure children get their shotS against diseases, and to help meet health needs of the elderly.

RWJF 1993-1994 Family Health Insurance Survey, which was sponsored by governors' offices or state health departments in the surveyed states and emphasized the value of those surveys to "... design a better system, in which health care is more affordable and easier for people to obtain." We were not able to develop a universal message about health tracking for the CTS Household Survey that was effective with the general public. Instead, we developed a menu of messages to respond to different subgroups. We found that advance materials were not effective in increasing survey participation, unless they were accompanied by a monetary incentive. Interviewers believed that state and local health department endorsements were helpful in gaining the respondents' attention and encouraging cooperation. However, advance letters on health department letterhead did not increase survey participation. Monetary incentives, interviewer persistence, and extensive use of refusal conversions (discussed in Chapter IV) were the key factors in achieving acceptable response rates on the CTS Household Survey. This experience reinforces the importance of systematically testing the impact that various components of a survey introduction (sponsorship, purpose, and content), methods of presentation (interviewer introduction, advance materials), and incentives have on survey cooperation.

D. INTERVIEWER SELECTION AND TRAINING

1. RDD Sample

a. Recruitment

Most of the interviewing for the RDD sample was conducted by MPR, with assistance from Battelle and CODA, survey organizations that use the same CATI system as MPR. Altogether, 302 interviewers were trained for the household survey, 258 from MPR, 31 from Battelle, and 13 from CODA. MPR conducted 10 initial training sessions in July 1996, CODA conducted 1 session in mid-September, and MPR conducted 2 sessions with Battelle interviewers in late October. MPR conducted two additional training sessions in mid-November and early December to compensate for attrition, as many of the interviewers hired during the summer were college students who returned to school. MPR conducted three additional sessions in April 1996 and May 1997 to ensure completion of the survey within the designated field period.

b. Telephone Interviewer Training Program

New interviewers were given MPR's standard general interviewer training program, which lasted 12 hours and which was conducted in three 4-hour sessions. Topics included obtaining cooperation, understanding bias, using probing methods, using the CATI system, and 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. CODA and Battelle conducted their own general interviewer training sessions.

Training on the survey instrument also lasted 12 hours, conducted in three 4-hour sessions. The initial training agenda included:

- An introduction to the study, client, and sample
- Review of the instrument presented or video screen
- Review of special skills needed to conduct health surveys and interview elderly respondents
- Review of contact procedures and advance materials
- Hands-on practice with four scripted mock interviews

Appendix F contains the Household Survey Interviewer's Manual (Center for Studying Health System Change 1998). The Household Survey Trainer's Agenda is included in Appendix G.

Because of low initial cooperation rates during the first month of interviewing, we reviewed the survey introduction (discussed previously) and interviewer training on refusal avoidance. Interviewer and supervisor debriefings indicated that the introduction was too long, and that interviewers were not always adept at addressing respondents' concerns about survey participation. As a result, the introduction was shortened, and a separate set of "Follow-Up Statements" was developed to give interviewers more flexibility in providing information specific to the respondents' concerns. These statements also were color coded as a memory cue for interviewers, so that the interviewers could more easily identify the follow-up statements that best responded to various concerns.

All initially trained interviewers were given refusal-avoidance retraining based on the new materials (see Appendix H: Refusal Avoidance Retraining Program). The first segment of retraining covered the new introduction, use of the follow-up statements, and general advice on persuasion styles. During the second segment, the interviewer made three calls to a mock respondent (a supervisor or assistant supervisor) who played three different types of reluctant respondents. The contents of the refusal avoidance retraining was added to the core training for later training sessions.

2. Field Sample

a. Recruitment

MPR interviewers were responsible for all field interviewing. Initially, 22 listers were trained in the 12 high-intensity sites; 4 were replaced over the course of the field period. All listers were trained to make screening calls; five screening interviewers, who had participated in listings, were subsequently replaced.

b. Training

MPR mailed its listing manual to all listers in advance of the training, and listers were instructed to conduct two hours of home study prior to receiving training. Training was conducted during a two- to three-hour conference call, in which the MPR trainer reviewed listing and survey procedures with groups of four to six listers. Topics included an overview of the project; an explanation of listing materials in the sample packet; step-by-step instructions for listing and handling inquiries; review exercises; and administrative issues.

Training to conduct screening interviews lasted two hours. Screening interviewers were trained in the survey introduction and 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 completion of the training program, each trainee called the MPR telephone center and conducted a mock screening interview with a supervisor. Memoranda provided to field listers and screening interviewers that outlined procedures and approaches for avoiding refusals are included in Appendix I. (As noted, field interviewers were not trained to conduct the survey; instead, they called the MPR telephone center and then gave the respondent a cellular telephone to complete the interview).

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.[7] The CSM computer-assisted interviewing program, CASES, is one of the most widely used

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

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 controlling the distribution and timing of calls and developed specialized reports (discussed in Chapter IV) for monitoring the survey results.

IV. DATA COLLECTION

In the first part of this chapter, we provide an overview of sample sizes, response rate calculations, and data collection procedures. In subsequent sections, we describe organization of the random-digit-dialing (RDD) and field data collection efforts; response rates; efforts to reduce nonresponse, including refusal conversions, monetary incentives, use of Spanish-speaking interviewers, selective use of proxy respondents, and calls to telephone companies to ascertain residential status for telephone numbers that were difficult to contact (for the RDD sample only); and quality assurance procedures.

A. OVERVIEW

1. Sample Sizes and Response Rates

A total of 32,732 family insurance unit (FIU) interviews were completed, 32,097 from the RDD sample and 635 from the field sample. The FIUs included 49,807 eligible adults and 10,639 sampled children under the age of 18, for a total of 60,446 individuals. The FIU response rate is the product of the household enumeration response rate and the FIU interview response rate, the latter conditional on completing the household enumeration questions needed to determine eligibility for the survey and to form FIUs (see Part A of the survey instrument). Sixty-eight percent of the combined RDD and field samples completed the household enumeration questions. Among the households completing these questions, 95 percent of the FIUs that were formed completed the interview, for a cumulative response rate of 65 percent.

For the RDD sample, the household enumeration response rate is the ratio of the number of households completing the enumeration questions to the estimated number of households. The estimated number of households was the total of telephone numbers confirmed by interviewer contact to be those of households and an imputed estimate of residential status, for telephone numbers that did not result in a contact after 20 calls. The FIUs were formed after household enumeration. Some households included more than one FIU, and some of the secondary family informants in these households refused to complete the interviews or could not be successfully interviewed for other reasons.

For the field sample, the household enumeration response rate is analogous to the RDD rate; that is, it is the ratio of field-eligible screened households to the estimated number of eligible households in the field sample. For field interviews, the household informant acted as informant for all FIUs, and there was no additional nonresponse at the FIU level. Field interviews were conducted in the 12 high-intensity sites and were used in making site-level and national estimates.

Reported response rates for the RDD samples and for the combined RDD and field samples are unweighted, as they are intended as measures of performance. Because of the minimal use of oversampling in the Community Tracking Survey (CTS), unweighted response rates can also serve as reasonable proxy measures for potential nonresponse bias.

2. Call Scheduling and Follow-Up Efforts

Telephone numbers in the RDD frame released for interviewing were controlled by the computer-assisted telephone interviewing (CATI) scheduler, supplemented by hard-copy disposition sheets that were used to track selected subgroups, requiring different follow-up rules that could not be met by the scheduler. The scheduling program randomly assigned sampled telephone numbers to interviewers, with nonscheduled calls based on optimal calling patterns, dispersed over different times of the day and different days of the week. 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.

Refusals were assigned to a subgroup of particularly skilled interviewers (known as "refusal converters"); the period for reassigning refusals, typically four to eight weeks, was designed to minimize the impact of the prior refusals. Refusal converters used information about the reason and intensity of the prior refusals in planning their calls. Because initial refusal rates were higher than expected, we normally did not retire a case until household members had refused three times, with refusal conversion calls dispersed over several months. Thirty-one percent of FIUs refused at least once before agreeing to complete the interview; the refusal conversion rate was 52 percent.

A separate core of Spanish-speaking interviewers was assigned cases in which the family informant or other adults assigned a self-response module preferred to conduct the interview in Spanish. A total of 1,182 Spanish interviews were completed, representing 3.7 percent of all FIUs.

We initially planned to limit follow-up efforts to 12 calls to determine whether a telephone number was residential, and to 20 calls to complete an interview after an FIU was identified. Because many telephones could not be classified as residential or nonresidential after 12 calls, and efforts to obtain information on residential status from local telephone companies were generally unsuccessful, we increased this limit from 12 to 20 calls. In addition, we increased the limit on total calls to confirmed residential telephone numbers from 20 to 40 to allow sufficient time to complete refusal conversion and other follow-up efforts.

Initially, we limited monetary incentives to $15 per FIU and offered them only to informants who had previously refused or for whom accurate addresses were needed for a followup survey.[1]

[1]The CTS Household Survey was designed to be reported two years later (Metcalf et al. 1996). We assumed that a sample of households completing round one interviews would be surveyed again. Incentives were offered to family informants in households selected for reinterview that had unlisted telephone numbers.

However, as refusal rates were higher than expected, and the $15 incentive appeared to be relatively ineffective, we conducted an experiment to determine the impact of monetary incentives on survey cooperation. As a result of this experiment (discussed in greater detail later in this chapter), we offered $25 to each FIU.

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. Because the cost of completing field interviews is much higher than that of completing telephone interviews, we attempted to reduce the likelihood of refusals by offering each family informant $25 for completing the interview. Surveys conducted in person typically have lower refusal rates than those conducted by telephone; also, the field sample, which was limited to households with no or intermittent telephone service, was predominantly low income and responsive to cash incentives. Refusal rates were very low, and we did not need to mount a refusal conversion effort 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.

3. 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 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 for 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 Standard Industrial Classification (1987) (SIC) coding structure.[2] 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 were reviewed by a coder. Fifty-four percent of the codes were assigned by the program; the remaining 46 percent were coded manually.

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. Additional codes were assigned to text responses for the following questions:

2The *Federal Register* has indicated that the SIC will be replaced with the North American Industry Classification System; if replacement occurs, we may have to revise the codes used for the initial CTS to conform to the revised system for subsequent surveys.

b1i1: type of health plan (other)
b84: why coverage stopped (other)
b881: why changed plans (other)
c831: why postponed care (other)
d151: why changed usual source of care (other)
f521: why not participant (other)
f531: why ineligible (other)

Text files were provided, but additional codes were not assigned for these questions:

b1f1: type of military plan (other)
b121: type of health profile (other)
d201: reason for no usual source of care (other)
g221:race/ethnicity (other)

4. 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). Instrument cleaning, reformatting, and SAS programs used in the preparation of these files are maintained by Mathematica Policy Research, Inc. (MPR).

B. ORGANIZATION OF THE RDD AND FIELD SURVEYS

1. RDD Survey

Interviewing for the RDD sample was conducted primarily from MPR's Princeton telephone center, with assistance from two subcontractors--CODA and Battelle. All three 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 its subcontractors 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 field 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 households with published and unpublished telephone numbers; and on subgroups, including Spanish samples, primary and secondary FIUs, and refusal conversion samples

- ***Weekly 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 incentives, advance letters, and 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

Field reports were supplemented by regularly scheduled weekly conference calls with survey supervisors and by site visits to review changes in procedures.

The distribution of completed RDD FIU interviews, by data collection site, is shown here[3]:

	MPR	Battelle	CODA
Completed FIU Interviews	21,347	8,272	2,487
Percentage of Total	66.5	25.8	7.7

[3]Includes a small number of completed interviews deleted from the final data file during data cleaning.

Although 302 interviewers were trained, most of the interviewing was conducted by a cadre of 70 interviewers at MPR and 19 at Battelle, each of whom completed at least 100 interviews. These interviewers completed 76.5 percent of all the interviews. The main sources of interviewer attrition were (1) college students returning to school in the fall of 1996; and (2) high initial refusal rates, which required extensive follow-up efforts.

2. Field Survey

All the field listing, screening, and interviewing was conducted by MPR staff. We maintained a staff of 22 listers across the 12 high-intensity sites, varying the number assigned to listing and screening by site to reflect differences in sample allocation. Four staff members were replaced during listing, and five during screening. All field staff were supervised by MPR survey managers, located in Princeton. Reports were developed to monitor both listing 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

The following sections describe how we calculated response rates for the RDD, field, and combined samples.

a. RDD Sample--Residential Telephone Status

The first step was to determine residential telephone status for each sampled telephone number on the basis of its final disposition code, summarized in Table IV.1. Disposition codes are defined more completely in Table IV.2.

b. RDD--Household Enumeration Questions Rate

At the household level, we calculated a household enumeration questions response rate using disposition codes for all released telephone numbers; for households with more than one FIU, we used the disposition code for the primary FIU. In forming this file, we created three flag variables based on disposition class (see Table IV.1):

1. If class A, B, C, D, or E, then PHONDET=1 (residential or nonresidential telephone status was determined)

2. If class A, B, C, or D, then PHONELIG=1 (working residential telephone number)

3. If class A, B, C, or D and NFAM (number of FIUs variable) is nonmissing,the SCRCOMP=1 (primary FIU completed household enumeration questions in the instrument)

The next step was to count the number of sampled telephone numbers in each of the three categories, as well as the total number of telephone numbers released, by site (the supplemental sample is treated as a separate site) and by telephone number type (published or unpublished). The cumulative counts were as follows:

ALL = 70,936

PHONDET = 67,312(3,624 had undetermined residential telephone status)

PHONELIG = 38,291(29,021 were nonworking or nonresidential numbers)

SCRCOMP = 27,381(10,910 primary FIUs did not complete the household enumeration questions)

TABLE IV.1

DETERMINATION OF RDD SAMPLE ELIGIBILITY

.

Class and Description	Final Disposition Codes: Primary	Final Disposition Codes: Secondary
A Working residential telephone, FIU eligible, FIU responded	1, 2	1, 2
B Working residential telephone, FIU eligible, FIU did not respond	22	21, 22, 30
C Working residential telephone, FIU ineligible	41, 47, 49	41, 47, 49
D Working residential telephone, FIU eligibility undetermined	21, 30, 39, 64, 66, 67	44, 53, 54, 65, 67
E Not a working residential telephone number	40, 50	
F Working residential telephone status undetermined	65	

TABLE IV.2

CTS HOUSEHOLD SURVEY FINAL DISPOSITION CODES FOR THE RDD SAMPLE

Final Disposition Code	Description	Working Residential Telephone?	Primary FIU Eligible?	Secondary FIU Eligible?
1	Complete	Yes	Yes (n=24,903)	Yes (n=5,793)
2	Complete (but not all other adults)	Yes	Yes (n=1,377)	Yes (n=28)
21, 39, 64	Refusal	Yes	Unknown (n=6,687)	Yes[a] (n=283)
22	Refusal (breakoff)	Yes	Yes (n=2,156)	Yes (n=220)
30	Language/other barrier (inaccessible), no proxy available, proxy refused	Yes	Unknown (n=667)	Yes[a] (n=59)
40	Not a residence	No	Not applicable (n=13,839)	Not applicable
41	Not selected--ineligible (no one eligible to be informant)	Yes	No (n=325)	No (n=441)
44	Secondary FIU nonworking number (disconnect--no listing)	Yes	Not applicable	Unknown (n=59)
47	Ineligible FIU	Yes	No (n=3)	No (n=11)
49	Other ineligible--died or error in household composition; duplicates recorded	Yes	No (n=13)	No (n=70)
50	Nonworking number	No	Not applicable (n=15,182)	Not applicable
53	Secondary FIU moved; no forwarding number; no proxy available	Yes	Secondary only	Unknown (n=190)
54	Secondary FIU--changed to nonworking number during callbacks	Yes	Secondary only	Unknown (n=204)
65	Maximum calls (no contact)	Unknown	Unknown (n=3,624)	Unknown (n=5)

Final Disposition Code	Description	Working Residential Telephone?	Primary FIU Eligible?	Secondary FIU Eligible?
66	Maximum calls	Yes	Unknown (n=2,069)	Not applicable
67	Effort ended	Yes	Unknown (n=91)	Unknown (n=191)
Total			**70,936**	**7,554**

[a]These status codes can be assigned only to a secondary FIU, when the primary FIU was determined to be eligible.

Within each site and telephone type, we calculated the telephone eligibility rate among released telephones with known eligibility status. This rate was then applied to released telephone numbers with undetermined telephone eligibility status, within each site and telephone type, to impute eligibility.

The ***site-specific household enumeration RDD response rate*** can be computed as:

$$(1)\quad RR(H)_i = \frac{\sum_j SCRCOMP_{ij}}{\sum_j PHONELIG_{ij} + \sum_j [(ALL_{ij} - PHONDET_{ij}) \times (phone\ elig\ rate_{ij})]}$$

for site i (i=0,1,...,60) and telephone type j (j=published, unpublished).

The ***overall household enumeration RDD response rate*** can be computed as:

$$(2)\quad RR(H) = \frac{\sum_i \sum_j SCRCOMP_{ij}}{\sum_i \sum_j PHONELIG_{ij} + \sum_i \sum_j [(ALL_{ij} - PHONDET_{ij}) \times (phone\ elig\ rate_{ij})]}$$

$$= \frac{27{,}381}{38{,}291 + 1{,}815.2} = .6827.$$

c. RDD FIU Response Rate

To calculate the response rate at the FIU level, we created three flag variables, as follows:

4. If class A, B, or C, then FAMDET=1 (determined eligibility status of FIU)
5. If class A or B, then FAMELIG=1 (eligible FIU)
6. If class A, then FAMRESP=1 (FIU responded)

The next step was to count the number of FIUs in households completing the household enumeration questions in each of the three categories, as well as the total number of these FIUs, by site and FIU type (primary versus secondary). All secondary FIUs were in households in which the primary family informant completed the household enumeration questions and FIU interview; that is, it was not possible to generate a secondary FIU without first completing the household enumeration questions and primary FIU interview. The cumulative counts for FIUs in households completing the household enumeration questions were as follows:

	Primary FIU	Secondary FIU	Total
ALLFAM =	27,381	7,554	34,935
FAMDET =	27,079 (302 undetermined eligibility)	6,905 (649 undetermined eligibility)	33,984
FAMELIG =	26,747 (332 ineligible[4])	6,383 (522 ineligible)	33,130
FAMRESP =	26,277 (470 refused)	5,820 (563 refused/barrier)	32,097

Within each site and FIU type, we calculated the eligibility rate among FIUs in households completing enumeration with known eligibility status. This rate was applied to FIUs in households completing enumeration that had undetermined eligibility status, in order to impute eligibility to them, by site and FIU type.

The ***site-specific FIU RDD response rate***, among those in enumeration-complete households, can then be computed as:

$$(3)\quad RR(F)_i = \frac{\sum_k FAMRESP_{ik}}{\sum_k FAMELIG_{ik} + \sum_k [(ALLFAM_{ik} - FAMDET_{ik}) \times (family\ elig\ rate_{ik})]}$$

for site i (i=0,1,...,60) and FIU type k (k=primary, secondary).

[4]An ineligible FIU is generally one with no eligible (civilian adult) informant. Ineligible secondary FIUs could also include those generated erroneously by the primary FIU.

The ***overall FIU RDD response rate***, among those in enumeration-complete households, can be computed as:

$$(4)\quad RR(F) = \frac{\sum_i \sum_k FAMRESP_{ik}}{\sum_i \sum_k FAMELIG_{ik} + \sum_i \sum_k [(ALLFAM_{ik} - FAMDET_{ik}) \cdot (family\ elig\ rate_{ik})]}$$

$$= \frac{32{,}097}{33{,}130 + 293.5 + 600.2} = .9432.$$

The ***site-specific combined household-FIU RDD response rate*** can then be computed as:

$$(5)\quad RR_i = RR(H)_i \cdot RR(F)_i .$$

The ***overall combined household-FIU RDD response rate*** can then be computed as:

$$(6)\quad RR = RR(H) \cdot RR(F) = .6827 \cdot .9432 = .6440$$

d. Field Sample Response Rate

In the following sections, we describe the calculations of the response rates for the field component of the sample. As with the RDD sample, the first step was to determine eligibility of sampled addresses according to their final disposition codes (Table IV.3). Table IV.4 defines the disposition codes.

At the household level, we calculated a response rate to the household enumeration questions, using a household-level file that contained the status codes of all released addresses, by site. We created four flag variables:

1. If class A, B, C or D, then ADDRELIG=1 (residential address)
2. If class A, B, or C, then TSCRCOMP=1 (household completed enumeration questions)

TABLE IV.3

DETERMINATION OF FIELD SAMPLE ELIGIBILITY

Class and Description		Final Disposition Code
A	Address eligible, household eligible, household responded	1, 2
B	Address eligible, household eligible, household did not respond	22
C	Address eligible, household ineligible	45, 49
D	Address eligible, household's eligibility undetermined	20, 30, 39, 64, 65
E	Address ineligible (not a residence)	40, 48

TABLE IV.4

CTS HOUSEHOLD SURVEY FINAL DISPOSITION CODES
FOR THE FIELD SAMPLE

Final Disposition Code	Description	Residence?	Household Eligible?
1	Complete	Yes	Yes (n=467)
2	Complete (but missing one more adult self-response modules)	Yes	Yes (n=3)
20	Refused before completing household enumeration questions	Yes	Unknown (n=155)
22	Eligible, refused to complete interview	Yes	Yes (n=50)
30	Language/other barrier	Yes	Unknown (n=29)
39	Other, possibly eligible household	Yes	Unknown (n=1)
40	Not a residence	No	Not applicable (n=11)
45	Ineligible, has telephone service	Yes	No (n=4,039)
48	No housing unit existed	No	Not applicable (n=374)
49	Other ineligible	Yes	No (n=2)
64	Never contacted at home (six attempts)	Yes	Unknown (n=99)
65	Contacted but never completed household enumeration questions	Yes	Unknown (n=28)
Total			**5,258**

3. If class A or B, then HHELIG=1 (household meets eligibility criteria)

4. If class A, then INTCOMP=1 (household completed interview)

The cumulative counts were as follows:

ALL	=	5,258
ADDRELIG	=	4,873 (385 were not residences; codes 40, 48)
TSCRCOMP	=	4,561 (312 households did not complete the household enumeration questions; codes 20, 30, 39, 64, 65)
HHELIG	=	520 (4,041 households had no interruption in telephone service; codes 45, 49)
INTCOMP	=	470 (50 eligible households did not complete the interview; code 22)

Within each site, we calculated the household eligibility rate among households with known eligibility statuses. This rate was applied to households with undetermined eligibility status, within each site, to impute eligibility.

The ***site-specific household interview field response rate*** can be computed as:

$$(7)\quad RR(H)_i = \frac{INTCOMP_i}{HHELIG_i + [(ADDRELIG_i - TSCRCOMP_i)\cdot(hhold\ elig\ rate_i)]}$$

for site i (i=1,...,12).

The ***overall household interview field response rate*** can be computed as:

$$(8)\quad RR(H) = \frac{\sum_i INTCOMP_i}{\sum_i HHELIG_i + \sum_i [(ADDRELIG_i - TSCRCOMP_i)\cdot(hhold\ elig\ rate_i)]}$$

$$= \frac{470}{520 + 53.1} = .8201 .$$

Because a single household informant responded on behalf of all eligible FIUs for the field sample, the FIU response rate was the same as the household response rate. The 470 responding households generated 656 FIUs, of which 635 were eligible for the survey.

e. Combined RDD/Field Response Rate

We combined the RDD response rate and field response rate into a single response rate, as follows.

The ***site-specific household enumeration RDD+field response rate*** can be computed as:

$$(9) \quad RR'(H)_i = \frac{\sum_j SCRCOMP_{ij} + INTCOMP_i}{estimated\ eligible\ households_i},$$

where:

$$\begin{aligned} estimated\ eligible\ households_i = & \sum_j PHONELIG_{ij} \\ & + \sum_j [(ALL_{ij} - PHONEDET_{ij}) \times (phone\ elig\ rate_{ij})] \\ & + HHELIG_i \\ & + [(ADDRELIG_i - TSCRCOMP_i) \times (hhold\ elig\ rate_i)] \\ & \quad for\ site\ i = 1,2,...12. \end{aligned}$$

The overall household enumeration RDD+field response rate can be computed as:

$$(10) \quad RR'(H) = \frac{\sum_i \sum_j SCRCOMP_{ij} + \sum_i INTCOMP_i}{\sum_i estimated\ eligible\ households_i}$$

where *i* is summed over all 60 sites and the supplemental sample (combined sample) or other groupings of sites. (Of course, the field factors can be summed only over the 12 high-intensity sites.) For the combined sample, the household-level response rate is 68.5 percent.

Among households that completed the enumeration questions, the ***site-specific FIU RDD+field response rate*** can be computed as:

$$(11)\ RR'(F)_i = \frac{\sum_k FAMRESP_{ik} + ffc_i}{\sum_k FAMELIG_{ik} + \sum_k [(ALLFAM_{ik} - FAMDET_{ik}) \cdot (family\ elig\ rate_{ik})] + ffe_i},$$

where ffc_i is the number of field FIU completes in site i, and ffe_i is the number of eligible field FIUs in site i, for i=1,2,...,12.

The ***overall FIU RDD+field sample response rate***, among FIUs in households completing enumeration questions, can be computed as:

$$(12)\ RR'(F) = \frac{\sum_i \sum_k FAMRESP_{ik} + \sum_i ffc_i}{\sum_i \sum_k FAMELIG_{ik} + \sum_i \sum_k [(ALLFAM_{ik} - FAMDET_{ik}) \cdot (family\ elig\ rate_{ik})] + \sum_i ffe_i},$$

where i is summed either over the 12 high-intensity sites or over all 60 sites and the supplemental sample. (Of course, ffc and ffe can be summed only over the 12 sites.) For the combined sample, the response rate is 94.4 percent.

The site-specific combined FIU RDD+field response rate can then be computed as:

$$(13)\ RR'_i = RR'(H)_i \cdot RR'(F)_i .$$

For the combined sample (60 sites and supplemental sample), the ***overall combined-FIU RDD+field response rate*** can then be computed as:

$$(14)\ RR' = RR'(H) \cdot RR'(F) = .6846 \cdot .9443 = .6465 .$$

2. Household and FIU Response Rates, by Sample Type and Site

Tables IV.5 through IV.7 show overall household and FIU response rates and number of completed FIU interviews by sample type and site. Table IV.5 shows response rates for the combined RDD and field samples; Table IV.6, rates for the RDD samples; and Table IV.7, rates for the field sample.

The overall FIU-level response rate for the combined 60 sites and supplemental sample was 64.7 percent. For high-intensity sites, the response rate varied from a low of 52.0 percent in Miami to a high of 74.5 percent in Little Rock (Table IV.5). For the RDD sample, differences in response rates by site mainly reflected variability in response rates to the household enumeration questions. For high intensity sites, these response rates varied from 56.9 to 76.3 percent (Table IV.6). Family insurance units were formed after the household enumeration questions were completed. We observed less variability among FIU interview response rates for these sites (from 90.0 to 97.2 percent). In general, we obtained higher response rates in smaller metropolitan statistical areas (MSAs), such as Greenville, Indianapolis, Lansing, and Syracuse, and lower response rates in the largest MSAs, particularly Miami and Orange County. Similar patterns were observed for low-intensity sites. The overall response rate for the field sample (82.0 percent) was considerably higher than for the RDD sample (64.4 percent). Field sample response rates by site exceeded 69 percent for all sites except for Orange County (53.0 percent; see Table IV.7).

The sources of non-response (mainly non-contacts, refusals, and maximum contacts without a confirmed refusal) varied by type of sample and site. The sources of nonresponse are shown for the RDD sample at the household level in Table IV.8; for RDD primary and secondary FIU interviews, respectively, in Tables IV.9 and IV.10; and for the field sample in Table IV.11. For the RDD sample, 68.3 percent of households completed the household enumeration questions. Refusals dominated nonresponse (20.9 percent), followed by confirmed residential households that were retired after 40 call attempts (4.8

TABLE IV.5

CTS RESPONSE RATES, BY SAMPLE TYPE: COMBINED RDD AND FIELD SAMPLES

	FIU Completes	Household Response Rate	FIU Response Rate	Household + FIU Response Rate
All sites + national supplement	32,732	68.46%	94.43%	64.65%
All 60 sites in site sample	29,456	68.43%	94.39%	64.60%
12 high-intensity sites	14,945	67.95%	94.56%	64.25%
High-Intensity Sites				
Boston, MA	1,145	62.63%	92.88%	58.17%
Cleveland, OH	1,211	65.91%	95.28%	62.80%
Greenville, SC	1,285	74.09%	95.04%	70.41%
Indianapolis, IN	1,316	73.95%	97.37%	72.01%
Lansing, MI	1,232	73.19%	96.88%	70.91%
Little Rock, AR	1,412	76.90%	96.84%	74.47%
Miami, FL	1,171	57.61%	90.29%	52.02%
Newark, NJ	1,282	63.12%	92.01%	58.07%
Orange County, CA	1,157	58.02%	91.94%	53.35%
Phoenix, AZ	1,250	71.28%	96.35%	68.67%
Seattle, WA	1,181	68.11%	94.02%	64.04%
Syracuse, NY	1,303	72.28%	95.48%	69.02%

TABLE IV.6

CTS HOUSEHOLD SURVEY RESPONSE RATES, BY SAMPLE TYPE:
RDD SAMPLE

	Estimated Households	Household Enumeration Completes	Household Enumeration Response Rate	Estimated Eligible FIUs	FIU Interview Completes	FIU Interview Response Rate	Combined Response Rate
All sites + national supplement	40,106.2	27,381	68.27%	34,028.7	32,097	94.32%	64.40%
All 60 sites in site sample	36,012.4	24,566	68.22%	30,570.2	28,821	94.28%	64.31%
National supplement	4,093.8	2,815	68.76%	3,458.5	3,276	94.72%	65.13%
12 high-intensity sites	17,860.3	12,055	67.50%	15,169.0	14,310	94.34%	63.67%
48 low-intensity sites	18,152.1	12,511	68.92%	15,401.2	14,511	94.22%	64.94%
High-Intensity Sites							
Boston, MA	1,487.0	930	62.54%	1,215.8	1,128	92.78%	58.03%
Cleveland, OH	1,543.4	1,014	65.70%	1,229.0	1,169	95.12%	62.49%
Greenville, SC	1,437.5	1,059	73.67%	1,284.0	1,217	94.78%	69.83%
Indianapolis, IN	1,429.2	1,045	73.12%	1,255.5	1,220	97.17%	71.05%
Lansing, MI	1,446.6	1,059	73.21%	1,247.7	1,208	96.82%	70.88%
Little Rock, AR	1,505.1	1,148	76.27%	1,372.1	1,326	96.64%	73.71%
Miami, FL	1,571.0	894	56.91%	1,254.9	1,129	89.97%	51.20%
Newark, NJ	1,558.2	979	62.83%	1,333.4	1,222	91.65%	57.58%
Orange County, CA	1,606.8	933	58.07%	1,244.4	1,143	91.85%	53.33%
Phoenix, AZ	1,387.1	976	70.36%	1,217.4	1,170	96.11%	67.62%
Seattle, WA	1,425.5	970	68.05%	1,201.1	1,126	93.75%	63.79%
Syracuse, NY	1,462.9	1,048	71.64%	1,313.7	1,252	95.30%	68.27%
Low-Intensity Sites							
Atlanta, GA	371.1	254	68.45%	317.1	296	93.35%	63.89%
Augusta-Aiken, GA/SC	371.0	254	68.46%	306.7	291	94.88%	64.96%
Baltimore, MD	344.7	233	67.60%	299.0	285	95.32%	64.43%
Bridgeport, CT	412.4	247	59.89%	313.4	284	90.62%	54.27%
Chicago, IL	400.1	256	63.98%	317.8	293	92.20%	58.99%
Columbus, OH	371.4	265	71.35%	317.2	296	93.32%	66.58%
Denver, CO	377.1	256	67.89%	304.7	291	95.50%	64.83%
Detroit, MI	379.3	265	69.87%	337.1	309	91.66%	64.04%
Greensboro, NC	336.0	238	70.83%	284.7	271	95.19%	67.42%
Houston, TX	367.6	243	66.10%	302.2	280	92.65%	61.25%
Huntington-Ashland, WV-KY-OH	330.1	253	76.64%	317.9	307	96.57%	74.02%
Killeen, TX	385.5	299	77.56%	320.9	298	92.86%	72.03%
Knoxville, TN	357.5	263	73.57%	322.9	311	96.31%	70.86%
Las Vegas, NV/AZ	365.2	217	59.42%	289.2	267	92.32%	54.86%
Los Angeles, CA	375.8	216	57.48%	287.7	261	90.72%	52.14%
Middlesex, NJ	363.3	247	67.99%	326.7	311	95.19%	64.72%
Milwaukee, WI	371.3	261	70.29%	320.5	311	97.04%	68.21%
Minneapolis-St. Paul, MN/WI	351.9	276	78.43%	350.7	334	95.24%	74.70%
Modesto, CA	389.8	260	66.70%	320.7	306	95.42%	63.64%

	Estimated Households	Household Enumeration Completes	Household Enumeration Response Rate	Estimated Eligible FIUs	FIU Interview Completes	FIU Interview Response Rate	Combined Response Rate
Nassau, NY	451.9	280	61.96%	365.4	341	93.32%	57.82%
New York City, NY	493.1	246	49.89%	331.0	292	88.22%	44.01%
Philadelphia, PA/NJ	389.7	259	66.46%	324.9	309	95.11%	63.21%
Pittsburgh, PA	383.6	259	67.52%	315.8	299	94.68%	63.93%
Portland-Salem, OR/WA	379.2	267	70.41%	329.7	307	93.12%	65.56%
Riverside-San Bernardino, CA	378.4	258	68.18%	322.4	304	94.29%	64.29%
Low-Intensity Sites *(continued)*							
Rochester, NY	413.4	301	72.81%	371.6	355	95.53%	69.56%
San Antonio, TX	405.2	267	65.89%	316.3	299	94.53%	62.29%
San Francisco, CA	431.4	230	53.31%	310.4	281	90.53%	48.27%
Santa Rosa, CA	393.1	242	61.56%	310.8	285	91.70%	56.45%
Shreveport, LA	343.6	258	75.09%	321.2	298	92.78%	69.66%
St. Louis, MO/IL	375.7	280	74.53%	337.5	318	94.22%	70.22%
Tampa, FL	384.8	241	62.63%	288.9	268	92.77%	58.10%
Tulsa, OK	394.4	259	65.67%	306.9	292	95.15%	62.48%
Washington, DC/MD	375.4	256	68.19%	325.7	310	95.18%	64.91%
West Palm Beach, FL	361.1	206	57.05%	279.8	253	90.42%	51.58%
Worcester, MA	417.8	283	67.74%	332.7	310	93.18%	63.11%
Dothan, AL	371.2	275	74.08%	320.7	301	93.86%	69.53%
Terre Haute, IN	331.9	252	75.93%	305.5	293	95.91%	72.82%
Wilmington, NC	339.4	264	77.78%	319.6	303	94.81%	73.74%
West Central, Alabama	357.7	270	75.48%	350.5	329	93.87%	70.85%
Central Arkansas	443.7	342	77.08%	385.8	379	98.24%	75.72%
West Georgia	311.2	232	74.55%	288.8	273	94.53%	70.47%
North East Illinois	374.9	263	70.15%	303.0	294	97.03%	68.07%
North East Indiana	349.3	260	74.43%	300.8	286	95.08%	70.77%
East Maine	343.2	286	83.33%	332.0	319	96.08%	80.07%
East North Carolina	340.3	270	79.34%	320.3	304	94.91%	75.30%
West Utah	402.9	323	80.17%	388.2	377	97.11%	77.86%
North West Washington	393.5	279	70.90%	337.9	330	97.66%	69.24%

TABLE IV.7

CTS HOUSEHOLD SURVEY RESPONSE RATES, BY SAMPLE TYPE:
FIELD SAMPLE

	Estimated Households	Completed Household Interview	Household Response Rate[a]	Completed FIU Interview	FIU Interview Response Rate[a]	Combined Response Rate[a]
12 high-intensity sites	573.1	470	82.01%	635	100%	82.01%
High-Intensity Sites						
Boston, MA	10.6	8	75.42%	17	100%	75.42%
Cleveland, OH	45.1	33	73.25%	42	100%	73.25%
Greenville, SC	62.1	52	83.73%	68	100%	83.73%
Indianapolis, IN	78.6	70	89.10%	96	100%	89.10%
Lansing, MI	22.1	16	72.38%	24	100%	72.38%
Little Rock, AR	69.7	63	90.43%	86	100%	90.43%
Miami, FL	29.3	28	95.40%	42	100%	95.40%
Newark, NJ	68.9	48	69.68%	60	100%	69.68%
Orange County, CA	13.2	7	53.04%	14	100%	53.04%
Phoenix, AZ	58.0	54	93.14%	80	100%	93.14%
Seattle, WA	70.6	49	69.44%	55	100%	69.44%
Syracuse, NY	45.1	42	93.22%	51	100%	93.22%

[a] For the field sample, the household informant answered all questions, except for those in the adult self-response module and children's last visit module. Therefore, the household informant completed all FIU interviews.

TABLE IV.8

DISPOSITION OF THE RDD HOUSEHOLD-LEVEL SAMPLE,
BY CTS SITE, SUPPLEMENTAL SAMPLE,
AND TOTAL SAMPLE

				Estimated Households					
	Telephone Numbers Released	Ineligible Telephone Numbers	Non-contacts	Noncontacts Estimated to be Households	Household Refusals	Language/ Other Barriers	Maximum Contacts	Completed Household Enumeration	Total
Total RDD Sample									
	70,936	29,021	3,624	1,815.2	8,364	606	1,940	27,381	40,106.2
				4.53%	20.85%	1.51%	4.84%	68.27%	100%
High-Intensity Sites									
Boston, MA	2,373	807	168	89.0	359	39	70	930	1,487
				5.99%	24.14%	2.62%	4.71%	62.54%	100%
Cleveland, OH	2,641	1,024	143	69.4	391	17	52	1014	1543.4
				4.50%	25.33%	1.10%	3.37%	65.70%	100%
Greenville, SC	2,485	997	103	52.5	259	17	50	1059	1437.5
				3.65%	18.02%	1.18%	3.48%	73.67%	100%
Indianapolis, IN	2,299	816	113	59.2	274	8	43	1045	1429.2
				4.14%	19.17%	0.56%	3.01%	73.12%	100%
Lansing, MI	2,320	816	127	69.6	264	12	42	1059	1446.6
				4.81%	18.25%	0.83%	2.90%	73.21%	100%
Little Rock, AR	2,664	1,096	124	61.1	228	11	57	1148	1505.1
				4.06%	15.15%	0.73%	3.79%	76.27%	100%
Miami, FL	2,741	1,095	163	88.0	355	42	192	894	1571
				5.60%	22.60%	2.67%	12.22%	56.91%	100%
Newark, NJ	2,728	1,081	184	95.2	352	34	98	979	1558.2
				6.11%	22.59%	2.18%	6.29%	62.83%	100%
Orange County, CA	3,195	1,461	240	112.8	388	60	113	933	1606.8
				7.02%	24.15%	3.73%	7.03%	58.07%	100%
Phoenix-Mesa, AZ	2,673	1,221	120	55.1	284	11	61	976	1387.1
				3.97%	20.47%	0.79%	4.40%	70.36%	100%
Seattle, WA	2,785	1,279	143	62.5	294	41	58	970	1425.5
				4.38%	20.62%	2.88%	4.07%	68.05%	100%
Syracuse, NY	2,313	797	107	53.9	304	14	43	1048	1462.9
				3.68%	20.78%	0.96%	2.94%	71.64%	100%

	Telephone Numbers Released	Ineligible Telephone Numbers	Non-contacts	Estimated Households					
				Noncontacts Estimated to be Households	Household Refusals	Language/ Other Barriers	Maximum Contacts	Completed Household Enumeration	Total
Low-Intensity Sites									
Atlanta, GA	623	234	37	19.1 5.15%	68 18.32%	4 1.08%	26 7.01%	254 68.45%	371.1 100%
Augusta-Aiken, GA/SC	604	221	27	15.0 4.04%	86 23.18%	4 1.08%	12 3.23%	254 68.46%	371 100%
Baltimore, MD	557	196	36	19.7 5.72%	68 19.73%	6 1.74%	18 5.22%	233 67.60%	344.7 100%
Bridgeport, CT	782	337	57	24.4 5.92%	103 24.98%	7 1.70%	31 7.52%	247 59.89%	412.4 100%
Chicago, IL	763	335	49	21.1 5.27%	98 24.49%	5 1.25%	20 5.00%	256 63.98%	400.1 100%
Columbus, OH	619	235	28	15.4 4.15%	76 20.46%	2 0.54%	13 3.50%	265 71.35%	371.4 100%
Denver, CO	711	314	37	17.1 4.53%	76 20.15%	6 1.59%	22 5.83%	256 67.89%	377.1 100%
Detroit, MI	613	218	34	18.3 4.82%	78 20.56%	2 0.53%	16 4.22%	265 69.87%	379.3 100%
Greensboro, NC	605	255	27	13.0 3.87%	65 19.35%	7 2.08%	13 3.87%	238 70.83%	336 100%
Houston, TX	713	334	20	8.6 2.34%	76 20.67%	9 2.45%	31 8.43%	243 66.10%	367.6 100%
Huntington-Ashland, WV-KY-OH	510	173	16	9.1 2.76%	55 16.66%	6 1.82%	7 2.12%	253 76.64%	330.1 100%
Killeen, TX	713	316	22	10.5 2.72%	68 17.64%	0 0.00%	8 2.08%	299 77.56%	385.5 100%
Knoxville, TN	639	268	27	13.5 3.78%	71 19.86%	3 0.84%	7 1.96%	263 73.57%	357.5 100%
Las Vegas, NV/AZ	714	325	48	24.2 6.63%	91 24.92%	4 1.10%	29 7.94%	217 59.42%	365.2 100%
Los Angeles, CA	738	338	46	21.8 5.80%	85 22.62%	15 3.99%	381 0.11%	216 57.48%	375.8 100%
Middlesex, NJ	667	283	40	19.3 5.31%	76 20.92%	3 0.83%	18 4.95%	247 67.99%	363.3 100%
Milwaukee, WI	621	241	20	11.3 3.04%	82 22.08%	4 1.08%	13 3.50%	261 70.29%	371.3 100%

	Telephone Numbers Released	Ineligible Telephone Numbers	Non-contacts	Estimated Households: Noncontacts Estimated to be Households	Household Refusals	Language/ Other Barriers	Maximum Contacts	Completed Household Enumeration	Total
Low-Intensity Sites *(continued)*									
Minneapolis-St. Paul, MN/WI	620	256	23	10.9	57	3	5	276	351.9
				3.10%	16.20%	0.85%	1.42%	78.43%	100%
Modesto, CA	719	318	23	11.8	100	7	11	260	389.8
				3.03%	25.65%	1.80%	2.82%	66.70%	100%
Nassau, NY	689	223	35	20.9	118	6	27	280	451.9
				4.62%	26.11%	1.33%	5.97%	61.96%	100%
New York City, NY	785	265	65	38.1	119	33	571	246	493.1
				7.73%	24.13%	6.69%	1.56%	49.89%	100%
Philadelphia, PA/NJ	609	202	39	21.7	87	4	18	259	389.7
				5.57%	22.32%	1.03%	4.62%	66.46%	100%
Pittsburgh, PA	591	192	33	17.6	92	3	12	259	383.6
				4.59%	23.98%	0.78%	3.13%	67.52%	100%
Portland-Salem, OR/WA	704	309	31	15.2	86	2	9	267	379.2
				4.01%	22.68%	0.53%	2.37%	70.41%	100%
Riverside-San Bernardino, CA	691	293	41	21.4	74	3	22	258	378.4
				5.66%	19.56%	0.79%	5.81%	68.18%	100%
Rochester, NY	696	264	37	18.4	74	3	17	301	413.4
				4.45%	17.90%	0.73%	4.11%	72.81%	100%
San Antonio, TX	711	289	34	17.2	85	3	33	267	405.2
				4.24%	20.98%	0.74%	8.14%	65.89%	100%
San Francisco, CA	891	420	71	31.4	102	25	43	230	431.4
				7.28%	23.64%	5.80%	9.97%	53.31%	100%
Santa Rosa, CA	640	231	39	23.1	97	4	27	242	393.1
				5.88%	24.68%	1.02%	6.87%	61.56%	100%
Shreveport, LA	639	278	31	13.6	53	2	17	258	343.6
				3.96%	15.42%	0.58%	4.95%	75.09%	100%
St. Louis, MO/IL	652	260	33	16.7	68	2	9	280	375.7
				4.45%	18.10%	0.53%	2.40%	74.53%	100%
Tampa, FL	729	328	35	18.8	104	3	18	241	384.8
				4.89%	27.03%	0.78%	4.68%	62.63%	100%
Tulsa, OK	748	339	30	15.4	98	2	20	259	394.4
				3.90%	24.85%	0.51%	5.07%	65.67%	100%
Washington, DC/MD	703	309	35	16.4	71	6	26	256	375.4
				4.37%	18.91%	1.60%	6.93%	68.19%	100%

	Telephone Numbers Released	Ineligible Telephone Numbers	Non-contacts	Estimated Households: Noncontacts Estimated to be Households	Estimated Households: Household Refusals	Estimated Households: Language/ Other Barriers	Estimated Households: Maximum Contacts	Estimated Households: Completed Household Enumeration	Estimated Households: Total
Low-Intensity Sites *(continued)*									
West Palm Beach, FL	673	307	11	6.1 1.69%	105 29.08%	6 1.66%	381 0.52%	206 57.05%	361.1 100%
Worcester, MA	602	173	29	17.8 4.26%	98 23.46%	1 0.24%	18 4.31%	283 67.74%	417.8 100%
Dothan, AL	630	251	16	8.2 2.21%	78 21.01%	5 1.35%	5 1.35%	275 74.08%	371.2 100%
Terre Haute, IN	782	440	16	5.9 1.78%	68 20.49%	0 0.00%	6 1.81%	252 75.93%	331.9 100%
Wilmington, NC	596	245	26	14.4 4.24%	40 11.79%	0 0.00%	21 6.19%	264 77.78%	339.4 100%
West Central Alabama	656	283	28	12.7 3.55%	57 15.94%	9 2.52%	9 2.52%	270 75.48%	357.7 100%
Central Arkansas	750	293	27	13.7 3.09%	70 15.78%	2 0.45%	16 3.61%	342 77.08%	443.7 100%
West Georgia	529	206	25	13.2 4.24%	53 17.03%	2 0.64%	11 3.53%	232 74.55%	311.2 100%
North East Illinois	805	414	25	8.9 2.37%	98 26.14%	1 0.27%	4 1.07%	263 70.15%	374.9 100%
North East Indiana	613	258	17	11.3 3.24%	66 18.89%	2 0.57%	10 2.86%	260 74.43%	349.3 100%
East Maine	620	263	25	11.2 3.26%	35 10.20%	2 0.58%	9 2.62%	286 83.33%	343.2 100%
East North Carolina	754	396	37	19.3 5.67%	38 11.17%	4 1.18%	9 2.64%	270 79.34%	340.3 100%
West Utah	759	335	36	14.9 3.70%	54 13.40%	3 0.74%	8 1.99%	323 80.17%	402.9 100%
North West Washington	766	356	32	15.5 3.94%	83 21.09%	4 1.02%	12 3.05%	279 70.90%	393.5 100%
Supplemental Sample									
	7,175	2,912	333	163.8 4.00%	862 21.06%	61 1.49%	192 4.69%	2815 68.76%	4093.8 100%

TABLE IV.9

DISPOSITION OF THE RDD FIU INTERVIEW SAMPLE, BY CTS SITE, SUPPLEMENTAL SAMPLE, AND TOTAL SAMPLE: PRIMARY FIUs

			Undetermined FIU Eligibility[a]			Eligible Primary FIUs			
	All Primary FIUs	Ineligible	Refusal	Language/ Other Barrier	Maximum Calls	Undetermined Estimated Eligible	Eligible Refusal	Completed FIU Interviews	Total
Total RDD Sample									
	27,381	332	11	61	230	298.5 1.1%	470 1.74%	26,277 97.16%	27,045.5 100%
High-Intensity Sites									
Boston, MA	930	12	0	2	9	10.9 1.19%	16 1.74%	891 97.07%	917.9 100%
Cleveland, OH	1,014	13	1	5	3	8.9 0.89%	14 1.40%	978 97.71%	1,000.9 100%
Greenville, SC	1,059	5	1	2	10	12.9 1.22%	24 2.28%	1,017 96.50%	1,053.9 100%
Indianapolis, IN	1,045	5	0	1	3	4.0 0.38%	14 1.35%	1,022 98.27%	1,040.0 100%
Lansing, MI	1,059	28	0	0	4	3.9 0.38%	5 0.49%	1,022 99.14%	1,030.9 100%
Little Rock, AR	1,148	14	0	2	8	9.9 0.87%	13 1.15%	1,111 97.98%	1,133.9 100%
Miami, FL	894	9	1	6	23	29.7 3.36%	24 2.71%	831 93.93%	884.7 100%
Newark, NJ	979	8	1	3	18	21.8 2.25%	25 2.58%	924 95.18%	970.8 100%
Orange County, CA	933	13	0	2	12	13.8 1.50%	16 1.74%	890 96.76%	919.8 100%
Phoenix, AZ	976	14	0	2	5	6.9 0.72%	13 1.35%	942 97.93%	961.9 100%
Seattle, WA	970	15	0	3	7	9.8 1.03%	13 1.36%	932 97.61%	954.8 100%
Syracuse, NY	1,048	8	1	1	6	7.9 0.76%	20 1.92%	1,012 97.32%	1,039.9 100%

	All Primary FIUs	Ineligible	Undetermined FIU Eligibility[a] Refusal	Language/ Other Barrier	Maximum Calls	Eligible Primary FIUs Undetermined Estimated Eligible	Eligible Refusal	Completed FIU Interviews	Total
Low-Intensity Sites									
Atlanta, GA	254	2	0	0	5	5.0 1.98%	4 1.59%	243 96.43%	252 100%
Augusta-Aiken, GA/SC	254	6	0	0	0	0.0 0.00%	7 2.82%	241 97.18%	248 100%
Baltimore, MD	233	1	0	1	3	4.0 1.72%	3 1.29%	225 96.98%	232 100%
Bridgeport, CT	247	2	0	3	7	9.9 4.04%	8 3.27%	227 92.69%	244.9 100%
Chicago, IL	256	0	0	0	0	0.0 0.00%	9 3.52%	247 96.48%	256 100%
Columbus, OH	265	5	0	0	3	2.9 1.12%	2 0.77%	255 98.11%	259.9 100%
Denver CO	256	1	0	1	1	2.0 0.78%	3 1.18%	250 98.04%	255 100%
Detroit, MI	265	1	0	0	3	3.0 1.14%	8 3.03%	253 95.83%	264 100%
Greensboro, NC	238	0	0	1	3	4.0 1.68%	3 1.26%	231 97.06%	238 100%
Houston, TX	243	2	0	0	3	3.0 1.24%	3 1.24%	235 97.51%	241 100%
Huntington-Ashland, WV-KY-OH	253	1	0	0	0	0.0 0.00%	4 1.59%	248 98.41%	252 100%
Killeen, TX	299	25	0	1	4	4.6 1.68%	9 3.29%	260 95.03%	273.6 100%
Knoxville, TN	263	2	0	0	0	0.0 0.00%	5 1.92%	256 98.08%	261 100%
Las Vegas, NV/AZ	217	1	0	0	0	0.0 0.00%	3 1.39%	213 98.61%	216 100%
Los Angeles, CA	216	1	0	3	5	8.0 3.72%	5 2.33%	202 93.95%	215 100%
Middlesex, NJ	247	3	0	0	2	2.0 0.82%	3 1.23%	239 97.95%	244 100%

	All Primary FIUs	Ineligible	Undetermined FIU Eligibility[a]: Refusal	Language/ Other Barrier	Maximum Calls	Eligible Primary FIUs: Undetermined Estimated Eligible	Eligible Refusal	Completed FIU Interviews	Total
Low-Intensity Sites *(continued)*									
Milwaukee, WI	261	3	0	1	1	2.0 0.78%	3 1.16%	253 98.06%	258 100%
Minneapolis-St. Paul, MN/WI	276	4	0	1	1	2.0 0.74%	2 0.74%	268 98.53%	272 100%
Modesto, CA	260	2	0	0	1	1.0 0.39%	6 2.33%	251 97.29%	258 100%
Nassau, NY	280	3	0	0	3	3.0 1.08%	6 2.17%	268 96.75%	277 100%
New York City, NY	246	4	0	1	7	7.9 3.27%	8 3.31%	226 93.43%	241.9 100%
Philadelphia, PA/NJ	259	3	0	0	2	2.0 0.78%	2 0.78%	252 98.44%	256 100%
Pittsburgh, PA	259	1	0	0	1	1.0 0.39%	6 2.33%	251 97.29%	258 100%
Portland-Salem, OR/WA	267	5	0	1	3	3.9 1.49%	2 0.76%	256 97.75%	261.9 100%
Riverside-San Bernardino, CA	258	3	0	0	5	4.9 1.92%	4 1.57%	246 96.51%	254.9 100%
Rochester, NY	301	0	0	0	1	1.0 0.33%	6 1.99%	294 97.67%	301 100%
San Antonio, TX	267	6	0	0	4	3.9 1.49%	4 1.53%	253 96.97%	260.9 100%
San Francisco, CA	230	3	0	0	2	2.0 0.88%	6 2.64%	219 96.48%	227 100%
Santa Rosa, CA	242	0	0	0	3	3.0 1.24%	2 0.83%	237 97.93%	242 100%
Shreveport, LA	258	6	0	0	1	1.0 0.40%	8 3.17%	243 96.43%	252 100%
St. Louis, MO/IL	280	1	0	0	1	1.0 0.36%	4 1.43%	274 98.21%	279 100%
Tampa, FL	241	1	0	0	5	5.0 2.08%	4 1.67%	231 96.25%	240 100%

		Undetermined FIU Eligibility[a]				Eligible Primary FIUs			
	All Primary FIUs	Ineligible	Refusal	Language/ Other Barrier	Maximum Calls	Undetermined Estimated Eligible	Eligible Refusal	Completed FIU Interviews	Total
Low-Intensity Sites *(continued)*									
Tulsa, OK	259	1	1	0	1	2.0 0.78%	5 1.94%	251 97.29%	258 100%
Washington, DC	256	3	0	1	2	3.0 1.19%	5 1.98%	245 96.84%	253 100%
West Palm Beach, FL	206	0	0	0	3	3.0 1.46%	3 1.46%	200 97.09%	206 100%
Worcester, MA	283	4	0	2	4	5.9 2.12%	8 2.87%	265 95.02%	278.9 100%
Dothan, AL	275	9	0	1	2	2.9 1.09%	11 4.14%	252 94.77%	265.9 100%
Terre Haute, IN	252	8	0	2	0	1.9 0.78%	2 0.82%	240 98.40%	243.9 100%
Wilmington, NC	264	2	0	0	0	0.0 0.00%	5 1.91%	257 98.09%	262 100%
West Central Alabama	270	2	0	1	1	2.0 0.75%	10 3.73%	256 95.52%	268 100%
Central Arkansas	342	3	0	1	0	1.0 0.29%	3 0.88%	335 98.82%	339 100%
West Georgia	232	1	0	1	0	1.0 0.43%	3 1.30%	227 98.27%	231 100%
North East Illinois	263	0	1	0	0	1.0 0.38%	6 2.28%	256 97.34%	263 100%
North East Indiana	260	2	0	0	4	4.0 1.55%	3 1.16%	251 97.29%	258 100%
East Maine	286	4	0	0	2	2.0 0.71%	6 2.13%	274 97.16%	282 100%
East North Carolina	270	4	0	1	4	4.9 1.84%	4 1.50%	257 96.65%	265.9 100%
West Utah	323	1	1	0	0	1.0 0.31%	1 0.31%	320 99.38%	322 100%
North West Washington	279	3	0	0	0	0.0 0.00%	1 0.36%	275 99.64%	276 100%

	All Primary FIUs	Undetermined FIU Eligibility[a]				Eligible Primary FIUs			
		Ineligible	Refusal	Language/ Other Barrier	Maximum Calls	Undetermined Estimated Eligible	Eligible Refusal	Completed FIU Interviews	Total
Supplemental Sample									
	2,815	43	3	8	19	29.5	45	2697	2771.5
						1.06%	1.62%	97.31%	100%

[a]Household composition questions needed to determine eligibility for the survey were not completed.

TABLE IV.10

DISPOSITION OF THE RDD FIU INTERVIEW SAMPLE, BY CTS SITE, SUPPLEMENTAL SAMPLE, AND TOTAL SAMPLE: SECONDARY FIUs

			Undetermined FIU Eligibility[a]		Eligible Secondary FIUs				
	All Secondary FIUs	Ineligible	Moved/ Nonworking Number	Maximum Calls	Undetermined Estimated Eligible	Eligible Refusal	Language/ Other Barrier	Completed FIU Interviews	Total
Total RDD Sample									
	7,554	522	453	196	600.2	504	59	5,820	6,983.2
					8.59%	7.22%	0.84%	83.34%	100%
High-Intensity Sites									
Boston, MA	328	27	22	12	30.9	24	6	237	297.9
					10.37%	8.06%	2.01%	79.56%	100%
Cleveland, OH	252	22	16	4	18.1	19	0	191	228.1
					7.94%	8.33%	0.00%	83.74%	100%
Greenville, SC	257	25	15	3	16.1	13	1	200	230.1
					7.00%	5.65%	0.43%	86.92%	100%
Indianapolis, IN	229	13	4	5	8.5	7	2	198	215.5
					3.94%	3.25%	0.93%	91.88%	100%
Lansing, MI	235	17	13	3	14.8	14	2	186	216.8
					6.83%	6.46%	0.92%	85.79%	100%
Little Rock, AR	268	28	12	4	14.2	9	0	215	238.2
					5.96%	3.78%	0.00%	90.26%	100%
Miami, FL	394	21	32	15	44.2	24	4	298	370.2
					11.94%	6.48%	1.08%	80.50%	100%
Newark, NJ	384	20	18	7	23.6	34	7	298	362.6
					6.51%	9.38%	1.93%	82.18%	100%
Orange County, CA	359	30	30	16	41.6	25	5	253	324.6
					12.82%	7.70%	1.54%	77.94%	100%
Phoenix, AZ	280	23	9	8	15.5	12	0	228	255.5
					6.07%	4.70%	0.00%	89.24%	100%
Seattle, WA	264	16	17	9	24.3	25	3	194	246.3
					9.87%	10.15%	1.22%	78.77%	100%
Syracuse, NY	295	20	13	4	15.8	17	1	240	273.8
					5.77%	6.21%	0.37%	87.66%	100%

	All Secondary FIUs	Ineligible	Undetermined FIU Eligibility[a]: Moved/ Nonworking Number	Undetermined FIU Eligibility[a]: Maximum Calls	Eligible Secondary FIUs: Undetermined Estimated Eligible	Eligible Secondary FIUs: Eligible Refusal	Eligible Secondary FIUs: Language/ Other Barrier	Eligible Secondary FIUs: Completed FIU Interviews	Eligible Secondary FIUs: Total
Low-Intensity Sites									
Atlanta, GA	73	7	7	1	7.1 10.91%	5 7.68%	0 0.00%	53 81.41%	65.1 100%
Augusta-Aiken, GA/SC	62	3	5	1	5.7 9.71%	3 5.11%	0 0.00%	50 85.18%	58.7 100%
Baltimore, MD	69	2	1	0	1.0 1.49%	6 8.96%	0 0.00%	60 89.55%	67 100%
Bridgeport, CT	75	6	3	3	5.5 8.03%	6 8.76%	0 0.00%	57 83.21%	68.5 100%
Chicago, IL	64	2	5	0	4.8 7.77%	10 16.18%	1 1.62%	46 74.43%	61.8 100%
Columbus, OH	62	4	8	1	8.3 14.49%	7 12.22%	1 1.75%	41 71.55%	57.3 100%
Denver, CO	52	2	6	0	5.7 11.47%	3 6.04%	0 0.00%	41 82.49%	49.7 100%
Detroit, MI	79	5	9	3	11.1 15.18%	6 8.21%	0 0.00%	56 76.61%	73.1 100%
Greensboro, NC	50	3	4	1	4.7 10.06%	2 4.28%	0 0.00%	40 85.65%	46.7 100%
Houston, TX	65	3	10	4	13.2 21.57%	3 4.90%	0 0.00%	45 73.53%	61.2 100%
Huntington-Ashland, WV-KY-OH	69	3	3	0	2.9 4.40%	4 6.07%	0 0.00%	59 89.53%	65.9 100%
Killeen, TX	55	7	3	2	4.3 9.09%	4 8.46%	1 2.11%	38 80.34%	47.3 100%
Knoxville, TN	64	2	4	0	3.9 6.30%	3 4.85%	0 0.00%	55 88.85%	61.9 100%
Las Vegas, NV/AZ	77	3	12	4	15.2 20.77%	4 5.46%	0 0.00%	54 73.77%	73.2 100%
Los Angeles, CA	75	2	4	5	8.7 11.97%	5 6.88%	0 0.00%	59 81.16%	72.7 100%
Middlesex, NJ	88	5	2	3	4.7 5.68%	6 7.26%	0 0.00%	72 87.06%	82.7 100%
Milwaukee, WI	72	9	3	1	3.5 5.60%	1 1.60%	0 0.00%	58 92.80%	62.5 100%

	All Secondary FIUs	Ineligible	Undetermined FIU Eligibility[a]: Moved/ Nonworking Number	Undetermined FIU Eligibility[a]: Maximum Calls	Eligible Secondary FIUs: Undetermined Estimated Eligible	Eligible Secondary FIUs: Eligible Refusal	Eligible Secondary FIUs: Language/ Other Barrier	Eligible Secondary FIUs: Completed FIU Interviews	Eligible Secondary FIUs: Total
Low-Intensity Sites *(continued)*									
Minneapolis-St. Paul, MN/WI	82	3	6	1	6.7 8.51%	5 6.35%	1 1.27%	66 83.86%	78.7 100%
Modesto, CA	67	4	4	1	4.7 7.50%	2 3.19%	1 1.59%	55 87.72%	62.7 100%
Nassau, NY	98	9	5	1	5.4 6.11%	9 10.18%	1 1.13%	73 82.58%	88.4 100%
New York City, NY	96	6	9	4	12.1 13.58%	9 10.10%	2 2.24%	66 74.07%	89.1 100%
Philadelphia, PA/NJ	70	1	1	5	5.9 8.56%	6 8.71%	0 0.00%	57 82.73%	68.9 100%
Pittsburgh, PA	62	4	3	0	2.8 4.84%	7 12.11%	0 0.00%	48 83.04%	57.8 100%
Portland-Salem, OR/WA	70	2	4	3	6.8 10.03%	8 11.80%	2 2.95%	51 75.22%	67.8 100%
Riverside-San Bernardino, CA	75	7	2	3	4.5 6.67%	4 5.93%	1 1.48%	58 85.93%	67.5 100%
Rochester, NY	79	8	2	2	3.6 5.10%	6 8.50%	0 0.00%	61 86.40%	70.6 100%
San Antonio, TX	68	12	2	1	2.4 4.33%	6 10.83%	1 1.81%	46 83.03%	55.4 100%
San Francisco, CA	89	5	6	4	9.4 11.27%	10 11.99%	2 2.40%	62 74.34%	83.4 100%
Santa Rosa, CA	75	5	10	5	13.8 20.06%	7 10.17%	0 0.00%	48 69.77%	68.8 100%
Shreveport, LA	77	7	6	2	7.2 10.40%	6 8.67%	1 1.45%	55 79.48%	69.2 100%
St. Louis, MO/IL	68	8	6	5	9.5 16.24%	4 6.84%	1 1.71%	44 75.21%	58.5 100%
Tampa, FL	51	2	3	0	2.9 5.93%	9 18.40%	0 0.00%	37 75.66%	48.9 100%
Tulsa, OK	52	3	0	2	1.9 3.89%	5 10.22%	1 2.04%	41 83.84%	48.9 100%
Washington, DC/MD	79	6	2	2	3.7 5.09%	4 5.50%	0 0.00%	65 89.41%	72.7 100%

	All Secondary FIUs	Ineligible	Undetermined FIU Eligibility[a]: Moved/ Nonworking Number	Undetermined FIU Eligibility[a]: Maximum Calls	Eligible Secondary FIUs: Undetermined Estimated Eligible	Eligible Secondary FIUs: Eligible Refusal	Eligible Secondary FIUs: Language/ Other Barrier	Eligible Secondary FIUs: Completed FIU Interviews	Eligible Secondary FIUs: Total
Low-Intensity Sites *(continued)*									
West Palm Beach, FL	80	5	13	2	13.8 18.70%	7 9.49%	0 0.00%	53 71.82%	73.8 100%
Worcester, MA	56	2	2	3	4.8 8.92%	4 7.43%	0 0.00%	45 83.64%	53.8 100%
Dothan, AL	59	4	2	1	2.8 5.11%	3 5.47%	0 0.00%	49 89.42%	54.8 100%
Terre Haute, IN	68	6	3	1	3.6 5.84%	4 6.49%	1 1.62%	53 86.04%	61.6 100%
Wilmington, NC	63	5	3	2	4.6 7.99%	7 12.15%	0 0.00%	46 79.86%	57.6 100%
West Central Alabama	88	5	2	6	7.5 9.09%	2 2.42%	0 0.00%	73 88.48%	82.5 100%
Central Arkansas	51	4	2	0	1.8 3.85%	1 2.14%	0 0.00%	44 94.02%	46.8 100%
West Georgia	60	2	4	2	5.8 10.03%	5 8.65%	1 1.73%	46 79.58%	57.8 100%
North East Illinois	43	3	0	0	0.0 0.00%	2 5.00%	0 0.00%	38 95.00%	40 100%
North East Indiana	46	3	1	2	2.8 6.54%	4 9.35%	1 2.34%	35 81.78%	42.8 100%
East Maine	51	1	0	0	0.0 0.00%	3 6.00%	2 4.00%	45 90.00%	50 100%
East North Carolina	62	7	2	3	4.4 8.09%	3 5.51%	0 0.00%	47 86.40%	54.4 100%
West Utah	72	5	10	0	9.2 13.90%	0 0.00%	0 0.00%	57 86.10%	66.2 100%
North West Washington	64	2	3	0	2.9 4.68%	3 4.85%	1 1.62%	55 88.85%	61.9 100%
Supplemental Sample									
	737	46	45	14	55 8.01%	48 6.99%	5 0.73%	579 84.28%	687 100%

[a] Secondary FIUs could not be formed until the primary FIU interview was completed. However, persons in some secondary FIUs moved o ut of the household or would not ne contacted for other reasons before their eligibility in the survey could be verified.

TABLE IV.11

DISPOSITION OF THE FIELD SAMPLE

	Total Released	Not a Residence	Household Enumeration Completed, Ineligible Residence	Household Enumeration Not Completed: Never at Home	Household Enumeration Not Completed: Refused Household Enumeration	Household Enumeration Not Completed: Language/ Other Barrier	Household Enumeration Not Completed: Other Household Enumeration Non-complete	Household Enumeration Not Completed: Total Household Enumeration Non-completes	Eligible Households: Household Enumeration Non-completes Estimated Eligible	Eligible Households: Household Enumeration Completed, Eligible, Refused Interview	Eligible Households: Completed Interview[a]	Eligible Households: Total Estimated Eligible Households	Ineligible FIUs	Eligible FIU Interviews
All Sites														
	5,258	385	4,041	99	155	29	29	312	53.1 9.27%	50 8.72%	470 82.01%	573.1 100%	21	635
High-Intensity Sites														
Boston, MA	162	30	103	14	5	1	0	20	1.6 15.15%	1 9.43%	8 75.42%	10.6 100%	1	17
Cleveland, OH	237	14	158	7	13	3	2	25	5.1 11.21%	7 15.54%	33 73.25%	45.1 100%	0	42
Greenville, NC	1,374	74	1,196	13	26	1	4	44	2.1 3.38%	8 12.88%	52 83.73%	62.1 100%	2	68
Indianapolis, IN	267	27	150	9	4	0	4	17	5.6 7.08%	3 3.82%	70 89.10%	78.6 100%	4	96
Lansing, MI	219	19	169	3	6	0	1	10	1.1 5.00%	5 22.62%	16 72.38%	22.1 100%	1	24
Little Rock, AR	494	34	381	7	4	0	0	11	1.7 2.39%	5 7.18%	63 90.43%	69.7 100%	2	86
Miami, FL	540	35	470	0	4	2	0	6	0.3 1.19%	1 3.41%	28 95.40%	29.3 100%	1	42
Newark, NJ	215	29	85	5	43	3	0	51	19 27.42%	2 2.90%	48 69.68%	68.9 100%	0	60
Orange County, CA	709	26	609	15	35	11	1	62	1.2 9.08%	5 37.88%	7 53.04%	13.2 100%	0	14
Phoenix, AZ	563	28	469	1	2	5	1	9	1 1.68%	3 5.17%	54 93.14%	58 100%	7	80
Seattle, WA	303	22	170	23	12	3	16	54	14 19.22%	8 11.34%	49 69.44%	70.6 100%	2	55
Syracuse, NY	175	47	81	2	1	0	0	3	1.1 2.34%	2 4.44%	42 93.22%	45.1 100%	1	51

[a] For the field sample, the household informant completed all FIU interviews, except for self-response modules. Therefore, the FIU response rate is equal to the household interview response rate.

percent), noncontacts estimated to be residential households (4.5 percent), and language and other barriers (1.5 percent) (Table IV.8). Interviews were conducted in English or Spanish; no interviews were attempted in households in which no one spoke either of those languages. In addition, we were unable to complete interviews in a few FIUs because identified family informants were too impaired to be interviewed.

We observed moderate differences among sites in household-level refusal rates (which ranged from 15.2 to 25.3 percent in high-intensity sites and from 10.2 to 29.0 percent in low-intensity sites), reflecting intensive efforts to minimize refusals through follow-up efforts and incentives (Table IV.8). However, among some sites, the percentage of other reasons for nonresponse varied considerably. The low response rate in Miami was mainly due to a very high percentage (12.2 percent) of retired residential households (that is, confirmed residential households that neither completed nor refused interviews after 40 calls); in contrast, the average for the entire RDD sample was 4.8 percent. We observed a similar pattern in other areas of Florida and thought that it might have resulted, in part, from a "snowbird" effect (that is, household informants may have been contacted in the winter, delayed the interview, and then returned to residences in other parts of the country). In these cases, the households would have been coded as nonresponses, based on the initial contact. For the second round of the household survey, this problem will be mitigated by adding screening questions on year-round residence. The problem may also have been exacerbated by households that neither refused nor completed interviews after 40 attempts. Some of these nonrespondents may have been reluctant to be interviewed but also did not want to offend interviewers by refusing.

We also observed relatively high rates of retired residential households in Newark (6.3 percent) and Orange County (7.0 percent) (Table IV.8). Here, the problem was due mainly to increased difficulty achieving contact. Because we left messages on answering machines and called at various times, it is unlikely that this source of nonresponse could be significantly reduced by making changes in field procedures.

Overall, we lost only 1.5 percent of the household-level RDD sample to language and other barriers (Table IV.8). This source of nonresponse was slightly higher (two to four percent) in Boston, Miami, Newark, Orange County, and Seattle (all high-intensity sites), as well as in some low-intensity sites. Languages other than English and Spanish may have been more common in those areas; it also is possible that some interviewers did not recognize some Spanish dialects and assumed that other languages were spoken. Households in which languages other than English or Spanish were spoken were reviewed by supervisors, and interviewers were instructed to ask to speak to someone in the household who spoke English or Spanish. It would be necessary to translate the survey into other languages, at very high cost, to reduce this source of nonresponse further.

The other source of nonresponse was noncontacted telephone numbers estimated to be residential households (4.5 percent of the estimated RDD household screening sample; Table IV.8). We made up to 12 calls to contact a telephone number and then called local telephone companies to determine whether these telephone numbers were residential. Most of these efforts were unsuccessful, and we attempted another 8 calls before retiring a telephone number as unresolved if no contact was made after 20 calls. In most sites, we were able to limit this source of nonresponse to less than five percent; however, some sites (for example, Boston, Miami, Newark, and Orange County) were in the range of five to seven percent. We also observed higher noncontact rates among very large metropolitan areas selected as low-intensity sites. It is unlikely that we could reduce this rate significantly without obtaining more cooperation from telephone companies. As discussed later in this chapter, this added cooperation seems unlikely.

Overall, 94.4 percent of FIUs formed from eligible households completed interviews (see Table IV.5). Among high-intensity sites, there was some variability, from a low of 90.3 percent in Miami to a high of 97.4 percent in Indianapolis.

For the RDD sample, the "primary" FIU was the one to which the person completes the household enumeration questions belonged. Usually, this was the householder or householder's spouse, but it could be eligible adults forming other FIUs. Interviews with "secondary" FIUs were scheduled after the primary FIU interview was completed.

Nearly all (97.2 percent) primary FIU interviews in the RDD sample were completed (see Table IV.9). Because interviews with secondary FIUs had to be scheduled after the primary FIU interview was completed, a smaller percentage of these FIUs (83.3 percent) completed their interviews, with the main reasons divided between refusals (7.2 percent) and persons who moved ("movers") or who did not respond after many calls (8.6 percent estimated to be eligible (see Table IV.10). We also include as movers secondary FIUs that were retired because their household's telephone numbers were disconnected before the interviews could be completed. Mobility was a problem because eligibility was determined at household enumeration (except for errors in initial enumeration that were discovered later). We believe that nonresponse by individuals in secondary FIUs can be reduced in the future by making greater efforts, through changes in CATI program design, to minimize the length between household enumeration and secondary FIU interviewing.

As noted, response rates to the field survey were generally higher than for the RDD survey, with an overall response rate of 82.0 percent (Table IV.11). Refusals, followed by chronic noncontacts, were the main sources of nonresponse to interviews completed from the field sample. Many of the initial nonresponses were located in gated communities. In most buildings, we were unable to obtain access after efforts to contact the building manager or owner by telephone and mail. Where we could not obtain access, we used reverse address directories to contact households with published telephone numbers. Only one site, Orange County (53.0 percent), had a response rate below 69 percent; this site also had a very low sample allocation resulting from a very high telephone penetration rate. Thus, the low response rate may have been a function of sampling variability.

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, including subjective assessments of health, tobacco use, satisfaction with care, and aspects of the physician–patient interaction. Efforts to obtain self-responses were successful, as 95.6 percent of adults in the total sample, including 95.6 percent in the RDD sample and 97.8 percent in the field sample, completed these questions (Table IV.12). The family informant was allowed to complete the self-response module in certain circumstances--when an adult FIU member was too ill to respond, was temporarily unavailable, or was unwilling to respond after several interviewing efforts had been made. Overall, only 1.2 percent of the self-response modules were completed by proxy respondents.

TABLE IV.12

RESPONSE RATES FOR THE ADULT SELF-RESPONSE MODULE

	RDD Sample	Field Sample	Total
Completed Module (Percent)	95.6	97.8	95.6
Proxy Accepted (Percent)	1.2	1.8	1.2
Refusal/Not Available (Percent)	3.2	0.4	3.1
Total (Percent)	100	100	100
Number of Adults	49,077	730	49,807

We also asked the adult who took the sampled child to the last physician visit before the interview to answer questions about that visit; those questions are similar to items included in the adult self-response module. The adult who took the child to the physician may not have been the FIU informant, and we were not always able to obtain these data through follow-up calls to the adult who accompanied the child. Altogether, 84.5 percent of sampled children had one or more physician visits in the last year (data not shown). We obtained information on the last visit for 92.3 percent of children who had such a visit. Completion rates for these questions were virtually identical for the RDD and field samples (Table IV.13). The main reasons for missing data were that the person accompanying the child to the physician was not an FIU member or was not identified by the family informant.

TABLE IV.13

RESPONSE RATES FOR THE MODULE ON
THE CHILD'S LAST PHYSICIAN VISIT

	RDD Sample	Field Sample	Total
Completed Module (Percent)	92.3	92.3	92.3
Person Accompanying Child Unknown or Not in FIU (Percent)	6.5	7.7	6.5
Refusal/Not Available (Percent)	1.2	0.0	1.2
Total (Percent)	100	100	100
Number of Children	8,824	168	8,992

NOTE: Includes children who had one or more physician visits during the 12 months before the interview.

D. EFFORTS TO REDUCE NONRESPONSE

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 reduce nonresponse, including:

- Making up to 20 calls to determine residential status, and up to 40 calls to complete an interview with an FIU
- Making multiple rounds of refusal conversion calls
- Offering monetary incentives
- Using Spanish-speaking interviewers
- Leaving messages on answering machines
- Making calls to telephone companies to ascertain residential status for telephone numbers that were difficult to contact

We have described call rules. In this section, we focus on the other efforts designed to reduce specific sources of nonresponse.

1. Multiple Rounds of Refusal Conversion Calls

During the first few weeks of data collection, cooperation rates (ratios of completed interviews to the sum of completed interviews and initial refusals) averaged about 40 percent. Because these cooperation rates were much lower than we had experienced in other health surveys, we reevaluated our data collection procedures, including:

- A review of response rates in related studies
- An assessment of the effectiveness of survey messages and advance materials (see Chapter III)
- Development of revised survey messages and training materials (see Chapter III)

- Increasing limits on calls (discussed previously in this chapter)

- Incorporating a systematic test of the effect of varied respondent incentives ($0, $15, $25, and $35) on response rates (discussed in the next section)

After testing revised survey messages and various monetary incentive levels, we concluded that intensive refusal conversion efforts would be required to achieve an acceptable response rate. Efforts to convert refusals focused on identifying effective refusal converters, training these individuals to use information on reasons for prior refusals and personal interactions between prior interviewers and informants, and the use of varied messages to respond to specific concerns of potential respondents. We also developed an interviewer bonus plan, based on interview difficulty and longevity, to reduce attrition and to retain the most effective interviewers. Finally, interviewers were allowed to give the name of an official of The Robert Wood Johnson Foundation (RWJF) to respondents who wanted to contact the organization sponsoring the survey.

These efforts were generally successful. For the RDD sample, final refusal rates represented 20.9 percent of the estimated household sample (see Table IV.8). Among FIUs identified by household enumeration questions, only 1.7 percent of the primary family informants and 7.2 percent of the secondary family informants refused to be interviewed (see Tables IV.9 and IV.10, respectively). These results were largely due to refusal conversion efforts for the RDD sample. Overall, of the 32,097 FIU interviews completed from the RDD frame, 9,921 (30.9 percent) refused at least once. The refusal conversion rate (ratio of completed FIU interviews that ever refused to the sum of final refusals plus completed FIU interviews that had ever refused) was 51.5 percent.

Efforts to reduce interviewer attrition were successful. We experienced minimal attrition after introducing the interviewer bonus plan.

Staff from the RWJF recorded nearly 500 calls from persons in sampled households; callers included persons who had completed interviews, as well as those who wanted more information before participating. These calls included requests both for verification about the legitimacy of the survey and for more information about study objectives.

2. Monetary Incentives

One of the actions taken early in the survey was to design an experiment to test the effect of respondent incentives on response rates. The results of this experiment and of a prior experiment conducted for the RWJF Family Health Insurance Survey were presented at the 1997 meeting of the American Association of Public Opinion Research and appeared in the 1997 *ASA Proceedings of the Survey Research Methods Section* (Strouse and Hall, 1997). For the CTS Household Survey experiment, treatments were randomized equally across households in four cells: $0, $15, $25, and $35. The experimental sample was selected to represent all 60 communities and the supplemental sample. Households in the $0, $15, and $25 cells that refused were offered $25 during refusal conversion calls; those initially offered $35 were offered $35 during refusal conversion calls. For households with unpublished telephone numbers, respondent incentives were promised at the initial call and during refusal conversion calls. Households with published addresses that refused were mailed letters and checks before interviewers called. A minimum of eight weeks was allowed between the initial refusal and each round of refusal conversion calls, and two rounds of refusal conversion calls were made.

The experiment showed that incentives had a large impact on initial cooperation, and that the impact was still significant after refusal conversion efforts were made. Initial cooperation rates increased from 38 percent for the nonincentive group to 46 percent for those offered $15, to 52 percent for those offered $25, and to 49 percent for those offered $35.[5] Final cooperation rates after refusal conversions were 61 percent (no payment), 64 percent ($15), and 67 percent ($25). Increasing the incentive to $35 had no additional impact, as the cooperation rate after refusal conversions for that group also was 67 percent. As a result of this experiment, we decided to offer all families $25 to complete the survey.[6]

Some individuals participating in the first round of the Household Survey are (or will be) in households selected for the second round of the survey. In addition, the Household Survey sample is being used as a frame for other surveys. Compensating respondents is expected to increase the likelihood of participation in the followup surveys and in collateral studies.

3. Spanish-Speaking Interviews

We prepared a Spanish version of the CATI instrument and trained bilingual 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. Overall, we interviewed 1,182 FIUs in Spanish (3.7 percent of the total) and completed an additional 54 Spanish self-response modules in FIUs for which the core interview was completed in English (Table IV.14). Spanish interviews were critical in the Miami site, where

---

[5]Here, the cooperation rate is defined as the ratio of completed interviews to all confirmed residential households; telephone numbers for which no contact was made after 20 calls and nonresidential numbers were excluded.

[6]We also tested the impact of including a $25 check with an advance letter on efforts to convert refusals for FIUs with known addresses. Our goal was primarily to reduce the time required to resolve refusals and to maintain the survey schedule. Preliminary results indicated that this goal was met, as household respondents who received $25 checks usually had read the letter and responded to the interviewers' calls. Because relatively few respondents who refused the offer cashed checks, we did not have to bear significant added costs with this method. Final results are being used in round two refusal conversion efforts.

TABLE IV.14

SPANISH-SPEAKING FIU INTERVIEWS, BY SITE AND OVERALL

	Number of Spanish FIUs	Percentage of FIUs Completed in Spanish
Total Overall	**1,182**	**3.7**
High-Intensity Sites		
Boston, MA	31	2.8
Cleveland, OH	7	0.6
Greenville, SC	5	0.4
Indianapolis, IN	1	0.1
Lansing, MI	2	0.2
Little Rock, AR	1	0.1
Miami, FL	342	30.3
Newark, NJ	61	4.5
Orange County, CA	148	13.0
Phoenix, AZ	75	6.4
Seattle, WA	6	0.5
Syracuse, NY	4	0.3
Low-Intensity Sites		
Atlanta, GA	4	1.35
Augusta-Aiken, GA/SC	1	0.34
Baltimore, MD	.	.
Bridgeport, CT	8	2.82
Chicago, IL	10	3.41
Columbus, OH	.	.
Denver, CO	14	4.81
Detroit, MI	2	0.65
Greensboro, NC	.	.
Houston, TX	2	0.74
Huntington-Ashland, WV-KY-OH	30	10.71
Killeen, TX	.	.

	Number of Spanish FIUs	Percentage of FIUs Completed in Spanish
Low-Intensity Sites *(continued)*		
Knoxville, TN	6	2.01
Las Vegas, NV/AZ	.	.
Los Angeles, CA	14	5.24
Middlesex, NJ	50	19.16
Milwaukee, WI	7	2.25
Minneapolis-St. Paul, MN/WI	2	0.64
Modesto, CA	2	0.60
Nassau, NY	39	12.75
New York City, NY	16	4.69
Philadelphia, PA/NJ	39	13.36
Pittsburgh, PA	10	3.24
Portland-Salem, OR/WA	.	.
Riverside, CA	11	3.58
Rochester, NY	26	8.55
San Antonio, TX	3	0.85
San Francisco, CA	16	5.35
Santa Rosa, CA	13	4.63
Shreveport, LA	14	4.91
St. Louis, MO/IL	.	.
Tampa, FL	.	.
Tulsa, OK	.	.
Washington, DC/MD	.	.
West Palm Beach, FL	.	.
Worcester, MA	12	3.87
Dothan, AL	10	3.95
Terre Haute, IN	.	.
Wilmington, NC	3	0.97
West Central Alabama	.	.

	Number of Spanish FIUs	Percentage of FIUs Completed in Spanish
Low-Intensity Sites *(continued)*		
Central Arkansas	1	0.26
Northern Georgia	6	2.20
Northeast Illinois	1	0.34
Northeast Indiana	.	.
Eastern Maine	.	.
Eastern North Carolina	4	1.32
Northern Utah	8	2.12
Northwest Washington	4	1.21
Supplemental Sample	**111**	**3.39**

NOTE: In addition, 54 Spanish self-response modules were completed by FIUs whose core interviews were conducted in English.

PMSA = Primary Metropolitan Statistical Area.

where 30.3 percent of the interviews were conducted in that language. Spanish interviews were also important in three other sites: (1) Newark (5.0 percent), (2) Orange County (13.0 percent), and (3) Phoenix (6.4 percent). Review of CATI reports showing cooperation rates for Spanish-speaking samples found that these cooperation rates did not vary appreciably from rates in English-speaking samples, either overall or within sites.

4. Messages on Answering Machines

Some residential households were difficult to contact because answering machines were used to screen calls. Initially, we instructed interviewers not to leave messages, fearing that this action would increase refusals; however, we later revised the procedure to counter chronic no answers and to offer the $25 incentive to all FIUs. Interviewers were instructed to leave the following message:

> My name is ______. I'm calling on behalf of (fill in name of state health department if endorsement obtained) and a nonprofit foundation. We would like your family to take part in a telephone interview for a major health study. We know how busy you are, so we will pay you [amount] for helping us. I want to assure you that we're not selling anything or asking for money. I'll call back another time to explain the study and see if you can set up a time for the interview. Thanks.

The interviewer was instructed to leave notes in the CATI system indicating that the message had been left on the machine. The interviewer also was instructed 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.

We did not systematically test this procedure. However, we believe it was a useful, low-cost approach to increase cooperation for an RDD survey that included a monetary incentive.

5. Local Telephone Companies

Overall, we were unable to verify residential status by speaking to a person or obtaining confirmation from telephone companies for 5.1 percent of the RDD sample of released telephone numbers (3,624/70,936). Using the imputation procedures described previously, we estimated that 50.1 percent (1,815/3,624) of these telephone numbers were residential. These sample points represented 4.5 percent of the estimated number of residential households in the RDD sample frame (see Table IV.8).

Studies based on other RDD surveys have indicated that the actual percentage of occupied residential households represented by telephone numbers that were chronic no answers after 12 to 20 calls may be lower than 50 percent. Communication with staff at Abt Associates, who investigated this problem for the National Immunization Survey (NIS) for the Centers for Disease Control, reinforced our assumption. Unfortunately, local telephone companies are the only source that can verify the residential status of chronic no answers. Staff at Abt indicated that it had become very difficult to obtain cooperation from telephone company business offices. However, the CTS was requesting verification for far fewer telephone numbers then was the NIS, so we were hopeful about our ability to obtain this information.

After we had made at least 12 calls for about two-thirds of the RDD sample, we produced listings of telephone numbers with chronic no answers (12 calls), sorted by local telephone company. We contacted local telephone company business offices, explained our objectives, responded to any confidentiality concerns, expressed our willingness to compensate staff for time required to obtain the information, and mailed or faxed customized letters and descriptions of the study to designated personnel. Over a two- to three-month period, an MPR survey manager followed up with additional calls and responded to concerns.

Altogether, we contacted 19 telephone companies representing 2,467 telephone numbers. (We eliminated companies associated with few chronic no answers.) Unfortunately, we were able to obtain results from only six companies, representing 319 telephone numbers. Of the 292 telephone numbers for which the companies could verify status, only 32.4 percent were considered to be active residential numbers. The rest were assigned to businesses or pay telephones or were nonworking or inactive. Although these results supported our expectations, the samples were small and unrepresentative and were not used in computing response rates.

E. QUALITY ASSURANCE

1. RDD Sample

Interviewer performance was evaluated on the basis of production reports and regular on-line monitoring. Daily production reports provided information on several performance indicators, including completed interviews and self-response modules, calls made, 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 supervisory monitoring of actual calls, supplemented by review of interviewer notes maintained in the CATI system (all calls and notes recorded about those 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 the 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 incorrect study presentation; 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. Overall, 11.8 percent of the interviews were monitored.

2. Field Sample

The accuracy of listing was verified by assigning a different staff member to screen the listed segments; thus, each segment was listed and verified. Errors were corrected through supplemental listing forms used to augment or delete initially listed housing units. 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.

V. WEIGHTING AND ESTIMATION

A. OVERVIEW

The sample design was complex, employing stratification, clustering, and oversampling Weights were designed to restore proportionality to the sample and were adjusted to compensate for nonresponse at the household, family insurance unit (FIU), and person levels. The use of unweighted data is likely to result in seriously biased estimates because the unweighted samples are distributed differently than are the populations they represent. This occurred for the following reasons:

- Design decisions, such as setting fixed sample sizes for sites, restricting the high-intensity sites to metropolitan statistical areas (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 random-digit-dialing (RDD) frame omitted numbers in banks of 100 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.

The proper use of weights in analyzing Community Tracking Study (CTS) Household Survey data will substantially reduce the bias of estimates due to the sample design and survey nonresponse. However, the weights do not address the potential for bias resulting from item nonresponse or response errors. Procedures used to impute missing data for individual variables are discussed in the Household Survey Public Use File Technical Publication Number 7, (Center for Studying Health System Change 1998). Furthermore, 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 to use this software are also included in Technical Publication No. 7.

1. Weights Provided for Public Use

Thirteen weighting variables, summarized in Table V.1, are available for researchers' use. Weights were constructed to allow for both site-specific and national estimates for individuals, FIUs, and sites.[1] Site-specific estimates are made for a site or involve comparisons of sites. In contrast, national estimates involve inferences to a broader population, unrestricted to 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 four classes of estimates, defined as follows:

1. ***Augmented Site Sample.*** Weights for site-specific estimates that use data from the site's sample, augmented with observations from the supplemental sample that are located in the site

2. ***Site Sample.*** Weights for national estimates that use data from the 60-site sample

3. ***Supplemental Sample.*** Weights for national estimates that use the supplemental sample

4. ***Combined Sample.*** Weights used for national estimates that combine data from the 60-site sample and supplemental sample

[1]Throughout this document, "national" is used to refer to the population of the 48 contiguous states. It does not include Alaska and Hawaii

TABLE V.1

NAMES OF CTS HOUSEHOLD SURVEY WEIGHTS

	Analytical Sample and Estimate Type			
	Site-Specific Estimate	National Estimate		
Level of Analysis	Augmented Site Sample	Site Sample	Supplemental Sample	Combined Sample
Person	WTPER1 (winttpp3)	WTPER2 (winttpp1)	WTPER3 (wteltpp4)	WTPER4 (winttppm)
Self-Response Module Respondent	WTSRM1 (winttps3)	WTSRM2 (winttps1)	WTSRM3 (wteltps4)	WTSRM4 (winttpsm)
FIU	WTFAM1 (wintuif3)	WTFAM2 (wintuif1)	WTFAM3 (wteluaf4)	WTFAM4
Site	- -	WTSITE		

NOTE: The original variable names for the weights, before the names were modified for the public use file, are in parentheses.

For each of the four classes of estimates, three weights are provided: one for analysis of FIU data and two for person-level analyses. The person-level weights include:

- Weights for respondents to the core person-level survey questions
- Weights for respondents to the adult self-response module and the child's last physician visit (described as the "self-response" module)

We also include a weight to be used when the analytical unit is the site itself.

The person-level and self-response module weights underwent a trimming procedure. Trimming weights reduced 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 was quite small (the extent of trimming was not great), the trimmed weights result in better precision.

The combined weights include two individual-level weights and one FIU weight for national estimates designed to combine data from the 60-site and supplemental samples. The individual-level weights include one weight for core questionnaire data, and one for the self-response module. These weights are based on the relative variances of the two samples and allow researchers to more easily take advantage of the increased precision of the combined samples.

Augmented site sample weights, combined sample weights, and site sample weights for the high-intensity sites include cases from both the RDD and field components. The supplemental weights and site weights for the low-intensity sites include only RDD cases. It is 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 include both the 60-site and supplemental samples. The precision of such estimates is substantially greater than that of estimates obtained for either sample by itself, especially for estimates about subgroups. However, either sample 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

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 60-site and supplemental samples to be used together for national estimates

3. Sampling Error Estimation

Some element of uncertainty is always associated with sample-based estimates of population characteristics because the estimates are not based on the full population. Known as "sampling error," this element of uncertainty 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.[2] Among other things, the standard error can be used to construct confidence intervals around estimates; for example, one can produce a range of numbers surrounding an estimate within which one has 95 percent confidence that the true value lies, given the standard error of the estimate.

The complexities of the CTS Household Survey design (stratification, clustering, and oversampling) preclude the use of common statistical packages (such as SAS or SPSS) for variance estimation. The variance estimates from these statistical packages may severely underestimate the sampling variance. The CTS data therefore require the use of specialized techniques for estimating sampling variances; that is, it is necessary to use survey data analysis software or specially developed programs designed to accommodate the statistic being estimated and the sampling design.

For the CTS Household Survey, the sampling variance is a function of the sampling design and the population parameter being estimated; it is called a "design-based sampling variance." The CTS data base 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 for estimation of 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

[2]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 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 becomes smaller as the sample size increases. The sampling variance is zero when the full population is observed, as in a census.

multistage design, joint inclusion probabilities, and variance components in the CTS Household Survey design.

The remainder of this chapter discusses weighting procedures in more detail and gives a more complete explication of sampling error estimation for the CTS Household Survey. 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. Sections E and F present two topics that overlay all the weights; Section E describes the procedures to identify and trim extremely large sampling weights, and Section F covers separate weights for the self-response module. Section G discusses the weights for combining the 60-site and supplemental samples for national estimates. Finally, Section H discusses appropriate methods for estimating sampling error for the CTS 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 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. The first stage was to select the 60 sites, and then to randomly select the high-intensity sites from among them. For the RDD sample, we then selected telephone numbers, identified households, defined FIUs within households, and collected data on FIUs and persons (adults and a sample of children) within FIUs. All these stages were considered in weighting. The initial weight of a unit (whether it is 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 then adjusted for nonresponse at this stage. Then, we computed initial weights for households whose telephone numbers were sampled and adjusted this weight for nonresponse at the household level. The sum of the household weights was compared with an external estimate of households, at which point a poststratification adjustment factor was applied to the weights.

For FIUs within sampled households, we constructed initial weights, a nonresponse adjustment factor, and then poststratified using the household poststratification factor; the result was an FIU-level weight. Finally, analogous steps were used to construct weights at the person level.

a. Telephone Number Initial Weight

The telephone number was the second stage of selection for the 60-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 or stratum (that is, the number of telephone numbers released out of the total number of telephone numbers in working banks). A telephone number "bank" was defined as the first 8 digits of a 10-digit telephone number; a bank can have 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." Although this formula differed slightly depending on the type of estimate for which the weight was designed (described later in more detail), the general form of this probability is:

(1) $$P(phone\ p\ in\ stratum\ h) = \frac{n_h}{N_h} \cdot \frac{nrel_h}{n_h - nbad_h},$$

where:

n_h = the number of telephone numbers selected in stratum h[3]

N_h = the number of working telephone banks in stratum h, times 100

$nrel_h$ = the number of telephone numbers released in stratum h

$nbad_h$ = the number of telephone numbers found to be nonworking or business numbers in stratum h, using Genesys ID.

All *released* telephone numbers were assigned this probability.

The sampling weight for phone p in stratum h is:

(2) $$SW(phone_{hp}) = \frac{1}{P(phone\ p\ in\ stratum\ h)}.$$

b. Nonresponse Adjustment to Telephone Weight

For the telephone number weight, an adjustment was made for nonresponse. Nonresponse could have occurred at the telephone-number level if we could not determine whether a telephone number was a working residential number.

We formed weighting cells to make the adjustment. In defining weighting cells, we tried to group respondents who were similar with respect to the most important analytical variables, as well as the

[3]Throughout this document, we refer to stratum h. In sites where substratification was not used (low-intensity sites), it will refer to the entire site. For the high-intensity sites, it will refer to the substrata used in selecting the sample. For the supplemental sample, it will refer to the five strata used in selecting the sample. Strata and substrata are defined in Chapter I.

likelihood of each type of nonresponse. The information used to form these cells must be known for both nonrespondents and respondents. The cell definitions may be the same or different for each type of nonresponse adjustment. Based on generally accepted guidelines, we decided that each cell should also 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.

For the telephone number adjustment, the primary weighting cells were the site and sampling strata; for the supplemental sample, the cell was defined by stratum. The following nonresponse adjustment factor at the telephone number level adjusted for failure to determine the eligibility of a telephone number:

$$(3) \qquad A_{nr}(phone_c) = \frac{\sum_{released\ ph \in c} SW(phone_{hp})}{\sum_{determ\ ph \in c} SW(phone_{hp})} ,$$

where the summation in the numerator is over all telephone numbers released in cell c, and the denominator is summed over all telephone numbers in cell c for which an eligibility determination was made.

A nonresponse-adjusted telephone number weight was then calculated:

$$(4) \qquad W1(phone_{hp} \in c) = SW(phone_{hp}) \cdot A_{nr}(phone_c), \text{if eligibility of telephone number determined}$$

$$W1(phone_{hp} \in c) = 0, \text{otherwise.}$$

c. Initial Household Weight

Because some households have multiple telephone numbers, a household multiplicity factor was used to adjust for the number of telephone numbers in the household. The initial household weight was the inverse of the product of the probability of selection of the household's telephone number and the household multiplicity factor:

$$P(household\ i\ in\ stratum\ h) = P(phone\ p\ in\ stratum\ h) \cdot nphone_{hpi}, \tag{5}$$

where $nphone_{hpi}$ is the reported number of telephone numbers at which household *i* with telephone number *p* in stratum *h* receives residential calls.

All households associated with telephone numbers determined to be eligible had this probability assigned. The last term in equation (5) was not available for households in which no FIUs completed interviews. For these cases, $nphone_{hpi}$ was set equal to 1. For all other cases, $nphone_{hpi}$ was set equal to the total number of telephone numbers in the household.[4]

The initial weight for household *i* in stratum *h* was:

$$SW(hh_{hi}) = \frac{1}{P(household\ i\ in\ stratum\ h)}. \tag{6}$$

d. Nonresponse Adjustments to Household Weight

The household-level nonresponse adjustment incorporated the telephone number nonresponse adjustment. Thus, we first defined:

$$W1(hh_{hi},c) = SW(hh_{hi}) \cdot A_{nr}(phone_c), \tag{7}$$

---

[4]Question h30 in the interview asked whether the household had any additional telephone numbers and, if so, how many; in the case of one or more, question h31 asks whether the additional number(s) is(are) for home or business use. If h30 = 1, 2, 3, or 4 and h31 = 1 or 2 (home use or both), then we set $nphone_{hi}$ equal to h30 plus 1. For any remaining cases (h30 = 9 or h31 >2), $nphone_{hi}$ was set equal to 1. Because questions h30 and h31 were asked of each family informant, some households had discrepant reports from two or more FIUs. In such cases, we set $nphone_{hi}$ to the average number of telephone numbers reported (among those reporting additional numbers).

where $A_{nr}(phone_c)$ is defined in equation (3).

Next, an adjustment was made for household-level nonresponse (that is, households that did not complete the household enumeration questions needed to form eligible FIUs).[5] We used the same weighting cells that were used for the previous adjustment (equation 3). The nonresponse adjustment factor at the household level adjusted for screener nonresponse among known households:

$$(8) \qquad A_{nr}(household_c) = \frac{\sum_{hh \in c} W1(hh_{hi})}{\sum_{resp\ hh \in c} W1(hh_{hi})},$$

where the summation in the numerator was over all households in cell c, and the denominator was summed over all households completing the screener in cell c.

A second interim household weight was then calculated:

$$(9) \qquad W2(hh_{hi} \in c) = W1(hh_{hi}) \cdot A_{nr}(household_c), \text{ for responding households}^{6}$$

$$W2(hh_{hi} \in c) = 0, \text{ for nonresponding households.}$$

---

[5]An FIU was ineligible if all adult members were on active duty in the military. In addition, an FIU could not be formed from unmarried, full-time students, less than 23 years old, who were not the children or wards of someone in the household.

[6]Responding households are those in which the enumeration questions (Part A of the survey) were completed.

e. Poststratification Adjustment to Household Weight

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 July 1996 estimates from Genesys[7] of the number of households in each site (and in each stratum for the 12 high-intensity sites) and nationally (by whether or not in an MSA).[8] For site-specific estimates, we then used 1990 Census data to estimate the proportion of telephone households in the same sites and strata and adjusted the 1996 estimates of the total number of households by these proportions to develop an estimate of telephone households in 1996.[9] For national estimates of telephone and nontelephone households (by metropolitan status), we used July 1996 estimates from Genesys. The poststratification adjustment factor for telephone households was:

$$(10) \qquad A_{ps\text{-}tel}(stratum\ \ h) = \frac{TELHH_h}{\sum_{resp\ hh_i \in h} W2(hh_{hi})} ,$$

where $TELHH_h$ is the estimated number of telephone households in July 1996. This adjustment was used as a poststratification factor for FIU and person weights but not to produce a final household weight, because (1) household-level analysis was not of interest, (2) external estimates of FIUs and persons within

---

[7]Genesys uses intercensal estimates developed by Claritas.

[8]Genesys defines a household according to the Census definition, which "... includes all the persons who occupy a housing unit," and a housing unit is defined 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 exclude households containing only unmarried students under the age of 23 or persons in the military. Unmarried students under the age of 23 are eligible for the CTS Household Survey, but only through their parents' households.

[9]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 1990 through 1997 (U.S. Bureau of the Census 1994, 1997).

FIUs were not available, and (3) the household poststratification adjustment was deemed a reasonable proxy. The household-level weight poststratified to telephone households was:

(11) $$WT_{tel}(hh_{hi}) = W2(hh_{hi}) \cdot A_{ps\text{-}tel}(stratum\ h).$$

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 in the year preceding the interview.[10] 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 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:

(12) $$WT_{interruption}(hh_{hi}) = \frac{W2(hh_{hi})}{prop.\ of\ year\ with\ phone}.$$

The poststratification adjustment factor for total households would be:

(13) $$A_{ps\text{-}all}(phone\ status\ g,\ stratum\ h) = \frac{TOTHH_{gh}}{\sum_{resp\ hh_i \in h\ with\ phone\ status\ g} WT_{interruption}(hh_{hi})},$$

[10]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....?"). In households with discrepant reports from two or more FIUs, we set the number of months of interruption in service to the average number of months reported (among those reporting an interruption in service).

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

A household-level weight poststratified to all households is:

$$(14) \qquad WT_{all}(hh_{ghi}) = WT_{interruption}(hh_{hi}) \cdot A_{ps\text{-}all}(phone\ status\ g,\ stratum\ h).$$

f. FIU Initial Weight

The probability of selection for each FIU is equal to the probability of selection for its household; that is, all FIUs within each selected household were selected for the interview. All *eligible* FIUs in responding households will have this probability assigned, regardless of whether the FIU itself responded:

$$(15) \qquad P(FIU\ j\ in\ household\ i\ in\ stratum\ h) = P(household\ i\ in\ stratum\ h).$$

The initial weight for FIU j in household i in stratum h is:

$$(16) \qquad SW(FIU_{hij}) = \frac{1}{P(FIU\ j\ in\ household\ i\ in\ stratum\ h)}.$$

g. Nonresponse Adjustment to FIU Weight

The first step in the FIU nonresponse adjustment was to adjust the FIU initial weight for telephone- and household-level nonresponse, using the factors defined in equations (4) and (9). FIU weights were then calculated:

$$(17) \qquad W1(FIU_{hij} | c) = SW(FIU_{hij}) \cdot A_{nr}(phone_c) \cdot A_{nr}(household).$$

The adjustment for FIU nonresponse was made among FIUs known to be eligible. This weighting adjustment used FIU composition characteristics (the number of persons, the presence of seniors, the presence of children) to form the weighting cells:

$$(18) \qquad A_{nr}(FIU_{hij}) = \frac{\sum_{elig\ FIU \in c} W2(FIU_{hij})}{\sum_{responding FIU \in c} W2(FIU_{hij})},$$

where the summation in the numerator was over all eligible FIUs[11] in cell *c*, and the denominator was summed over all responding FIUs in cell *c*. A third interim FIU weight was then calculated:

(19) $W3(FIU_{hij}) = W2(FIU_{hij}) \cdot A_{nr}(FIU_{hij})$, for eligible responding FIUs

$W3(FIU_{hij}) = 0$, for eligible nonresponding FIUs.

h. Poststratification Adjustment to FIU Weight

Because we had no external estimates of the number of FIUs, we applied the poststratification adjustments used for the household-level weight: $A_{ps\text{-}tel}(stratum\ h)$ and $A_{ps\text{-}all}(phone\ status\ g,\ stratum\ h)$. The poststratified FIU-level weights are:

(20a) $WT_{tel}(FIU_{hij}) = W3(FIU_{hij}) \cdot A_{ps\text{-}tel}(stratum\ \ h)$,

and

(20b) $WT_{all}(FIU_{ghij}) = W3(FIU_{hij}) \cdot A_{ps\text{-}all}(phone\ \ status\ \ g,\ \ stratum\ \ h)$.

[11]Secondary FIUs determined to be eligible at screening were considered eligible at this stage, regardless of later telephone status. Some FIUs were associated with telephones that were nonworking when callbacks were made. These FIUs were considered eligible nonrespondents.

The poststratification represented in equation 20b also incorporated the adjustment for telephone service interruption.

i. Initial Person Weight

The probability of selection for each adult is equal to the probability of selection of the FIU because all adults within an FIU were selected for the interview. However, only one child was selected at random per FIU, so the within-FIU probability of selection for a child is 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:

(21) $$P(\textit{person } k \textit{ in FIU } j \textit{ in household } i \textit{ in stratum } h) = \frac{P(FIU\ j\ \ in\ \ household\ \ i\ \ in\ \ stratum\ \ h)}{(\delta \cdot numkids_{hij}) + (1 - \delta)},$$

where $numkids_{hij}$ is the number of children in FIU *j* in household *i* in stratum *h*, and δ is equal to 0 for adults and 1 for children.

The initial weight for person *k in* FIU *j* in household *i* in stratum *h* is the inverse of the probability of selection:

(22) $$SW(person_{hijk}) = \frac{1}{P(person\ \ k\ \ in\ \ FIU\ \ j\ \ in\ \ household\ \ i\ \ in\ \ stratum\ \ h)}.$$

All eligible persons in all *responding FIUs* will have this weight assigned, regardless of whether data on the person were collected.

j. Nonresponse Adjustment to Person Weight

An editing program was used to determine whether a person record contained too many missing items to be usable. Only eight adults and two children were deleted because of high levels of missing information.

The next two steps adjust for this small amount of unit nonresponse at the person level. (Persons for whom the self-response module was incomplete are discussed in Section F.) The nonresponse adjustment involved several steps, adjusting for telephone, household, and FIU nonresponse prior to the person-level adjustment:

$$(23a) \quad W1(person_{hijk} \in c) = SW(person_{hijk}) \cdot A_{nr}(phone_c)$$

$$(23b) \quad W2(person_{hijk} \in c) = W1(person_{hijk}) \cdot A_{nr}(household_c)$$

$$(23c) \quad W3(person_{hijk} \in i) = W2(person_{hijk}) \cdot A_{nr}(FIU_{hij}),$$

where the adjustment factors are defined in equations (3), (8), and (18), respectively.

The final person-level nonresponse weighting adjustment uses the FIU as the weighting cell:

$$(24) \qquad A_{nr}(person_{hijk}) = \frac{\sum_{person \in cell_j} W3(person_{hijk})}{\sum_{responding\ person \in cell_j} W3(person_{hijk})},$$

where the summation in the numerator is over all *selected* persons in cell j, and the denominator is summed over all responding persons in cell j.

A fourth interim-person weight can then be calculated:

$$(25) \qquad W4(person_{hijk}) = W3(person_{hijk}) \cdot A_{nr}(person_{hijk}), \text{ for responding persons}$$
$$W4(person_{hijk}) = 0 \text{ for nonresponding persons.}$$

k. Poststratification Adjustment to Person Weight

Prior to combining the RDD sample with the field sample, we applied the poststratification adjustments used for the household-level weight: $A_{ps\text{-}tel}(stratum\ h)$ and $A_{ps\text{-}all}(phone\ status\ g,\ stratum\ h)$. The poststratified person-level weights would then be:

$$\text{(26a)} \quad WT_{tel}(person_{hijk}) = W4(person_{hijk}) \cdot A_{ps\text{-}tel}(stratum\ \ h)$$

$$\text{(26b)} \quad WT_{all}(person_{ghijk}) = W4(person_{hijk}) \cdot A_{ps\text{-}all}(phone\ \ status\ \ g,\ \ stratum\ \ h).$$

Poststratification of the person weights to external counts of individuals took place as part of integration of the RDD and field telephone sampling weights (see Section D.)

2. Using the 60-Site Sample to Make National Estimates

The *formulas* for the selection probabilities and weights at the household, FIU, and person levels (see equations [1] through [26b]) and the formulas and methodologies for the nonresponse and poststratification adjustments, are similar across three types of estimates: (1) site-specific estimates using the augmented site sample, (2) national estimates using the site sample, and (3) national estimates using the supplemental sample. However, the *values* of these weights and adjustment factors differ across the three types of RDD weights, because the telephone selection probabilities differ. Furthermore, weights to be used for making national estimates using the 60-site sample must also account for the probability of selection of the site, as well as 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 CTS Household Survey design included sites in three strata:

1. MSAs with 200,000 or more persons in 1992

2. MSAs with fewer than 200,000 persons in 1992
3. Nonmetropolitan areas (that is, a single county or a grouping of two or more counties contiguous counties)

Of the 48 site selections among the MSAs with 200,000 or more persons, 12 were randomly selected as high-intensity sites. None of the sites in the other two strata were eligible to be high-intensity sites. In the weighting formula, we must account for differences in sample allocation between the 12 high-intensity sites and the other 36 sites that were eligible to be high-intensity sites.

The sample size of telephone numbers in the RDD sample was k times larger (k is approximately equal to four) in the high-intensity sites than in the low-intensity sites. To account for the probability of selection of any telephone number, when making national estimates, we must use the expected number of selected telephone numbers, $E(n_{sh})$, in each site, rather than the actual number of selected telephone numbers, n_{sh}. For sites s in stratum h, where the site is an MSA with 200,000 or more persons, the expected number of selected telephone numbers is:

$$
\begin{aligned}
\text{(27a)}\quad E(n_{sh}) &= [n_{lo} \cdot k \cdot P(high\text{-}intensity)] + [n_{lo} \cdot P(low\text{-}intensity)] \\
&= [n_{lo} \cdot k \cdot 12/48] + [n_{lo} \cdot 36/48] \\
&= n_{lo} \cdot (k/4 + 3/4),
\end{aligned}
$$

where n_{lo} is the number selected for a low-intensity site. For sites that are MSAs with fewer than 200,000 persons, 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 can then be defined as:

(26b) $$P(telephone\ i\ in\ stratum\ h\ in\ PSU\ s) = PSUPROB_s \cdot \frac{E(n_{sh})}{N_{sh}} \cdot \frac{nrel_{sh}}{n_{sh} - nbad_{sh}},$$

where:

$PSUPROB_s$	=	the probability of selection of sites *s*[12]
$E(n_{sh})$	=	the expected number of telephone numbers selected in the 60-site sample in stratum *h* in site *s*
N_{sh}	=	the number of working telephone banks in stratum *h* in site *s*, times 100
$nrel_{sh}$	=	the number of telephone numbers released in the 60-site sample in stratum *h* in site *s*
n_{sh}	=	the actual number of telephone numbers selected in the 60-site sample in stratum *h* in site *s*
$nbad_{sh}$	=	the number of nonworking or business telephone numbers in the 60-site sample in stratum *h* in site *s,* using Genesys ID.

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

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

When combining the 60-site sample and the supplemental sample to make site-specific estimates, the probability of selection of the telephone number can be defined as:

(28) $$P(telephone\ i\ in\ stratum\ h) = \frac{n_h}{N_h} \cdot \frac{nrel_h}{n_h - nbad_h},$$

where:

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

n_h = the number of telephone numbers selected in the augmented sample in stratum *h*

N_h = the number of working telephone banks in stratum *h*, times 100

$nrel_h$ = the number of telephone numbers released in the augmented sample in stratum *h*

$nbad_h$ = the number of nonworking or business telephone numbers in the augmented sample in stratum *h,* using Genesys ID,

and where the term *augmented sample* refers to the 60-site sample combined with the supplemental sample cases that fell within the boundaries of 1 of the 60 sites.

4. Using the Supplemental Sample to Make National Estimates

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

$$P(telephone\ i\ in\ stratum\ h) = \frac{n_h}{N_h} \cdot \frac{nrel_h}{n_h - nbad_h}, \tag{29}$$

where:

n_h = the number of telephone numbers selected in the supplemental sample in stratum *h*

N_h = the number of working telephone banks in stratum *h*, times 100

$nrel_h$ = the number of telephone numbers released in the supplemental sample in stratum *h*

$nbad_h$ = the number of nonworking or business telephone numbers in the supplemental sample in stratum *h,* using Genesys ID.

C. WEIGHTS FOR THE FIELD SAMPLE

1. Introduction

This section describes the procedures used in constructing 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 interim weights and the interim weights representing the RDD sample were used to create integrated weights for making inferences about the entire population. Field sample weights for households, FIUs, and persons 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. The second set of weights are referred to as "national" weights. Each weight is the product of several factors, which reflect differences in probabilities of selection and nonresponse. The weights also include poststratification adjustments so that the sample matches external estimates of the relevant population.

2. Steps in the Weighting Process

The first weighting factor for a unit (household, FIU, or person) for any of the weights is the inverse of that unit's probability of selection. This factor differs 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 is the inverse of the overall probability of selection of a unit (housing unit, household, FIU, or person). Weights were computed for housing units, households, FIUs and persons. 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:

(30a) $SWN(LHU)_{abci} = 1/P(LHU_{abci}),$

where:

(30b) $P(LHU_{abci}) = P(PSU_a) \cdot P(SSU_b \backslash PSU_a) \cdot P(LA_c \backslash SSU_b) \cdot P(HU \backslash La_c)$

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 $P(HU \backslash LA_c)$ accounts for the fact that only a subsample of listed housing units were selected for interviewing in some listing areas. Note that for a household (*HH*), the initial weight is the same as for a listed housing unit. Thus, for national estimates:

(31a) $P(HH_{abci}) = P(LHU_{abci})$

(31b) $SWN(HH)_{abci} = SWN(LHU)_{abci}.$

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:

(32a) $SWS(HH)_{bci} = SWS(LHU)_{bci} = 1/P(LHU_{bci})$

(32b) $P(LHU_{abci}) = P(SSU_b \backslash PSU_a) \cdot P(IA_c \backslash SSU_b) \cdot P(HU \backslash LA_c).$

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 $numkids_{abcij}$:

(33) $P(Child_k \backslash FIU_j \backslash HH_{abci}) = P(HH_{abci}) \cdot 1/(numkids_{abcij})$.

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 should be the sites themselves for both national and site-based estimates.

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 (*SWN*) and site-based (*SWS*) weights. For cell *c*, we define a nonresponse-adjustment factor $A_{nr}(HU_c)$ and the weight $W_1(HH_{abci})$:

(34a) $$A_{nr}(HU_c) = \frac{\sum_{attempt \in k} SW1_{abci}}{\sum_{determ \in k} SW1_{abci}}$$

(34b) $W_1(HH_{abci}) = SW1_{abci} \cdot A_{nr}(HU_c)$, if eligibility of household is determined

$= 0$, otherwise.

As discussed in Chapter II, eligibility was imputed for some households in inaccessible buildings. Cases with imputed eligibility were treated in the same way as those whose eligibility had actually been determined.

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

$$(35) \qquad A_{nr}\,(HH_c) = \frac{\sum_{elig \in c} W_1\,(HH_{abci})}{\sum_{resp\ hh \in c} W_1\,(HH_{abci})},$$

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:

(36)	$W_2\,(Hh_{abci})$	$= W_1\,(HH_{abci}) \cdot A_{nr}\,(HH_c),$	if the household responded
		$= W_1\,(HH_{abci}),$	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 essentially the same as those described in the section on weighting the RDD sample data. 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 our 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:

(37a)
$$A_{ps1}(a) = \frac{Censuscount_{ga} \cdot g}{\sum_{i=1}^{n_{bc}} W2(HH_{abci})}$$

(37b) $PSW1 = W_2(HH_{abci}) \cdot A_{ps1}(a)$.

The second stage was a ratio adjustment of interviewed households to 1996 estimates[13] of all nontelephone households (including areas excluded from the sampling frame), nationally and for each site:

(38a)
$$A_{ps2}(site) = \frac{census\ np\ (site)}{\sum_{i \in resp}^{nsite} PSW1_i}$$

(38b)
$$A_{ps2}(nat) = \frac{census\ np\ (nat)}{\sum_{s=1}^{12} \sum_{i \in resp}^{nsite} PSW1_{is}}$$

(39a) $PSW2S_i = PSW1_i \cdot A_{ps2}(site)$

(39b) $PSW2N_i = PSW1_i \cdot A_{ps2}(nat)$.

A similar adjustment was made for individuals.

---

[13]These estimates were synthesized from the 1990 Census proportion of nontelephone households and the July 1996 Genesys estimate of total households

D. INTEGRATED WEIGHTS FOR HOUSEHOLD SURVEY

The integrated weights combine 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 account 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 persons, 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)[14]:

1. Poststratify 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 (IAF) 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 based on the site sample, the field component represents nontelephone households only in large MSAs (those with populations of 200,000 or more). For households in small MSA or nonmetropolitan strata, the "integrated" weights are 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, then poststratified

[14]For the national estimates from the site sample, we included only households from the site sample at this point in the process; for site-specific estimates, we also included households from the supplemental sample that are to be used for site-specific analyses.

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.[15]

For RDD households in the 48 large MSA sites, we began with the weights that represented the telephone portion of the population (WT_{tel}). Similarly, for the field households (all in the 12 high-intensity sites), we began with the weight that represented the nontelephone portion of the population (*PSWT2N*). Large MSA households in the RDD component that had intermittent telephone service were adjusted for dual selection probabilities (that is, 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.) Households in the field component that had some telephone service during the year preceding the survey were also adjusted for dual selection probabilities, while accounting for the length of interruption. Table V.2 illustrates how the RDD and field components were combined for national estimates based on the site sample.

For site-specific estimates, the field component represents nontelephone households in the 12 high-intensity sites only. For households in the low-intensity sites, the "integrated" weights are simply the weights that represented all households in those 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, then poststratified to the estimated number of nontelephone households (by site); those with no interruption

[15] For national estimates based on the supplemental sample, the "integrated" weights for all households are 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, than 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.

TABLE V.2

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

<table>
<tr><th></th><th>RDD Component</th><th>Field Component</th></tr>
<tr><td>High-intensity sites</td><td rowspan="2">Represents households in large MSAs in contiguous United States with continuous or intermittent telephone service</td><td>Represents households in large MSAs in contiguous United States with intermittent or no telephone service</td></tr>
<tr><td>Other large-MSA sites</td><td rowspan="2"></td></tr>
<tr><td>Small-MSA sites

Non-MSA sites</td><td>Represents all households in balance of contiguous United States.</td></tr>
</table>

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}). Similarly, for the field households (all in the 12 high-intensity sites), we began with the site-specific weight that represented the nontelephone portion of the population (*PSWT2S*). High-intensity-site households in the RDD component that had intermittent telephone service were adjusted for dual selection probabilities, while accounting for the length of interruption. Households in the field component that had some telephone service during the year preceding the survey were also adjusted for dual selection probabilities, while accounting for the length of interruption. Table V.3 illustrates how the RDD and field components were combined for site-specific estimates.

1. Telephone Service Interruption Adjustment Factor

A complicating factor in combining the RDD and field samples is that both components included households with interrupted telephone service during the year preceding the survey. The integrated weights assume that (1) those 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 sample, 3.4 percent of households completing interviews had an interruption in telephone service of two weeks or more during the year preceding the survey, but less than half of the households were in areas eligible for the field component. For the field sample, 57.9 percent (n = 272) of households completing interviews had at least one month during that year in which they had telephone service and could have been sampled for the RDD survey.

TABLE V.3

INTEGRATION OF RDD AND FIELD COMPONENTS
FOR SITE-SPECIFIC ESTIMATES

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

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, which would have been needed 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 are based on the 1990 Census and would not account for housing construction since then. Moreover, 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.

Instead, we constructed integrated weights that synthetically accounted for multiplicity by using a weighting adjustment that we call the telephone interruption adjustment factor. This factor was applied only to households in the "integration sites"; that is, those that are represented by both the RDD and field components. For national estimates, integration sites would include all large-MSA sites. For site-specific estimates, they would include the 12 high-intensity sites only. The interruption adjustment factor accounts for both length of telephone interruption and multiplicity. For the field component, households with no period of telephone service would have had no chance of selection into the RDD component, and so have interruption adjustment factor set to 1. For the RDD component, households in the integration sites with no period of telephone interruption would have been ineligible for the field component and also have IAF set to 1. 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 interruption adjustment factor, as described below, by the households' postratified weights; the weights were postratified to the populations their components represent (telephone or nontelephone). We calculated IAF_m as:

$$(40) \quad IAF_m = \frac{1/RelP_m}{1/MEDIAN(RelP)} \cdot k \qquad m = (1,2,...,12),$$

where

$$(41)\quad RelP_m = \left[PRatio \cdot \frac{(12-m)}{12}\right] + 1,$$

and

$$(42)\quad PRatio = \frac{HH\ in\ RDD\ sample\ /\ telephone\ HH\ in\ population}{HH\ in\ field\ sample\ /\ nontelephone\ HH\ in\ population},$$

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 those with an interruption in telephone service remains 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[16]; and *PRatio* is the probability of selection into the RDD component, relative to selection into the field component.

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

(43)	$WTINT_m = WT_{tel} \cdot IAF_m,$	for RDD households in integration sites
	$WTINT_m = PSWT2 \cdot IAF_m,$	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.

[16]In equation (41), the first term (in square brackets) represents the likelihood of selection into the RDD component, and the second term (1) reflects the likelihood of selection into the field component.

2. Poststratification of Person-Level Integrated Weights

For national estimates, person-level samples 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.[17] Weights were then poststratified to the estimate of the U.S. population.[18] For high-intensity sites, poststratification of site-specific weights was by age group, race, whether or not Hispanic or black, sex, and the estimated site population.[19] Weights for low-intensity, site-specific estimates were poststratified to site totals only.

E. TRIMMING PERSON WEIGHTS

In analyzing survey data, a few extremely large weights can result in inflated values of the sampling variance, resulting in less accurate point estimates. 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 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 integrated weights[20] and then assessed the effect of the trimming. We evaluated the extent of trimming and the inflation factor for the untrimmed weights needed to preserve the original sum of the weights, and then estimated the effect of the

---

[17]Based on intercensal estimates from the U.S. Bureau of the Census [www.census.gov/population/estimates/nation/intfile3-1.txt]. Age group by sex was from 2/97. Hispanic and black by sex was from 3/97. Education was from 3/96.

[18]Excluding Alaska and Hawaii. Genesys estimate from July 1996.

[19]Age, race, ethnicity, and total population by site were based on Genesys estimates from July 1996.

[20]FIU-level weights were not trimmed.

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 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. Use of unequal weighting has the potential to cause loss in 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.

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

$$(44) \qquad T_c = \left(k \sum_{i \in c} WT_i^2 \,/\, n_c\right)^{1/2},$$

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 cutpoint 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 cutpoint is recomputed and each weight is again compared with the cutpoint. If any weight exceeds the new cutpoint, the observation is assigned the value of the new cutpoint and the other weights are inflated to compensate for the trimming.

The cutpoint generated by the algorithm was generally used as the value of the trimmed weight. In a few trimming cells, the algorithm indicated a trimming level that was excessive, and a value larger than the computed cutpoint was used. Generally, this was done when the adjustment seemed excessive for the weights that were less than the cutpoint or when few observations were in a trimming class.

The weights designed to produce site-specific estimates were evaluated for adults and children within 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 weights for adults. The weights for trimming were identified using both the NAEP procedure and visual inspection for 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 core interview and self-response module weights for both site-sample and augmented site-sample estimates. The weights were trimmed for less than 0.4 percent of the adult and children observations.

For the weights designed to produce national estimates, similar weight trimming was conducted using the NAEP procedure and an assessment of 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 (metropolitan areas with more than 200,000 persons, metropolitan areas with fewer than 200,000 persons, and nonmetropolitan areas), 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 50 of the more than 56,000 weights of the site sample and fewer than 40 of the 5,600 weights in the national supplement.

F. WEIGHTING THE SELF-RESPONSE MODULE

As described in Chapter III, the self-response module contained questions on health status, personal attitudes, and experiences during the respondent's most recent physician visit. Each adult in an eligible FIU was asked to respond personally to this set of questions. For a randomly selected child, the FIU member who took the child to his or her most recent physician visit in the 12 months preceding the interview responded to a subset of these questions. (The questions are included in the survey instrument, Appendix A.) A separate status code was created for this module to indicate the completion status or reason for noncompletion for each person.

1. National Estimates Based on the 60-Site Sample

a. Self-Response-Module Nonresponse Adjustments

For national estimates based on the 60-site sample, nonresponse adjustment cells were defined by sex, age group, ethnicity (Hispanic, black non-Hispanic, other non-Hispanic), and education. Where necessary, we collapsed some cells by education or ethnicity to ensure sufficiently large cell sizes. The nonresponse adjustment was of the form of equation (24), except that response was defined as response to the self-response module.

After all these adjustments were made, children under age 18 were given their poststratified person-level weight as their "self-response-module" weight. Some of the questions included in the adult self-response module were also asked about the sampled child (see Chapter IV Section C.3). The informant for these questions was either the FIU informant or the person who took the child to the physician on his or her last visit. Although there might be some item nonresponse to these questions, all children with a positive person-level weight would have at least some data in this module.

Persons under age 18 who lived in the household or were spouses of people in the household were classified as adults in the survey and were given their self-response-module weights, adjusted so that the sum of the weights for all of those under 18 was the sum of their person-level weights. Adults age 18 or older were also given their self-response-module weights, adjusted so that the sum of their weights was the sum of their person-level weights.

b. Trimming and Poststratification

Self-response-module weights were trimmed and then poststratified to the totals that existed before trimming. In making this adjustment, we used the same cells (sex by age group, sex by whether or not Hispanic, sex by whether or not black, and education) as were used for the initial poststratification. This process was performed separately for children and adults. Finally, as for the pretrimmed weights, the self-response-module weights for children under age 18 were assigned the values of the children's person-level weights; householders or householder spouses under age 18 and adults 18 and older were given adjusted self-response-module weights.

2. National Estimates Based on the Supplemental Sample

a. Self-Response-Module Nonresponse Adjustment

The nonresponse adjustment for the self-response module was carried out for the supplemental sample as described for the 60-site sample national estimates. An adjustment was made to ensure that the sum of the self-response-module weights for children under age 18 matched the sum of the children's person-level weights.

After all these adjustments were made, children under age 18 were given their poststratified person-level weights as their self-response-module weights. Adults 18 and older and householders or householders' spouses under age 18 were given their self-response-module weights, adjusted so that the sum of their weights was the sum of their person-level weights.

b. Trimming and Poststratification

After the self-response-module weights for the supplemental sample were trimmed, they were adjusted to the pretrimming poststratified totals, using the same cells (sex by age group, sex by whether or not Hispanic, sex by whether or not black, and education) used for the original poststratification. This process was carried out in the same manner as for the site-sample national estimates.

3. Site-Specific Estimates Based on the Augmented Sample

a. Self-Response-Module Nonresponse Adjustment

In general, the nonresponse adjustment for the self-response module was carried out for the augmented site sample in the same manner as for the site-sample national estimates. For the 12 high-intensity sites, the nonresponse weighting cells were formed by site, sex, age group, and ethnicity (Hispanic, black non-Hispanic, other non-Hispanic) and were collapsed as needed. For the remaining sites, the cell was simply the site by two age groups.

After all these adjustments were made, children under age 18 were given their poststratified person-level weights as their self-response-module weights. Adults age 18 and older and persons under 18 who were treated as adults in the survey were given their self-response-module weights, adjusted so that the sum of their weights was the sum of their person-level weights.

b. Trimming and Poststratification

After the self-response-module weights for the augmented site sample were trimmed, they were again poststratified to the totals before trimming, using the same poststratification cells. For the 12 high-intensity sites, the cells were: site by age group, site by Hispanic/black, and sex. For the remaining sites, the cells were site and two age groups.

G. WEIGHTS FOR COMBINING THE 60-SITE AND SUPPLEMENTAL SAMPLE SURVEYS

The goal of the supplemental sample is to efficiently augment the site-based sample. The objective in combining the samples is to use the full sample (site and supplemental samples) to achieve the minimum variance for national estimates. 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 will 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:

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

where *Y(Site)* is the survey estimate from the site sample, *Y(Supp)* is the survey estimate from the supplemental sample, and lambda is an arbitrary constant between 0 and 1. For the sampling variance, *V(Y)*, the estimate is computed using the equation:

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

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 *ë* will result in an unbiased estimate of the survey estimate, but not necessarily an estimate with the minimum sampling variance. Two approaches were used to estimate *lambda*.

A value of *lambda* can be computed on the basis of the effective sample sizes for each sample component, where the effective sample size is computed using the design effects for the site and supplemental samples, *Deff(Site)* and *Deff(Supp)*, respectively. The design effect for an estimate from the site sample is computed as:

(47) $Deff(Site) = V(Y(Site)|Design) \mid V(Y(Site)|SRS)$,

where *V(Y(Site)|Design)* is the estimated sampling variance for *Y(Site)* using the full sampling design, and *V(Y(Site)|SRS)* is the estimated sampling variance for *Y(Site),* assuming a simple random sample of the same size. The design effect from the supplemental sample, *Deff(Supp)*, is computed analogously. The effective sample sizes, ($n_{eff}(Site)$ and $n_{eff}(Supp)$ for the site and supplemental samples, respectively) are then computed as:

(48a) $n_{eff}(Site) = n(Site) / Deff(Site)$

and

(48b) $n_{eff}(Supp) = n(Supp) / Deff(Supp)$,

where *n(Site)* and *n(Supp)* are the nominal sample sizes for the site and supplemental samples, respectively. The value for combining the estimates and variances using the effective sample sizes, $(lambda)_1$, is then:

(49) $(lambda)_1 \;=\; n_{eff}\,(Site) \,/\, [\, n_{eff}\,(Site) + n_{eff}\,(Supp)]$

$= [n(Site)\; V(Y(Site)|SRS) \,/ V(Y(Site)|Design)] \,/$

$[n(Site)\; V(Y(Site)|SRS) \,/\, V(Y(Site)|Design) \,+$

$n(Supp)\; V(Y(Supp)|SRS) \,/\, V(Y(Supp)|Design)]$.

Alternatively, the value associated with minimum variance, *lambda* $_2$, can be computed as:

(50) $(lambda)_2 \;=\; [1/V(Y(Site)|Design)] \,/\, [1 \,/\, V(Y(Site)|Design) \;+ 1/\, V(Y(Supp)|Design)]$

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

In this case, the minimum variance is:

(51) $V(Y) = [V(Y(Site)) \cdot V(Y(Supp))] \;/\; [V(Y(Site)) \;+ V(Y(Supp))]$

with the design designation omitted. If the estimated population variances are equal (that is, *n(Site)* C *V(Y(Site)|SRS)* equals *n(Supp)* C *V(Y(Supp)|SRS)*), then the two *lambdas* are equal and, likewise, the two variance estimates. Our analyses showed that either algorithm for computing *lambdas* resulted in essentially the same value for the variance estimate for the combined samples.

To compute the combined-sample estimate with minimum variance, a survey estimate is derived by first computing the estimate for each survey component, and computing a value of *lambda* 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 *lambda* (as in equation (45)). The sampling variance is estimated using the sampling variance estimate from each component survey and the computed value of *lambda* (as in equation (46)). Although it produces the minimum variance estimates, the process is computer intensive and results in some inconsistencies among estimates for percentages and proportions because of differing values of *lambda* among levels of a categorical variable. 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 *lambda*. In addition, this two-step process for computing estimates would likely pose analytical problems for regression analyses and other, more complex analyses. For these reasons, we explored the use of single or multiple values of *lambda* 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 *lambda* 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 this analysis, 14 analysis variables (10 categorical and 4 continuous) and nine populations (the full population and eight subpopulations) were identified. For dichotomous variables (for example, a yes/no variable), the sampling variances for both response options are equal and, therefore, redundant. After removing redundant and unstable estimates (estimates with a relative standard error of 0.30 or higher), 226 pairs of estimates and sampling variances were used in the analysis. We will confine the following discussion of *lambda* to the minimum variance approach $(lambda)_2$.

The mean value of the *lambdas* was 0.814 with a median of 0.840 and the distribution of the *ë*s was skewed, with 13.7 percent of the *lambdas* less than 0.70. The value of *lambda* is affected by design effects in the site sample, that is, by the average number of persons in a site and the correlation among responses within a site (that is, the intracluster correlation). As expected, the mean of the *lambdas* for estimates for the full population was the lowest (mean of 0.758 and median of 0.791) because of the number of persons in each site. For three key subpopulations (children, blacks, and Hispanics), the mean *lambda* value was approximately 0.83, and the mean of the median values for the three subpopulations was 0.837. The mean of the median values (0.837) was used as the *lambda* for combining the weights for three reasons. First, it was close to the median value for all *lambdas* (0.840). Second, the three subpopulations have relatively small sample sizes, and it was desirable to minimize the variance estimates for point estimates for these subpopulations. Third, the optimal *lambda* for the full population would result in less than optimal variances for subpopulations; however, a less than optimal *lambda* for the full population would not substantially increase the variance for that group. Our analysis across the 14 analysis variables indicated that the use of a single value of *lambda* (0.837) is likely to inflate the sampling variance by approximately 5 percent for most of the eight subpopulations used (and this varies by analysis variable), and up to 10 percent for the larger populations (for example, the full population or large subpopulations, such as working adults) and for some continuous variables.

Using the single value of *lambda*, the combined-sample weight was computed for persons in the site sample as:

(52a) *WT(Combined) = (lambda) WT(trimmed site sample weight)*,

and for persons in the supplemental sample as:

(52b) *WT(Combined) = (1 - lambda) WT(trimmed national supplement weight)*.

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

H. SAMPLING ERROR ESTIMATION

1. Background

The CTS Household Survey sample design is complex and requires specialized techniques for estimation of sampling variances. Standard statistical packages, such as SAS and SPSS, compute variances using formulas assuming the data are from a simple random sample from an infinite population. In some surveys, the simple random sample variance may approximate the sampling variance; with a design as complex as the CTS, the simple random sample variance is likely to substantially underestimate the sampling variance. 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:

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

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 becomes smaller 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. Also, an estimate with a relative standard error of greater than 30 percent may be reported but may be identified as potentially unreliable.

For the Household Survey, the sampling variance estimate is a function of the sampling design and the population parameter being estimated; it is called the design-based sampling variance. The design-based variance assumes the use of "fully adjusted" sampling weights. These weights 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 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 3.7. This means that the standard error is, on average, almost double what it would have been if the same number of cases had been selected using a simple random sample. With a design effect of 3.7, the Household Survey (with 60,446 observations) has the equivalent precision of a simple random sample with a size of about 16,300. Note that the design effect is generally lower for subclasses of the population, because there is less clustering of observations.

For the CTS Household Survey, the data files 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, the sampling weights and the sampling design features must be accounted for in order to compute unbiased estimates of population parameters and their associated sampling variances. The estimation of the sampling variance requires 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 (Technical Publication No. 7), available at *www.hschange.com*, contains tables of standard errors for various types of estimates.

Survey estimators fall into two general classes: (1) linear, and (2) nonlinear estimators. Linear estimates are weighted totals of the persons 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 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 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, PC-CARP and Stata), 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 or national estimates. The site-specific weights are designed for estimates that include units (either FIUs or persons) 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, and the combined weights that incorporate the site and supplemental samples. All three of the national weights are poststratified to the same population totals to ensure comparability; however, the three 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 Weights Based on the Augmented 60-Site 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. National Weights Based on the 60-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 CTS Household Survey, the sampling percentage of sites was sufficiently high among the

MSAs with more than 200,000 persons, and we used this concept to obtain more accurate and smaller sampling variance estimates. We also used the finite population correction concept for the MSAs with fewer than 200,000 persons, but not for the nonmetropolitan areas. For the nonmetropolitan areas, the sampling rate was sufficiently small that we assumed with-replacement sampling; thus, the finite population correction factor was not needed.

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 are 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 consists 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 MSAs with 200,000 or more persons in 1992 (to facilitate the computation of the joint selection probabilities)
- One stratum for MSAs with fewer than 200,000 persons in 1992
- One stratum for nonmetropolitan areas

In the nine analysis strata for the certainty selections, there is 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) accounts 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. National Weights Based on the Supplemental Sample

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

d. National Weights Based on the Combined Sample

The maximum precision for national survey estimates will be 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 (45). The sampling variance of this estimate, *V(Y)*, is computed using equation (46). Section G describes the value of *lambda* we derived to simplify processing without substantial loss in precision. The combined weights incorporate this value.

The variance estimation protocol treats the site survey sample and the supplemental sample as separate strata. The combined-sample variance estimation uses the full variance estimation protocols (as described) for each of the component designs. The combined sample weight achieves sampling variances that are slightly larger than the minimum variance (approximately 5 to 10 percent larger, depending on the population and variable).

REFERENCES

Anderson, G., R. Heyssel, and R. Dickler. "Competition Versus Regulation: Its Effect on Hospitals." *Health Affairs*, vol. 12, no. 1, 1993, pp. 70-80.

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.

Campbell, E.S., and G.M. Fournier. "Certificate of Need Deregulation and Indigent Hospital Care." *Journal of Health Politics, Policy and Law*, vol. 18, no. 4, 1993, pp. 905-925.

Cantor, Joel C., et al. "Private Employment-Based Health Insurance in Ten States." *Health Affairs*, summer 1995, Exhibit 3, p. 203, and Appendix Exhibit 1, p. 210.

1990 Census of Population and Housing, Supplementary Reports, Metropolitan Areas as Defined by the Office of Management and Budget, June 30, 1993 (1990 CPH-S-1).

Center for Studying Health System Change. *Tech Pub 7: Community Tracking Study Household Survey Public Use File: Users' Guide.* Washington, DC: HSC, June 1998.

Center for Studying Health System Change. *Health System Usage in 12 Community Chapters.* Washington, DC: HSC: 1997a.

Center for Studying Health System Change. *Tech Pub 4: Community Tracking Study Household Survey Instrument.* Washington, DC: HSC: July 1997b.

Chromy, J.R. "Sequential Sample Selection Methods." *Proceedings of the Annual Meeting of the American Statistical Association, Section on Survey Research Methods,* 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.

Cromwell, J. "Impact of State Hospital Rate Setting on Capital Formation." *Health Care Financing Review*, vol. 8, no. 3, 1987, pp. 69-82.

Felt-Lisk, Suzanne. "How HMOs Structure Primary Care Delivery." *Managed Care Quarterly*, vol. 4, no. 4, 1996, pp. 96-105.

Genesys Sampling Systems. *Genesys Sampling Systems Methodology.* Philadelphia: Genesys, 1994.

Gold, Marsha, Jack Hadley, Donna Eisenhower, Charles Metcalf, Lyle Nelson, Karyen Chu, Richard Strouse, and David Colby. "Design and Feasibility of a National Medicaid Access Survey with State-Specific Estimates." *Medical Care Research Review,* September 1995, pp. 405-429.

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 14, 1994.

Hayward R.A., A.M. Bernard, H.E. Freeman, and C.R. Corey. "Regular Source of Ambulatory Care and Access to Health Services." *American Journal of Public Health*, vol. 81, no. 4, 1991, pp. 434-438.

Kemper, P., D. Blumenthal, J.M. Corrigan, P.J. Cunningham, S.M. Felt, J.M. Grossman, L.T. Kohn, C.E. Metcalf, R.F. St. Peter, R.C. Strouse, and P.B. Ginsburg. "The Design of the Community Tracking Study: A Longitudinal Study of Health System Change and Its Effects on People." *Inquiry*, vol. 33, summer 1996, pp. 195-206.

Kenneth P. Johnson, "Redefinition of the BEA Economic Areas," *Survey of Current Business*, February 1995, pp.75-81.

Langa, K.M., and E.J. Sussman. "The Effect of Cost Containment Policies on Rates of Coronary Revascularization in California." *New England Journal of Medicine*, vol. 329, no. 24, 1993, pp. 1784-1789.

Luft, H.S., S.C. Maerki, and J.B. Trauner. "The Competitive Effects of Health Maintenance Organizations: Another Look at Evidence from Hawaii, Rochester and Minneapolis/St. Paul." *Journal of Health Politics, Policy and Law*, vol. 10, no. 4, 1986, pp. 625-658.

Makuc D., J. Madans, and J. Feldman. "Use of Last Physician Visit to Characterize Health Care Utilization." *Proceedings of the Annual Meeting of the American Statistical Association, Section on Survey Research Methods*, 1994, pp. 368-371.

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

Jay, G., T.E Raghunathan, and J. Lepkowski. "Variation in Doctor Visit Reports by Recall Period: A Comparison of Survey and Medical Record Data." In *Initial Studies of Issues Related to the Redesign of the 1995 National Health Interview Survey*. Ann Arbor, MI: University of Michigan, June 1995.

Melnick, G.A., and J. Zwanziger. "Hospital Behavior Under Competition and Cost Containment Policies, The California Experience, 1980-1985." *Journal of the American Medical Association*, vol. 260, no. 18, 1988, pp. 2669-2675.

Merrill, J., and C. McLaughlin. "Competition Versus Regulation: Some Empirical Evidence." *Journal of Health Politics, Policy and Law*, vol. 10, no. 4, 1986, pp. 613-623.

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

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

Research Triangle Institute. "Design of a Survey to Monitor Consumers' Access to Care, Use of Health Services, Health Outcomes, and Patient Satisfaction, Final Report." Washington, DC: 1995. (Submitted to Office of Program Development, Agency for Health Care Policy and Research, Rockville, MD.)

Robinson, J.C., and H.S. Luft. "Competition, Regulation and Hospital Costs, 1972-1982." *Journal of the American Medical Association*, vol. 260, no. 18, 1988, pp. 2676-2781.

Robinson, J.C., and C.S. Phibbs. "An Evaluation of Medicaid Selective Contracting in California." *Journal of Health Economics*, vol. 7, 1989, pp. 437-455.

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

Shortell, S.M., and E.F.X. Hughes. "The Effects of Regulation, Competition and Ownership on Mortality Rates Among Hospital Inpatients." *New England Journal of Medicine*, vol. 318, no. 17, 1988, pp. 1100-1107.

Strouse, Richard, and John Hall. "Incentives in Population Based Health Surveys." *American Statistical Association, Proceedings of the Section on Survey Research Methods*, 1997.

Strouse, R., J. Hall, B. Carlson, and J. Cheng. "Impact of the Nontelephone Sample of Family Health Insurance Estimates." Report to the Robert Wood Johnson Foundation, March 1997.

Thornberry, Owen T., and J. Massey, "Trends in United States Telephone Coverage Across Time and Subgroups." In *Telephone Survey Methodology,* edition by R. Groves et. al. New York: John Wiley, 1988.

U.S. Bureau of the Census. Unpublished Tables of Percentages of Households with Telephones in 1997. 1997.

U.S. Bureau of the Census. Unpublished Tables of Percentages of Households with Telephones for 1988-1993. 1994.

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.