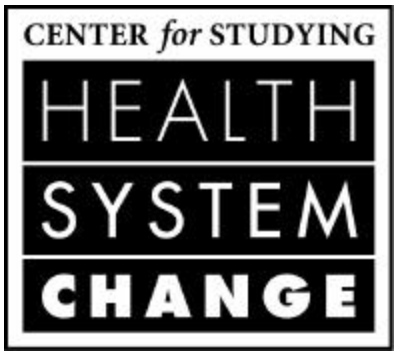
Community Tracking Study

Household Survey Methodology Report 2000-01 (Round Three)



Richard Strouse
Barbara Carlson
John Hall

600 Maryland Avenue, SW
Suite 550
Washington, DC 20024
www.hschange.org

Technical Publication No.

46

July 2003

AUTHORS

Richard Strouse, Center for Studying Health System Change

Barbara Carlson, Mathematica Policy Research, Inc.

John Hall, Mathematica Policy Research, Inc.

Center for Studying Health System Change
600 Maryland Ave., SW, Suite 550,
Washington, DC 20024
www.hschange.org

Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, NJ 08543-2393
www.mathematica-mpr.com

This is one in 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 (HSC), which is funded exclusively by The Robert Wood Johnson Foundation and is affiliated with Mathematica Policy Research, Inc.

Contract No.: 6005-99-05/6006-00-03
MPR Reference No.: 8689-340

MATHEMATICA
Policy Research, Inc.

Report on Survey Methods for the Community Tracking Study's 2000-2001 Round Three Household Survey

Final Report

July 2003

Richard Strouse
Barbara Carlson
John Hall

Submitted to:

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

Project Officer:
Lee Hargraves

Submitted by:

Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, NJ 08543-2393
Telephone: (609) 799-3535
Facsimile: (609) 799-0005

Project Director:
Richard Strouse

CONTENTS

Chapter **Page**

I OVERVIEW 1

A. OBJECTIVES OF THE COMMUNITY TRACKING STUDY 1

B. ANALYTIC COMPONENTS OF THE COMMUNITY TRACKING STUDY 1

C. THE ROUND THREE HOUSEHOLD SURVEY 3

II SITE SELECTION AND SAMPLE DESIGN 7

A. SITE SELECTION 7

1. Definition of Sites 8
2. Number of Sites 8
3. Site Selection 9
4. Additional Samples and Better National Estimates 11

B. HOUSEHOLD, FAMILY INSURANCE UNIT, AND INDIVIDUAL SELECTION 11

1. Households 11
2. Family Insurance Units 12
3. Individuals 14
4. Individuals Excluded from the Survey 15

C. SAMPLE SIZE CONSIDERATIONS FOR THE HOUSEHOLD SURVEY 16

1. Requirements for Site-Based Estimates 16
2. National Estimates, the Second-Tier Sample of Sites, and the Supplemental Sample 18
3. Actual Sample Sizes 19

D. SAMPLE TRACKING AND THE LONGITUDINAL COMPONENT 26

E. RANDOM-DIGIT-DIALING SAMPLE SELECTION 29

1. Sampling Frame 30
2. Stratification and Sample Allocation 31
3. Sample Selection and Release 34

CONTENTS *(continued)*

Chapter **Page**

F. FIELD SAMPLE SELECTION 37

1. Defining Eligibility 38
2. Allocation of the Sample Among the Sites 40
3. Selecting Sample for Round Three 40
4. Sample Release 43

III SURVEY DESIGN AND PREPARATION 45

A. OVERVIEW 45

B. INSTRUMENTATION 46

C. ADVANCE MATERIALS, SURVEY INTRODUCTION, AND INCENTIVES 53

1. Advance Letters 54
2. Survey Introduction 55
3. Incentives. 55

D. INTERVIEWER SELECTION AND TRAINING 56

1. RDD Sample 56
2. Field Sample 58

E. CATI SYSTEM 59

IV DATA COLLECTION 61

A. OVERVIEW 61

B. ORGANIZATION OF THE RDD AND FIELD SURVEYS 61

1. RDD Survey 61
2. Field Survey 63

C. RESPONSE RATES 63

1. Calculation of Response Rates 63
2. Patterns in Household and FIU Response Rates, by Sample Type 75
3. Response Rates for the Adult Self-Response Modules and Child's Physician Visit 91

CONTENTS *(continued)*

Chapter **Page**

D. EFFORTS TO INCREASE RESPONSE TO THE SURVEY 93

1. Follow-Up Calls for the RDD Sample 93
2. Follow-Up Calls for the Field Sample 94
3. Interviews Conducted in Spanish 95
4. Refusal Conversions 95
5. Monetary Incentives 100
6. Messages on Mechanical Answering Devices 105

E. QUALITY ASSURANCE 105

1. RDD Sample 105
2. Field Sample 106

F. DATA EDITING, CODING, AND CLEANING 107

G. REFORMATTING DATA FILES AND FILE DELIVERY 108

V WEIGHTING AND ESTIMATION 109

A. OVERVIEW 109

1. Weights Provided for Public Use 110
2. Constructing Weights 112
3. Sampling Error Estimation 113

B. WEIGHTING THE RDD COMPONENT 114

1. General Weighting Approach 114
2. Calculating the Base Weight for National Estimates Using the Site Sample 129
3. Calculating the Base Weight to Make National Estimates Using the Supplemental Sample 131
4. Using the Augmented Site Sample to Make National Estimates 131
5. Using the Augmented Site Sample to Make Site-Specific Estimates 133

C. WEIGHTS FOR THE FIELD SAMPLE 134

1. Steps in the Weighting Process 134

CONTENTS *(continued)*

Chapter **Page**

D. INTEGRATED WEIGHTS FOR THE HOUSEHOLD SURVEY 142

1. Telephone Service Interruption Adjustment Factor 145
2. Poststratification of Person-Level Integrated Weights 148

E. TRIMMING PERSON WEIGHTS 148

F. WEIGHTS FOR COMBINING THE SITE AND SUPPLEMENTAL SAMPLE SURVEYS 151

G. SAMPLING ERROR ESTIMATION 154

1. Background 155
2. Variance Estimation 157

APPENDIX A: ROUND THREE SURVEY INSTRUMENT A.1

APPENDIX B: ROUND THREE ADVANCE LETTERS B.1

APPENDIX C: ROUND THREE RDD TRAINING MANUAL C.1

APPENDIX D: ROUND THREE LOCATER SCREENING MANUAL D.1

APPENDIX E: ESTIMATING RESIDENCY E.1

APPENDIX F: DETAILED CTS ROUND THREE HOUSEHOLD SURVEY RESPONSE RATES F.1

TABLES

Table		Page
I.1	SITES SELECTED FOR THE COMMUNITY TRACKING STUDY	2
II.1	NUMBER OF FAMILY INSURANCE UNITS INTERVIEWED, BY SITE AND BY SAMPLE	20
II.2	NUMBER OF PEOPLE INTERVIEWED, BY SITE AND BY SAMPLE	23
II.3	RANDOM-DIGIT-DIALING SAMPLE STRATA FOR HIGH-INTENSITY SITES	32
II.4	RELEASE OF SAMPLE FOR ROUND THREE OF THE COMMUNITY TRACKING STUDY HOUSEHOLD SURVEY	35
II.5	TELEPHONE PENETRATION, ESTIMATED NUMBER OF NONTELEPHONE HOUSEHOLDS, AND ROUND ONE PRELIMINARY FIELD ALLOCATIONS	41
II.6	FIELD ALLOCATIONS AND COMPLETES FOR ROUNDS TWO AND THREE	42
III.1	CONTENT OF THE ROUND THREE HOUSEHOLD SURVEY	48
III.2	SOURCE OF DATA FOR INDIVIDUALS IN THE ROUND THREE HOUSEHOLD SURVEY, BY QUESTION TOPIC	52
IV.1	NUMBER OF INTERVIEWS COMPLETED WITH FIUs AND PERSONS BY ROUND OF THE CTS HOUSEHOLD SURVEY	62
IV.2	RINAL ROUND THREE CTS HOUSEHOLD-LEVEL SURVEY DISPOSITION (RDD SAMPLE)	65
IV.3	FINAL ROUND THREE CTS FIU-LEVEL SURVEY DISPOSITION (RDD SAMPLE)	67
IV.4	FINAL ROUND THREE CTS HOUSEHOLD SURVEY DISPOSITION (FIELD SAMPLE)	68
IV.5	ROUND THREE CTS HOUSEHOLD –LEVEL RESPONSE RATE, BY SAMPLE TYPE	76
IV.6	ROUND THREE CTS FIU-LEVEL RESPONSE RATE, BY SAMPLE TYPE	77

TABLES (*continued*)

Table		Page
IV.7	CTS SITE LEVEL UNWEIGHTED RESPONSE RATES BY ROUND	80
IV.8	CTS SITE LEVEL WEIGHTED RESPONSE RATES BY ROUND	83
IV.9	COMPARISON OF ROUND CTS TWO AND ROUND THREE FIU RESPONSE RATES BY SAMPLE GROUP	87
IV.10	IMPACT OF COLLATERAL STUDIES ON CTS HOUSEHOLD COOPERATION RATE FOR OVERLAP COMPLETES	90
IV.11	RESPONSE RATES FOR THE CTS ADULT SELF-RESPONSE MODULE, BY ROUND	92
IV.12	SPANISH-SPEAKING FIU INTERVIEWS, BY SITE AND OVERALL	96
IV.13	ROUND THREE HOUSEHOLD LEVEL UNWEIGHTED REFUSAL CONVERSION RATES, BY SAMPLE TYPE	99
IV.14	ROUNDS TWO AND THREE RDD UNWEIGHTED) CUMULATIVE HOUSEHOLD LEVEL RESPONSE RATES, BY NUMBER OF REFUSAL CONVERSIONS	101
IV.15	DISPOSITION OF PRIMARY FIUS IN HOUSEHOLDS OFFERED PREPAID INCENTIVES BY ROUND AND SAMPLE TYPE	104
V.1	NAMES OF ROUND THREE CTS HOUSEHOLD SURVEY WEIGHTS	111
V.2	INTEGRATION OF RDD AND FIELD COMPONENTS FOR NATIONAL ESTIMATES BASED ON SITE OR AUGMENTED SITE SAMPLE	144
V.3	INTEGRATION OF RDD AND FIELD COMPONENTS FOR SITE-SPECIFIC ESTIMATES BASED ON THE SITE OR AUGMENTED SITE SAMPLE	146

FIGURE

Figure **Page**

II.1 EXAMPLE OF FAMILY INSURANCE UNITS IN A HYPOTHETICAL HOUSEHOLD 13

I. OVERVIEW

A. OBJECTIVES OF THE COMMUNITY TRACKING STUDY

The Community Tracking Study (CTS) is the core research effort of the Center for Studying Health System Change (HSC), a nonpartisan policy research organization located in Washington, DC, and funded by the Robert Wood Johnson Foundation. HSC's mission is to inform health care decision makers about changes in the health care system at both the local and the national level, as well as about how such changes will affect people. HSC conducts surveys of those affected by changes in the health care system—households, physicians, and employers—and interviews with health care leaders in 12 communities.

The focus on markets is central to the design of the CTS. Understanding market changes requires a study both of local markets, including their culture and history, and of public policies relating to health care. To track change across the United States, we randomly selected 60 nationally representative communities stratified by region, community size, and whether metropolitan or nonmetropolitan (see Table I.1).[1]

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

B. ANALYTIC COMPONENTS OF THE COMMUNITY TRACKING STUDY

The CTS has qualitative and quantitative components, which are described below.

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

TABLE I.1

SITES SELECTED FOR THE COMMUNITY TRACKING STUDY

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

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

[a]Based on 1992 Census estimates.

1. ***Site Visits.*** Researchers examine the forces affecting health care organizations and how they are responding by interviewing 40 to 60 health care leaders in each of the 12 high-intensity sites. HSC conducts and manages the site visits, with assistance from outside researchers.

2. ***Household Survey.*** This survey, which comprises about 60,000 people in 33,000 families, focuses on assessing whether consumer access to the health care system is improving or declining over time. Particular areas of inquiry include access, satisfaction, use of services, and insurance coverage. The survey also collects information about health status and sociodemographic characteristics. To enhance the reliability of information on health plans, we conducted an "insurance followback" survey of the plans in which household respondents are enrolled in the first two rounds. HSC provides technical direction and oversight and Mathematica Policy Research (MPR) is responsible for sample design, data collection, and weights for the household and followback surveys.

3. ***Employer Survey.*** For the first round of the CTS, we interviewed 22,000 public and private employers to understand how the American population can access the health system nationally and locally. We asked these employers, which span size and industry sector, about the choice of plans they offer, how much their employees contribute to paying for their coverage, whether they participate in a purchasing alliance, and whether they provide high-quality information to their employees. HSC collaborated with RAND on the employer survey, which was not conducted for rounds two or three.

4. ***Physician Survey.*** A sample of more than 12,000 practicing physicians across the country offers perspective on how health care delivery is changing. Physicians respond to a series of questions about whether they are able to provide needed services for patients, how they are compensated, and what effect various care management strategies have on their practices, as well as questions about their practice arrangements. The Gallup Organization conducted the interviewing for the physician survey, and MPR is responsible for the sample design, sample weights, variance estimation, and tracing of physicians who could not be located.

Additional background on CTS is available at HSC's Web site (www.hschange.com).

C. THE ROUND THREE HOUSEHOLD SURVEY

This report describes the design and conduct of the third round of the household survey. Mathematic Policy Research, Inc. (MPR) was the primary contractor for survey and sample design, data collection, sample weights, and variance estimation. Social and Scientific Systems Inc. (SSS) was instrumental in converting the raw survey data into an analysis file. MPR and SSS collaborated to prepare the documentation for the Public and Restricted Use files.

Documentation of Rounds one and two are included on HSC's website (Technical Publications 15 and 34, respectively).

The survey asks about health insurance, use of health services, satisfaction with care and health plans, health status, and demographic information. A family informant provided information on most topics for each adult and one randomly selected child. In addition, each adult answered subjective questions that could not be reliably answered by a proxy respondent; these included unmet health needs, patient trust, satisfaction with physician choice and health plan, health status, risk behaviors, and details of the last physician visit. The adult family member who took the sampled child to his or her last physician visit was asked questions about that visit. (The adult may not have been the family informant.) A Spanish version of the instrument was used when appropriate. The survey was administered by computer assisted telephone interview (CATI).

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

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

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

each adult in the FIU. (The FIU is based on groupings of people typically used by insurance carriers. It includes an adult household member, spouse, and dependent children up to the age of 18 (or ages 18 to 22 if the child is in school).

Round Three interviews with 59,725 individuals in 32,669 FIUs were completed between September 2000 and September 2001. The weighted FIU-level response rate was 58.6 percent.

Reports describing the first two rounds of the household survey are included in Technical Publications #15 and #34 (www.hschange.com). In this report, we discuss the sampling design of the Round Three sample (Chapter II), survey design and preparation (Chapter III), data collection (Chapter IV), and sample weighting (Chapter V). The appendixes present the survey instruments (Appendix A), advance materials mailed to surveyed households (Appendix B), Training Manual (Appendix C), Locator/Screening Manual for field data collection (Appendix D), alternative procedures for estimating residuary (Appendix E), and detailed response rate tables (Appendix F).

Editing and imputation procedures will be described in The Household Survey Restricted Use File User's Guide for Round Three (forthcoming).

II. SITE SELECTION AND SAMPLE DESIGN

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

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

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

A. SITE SELECTION

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

1. Definition of Sites

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

2. Number of Sites

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

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

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

3. Site Selection

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

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

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

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

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

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

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

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

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

4. Additional Samples and Better National Estimates

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

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

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

B. HOUSEHOLD, FAMILY INSURANCE UNIT, AND INDIVIDUAL SELECTION

1. Households

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

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

2. Family Insurance Units

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

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

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

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

FIGURE II.1

EXAMPLE OF FAMILY INSURANCE UNITS IN A HYPOTHETICAL HOUSEHOLD

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

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

3. Individuals

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

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

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

4. Individuals Excluded from the Survey

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

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

Some FIUs (those listed by, but not including, the household informant) did not respond to the interview. Nonresponding FIUs were excluded from the file but were statistically represented by responding families in the weighting process. A small number of people had such high levels of nonresponse to individual survey questions that they were considered to be nonrespondents and so were excluded from the file. For Round Three, 30 people were excluded for this reason. Adult family members who did not respond to the self-response module were included on the file as long as the core interview contained responses for them.

C. SAMPLE SIZE CONSIDERATIONS FOR THE HOUSEHOLD SURVEY

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

1. Requirements for Site-Based Estimates

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

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

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

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

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

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

A goal of the CTS is tracking change over time and testing hypotheses about the causes of change. Measuring change over multiple interviewing rounds requires larger samples than does conducting cross-sectional surveys. For the CTS, we developed a mixed longitudinal/cross-sectional design for the second and subsequent rounds of the Household Survey that included a mix of households interviewed for the first time and households that had previously been interviewed in the prior round (Metcalf et al. 1996 and HSC Technical Publications 15 and 34). The goal in this approach has been that approximately 40 to 45 percent of the households interviewed in the current round would have been interviewed in the prior round. To measure

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

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

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

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

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

We also wanted to track changes in a way that would enable us to make statements about the nation, as well as about how individual sites compare with the nation. From this national

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

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 that size and would therefore result in poor precision for national estimates.

A sample of 60 sites would increase the precision for large metropolitan areas and would expand the generalizability of the household sample to small metropolitan areas and to nonmetropolitan areas. In addition, we decided to augment the clustered site sample with an unclustered telephone sample of the entire nation. The supplemental sample would not be subject to any site cluster-based design effects and was the most efficient method of expanding the effective size of the national sample. The initial unclustered sample size of the supplemental sample was approximately 3,500 FIUs, which we later reduced slightly; a total of 3,276 FIUs were interviewed from the first round, 3,251 from the second, and 3,095 from the third.

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

3. Actual Sample Sizes

The number of FIUs and the number of individuals interviewed, by site and type of sample, are shown in Tables II.1 and II.2, respectively. A total of 32,669 FIUs and 59,725 people were

TABLE II.1

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

	Sample			
Site/Geographic Area[a]	Site	Supplemental	Augmented Site	Combined
High-Intensity Sites				
01-Boston (MA)	1,196	57	1,253	1,253
02-Cleveland (OH)	1,199	21	1,220	1,220
03-Greenville (SC)	1,231	10	1,241	1,241
04-Indianapolis (IN)	1,279	20	1,299	1,299
05-Lansing (MI)	1,256	6	1,262	1,262
06-Little Rock (AR)	1,376	7	1,383	1,383
07-Miami (FL)	1,225	10	1,235	1,235
08-Newark (NJ)	1,261	19	1,280	1,280
09-Orange County (CA)	1,195	31	1,226	1,226
10-Phoenix(AZ)	1,180	28	1,208	1,208
11-Seattle (WA)	1,088	27	1,115	1,115
12-Syracuse (NY)	1,221	6	1,227	1,227
Low-Intensity Sites				
13-Atlanta (GA)	229	39	268	268
14-Augusta (GA/SC)	260	7	267	267
15-Baltimore (MD)	281	26	307	307
16-Bridgeport(CT)	280	8	288	288
17-Chicago (IL)	282	72	354	354
18-Columbus (OH)	328	17	345	345
19-Denver (CO)	277	42	319	319
20-Detroit (MI)	308	59	367	367
21-Greensboro (NC)	284	11	295	295
22-Houston (TX)	289	33	322	322
23-Huntington (WV/KY/OH)	304	7	311	311
24-Killeen (TX)	278	4	282	282
25-Knoxville (TN)	280	8	288	288
26-Las Vegas (NV/AZ)	285	17	302	302

TABLE II.1 *(continued)*

Site/Geographic Area[a]	Sample			
	Site	Supplemental	Augmented Site	Combined
27-Los Angeles (CA)	289	96	385	385
28-Middlesex (NJ)	291	15	306	306
29-Milwaukee (WI)	307	20	327	327
30-Minneapolis (MN/WI)	323	31	354	354
31-Modesto (CA)	336	9	345	345
32-Nassau (NY)	292	29	321	321
33-New York City (NY)	324	65	389	389
34-Philadelphia (PA/NJ)	335	58	393	393
35-Pittsburgh (PA)	290	25	315	315
36-Portland (OR/WA)	356	29	385	385
37-Riverside (CA)	323	25	348	348
38-Rochester (NY)	419	15	434	434
39-San Antonio (TX)	316	26	342	342
40-San Francisco (CA)	255	21	276	276
41-Santa Rosa (CA)	288	5	293	293
42-Shreveport (LA)	308	6	314	314
43-St. Louis (MO/IL)	359	23	382	382
44-Tampa (FL)	318	25	343	343
45-Tulsa (OK)	333	7	340	340
46-Washington (DC/MD)	298	72	370	370
47-W Palm Beach (FL)	278	17	295	295
48-Worcester (MA)	301	6	307	307
49-Dothan (AL)	335	4	339	339
50-Terre Haute (IN)	285	1	286	286
51-Wilmington (NC)	271	4	275	275
52-W-Cen Alabama	365	0	365	365
53-Cen Arkansas	411	11	422	422
54-N Georgia	257	15	272	272
55-NE Illinois	303	2	305	305
56-NE Indiana	290	3	293	293

TABLE II.1 *(continued)*

Site/Geographic Area[a]	Sample			
	Site	Supplemental	Augmented Site	Combined
57-E Maine	326	7	333	333
58-E North Carolina	337	12	349	349
59-N Utah	431	3	434	434
60-NW Washington	352	2	354	354
Areas Other than CTS Sites		1,814		1,814
Total	**29,574**	**3,095**	**30,855**	**32,669**

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

TABLE II.2

NUMBER OF PEOPLE INTERVIEWED, BY SITE AND BY SAMPLE

	Sample			
Site/Geographic Area[a]	Site	Supplemental	Augmented Site	Combined
High-Intensity Sites				
01-Boston (MA)	2,157	99	2,256	2,256
02-Cleveland (OH)	2,138	46	2,184	2,184
03-Greenville (SC)	2,280	18	2,298	2,298
04-Indianapolis (IN)	2,291	37	2,328	2,328
05-Lansing (MI)	2,307	15	2,322	2,322
06-Little Rock (AR)	2,525	14	2,539	2,539
07-Miami (FL)	2,115	22	2,137	2,137
08-Newark (NJ)	2,282	33	2,315	2,315
09-Orange County (CA)	2,171	44	2,215	2,215
10-Phoenix(AZ)	2,090	51	2,141	2,141
11-Seattle (WA)	1,931	46	1,977	1,977
12-Syracuse (NY)	2,277	6	2,283	2,283
Low-Intensity Sites				
13-Atlanta (GA)	416	68	484	484
14-Augusta (GA/SC)	484	10	494	494
15-Baltimore (MD)	516	51	567	567
16-Bridgeport(CT)	541	11	552	552
17-Chicago (IL)	516	133	649	649
18-Columbus (OH)	625	29	654	654
19-Denver (CO)	503	73	576	576
20-Detroit (MI)	585	101	686	686
21-Greensboro (NC)	516	23	539	539
22-Houston (TX)	542	71	613	613
23-Huntington (WV/KY/OH)	548	11	559	559
24-Killeen (TX)	517	6	523	523
25-Knoxville (TN)	501	15	516	516
26-Las Vegas (NV/AZ)	495	27	522	522

TABLE II.2 *(continued)*

Site/Geographic Area[a]	Sample			
	Site	Supplemental	Augmented Site	Combined
27-Los Angeles (CA)	497	163	660	660
28-Middlesex (NJ)	565	35	600	600
29-Milwaukee (WI)	557	43	600	600
30-Minneapolis (MN/WI)	605	56	661	661
31-Modesto (CA)	638	15	653	653
32-Nassau (NY)	550	58	608	608
33-New York City (NY)	537	108	645	645
34-Philadelphia (PA/NJ)	606	100	706	706
35-Pittsburgh (PA)	526	46	572	572
36-Portland (OR/WA)	663	51	714	714
37-Riverside (CA)	617	55	672	672
38-Rochester (NY)	786	25	811	811
39-San Antonio (TX)	577	39	616	616
40-San Francisco (CA)	394	35	429	429
41-Santa Rosa (CA)	535	8	543	543
42-Shreveport (LA)	561	10	571	571
43-St. Louis (MO/IL)	682	45	727	727
44-Tampa (FL)	546	43	589	589
45-Tulsa (OK)	611	12	623	623
46-Washington (DC/MD)	558	133	691	691
47-W Palm Beach (FL)	479	29	508	508
48-Worcester (MA)	579	8	587	587
49-Dothan (AL)	652	7	659	659
50-Terre Haute (IN)	538	3	541	541
51-Wilmington (NC)	474	7	481	481
52-W-Cen Alabama	658	0	658	658
53-Cen Arkansas	786	21	807	807
54-N Georgia	465	33	498	498
55-NE Illinois	572	2	574	574
56-NE Indiana	574	6	580	580

TABLE II.2 *(continued)*

Site/Geographic Area[a]	Sample			
	Site	Supplemental	Augmented Site	Combined
57-E Maine	594	11	605	605
58-E North Carolina	604	25	629	629
59-N Utah	937	9	946	946
60-NW Washington	645	5	650	650
Areas Other than CTS Sites		3,382		3,382
Total	**54,037**	**5,688**	**56,343**	**59,725**

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

interviewed in Round 3, compared with 32,047 FIUs and 58,596 in Round Two, and 32,732 FIUs and 60,446 people in Round One. The nominal number of FIUs per augmented high-intensity site varied in Round Three from 1,115 to 1,383. The variation among sites can be explained by the fact that sites are augmented by supplemental sample cases in proportion to their population, and by differential response rates. As shown in Table II.1, supplemental samples in large metropolitan low-intensity sites significantly increased the size of these site samples. In the low-intensity sites, the nominal augmented sample sizes of FIUs sites ranged from 267 to 434.

D. SAMPLE TRACKING AND THE LONGITUDINAL COMPONENT

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

Our approach to measuring changes in the population was to sample Round Two telephone numbers (for the RDD component) and addresses (for the field component) for the Round Three survey. This approach is relatively simple to implement, less costly than tracking individuals, and avoids attrition resulting from inability to locate sample members.

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

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

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

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

In designing the Round Three RDD sample, we incorporated Round Two results that showed large variations in response rates and cost by Round One disposition. Telephone

numbers where Round One interviews were completed were more likely to yield an interview than were those in any other category. The Round Two cost per interview for this group was thus relatively low. Conversely, telephone numbers where a refusal occurred in Round One yielded few interviews and the cost per Round Two interview was quite high. Using the principals of optimum allocation based on cost, we estimated optimal Round Three sampling rates for each Round Two response category.

To reduce complexity in the design we grouped categories and set three sampling rates for Round Two telephone numbers: (1) all the telephone numbers where a Round Two interview was completed were sampled; (2) telephone numbers where no contact was made at Round Two and hard refusers (those who refused two or more times) were subsampled at a rate of 33 percent; (3) telephone numbers where other refusals had occurred at Round Two and telephone numbers that were not household numbers at Round Two were sampled at a rate of 67 percent.

Although following telephone numbers for the RDD longitudinal component of the design may seem novel, the idea of following units other than the unit of observation (which, in the CTS, is the individual) is not new. The Current Population Survey retains housing units in its rotation groups, rather than following individuals or households (Robinson 1992; and U.S. Bureau of the Census 2000). Kish (1965) described two annual surveys (1951 and 1952) that used the same dwellings in both years, with good results for change estimates. Kish mentioned cost and practicality issues when deciding which unit to sample to achieve overlap for longitudinal studies, factors that we also considered. In addition to the CTS, another large, predominantly telephone survey—the National Survey of American Families (NSAF) (Ferraro et al. 2000)—has also used this approach.

Sampling the same telephone numbers is analogous to sampling the same dwellings. Because most people retain the same telephone number and address over a two-year interval,

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

E. RANDOM-DIGIT-DIALING SAMPLE SELECTION

In this section, we describe selection of the RDD samples for the Household Survey. The RDD samples for the three rounds were similar in that the same sites were included, there was a supplemental sample, and strata were defined using the same criteria in both rounds. However, because the design called for a partial overlap, the second and third round RDD samples had new components. In the first round, all telephone numbers were selected for the first time, whereas the RDD samples for Rounds Two and Three included three groups of telephone numbers: (1) those that had been selected in the prior round *(overlap sample)*, (2) those that had no chance of selection in the prior round *(new sample—new working banks)*, and (3) those that had a chance of selection in the prior round but had not actually been selected *(new sample—old working banks)*. (Working banks are defined below.)

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

1. Sampling Frame

We used the Genesys Sampling System to select all the RDD household samples. The entire Round One sample and old and new working banks in Rounds Two and Three were selected by Genesys. The overlap samples for Round Two and Three were selected from a list of the telephone numbers that had been attempted in the prior round.

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

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

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

2. Stratification and Sample Allocation

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

In high-intensity sites containing more than one county, we first stratified by county, assigning the county containing the central city of the MSA in one stratum and the other county or counties in another stratum. Next, we stratified the county containing the central city by race/ethnicity or income distributions. If that county included large black and Hispanic populations, we used both variables for stratification. If the county contained a significant fraction of only one of these population groups, or if one of these groups was dominant, we stratified by the percentage belonging to that group. For example, although Miami was approximately 18 percent black, a majority of the population was Hispanic. Therefore, we stratified on the percentage Hispanic. For sites in which neither the black nor Hispanic population was large enough to stratify on race or ethnicity, we stratified on income. Table II.3 shows the stratification variables for the high-intensity sites.

Although the same stratification criteria were used for all rounds, some exchanges could have "switched" strata or even sites between rounds. For example, the prevalence of Hispanic households in a Miami exchange may have changed between rounds. In practice, few such changes occurred. In Rounds Two and Three fewer than one per cent of all exchanges changed

TABLE II.3

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

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

stratum from the previous round, and fewer than 5 of the more than 15,000 exchanges changed site. Two rules were devised to deal with telephone numbers in exchanges that switched strata or sites:

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

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

We also varied the allocation of sample among overlap sample and new sample from the old and new working banks. For each low-intensity site, each stratum in a high-intensity site, and each stratum of the supplemental sample, we:

- Estimated the expected number of completed interviews (FIUs) from the overlap sample
- Estimated the sample size that should be generated from old working banks, given our goal of approximately equal probabilities of selection, and estimated the number of completed interviews from that sample

- Estimated the sample size that should be generated from new working banks to give numbers in these new banks the same probability of selection for Round Three as for cases in old working banks
- Estimated the number of completed interviews expected from the new working bank sample
- If the estimated number of completed interviews from the overlap and new working bank sample was less than the target number of interviews, calculated the additional amount of sampled telephone numbers needed from the old working banks[9]

3. Sample Selection and Release

The initial sample release was set at 60 percent of the total number of projected telephone numbers. The initial sample was released during August, September, and October of 2000. Subsequent sample releases were made for all sites and the supplemental sample to meet sample size and response rate targets. In the third round, release of reinterview sample (Round Two completes) had to be coordinated with the use of these cases as a frame for surveys by other organizations. This led to some Round Two completes being released late in the data collection period. (See Table II.4 for sample releases.) Toward the end of the survey, sample selection was tailored to meet interviewing targets in specific sites or groups of sites. The steps taken in selecting and releasing the sample included:

- Generating samples of telephone numbers
- Removing known business and nonworking numbers from the sample, using Genesys identification procedures
- Checking against prior releases for duplicates
- Randomly sorting the sample

---

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

TABLE II.4

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

Date	Total RDD	Round Two Completes	Round Two Hard Refuser and Other Nonrespondent	Round Two Other*	New or Old Working Banks	Total Field Sample
August 2000	13,279	0	0	0	13,279	0
September 2000	14,066	6,296	887	6,883	0	0
October 2000	13,141	5,032	2,784	5,325	0	0
November 2000	6,172	6,172	0	0	0	0
December 2000	11,359	2,836	0	0	8,523	4,532
January 2001	0	0	0	0	0	356
February 2001	2,570	2,570	0	0	0	0
March 2001	439	439	0	0	0	0
April 2001	5,292	1,033	0	0	4,259	0
May 2001	938	938	0	0	0	891
June 2001	0	0	0	0	0	157
Total	**67,255**	**25,316**	**3,671**	**12,208**	**26,061**	**5,936**

*Nonhousehold, no answer, mechanical answering device.

- Releasing sample to the automated call scheduler
- Using data collection reports to reestimate the size of future releases

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

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

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

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

F. FIELD SAMPLE SELECTION

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

Restricting the field sample to the 12 high-intensity sites reduced coverage bias for estimates made for the 48 large metropolitan sites as a whole and for estimates made for each of the 12

high-intensity sites. This option was far less expensive than collecting data through field interviewing in all 60 sites. However, limiting the field sample to the 12 high-intensity sites meant that families and people who did not have telephones and who lived in nonmetropolitan areas or in metropolitan areas with populations of fewer than 200,000 were not represented. (Weighting procedures to adjust for the absence of these households in national and other estimates are discussed in Chapter V.)

Within the 12 high-intensity sites, the strategy was (1) to sample geographic clusters with probability proportional to size, where size was the estimated number of nontelephone households; (2) count, list, and select housing units within these clusters; and (3) screen this sample for eligible households. Respondents in eligible households were then interviewed over cellular telephones, provided by MPR field staff to survey respondents. Thus, all interviews were conducted by CATI, which prevented differential response resulting from different interviewing modes.

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

1. Defining Eligibility

In defining eligibility, the term *nontelephone household* meant that the household was always or intermittently without telephone service. The field component was designed to include these households. In contrast, in the approach used by the decennial Census and the Current

Population Survey, households were classified as telephone or nontelephone on the basis of the presence or absence of a telephone at the time of interview.[10]

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

In Round Three, we defined *significant interruption* to mean two weeks or more of interrupted service in the 12 months prior to the screening interview (or since the date the household moved into the area defined by our site, if the move occurred after we started data collection for the RDD sample) and used questions about the length of interruptions to adjust sample weights.[11] The only exception to the two-week rule was that households also were eligible for the field survey if members had moved to the listed address within the two weeks preceding the interview and had been without a telephone since moving in.

[10]The 1990 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?"

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

2. Allocation of the Sample Among the Sites

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

We reviewed our experience from Rounds One and Two and retained the second round allocations for Round Three. The Round Two and Three target allocations and completed interviews are shown in Table II.6.[12]

3. Selecting Sample for Round Three

For Round Three, we contacted 5,936 addresses, all but 157 of which had been listed during Round One.[13] The initial release included either half or all of the addresses contacted in Round Two. If additional addresses were needed, we first released additional addresses contacted in Round Two. In two sites (Boston and Seattle) we also released addresses that had been listed in Round One, but that had not been contacted. We did not include new areas in the Round Three sample. We assumed that dwellings found in areas that had no chance of selection in Round One would most likely be housing constructed since Round One, and that they would have a low likelihood of including households without telephones.

During Round One, each sampled cluster had been divided into 10 replicates, each of which contained approximately one-tenth of the cluster's estimated nontelephone households. A replicate might contain several blocks; at the other extreme, one large block might comprise several replicates. The Round One release of blocks for listing consisted of a number of

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

[13]These 157 were "supplemental listings," housing units (as defined by the Census Bureau) in listed areas that had been missed in prior rounds or were perhaps newly constructed.

TABLE II.5

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

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

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

[b]Expected FIU interviews.

TABLE II.6

FIELD ALLOCATIONS AND COMPLETES FOR ROUNDS TWO AND THREE

	Household Interviews		
Site	Target (Each Round)	Completed Round Two	Completed Round Three
Boston (MA)	9	5	21
Cleveland (OH)	36	55	46
Greenville (SC)	65	96	69
Indianapolis (IN)	70	74	97
Lansing (MI)	22	26	20
Little Rock (AR)	64	69	61
Miami (FL)	34	19	37
Newark (NJ)	48	53	80
Orange County (CA)	7	7	11
Phoenix (AR)	54	108	76
Seattle (WA)	49	15	47
Syracuse (NY)	42	35	30
Total	**500**	**562**	**595**

replicates chosen to minimize variation of overall probability of selection within a site, subject to listing enough housing units overall to meet projected targets.

4. Sample Release

The initial release took place in December 2000 and was based on the estimated amount of sample required to meet our targets. Subsequent releases were based on shortfalls in specific sites. The sample release of the field component is summarized in Table II.4.

III. SURVEY DESIGN AND PREPARATION

A. OVERVIEW

The CTS Household Survey is the primary instrument for assessing the effects of health system change on individuals, including changes in health insurance coverage, access to care, use of health services, and satisfaction with health care. As described in Chapter II, the FIU is the primary interviewing unit for the survey, with selected subjective questions also asked of each adult FIU member. Within each FIU, questions are asked about all adults and about one randomly selected child. An adult familiar with the health care experiences of the individuals in the FIU serves as the informant for other adults on questions about health insurance, employment, demographics, and health services use during the 12 months preceding the survey. Each adult in the FIU (including the informant) also is asked to self-respond to questions about difficulty obtaining needed health services, health status, chronic diseases, tobacco use, details about her or his last physician visit, level of satisfaction with that visit, satisfaction with health plan features, level of physician trust, consumer preferences, attitudes toward risk, and consumer information about health. The adult who took the randomly selected child to the last physician visit before the survey was asked questions about that visit, the child's health status, and any chronic disease the child had.

The length of the interview varied with the number of individuals in the FIU and the complexity of the individuals' experiences with health care. The core interview, which is asked of the family informant, averaged 29.2 minutes and the self response module averaged 18.7 minutes. Thus, the core interview generally averages about 30 minutes for the family informant

and 15 to 20 minutes for each additional adult in the family. Since the FIU is designed to mimic the structure of health plans, few FIUs have more than two adults.

B. INSTRUMENTATION

The survey instruments for all three rounds were developed by staff at HSC and MPR, with consultation and review by several experts.[1] Respondents to the Round Three survey were questioned about the following topics:

- Household composition
- Health insurance coverage
- Use of health services
- Unmet needs and expenses
- Usual source of care
- Patient trust and satisfaction
- Plan satisfaction
- Last visit to a physician or other healthcare provider
- Health status (including SF12 ™ Health Survey)[2] and overall satisfaction with life
- Presence of chronic diseases
- Risk behaviors and smoking
- Consumer information
- Employment, earnings, and income
- Demographic characteristics

[1]See Chapter III in Technical Publications 15 and 34, respectively, for a discussion of the initial instrument design and changes made for Round Two.

[2] To maintain continuity with prior rounds of the survey, we used the SF12 ™ Health Survey, standard U.S. Version 1.0, 1994.

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

Different respondents were asked different questions, and not all questions were asked of all respondents (see Table III.2). For example, only the household informant was asked about household composition. Family informants were asked to answer questions about the family and individual family members. Each adult also provided information on topics that the informant could not provide, such as unmet need for medical care, patient trust, satisfaction, health status, chronic diseases, and risk behaviors. If the family had children under 18, the adult who took the sampled child to the doctor on her or his last visit was asked to answer questions about that visit.

Most of the CTS Household Survey interviews were obtained from the RDD sampling frame. As described above, 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. Households in the area probability sample were administered a screening interview to identify eligible households who were then interviewed by cellular telephone (see Appendix A).

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.

TABLE III.1

CONTENT OF THE ROUND THREE HOUSEHOLD SURVEY

Health Insurance	
Private insurance coverage (Section B)	Covered by employer- or union-related private insurance Covered by other private insurance: Purchased directly Premium for directly purchased private insurance Provided by someone not in household
Public insurance coverage (Section B)	Covered by Medicare Covered by both Medicare and supplemental private insurance Premium for supplemental private insurance Covered by both Medicare and Medicaid Covered by Medicaid Covered by other public insurance (military, Indian Health Service, other state and local)
Uninsured (Section B)	Not covered by public or private insurance
Continuity of coverage/changes in coverage (Section B)	Currently insured; lost coverage during previous 12 months Currently uninsured; obtained coverage during previous 12 months Uninsured during all of previous 12 months Uninsured at some point during previous 12 months Reasons for losing health insurance coverage Any type of change in health coverage: Changed private insurance plans Reasons for changing private plans Whether previous plan was HMO/non-HMO Changed from public or private plans Obtained or lost coverage
Insurance plan attributes (Section B)	Whether plan requires signing up with primary care physician or clinic for routine care Whether plan requires approval or referral to see a specialist Whether plan requires choosing a physician or clinic from a book, directory, or list Whether plan is an HMO Whether plan will pay any costs for out-of-network care
Other insurance variables (Section B)	Ever enrolled in an HMO Number of total years enrolled in an HMO
Access to Health Care	
Usual source of care (Section D)	Currently has/does not have a usual source of care Type of place of usual source of care Type of professional seen at usual source of care Reason for changing usual source of care
Travel/waiting time for physician visit (Section E)	Lag time between making appointment and seeing physician at last physician visit[a] Travel time to physician's office for last visit[a] Time spent in waiting room before seeing medical person at last physician visit[a]
Difficulty getting needed services in previous year (Section C)	Did not receive needed services[a] Delay in receiving needed services[a] Reasons for delay or for not receiving needed services[a] *Did not get needed prescriptions*[a]

TABLE III.1 *(continued)*

Resource Use	
Use of ambulatory services in previous 12 months (Section C)	Number of physician visits Number of emergency room visits Number of visits to nonphysician providers (nurse practitioner, physician assistant, midwife) Whether had any mental health visits Number of surgical procedures
Use of inpatient services in previous 12 months (Section C)	Number of overnight hospital stays Number of overnight hospital stays excluding delivery/birth Number of inpatient surgical procedures Total number of nights spent in hospital
Preventive services use (Section C)	Whether person had flu shot during previous 12 months Whether person ever had mammogram (asked of women) If yes, time elapsed since last mammogram
Nature of last physician visit (Section E)	Reason for last visit: Illness or injury[a] Checkup, physical exam, other preventive care[a] Type of physician seen at last visit (PCP or specialist)[a] Whether last visit was to usual source of care[a] Whether last visit was to an emergency room[a] Whether last visit was with appointment or walk-in[a]
Costs (Section C)	Total family out-of-pocket expenses for health care during previous 12 months
Satisfaction and Patient Trust	
General satisfaction (Section E)	Overall satisfaction with health care received by family Satisfaction with choice of primary care physicians[a] Satisfaction with choice of specialists[a]
Satisfaction with last physician visit (Section E)	Satisfaction with thoroughness and carefulness of exam[a] Satisfaction with how well physician listened[a] Satisfaction with how well physician explained things[a]
Satisfaction with health plan (Section E)	*Satisfaction with referrals (CAHPS)*[a] *Satisfaction with health plan approvals (CAHPS)*[a] *Satisfaction with paperwork (CAHPS)*[a] *Satisfaction with amount paid for health care (NHIS)*[a] *Overall plan satisfaction (CAHPS)*[a]
Patient's trust in physicians (Section D)	Agree/disagree that physician may not refer to specialist when needed[a] Agree/disagree that physician may perform unnecessary tests or procedures Agree/disagree that physician is influenced by health insurance company rules[a] Agree/disagree that physician puts patient's medical needs above all other considerations[a]

TABLE III.1 *(continued)*

Employment and Earnings	
Employment status and characteristics (Section F)	Whether adult respondent has the following characteristics: Owned a business or farm Worked for pay or profit during previous week Had more than one job or business Worked for private company/government/self-employed/family business Average hours worked per week, at primary job and at other jobs Size of firm (number employees), at site where respondent works, and at all sites Type of industry
Earnings (Section F)	Earnings, from primary job and from all jobs
Health insurance options at place of employment (Sections B and F)	Whether eligible for health insurance coverage by employer Reasons for ineligibility Whether offered health insurance coverage by employer Reasons for declining coverage (if eligible but not covered) Whether offered multiple plans Whether offered HMO plan Whether offered non-HMO plan
Other Variables	
Demographics (Section A)	Age Gender Highest education level completed Whether interview was administered in Spanish CTS site State County[b]
Health status (Section E)	Overall health status (five-point scale, from excellent to poor)[a] Limited in moderate activity[a] Limited in climbing stairs[a] Accomplished less because of physical health[a] Limited in kind of work because of physical health[a] Accomplished less because of emotional health[a] Less careful in work because of emotional health[a] Pain interfered with work[a] How much time have health problems interfered with social activities[a] How much time calm and peaceful[a] How much time have energy[a] How much time downhearted/blue[a] SF-12 scores: Physical Component Summary; Mental Component Summary[a]
Global satisfaction (Section E)	*Taken altogether, how would you say things are these days? Would you say that you are very happy, pretty happy, or not too happy? (GSS)* [a]
Chronic conditions- modified for Round 3 (Section E)	Presence of chronic conditions, including recent childbirth, abnormal uterine bleeding, alcohol-related problems, cataracts, diabetes, arthritis, asthma, pulmonary disease, hypertension, *coronary heart diseases*, skin cancer, other cancers, depression, *other health problem limiting normal activities* [a] [b]
Family income (Section G)	Family income Race, ethnicity
Consumer preferences (Section B)	Whether person would be willing to accept limited provider choice to save on out-of-pocket expenses[a]

TABLE III.1 *(continued)*

Risk behaviors (Section E)	Whether person agrees that he/she is more likely to take risks than the average person[a]
Smoking cessation interventions (Section E)	Whether person has smoked at least 100 cigarettes in lifetime[a] Whether currently smoking cigarettes every day, some days, or not at all[a] Average number of cigarettes smoked per day during previous 30 days[a] How long since quit smoking[a] Whether stopped smoking one day or longer during previous 12 months, in effort to quit[a] Whether physician advised smoker to stop smoking during previous 12 months[ac]
Consumer information (Section E)	*Looked for obtained information about health concerns from Internet, friends and relatives, TV or radio, books and magazines, other (specify)*[a] *During last 12 months, mentioned or shown doctor information about condition or treatment found or were told by others* *Did doctor order test, procedure, or prescription mainly because of information shown by respondent*[a]

Note: New or changed questions shown in italics.

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

[b] Available on the Restricted Use File only.

[c] Due to an error in CATI programming, this question was missing for a significant fraction of the sample and should not be used in analyses.

TABLE III.2

SOURCE OF DATA FOR INDIVIDUALS IN THE ROUND THREE HOUSEHOLD SURVEY, BY QUESTION TOPIC

(Illustrative Household Described in Figure II.1)

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

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

The Round Three instrument included nearly all of the Round Two questions (see Table III.1); changes to the instrument are summarized below:

- **Unmet need for prescriptions (c84).** A question was added on unmet need for prescription services (source: National Health Interview Survey (NHIS)).
- **Plan satisfaction (CAHPS10,23,37,38;SP14*).*** Questions on several dimensions of plan satisfaction were added to the survey, including referrals, approvals, paperwork, amount paid for health care, and overall plan satisfaction. Except for the amount paid for health care, which was selected from the NHIS, all other items were taken from the Consumer Assessment of Health Plans Survey (CAHPS®).
- **Global satisfaction (GS1571).** Market variation in subjective measures may be due to non-health community level factors. We added a question on overall satisfaction from the General Social Survey as a control variable.
- **Consumer information (ra34, ra36,ranew).** Consumer awareness and use of the Internet and direct-to-consumer advertising may have a significant impact on the use of health resources and patient-physician encounters. We modified questions from the RAND (2001) Consumer and Health Care Quality Information Survey on information sources used, whether the respondent provided his or her physician with information about a condition or treatment, and whether the physician ordered a test, procedure, or prescription because of this information. We obtained the provider's perspective by including related questions in the CTS Physician Survey.
- **Chronic conditions. (cc1-cc7).** The chronic condition check list used in Round 2 was reduced by excluding several items and revising others; the revised checklist is shown in Table III.1.

C. ADVANCE MATERIALS, SURVEY INTRODUCTION, AND INCENTIVES

Notifying potential respondents by mail before an initial call is made can reassure them about a survey's authenticity and purpose. The willingness of the general public to participate in a survey may also be increased by obtaining sponsorship or endorsement from well-known public organizations (usually a government agency), and by designing a convincing survey introduction describing the survey's purpose and value. Monetary incentives also can be effective in increasing response rates and retaining participants in a longitudinal survey. For Rounds One and Two, we tested the content of the survey introduction, the effectiveness of advance information about the study, and the amount and form of monetary incentives (results

discussed in Chapter IV). Based on findings from prior rounds of the survey, we standardized advance letters (see Appendix B), the text of the survey introduction, and incentives.

1. Advance Letters

Different approaches were used to notify households in the overlap complete component (Round Two completions sampled for Round Three) than for those in the Round Two noncomplete or new sample components. We also decided to drop endorsements from state health agencies because they were costly to obtain for a national sample and were not effective in prior rounds.

a. Overlap Complete

Since respondent incentives were used in Round Two, we had accurate information on the names and addresses of check recipients. Consequently, all households in the overlap complete sample were mailed an advance letter thanking them for participating in the last round and providing background on the upcoming survey. Since we were sampling telephone numbers rather than individuals or families, a small fraction of these letters were mailed to individuals who had moved; in most of these cases, the letter was returned as undeliverable.

b. Overlap Noncomplete and New Sample

Based on the results of an experiment conducted in Round One with households that had published addresses, mailing advance letters without monetary incentives did not significantly increase the initial cooperation rate. Consequently, we decided against mailing advance letters to these households for Round Three. However, interviewers were more comfortable knowing that advance letters were mailed and we subsequently mailed letters to nonresponding households with published addresses, except for those in the overlap sample that refused to participate in Round Two. The packet also included a brochure describing the study. (Since households in

the overlap complete component received this brochure in Round Two, we did not mail it to them again.)

2. Survey Introduction

We used different survey introductions for (1) overlap complete households (all of whom were mailed advance letters), (2) other households with published addresses, and (3) households for whom we did not have published addresses. A separate introduction also was used for field sample (see Appendix B). The content of survey introductions was similar to Round Two, briefly explaining the survey's purpose, prior contact (for overlap complete sample), the advance letter (if one was mailed), and the promised incentive. We provided interviewers with additional text to answer respondents' questions, including an explanation of why health tracking is important, examples of the types of questions included in the survey, a contact at The Foundation to verify the survey's authenticity, and additional background on sponsorship, length, and respondent selection. Based on prior experience and experiments, we concluded that a brief introduction with flexible responses to respondent's questions was more effective than a lengthy one.

3. Incentives

Unlike Round Two, where we only compensated the family informant, we gave $25 to each adult participating in Round Three. The primary reason for this change was that the household survey was being used as a sampling frame for two other surveys that offered *individuals* $25 for participating. These studies were the Health Care for Communities (HCC) Survey[3], which focused on mental health and substance abuse issues, and the Community Quality Index (CQI)

[3] The HCC Survey was conducted by the UCLA-NPI/RAND Research Center on Managed Care for Psychiatric Disorders.

Survey[4], which requested medical records from sampled participants. Many Round Three survey participants had participated in those surveys and some would be asked to do so again. Also, the self response module was more than half the length of the core interview, so non-informant adults responding to the survey were more likely to expect comparable compensation. Since it was very important for the survey to obtain a high cooperation rate from all adults in sampled families, we felt that the added expense was justified.

All sampled respondents were promised incentives with their initial letter and call. In addition, we promised incentives for most followup calls. However, near the end of data collection, we mailed $25 checks to individuals in nonresponding households for whom we had names and addresses, rather than promising checks. These included informants for the Round Two survey who were selected for Round Three, as well as part of the sample interviewed for the first time for whom we had also had complete names and addresses. Experience in Round Two demonstrated that shifting from promised to prepaid incentives resulted in faster responses and slightly higher cooperation rates than continuing with promised incentives. On the other hand, this procedure was costly since some nonrespondents cashed checks. The results of this effort are described in Chapter IV.

D. INTERVIEWER SELECTION AND TRAINING

1. RDD Sample

a. Recruitment

Interviewing for the RDD sample was conducted by MPR in its Columbia and Princeton telephone centers. Altogether, 243 telephone interviewers were trained for the Round Three household survey, virtually the same number as in Round Two (238). Interviewing supervisors

[4] The CQI Survey was conducted by RAND.

were given a more detailed manual that contained additional information enabling them to respond to interviewers' questions, and to resolve routine problems; this manual is included in Appendix C. Interviewers received a question-by-question review of the survey, approaches to contacting respondents, disposition coding, a summary of the interviewer bonus plan, and follow-up training on interviewing problems and refusal avoidance (sections V-VII and IX-XIII of the training manual).

b. Telephone Interviewer Training Program

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

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

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

- Exercises to test respondents' skills in obtaining cooperation (see Chapter VIII of the *Training Manual*)
- Review of disposition coding and call scheduling (see Chapter IX of the *Training Manual*)
- Review of the interviewer bonus plan[5]

At the end of the data collection period, interviewers were debriefed on the effectiveness of training, interviewing problems (especially refusal conversion procedures), the interviewer bonus plan, and respondent incentives.

Because initial refusal rates for all rounds of the survey were high, considerable efforts were devoted to preparing interviewers for follow-up calls during which they attempted to convert refusals. Supervisors reinforced training techniques throughout the survey by monitoring calls and providing regular feedback; approximately 10 percent of the interviews were monitored. In addition, we conducted refusal conversion training sessions, during which trainers reviewed effective approaches and interviewers shared experiences about the success or failure of various techniques.

2. Field Sample

a. Recruitment

A total of 14 MPR staff, all of whom worked on the Round Two survey, were trained to screen households in the 12 high-intensity sites for Round Three; one person quit and was replaced. We did not list any new segments for Round Two; however, interviewers were

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

responsible for listing new dwellings within existing segments (described as supplemental listing).

b. Training

All of the trainees had participated in Round Two; consequently, the training session was limited to a review of data collection procedures. Training was conducted during a two-hour conference call in which the MPR trainer reviewed screening procedures with trainees, including the survey introduction, refusal avoidance, the telephone status screener, operation of the cellular telephone, and follow-up interviewing methods (such as attempting contacts at varying times of the day and gaining entry to apartment buildings). After completing the training program, each trainee called the MPR telephone center and conducted a mock screening interview with a supervisor. The manual provided to field listers and screening interviewers is included in Appendix D. Because field staff called the MPR telephone center and then gave the respondent a cellular telephone to complete the interview, they did not have to be trained on how to conduct the survey.

E. CATI SYSTEM

All data collected for the CTS Household Survey were produced using computer programs made available through the Computer Assisted Survey Methods Program (CSM), University of California, Berkeley.[6]

MPR used CASES to develop instruments and data cleaning programs for the CTS. In addition, we developed customized programs for allocating the sample and for controlling the

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

distribution and timing of calls and developed specialized reports for monitoring the survey results (discussed in the next chapter).

IV. DATA COLLECTION

A. OVERVIEW

For Round Three, we interviewed 32,669 FIUs—31,744 from the RDD sample and 925 from the field sample. The FIUs included 49,603 eligible adults and 10,122 sampled children younger than age 18, for a total of 59,725 individuals (see Table IV.1) The unweighted Round Three household-level response rate was 64.2 percent and the unweighted FIU-level response rate was 60.7 percent. The weighted response rates for Round Three were 62.0 percent (household level) and 58.6 percent (FIU level).

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

B. ORGANIZATION OF THE RDD AND FIELD SURVEYS

1. RDD Survey

Interviewing for the RDD sample was conducted from September 2000 to September 2001 in MPR's Princeton, New Jersey, and Columbia, Maryland telephone centers by 243 interviewers.

Reports on field progress were transmitted daily to the telephone centers. The survey reports enabled project managers and interviewing supervisors to monitor production and performance continuously. Several reports were produced, including:

Status disposition reports showing daily and cumulative distributions of interim and final survey disposition codes (completions, various nonresponse and ineligibility dispositions,

TABLE IV.1

NUMBER OF INTERVIEWS COMPLETED WITH FIUs AND PERSONS BY ROUND OF THE CTS HOUSEHOLD SURVEY
(Numbers)

	Round One	Round Two	Round Three
Number of FIUs			
RDD	32,079	31,278	31,744
Field	635	769	925
Total	**32,732**	**32,047**	**32,669**
Number of Persons			
Adults	49,807	48,724	49,603
Children	10,639	10,232	10,122
Total	**60,446**	**58,956**	**59,725**

Site status disposition reports showing cumulative distributions of interim and final survey disposition codes, by site

Daily interviewer performance reports to monitor last-day and cumulative performance statistics, including completions, separate self-response modules, first refusals, final refusals, number of calls, time per call, and time per completed interview

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

2. Field Survey

A total of 14 MPR field interviewers, who were supervised by MPR survey managers in the Princeton office, screened addresses to identify households without telephone service or with interrupted telephone service. Reports were developed to monitor field costs and screening outcomes. Because interviews with eligible households were conducted via cellular telephone calls to MPR's Princeton telephone center, the CATI reports were used to monitor interview production and sample dispositions, by site. Field interviewers reported to the MPR supervisor on a weekly basis.

C. RESPONSE RATES

1. Calculation of Response Rates

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

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

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

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

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

The following sections describe how we calculated response rates for the RDD and field samples, as well as for the combination of the two samples. Table IV.2 shows the disposition of the RDD household sample, by sample type; Table IV.3 shows the disposition of the RDD sample at the FIU level; Table IV. 4 shows the final disposition of the field sample.

a. Determining Residency for the RDD Sample

When calculating a response rate, the denominator should reflect all eligible cases sampled. However, in many surveys, eligibility status is not determined for all cases and must be

TABLE IV.2

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

		Round Two Complete	Round Two Hard Refusal	Round Two Other Nonresponse	Round Two Non-Household	Round Two No Answer	Round Two Mech. Ans. Dev.	Old Working Banks	New Working Banks	Total
Complete										
1	Complete (all components)	13,622	382	103	1,503	24	59	6,762	275	22,730
2	Core complete self response missing	1,211	52	16	162	5	8	641	23	2,118
3	Core Complete secondary FIU missing	902	45	23	138	3	5	628	14	1,758
Ineligible										
41	No eligible person in hhold	69	8	14	77	4	1	111	8	292
42	Computer/fax/ modem	263	50	8	664	16	6	592	44	1,643
43	Disconnected, out of service	3,706	427	205	2,388	98	89	4,240	373	11,526
44	Cell phone, pager	45	6	3	62	0	3	76	7	202
45	Non-residence	622	72	46	2,613	49	26	2,652	177	6,257
48	Duplicate	1	0	0	0	0	0	1	0	2
Residential Nonresponding Household										
22	Breakoff	135	28	11	40	1	1	144	6	366
20	Hung up during introduction	495	271	58	173	7	20	838	41	1,903
21	Household refusal	3,082	1,317	200	639	22	39	4,135	128	9,562
23	Privacy manager	99	27	8	29	0	3	171	7	344
30	Language barrier	74	6	28	31	1	1	78	5	224
31	Illness barrier	40	13	7	6	0	1	51	1	119
34	Maximum calls (residential)	215	2	3	18	2	3	41	3	287
64	Answering service	1	0	0	0	0	0	1	0	2
Ring, No Answer										
65	Ring, no answer	356	66	85	2,005	659	131	2,878	170	6,350

Table IV.2 (c*ontinued*)

		Round Two Complete	Round Two Hard Refusal	Round Two Other Nonresponse	Round Two Non-Household	Round Two No Answer	Round Two Mech. Ans. Dev.	Old Working Banks	New Working Banks	Total
Mechanical Answering Device (MAD)										
67	Mechanical answering device	12	22	0	3	1	0	39	2	79
Contact/ Undetermined Residential										
36	Maximum calls (probable residence)[a]	358	51	69	251	44	69	645	39	1,526
66	Effort ended	7	3	0	5	0	0	13	1	29
Total		**25,315**	**2,848**	**887**	**10,807**	**936**	**465**	**24,737**	**1,324**	**67,319**

[a]Based on information provided during one or more calls, interviewer indicated that the telephone number is linked to a probable residence.

TABLE IV.3

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

	Round Two Complete	Round Two Hard Refusal	Round Two Other Nonresponse	Round Two Non-Household	Round Two No Answer	Round Two Mech. Ans. Dev.	Old Working Banks	New Working Banks	Totals
Responding Eligible FIU	18,850	561	162	2,169	33	78	9,525	366	31,744
Nonresponding Eligible FIU	919	61	30	153	7	5	689	14	1,878
Ineligible FIU (no civilian adults)	189	5	3	45	0	0	126	8	376
Total	**19,958**	**627**	**195**	**2,367**	**40**	**83**	**10,340**	**388**	**33,998**

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

TABLE IV.4

FINAL ROUND THREE CTS HOUSEHOLD SURVEY DISPOSITION (FIELD SAMPLE)
(Numbers)

Disposition		Number
Complete		
1,2,3	Core complete–eligible[a]	595
Eligible Nonresponse (Telephone Interruption)		
21	Refused survey screener	63
22	Breakoff during main interview	6
Ineligible Household		
40	No interruption in telephone service	3,843
41	Not selected (all military or children)	5
Household with Unknown Telephone Status Eligibility		
20	Refused telephone screener at doorstep	331
30	Language/other barrier	54
65	Effort ended—no contact	50
Not a Residential Household		
45	Not a residence	30
46	No housing unit	193
47	Vacant unit	462
Unknown Whether a Household		
67	Effort ended—locked building	304
Total		**5,936**

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

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

estimated. For random digit dialing (RDD) surveys, residency is typically not established for all sampled telephone numbers, even after many calls have been made. For example, some telephone numbers ring when dialed even though the telephone number is not in use. Consequently, the first step in computing the RDD response rate was to estimate residency for sampled telephone numbers. Residency was determined for 91.8 percent of the 74,348 sampled telephone numbers (Table IV.2). Residency was not confirmed for the remaining sample, which included 4.5 percent ring, no answers; 2.2 percent mechanical answering devices, and 1.5 percent personal contact, but with no confirmation of residency.

We evaluated several procedures for estimating residency (see Appendix E) and decided to us use a method commonly referred to as the CASRO method. [1] First we classified each sampled telephone number into one of four groups:

1. Telephone number resolved as a working residential number
2. Telephone number resolved as a business or non-working number by Genesys-ID or Genesys-ID-Plus, which excludes many business and non-working telephone numbers before the telephone number is called by an interviewer
3. Telephone number resolved as a business number or a non-working number after being called by an interviewer
4. Telephone number status never resolved

We then calculated a residency rate among the resolved cases and applied this rate to the unresolved telephone numbers (Group 4) to estimate how many of them are residential telephone

[1]CASRO stands for the Council of American Survey Research Organizations, and their special report, "On the definition of Response Rates," [LR Frankel, Chairman, A Special Report of the CASRO Task Force on Completion Rates, June 1982]. We refer to this method as CASRO, because one option in their recommendations is to apply the eligibility rate for cases with determined eligibility status to those with undetermined eligibility status.

numbers. The residency rate among the resolved cases was calculated as:

$$(2) \qquad \text{Residency Rate} = \frac{\text{Group 1}}{\text{Group1} + \text{Group2} + \text{Group3}}$$

b. Household Response Rate for the RDD Sample

To calculate an interview response rate at the household level, we first determined whether each telephone number was residential (using the method described above) and then determined whether each household completed at least one FIU interview.

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

a. At least one eligible responding FIU in the household—codes 1, 2, 3 (n = 26,606)

b. Eligible nonresponding household—code 22 (n = 366)

c. Nonresponding residential household, with insufficient information to determine whether there is an eligible FIU—codes 20, 21, 23, 30, 31, 66 (n = 12,181)

d. Residential household, where all FIUs in the household are ineligible—codes 41, 48 (n = 294)

e. Telephone number was coded by the interviewer as nonresidential or nonworking—codes 42, 43, 44, 45, (n = 19,628) or screened out as nonresidential or nonworking by Genesys ID or ID Plus (n=32,868)

f. Unable to determine whether telephone number was residential (n = 8,244)

- f1. Ring, no answer—codes 65 (n = 6,350)
- f2. Mechanical answering device—codes 64, 67 (n = 81)
- f3. Voice contact made—codes 34, 46 (n = 1,813)

Within each site (with the supplemental sample being treated as a separate site) and sampling groups, [2] we calculated a residency rate among telephone numbers with resolved residency status, and a survey eligibility rate among residential households with known survey eligibility:

(3) $$RSDR_s = (A_s + B_s + C_s + D_s)/(A_s + B_s + C_s + D_s + E_s)$$

(4) $$SER_s = (A_s + B_s)/(A_s + B_s + D_s).$$

We then calculated within each site and sampling group the estimated number of eligible households as:

$$HH_s = A_s + B_s + \left(C_s + \left(F_s \cdot RSDR_s\right)\right) \cdot SER_s \quad (0.1)$$

Finally, we calculated a household response rate within each site as follows:

(5) $$HRR_s = \frac{A_s}{HH_s}.$$

To compute response rates involving more than one site or sampling group, we summed the number of completes and the estimated number of eligible households across sites or sampling groups, and divided the number of completes by the estimated number of eligible households.

Weighted response rates were calculated similarly, except that we used counts weighted by

[2]Sampling group refers to the six overlap categories (complete, hard refusal, other nonresponse, nonhousehold, no answer, mechanical answering device) plus the two new sample categories (old working banks and new working banks).

sampling weights, by which we mean the inverse of the probability of selection.[3]

c. Household Response Rate for the Field Sample

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

a. Eligible responding household—codes 1, 2, 3 (n = 595)

b. Eligible nonresponding household—code 22 (n = 6)

c. Nontelephone household ineligible for survey—code 41 (n=5)

d. Nontelephone household with insufficient information to determine whether eligible for survey—code 21 (n = 63)

e. Ineligible household (no interruption in phone service)—code 40 (n = 3,843)

f. Unable to determine telephone status of household—codes 20, 30, 65 (n=435)

g. Not a household or vacant—codes 45, 46, 47 (n = 685)

h. Unable to determine whether address was residential—code 67 (n = 304)

Within each high-intensity site *s*, we calculated three eligibility rates: (1) a household eligibility rate (proportion of addresses known to be occupied residences), (2) a field component eligibility rate (proportion of residences known to have had telephone interruption), and (3) a

[3]The probability of selection cited here included adjustments for site selection and for alternate probabilities of selection discussed below in chapter V.

survey eligibility rate (proportion of residences with telephone interruption known to be eligible for the survey):

(6) $$HER_s = (a_s + b_s + c_s + d_s + e_s + f_s)/(a_s + b_s + c_s + d_s + e_s + f_s + g_s)$$

(7) $$NER_s = (a_s + b_s + c_s + d_s)/(a_s + b_s + c_s + d_s + e_s)$$

(8) $$SER_s = (a_s + b_s)/(a_s + b_s + d_s)$$

Within each high-intensity site *s*, we calculated the estimated number of eligible households as:

(9) $$HH_s = a_s + b_s + \left(d_s + \left(f_s + \left(h_s \cdot HER_s\right)\right) \cdot NER_s\right) \cdot SER_s$$

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

(10) $$HRR_s = \frac{A_s}{HH_s}$$

To compute response rates involving more than one site, we summed the number of completes and the estimated number of eligible households across sites, and divided the number of completes by the estimated number of eligible households.

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

d. Combinations of Household Response Rates

When calculating a response rate for combinations of various sample components (such as the site sample and supplemental sample combined, or the RDD sample and the field sample combined), we summed the number of completes and the estimated number of eligible

households across sample components, and divided the number of completes by the estimated number of eligible households.

e. Family Interview Response Rate

To calculate an interview response rate at the FIU level,[4] we began with all FIUs in responding households (that is, households with at least one eligible responding FIU). We classified each FIU according to the categories in Table IV.3 as follows:

A. FIU is eligible for the survey and responded to interview— (n = 31,744)

B. FIU is eligible for the survey but did not respond to interview— (n = 1,878)

C. FIU is ineligible for survey (n=376)

For each site and sampling group, we then calculated an FIU-level response rate conditioned on being in a household with at least one completed FIU interview:

$$(11) \qquad FRR_s = \frac{A_s}{A_s + B_s}.$$

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

$$(12) \qquad RR_s = HRR_s \cdot FRR_s.$$

[4]For the field component, the household informant was allowed to respond for each FIU; consequently, the FIU response rate is approximately equal to the household response rate. Among the 963 FIUs in the responding field households, 925 were completes, 21 were coded as ineligible and 17 were eligible nonresponding FIUs. Among the 17, nine had an FIU level status code of "refusal," one had a code of "illness barrier," and seven had a code of "respondent away, no proxy."

For any conditional FIU-level response rates involving more than one site or sampling group, we first summed the number of cases in categories A and B listed above (for example, $A = \sum_s A_s$, $B = \sum_s B_s$, if summing across the entire sample) and then calculated the conditional response rate.

$$(13) \qquad FRR = \frac{A}{A + B}$$

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

$$(14) \qquad RR = HRR \cdot FRR$$

Weighted response rates at the FIU level were calculated similarly, except that we used counts weighted by sampling weights.

Conditional FIU response rates for the RDD and in-person components and the combined site and supplemental samples were calculated in the same way as the household response rates.

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

a. Response Rates, by Sample Type

Tables IV.5-6 show the unweighted and weighted household- and FIU-level response rates for the Round Three sample, by sample type. (Appendix F provides additional details on the computation of response rates for subgroups of the sample.) Household-level response rates were higher than FIU-level response rates because some households included multiple FIUs and some of these FIUs did not complete interviews. Although both unweighted and weighted household and FIU response rates are shown, we will generally limit our discussion to weighted FIU response rates, since weighted data will be used for most analyses and the FIU was the

TABLE IV.5

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

	Unweighted	Weighted
RDD		
Round Two Overlap Sample		
Completed interviews	77.17	78.64
Hard refusal	21.56	19.24
Nonresponse (other than hard refusal)	32.59	24.38
No answer	26.95	18.68
Mechanical answering device	37.04	30.11
Not a household	59.88	54.29
Total Round Two Overlap Sample	69.25	63.89
New Sample		
Old working banks[a]	54.89	53.90
New working banks[b]	57.96	52.95
Total New Sample	55.00	53.84
Total RDD	64.05	60.74
Field	71.44	73.50
Total Sample	**64.19**	**61.95**

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

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

TABLE IV.6

ROUND THREE CTS FIU-LEVEL[a] RESPONSE RATE, BY SAMPLE TYPE
(Percents)

	Unweighted	Weighted
RDD		
Round Two Overlap Sample		
Completed interviews	73.58	75.46
Hard refusal	19.44	18.08
Nonresponse (other than hard refusal)	27.50	21.33
No answer	22.24	12.99
Mechanical answering device	34.81	27.68
Not a household	55.93	50.05
Total Round Two Overlap Sample	65.72	60.82
New Sample		
Old working banks[b]	51.19	49.76
New working banks[c]	55.82	50.04
Total New Sample	51.35	49.77
Total RDD	60.47	57.35
Field	70.15	72.60
Total Sample	**60.67**	**58.56**

[a]Combined household-level response rate and FIU-level response rate.

[b]Working banks in existence at the time the Round Two sample was selected.

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

primary interviewing unit for the Household Survey. In addition, patterns in response rates by sample type and geographic units were the same for households and FIUs.

Weighted response rates varied by type of sample. The Round Three weighted RDD FIU response rate for the overlap sample (all telephone numbers sampled from Round Two) was 60.8 percent, compared with 49.8 percent for new sample (old and new working banks combined). The higher response rate for the overlap sample was due to the high level of cooperation among households whose telephone numbers were selected from Round Two completed interviews (75.5 percent). The interval between rounds was only two years, so most of the families and persons interviewed for Round Two were at the same telephone number for Round Three. Because Round Two families had received monetary incentives of $25[5], most also remembered the interview and were aware that they would be compensated for participation, a factor that may have contributed to the high response rate.

Not surprisingly, the weighted FIU response rate was poor for Round Two hard refusals (18.1 percent) and for other nonresponses (21.3 percent).[6] These households had been contacted many times in both rounds and most had refused, many several times. The FIU response rate for Round Two non-households (50.0 percent) was comparable to the rate for new working banks. This result is not surprising, as the telephone numbers linked to these households had been nonresidential at the time of the Round Two survey and therefore were contacted for the first time in Round Three. The low weighted FIU response rate for Round Two telephone numbers that had final dispositions of no answer (13.0 percent) or mechanical answering device (27.7

[5]Nearly all FIUs participating in Round Two received $25 for completing that survey; however, a small number selected for a Round Two experiment received $50.

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

percent) was due to the very large fraction of telephone numbers in these subsamples with undetermined residency in both rounds, which resulted in a large fraction having residency imputed.

b. Patterns in Response Rates

Site level unweighted and weighted RDD and field response rates by round are shown in Tables IV.7-8, respectively. RDD response rates continued to decline in virtually all sites. Across the three rounds, the RDD FIU response rate for site and supplemental samples declined from 64.4 percent in Round One, to 62.2 percent in Round Two, and to 57.4 percent in Round Three. While individual site response rates for the field sample vary considerably by round due to small sample sizes, the overall field FIU response rate remained stable between Round Two (73.3 percent) and Round Three (72.6 percent), after declining somewhat from Round One (83.2 percent)

For all three rounds, RDD response rates varied inversely with the size of the site population. For Round One, the weighted FIU RDD response rate ranged from a low of 56.6 percent in MSAs of 3 million or more people to a high of 72.6 percent in nonmetropolitan areas. The trend in Round Two response rates was similar, ranging from 55.3 percent for MSAs with 3 million or more to 71.0 percent for nonmetropolitan areas, and in Round Three with a range from 50.5 percent in the largest MSAs to 64.9 percent in nonmetropolitan areas. The larger MSAs may have lower response rates because they correspond to the largest media markets, whose residents are subject to greater telemarketing and market research penetration.

Although relative response rates by market size remained consistent, we observed a decline between Rounds Two and Three in RDD response rates across sites, regardless of population size. The RDD FIU response rate for MSAs of three million or more declined from 55.3 percent to 50.5 percent, while the response rate for nonmetropolitan areas dropped from 71.0 to 64.9

TABLE IV.7

CTS SITE LEVEL UNWEIGHTED RESPONSE RATES BY ROUND
(Percents)

		Unweighted					
		Round One		Round Two		Round Three	
		Household	FIU	Household	FIU	Household	FIU
	RDD						
0	Supplemental Sample	66.46	64.26	67.18	64.66	68.04	64.22
1	Boston, MA Portion	59.94	56.72	61.51	57.47	57.93	53.89
2	Cleveland-Lorain-Elyria, OH PMSA	63.52	61.47	64.12	62.10	61.80	58.07
3	Greenville-Spartanburg-Anderson, SC MSA	70.77	68.96	69.39	67.19	64.01	60.95
4	Indianapolis, IN MSA	71.25	70.22	70.22	67.58	63.05	60.69
5	Lansing-East Lansing, MI MSA	71.97	70.12	70.64	68.34	67.37	64.54
6	Little Rock-North Little Rock, AR MSA	74.29	72.91	72.18	69.64	71.29	68.17
7	Miami, FL PMSA	53.21	49.89	54.94	50.58	52.15	46.66
8	Newark, NJ PMSA	59.36	56.32	56.76	53.03	58.72	53.77
9	Orange County, CA PMSA	55.83	52.35	55.37	50.26	54.95	50.21
10	Phoenix-Mesa, AZ MSA	68.54	66.88	65.29	62.06	66.69	62.61
11	Seattle-Bellevue-Everett, WA PMSA	65.90	62.89	62.38	59.23	60.07	56.02
12	Syracuse, NY MSA	69.00	67.12	71.98	70.46	64.46	61.47
13	Atlanta, GA MSA	65.44	62.68	71.19	67.86	62.38	59.77
14	Augusta-Aiken, GA-SC MSA	66.06	64.08	69.03	66.93	67.92	63.98
15	Baltimore, MD PMSA	65.03	63.47	67.07	64.86	65.26	60.52
16	Bridgeport-Danbury-Stamford, CT Portion	54.65	52.44	58.54	53.82	63.08	59.27
17	Chicago-Kenosha-Kankakee, IL-WI PMSA	60.99	57.83	57.40	53.48	61.18	58.88
18	Columbus, OH MSA	69.67	65.88	68.49	66.31	69.89	66.06
19	Denver-Boulder-Greeley, CO PMSA	66.08	64.10	65.89	61.42	61.79	59.23
20	Detroit, MI PMSA	66.44	62.78	60.83	59.76	60.19	56.18
21	Greensboro-Winston Salem-High Point, NC MSA	68.33	66.60	70.93	68.74	66.81	63.89
22	Houston-Galveston-Brazoria, TX PMSA	64.14	60.47	59.83	57.18	57.66	54.63
23	Huntington-Ashland, WV-KY-OH MSA	75.13	73.45	74.45	71.96	71.44	69.16
24	Killeen-Temple, TX MSA	73.17	70.80	73.34	71.64	67.82	65.47
25	Knoxville, TN MSA	72.06	70.48	68.74	67.37	66.84	65.21
26	Las Vegas, NV-AZ MSA	58.52	54.44	61.63	58.58	63.95	59.17
27	Los Angeles-Long Beach, CA PMSA	53.99	51.25	56.16	52.31	56.92	52.73
28	Middlesex-Trenton, NJ PMSA	66.07	63.82	65.44	62.83	63.85	60.72
29	Milwaukee-Racine, WI PMSA	68.77	67.68	70.95	68.43	72.79	71.17
30	Minneapolis-St Paul, MN-WI MSA	76.74	73.87	76.03	73.89	81.87	79.41
31	Modesto, CA MSA	64.76	63.11	66.97	62.90	66.29	62.56
32	Nassau-Suffolk, NY PMSA	59.63	56.95	61.15	55.27	57.48	53.79
33	New York City, NY PMSA	46.27	42.76	57.06	50.52	50.57	45.76
34	Philadelphia, PA-NJ PMSA	64.80	62.38	59.52	56.80	57.11	51.85
35	Pittsburgh, PA MSA	64.99	62.88	67.83	66.03	66.67	61.96

TABLE IV.7 (*Continued*)

		Unweighted					
		Round One		Round Two		Round Three	
		Household	FIU	Household	FIU	Household	FIU
36	Portland-Salem, OR-WA PMSA	68.46	64.87	72.14	69.11	71.57	68.31
37	Riverside-San Bernardino, CA PMSA	65.57	63.48	64.83	61.59	63.60	59.89
38	Rochester, NY MSA	70.47	68.54	70.77	68.28	69.82	66.19
39	San Antonio, TX MSA	63.45	61.39	67.27	63.22	66.91	63.12
40	San Francisco, CA PMSA	51.24	47.52	54.78	50.44	55.98	48.89
41	Santa Rosa, CA PMSA	60.15	55.84	62.21	59.38	60.65	58.42
42	Shreveport-Bossier City, LA MSA	71.65	68.22	71.25	68.48	68.50	65.11
43	St. Louis, MO-IL MSA	72.71	69.23	74.68	72.96	70.73	68.81
44	Tampa-St Petersburg-Clearwater, FL MSA	60.32	57.74	57.98	54.94	60.30	57.76
45	Tulsa, OK MSA	63.80	62.10	70.12	65.78	67.00	63.93
46	Washington-Hagerstown, DC-MD-VA-WV PMSA	65.59	63.94	68.23	65.05	63.82	59.25
47	West Palm Beach-Boca Raton, FL MSA	55.42	50.98	53.93	49.99	47.03	44.78
48	Worcester-Fitchburg, MA Portion	63.91	62.11	66.13	62.38	63.31	60.12
49	Dothan, AL MSA	69.81	68.45	72.75	71.67	67.48	64.59
50	Terre Haute, IN MSA	74.38	72.17	67.64	64.98	67.00	65.62
51	Wilmington, NC MSA	76.18	73.28	71.96	69.22	66.56	63.97
52	West-Central Alabama	71.54	69.43	70.48	68.62	71.90	67.29
53	Central Arkansas	75.94	75.34	74.61	73.45	76.88	75.23
54	Northern Georgia	72.87	69.81	70.48	67.88	64.18	60.42
55	Northeast Illinois	67.77	67.32	71.32	68.93	70.75	69.15
56	Northeast Indiana	72.74	70.76	71.05	69.52	68.89	67.72
57	Eastern Maine	80.32	79.08	81.82	81.57	78.87	76.98
58	Eastern North Carolina	77.24	75.26	75.04	72.08	72.10	69.42
59	Northern Utah	78.79	76.76	83.19	82.19	80.40	78.22
60	Northwest Washington	70.43	68.97	70.90	67.37	70.76	68.43
	All high-intensity sites	65.07	62.69	64.16	60.97	61.57	57.70
	All low-intensity sites	66.59	64.13	67.38	64.45	65.67	62.44
	Total RDD site sample	65.84	63.42	65.77	62.71	63.64	60.09
	Total RDD (supplemental + site)	65.90	63.50	65.92	62.91	64.05	60.47
	Field						
1	Boston, MA Portion	75.42	75.42	74.07	74.07	58.74	57.16
2	Cleveland-Lorain-Elyria, OH PMSA	73.25	73.25	84.71	84.71	86.16	86.16
3	Greenville-Spartanburg-Anderson, SC MSA	83.73	83.73	83.41	83.41	89.97	89.03
4	Indianapolis, IN MSA	89.10	89.10	82.99	82.99	69.20	66.54
5	Lansing-East Lansing, MI MSA	72.38	72.38	70.42	70.42	67.55	67.55
6	Little Rock-North Little Rock, AR MSA	90.43	90.43	80.97	80.97	86.72	84.80
7	Miami, FL PMSA	95.40	95.40	71.00	71.00	90.34	90.34
8	Newark, NJ PMSA	69.68	69.68	56.26	56.26	62.61	61.10
9	Orange County, CA PMSA	53.04	53.04	61.52	61.52	79.15	79.15
10	Phoenix-Mesa, AZ MSA	93.14	93.14	96.28	96.28	99.29	98.57

TABLE IV.7 (*Continued*)

		Unweighted					
		Round One		Round Two		Round Three	
		Household	FIU	Household	FIU	Household	FIU
11	Seattle-Bellevue-Everett, WA PMSA	69.44	69.44	26.53	26.53	35.34	35.34
12	Syracuse, NY MSA	93.22	93.22	67.38	67.38	86.27	82.89
	Total field sample	84.60	84.60	79.17	79.17	71.44	70.15
	RDD Site Sample + Field						
1	Boston, MA Portion	60.05	56.87	61.56	57.55	57.94	53.97
2	Cleveland-Lorain-Elyria, OH PMSA	63.80	61.81	65.02	63.07	62.62	59.02
3	Greenville-Spartanburg-Anderson, SC MSA	71.31	69.57	70.40	68.37	65.26	62.33
4	Indianapolis, IN MSA	72.18	71.21	71.03	68.59	63.59	61.20
5	Lansing-East Lansing, MI MSA	71.98	70.16	70.64	68.40	67.38	64.60
6	Little Rock-North Little Rock, AR MSA	75.00	73.70	72.67	70.30	71.97	68.91
7	Miami, FL PMSA	53.99	50.74	55.19	50.90	53.00	47.71
8	Newark, NJ PMSA	59.80	56.87	56.74	53.17	59.01	54.36
9	Orange County, CA PMSA	55.81	52.37	55.41	50.34	55.14	50.45
10	Phoenix-Mesa, AZ MSA	69.53	67.95	67.61	64.69	68.55	64.76
11	Seattle-Bellevue-Everett, WA PMSA	66.07	63.18	60.94	57.91	58.05	54.30
12	Syracuse, NY MSA	69.72	67.89	71.81	70.36	64.93	61.94
	High-intensity sites (with field)	65.61	63.30	64.59	61.54	62.00	58.26
	Sites Grouped by Population Size (RDD Site Sample)						
	3 million+	60.32	57.26	61.04	57.34	58.13	54.13
	2-3 million	62.55	59.87	61.98	58.43	60.59	56.42
	1-2 million	65.26	62.90	65.04	61.90	64.13	60.38
	<1 million large MSA	68.56	66.43	68.91	66.30	65.54	62.52
	Small MSA (<200,000)	73.41	71.27	70.88	68.77	67.05	64.73
	Non-MSA	74.19	72.56	74.42	72.51	73.07	70.61
	Total (RDD site + supplemental + field)	**66.16**	**63.80**	**66.15**	**63.21**	**64.19**	**60.67**

TABLE IV.8

CTS SITE LEVEL WEIGHTED RESPONSE RATES BY ROUND
(Percentages)

		Weighted					
		Round One		Round Two		Round Three	
		Household	FIU	Household	FIU	Household	FIU
	RDD						
0	Supplemental Sample	66.89	64.73	65.61	63.06	61.85	58.27
1	Boston, MA Portion	59.94	56.72	59.88	55.97	52.85	49.04
2	Cleveland-Lorain-Elyria, OH PMSA	63.53	61.47	61.89	59.87	56.16	52.78
3	Greenville-Spartanburg-Anderson, SC MSA	70.77	68.96	67.60	65.33	59.37	56.23
4	Indianapolis, IN MSA	71.25	70.22	68.50	65.96	58.62	56.31
5	Lansing-East Lansing, MI MSA	71.97	70.11	68.32	66.10	61.99	59.24
6	Little Rock-North Little Rock, AR MSA	74.29	72.91	70.57	68.09	65.93	62.72
7	Miami, FL PMSA	53.22	49.90	54.04	49.74	47.83	42.78
8	Newark, NJ PMSA	59.36	56.32	55.48	51.75	52.77	48.32
9	Orange County, CA PMSA	55.83	52.35	53.90	48.66	48.93	44.79
10	Phoenix-Mesa, AZ MSA	68.53	66.87	62.98	59.80	61.46	57.75
11	Seattle-Bellevue-Everett, WA PMSA	65.90	62.88	60.53	57.52	55.10	51.39
12	Syracuse, NY MSA	69.00	67.12	69.78	68.35	59.56	56.89
13	Atlanta, GA MSA	65.44	62.68	69.08	65.81	60.22	57.55
14	Augusta-Aiken, GA-SC MSA	66.06	64.08	67.22	65.10	62.86	59.53
15	Baltimore, MD PMSA	65.03	63.47	66.06	63.83	60.27	55.74
16	Bridgeport-Danbury-Stamford, CT Portion	54.65	52.44	57.01	52.48	57.79	54.16
17	Chicago-Kenosha-Kankakee, IL-WI PMSA	60.99	57.83	55.78	51.87	55.69	53.63
18	Columbus, OH MSA	69.67	65.88	67.36	65.24	66.81	62.97
19	Denver-Boulder-Greeley, CO PMSA	66.08	64.10	64.11	59.28	58.18	55.54
20	Detroit, MI PMSA	66.44	62.78	59.30	58.27	56.91	53.15
21	Greensboro-Winston Salem-High Point, NC MSA	68.33	66.60	68.64	66.56	61.77	59.12
22	Houston-Galveston-Brazoria, TX PMSA	64.14	60.47	58.79	56.23	55.54	52.45
23	Huntington-Ashland, WV-KY-OH MSA	75.13	73.45	72.70	70.26	66.48	64.49
24	Killeen-Temple, TX MSA	73.17	70.80	72.08	70.29	64.48	62.20
25	Knoxville, TN MSA	72.06	70.48	66.91	65.60	62.41	60.99
26	Las Vegas, NV-AZ MSA	58.52	54.44	60.22	57.26	58.96	54.46
27	Los Angeles-Long Beach, CA PMSA	53.99	51.25	55.22	51.39	53.31	49.37
28	Middlesex-Trenton, NJ PMSA	66.07	63.82	64.05	61.48	59.01	56.11
29	Milwaukee-Racine, WI PMSA	68.77	67.68	67.83	65.42	69.54	67.67
30	Minneapolis-St Paul, MN-WI MSA	76.74	73.87	74.36	72.32	79.06	76.77
31	Modesto, CA MSA	64.76	63.11	65.09	60.77	60.62	56.98
32	Nassau-Suffolk, NY PMSA	59.63	56.95	59.14	53.33	52.35	49.10
33	New York City, NY PMSA	46.27	42.76	55.17	48.58	48.22	43.60
34	Philadelphia, PA-NJ PMSA	64.80	62.38	58.38	55.61	55.34	50.16
35	Pittsburgh, PA MSA	64.99	62.88	65.61	63.91	62.65	58.24
36	Portland-Salem, OR-WA PMSA	68.46	64.87	70.02	67.03	67.37	64.30

TABLE IV.8 (*Continued*)

		Weighted					
		Round One		Round Two		Round Three	
		Household	FIU	Household	FIU	Household	FIU
37	Riverside-San Bernardino, CA PMSA	65.57	63.48	63.58	60.30	60.70	57.06
38	Rochester, NY MSA	70.47	68.54	68.90	66.41	64.99	61.60
39	San Antonio, TX MSA	63.45	61.39	65.17	60.95	63.52	60.01
40	San Francisco, CA PMSA	51.24	47.52	52.77	48.45	50.19	43.85
41	Santa Rosa, CA PMSA	60.15	55.84	60.37	57.62	55.55	53.54
42	Shreveport-Bossier City, LA MSA	71.65	68.22	69.98	67.21	63.76	60.55
43	St. Louis, MO-IL MSA	72.71	69.23	73.26	71.62	67.29	65.50
44	Tampa-St Petersburg-Clearwater, FL MSA	60.32	57.74	54.99	52.20	56.16	53.87
45	Tulsa, OK MSA	63.80	62.10	67.87	63.68	63.04	60.24
46	Washington-Hagerstown, DC-MD-VA-WV PMSA	65.59	63.94	67.06	64.02	60.61	55.81
47	West Palm Beach-Boca Raton, FL MSA	55.42	50.98	51.82	48.05	42.50	40.53
48	Worcester-Fitchburg, MA Portion	63.91	62.11	64.88	60.69	58.28	55.33
49	Dothan, AL MSA	69.81	68.45	70.88	69.84	63.47	60.84
50	Terre Haute, IN MSA	74.38	72.17	66.19	63.60	61.33	60.07
51	Wilmington, NC MSA	76.18	73.28	70.57	67.81	60.57	57.91
52	West-Central Alabama	71.54	69.43	69.14	67.20	64.84	60.45
53	Central Arkansas	75.94	75.34	72.85	71.72	71.35	69.52
54	Northern Georgia	72.87	69.81	68.70	66.24	58.26	54.72
55	Northeast Illinois	67.77	67.32	68.07	65.80	64.82	63.33
56	Northeast Indiana	72.74	70.76	69.02	66.68	62.13	61.05
57	Eastern Maine	80.32	79.08	80.30	80.06	74.80	73.04
58	Eastern North Carolina	77.24	75.26	73.99	71.08	66.82	64.41
59	Northern Utah	78.79	76.76	82.73	81.76	75.33	73.28
60	Northwest Washington	70.43	68.97	69.35	65.81	63.75	61.57
	All high-intensity sites	64.96	62.54	62.44	59.29	56.30	52.72
	All low-intensity sites	67.79	65.43	66.75	63.97	61.99	59.11
	Total RDD site sample	66.59	64.21	64.95	62.00	59.56	56.37
	Total RDD (supplemental + site)	66.64	64.37	65.04	62.30	60.74	57.35
	Field						
1	Boston, MA Portion	90.39	90.39	63.96	63.96	41.20	40.18
2	Cleveland-Lorain-Elyria, OH PMSA	72.15	72.15	78.68	78.68	76.74	76.74
3	Greenville-Spartanburg-Anderson, SC MSA	82.65	82.65	83.82	83.82	91.07	90.47
4	Indianapolis, IN MSA	89.74	89.74	83.12	83.12	74.56	72.58
5	Lansing-East Lansing, MI MSA	68.96	68.96	67.01	67.01	65.83	65.83
6	Little Rock-North Little Rock, AR MSA	91.04	91.04	83.44	83.44	87.68	86.63
7	Miami, FL PMSA	94.01	94.01	72.44	72.44	88.45	88.45
8	Newark, NJ PMSA	68.93	68.93	46.86	46.86	63.10	61.73
9	Orange County, CA PMSA	56.61	56.61	66.75	66.75	80.63	80.63
10	Phoenix-Mesa, AZ MSA	89.26	89.26	97.24	97.24	98.95	98.59
11	Seattle-Bellevue-Everett, WA PMSA	72.28	72.28	31.33	31.33	41.58	41.58
12	Syracuse, NY MSA	92.53	92.53	78.28	78.28	88.43	83.49
	Total field sample	83.20	83.20	73.30	73.30	73.50	72.60

TABLE IV.8 (*Continued*)

		Weighted					
		Round One		Round Two		Round Three	
		Household	FIU	Household	FIU	Household	FIU
	RDD Site Sample + Field						
1	Boston, MA Portion	60.46	57.36	59.90	56.04	53.55	49.73
2	Cleveland-Lorain-Elyria, OH PMSA	63.97	62.09	63.06	61.22	58.80	55.43
3	Greenville-Spartanburg-Anderson, SC MSA	71.34	69.66	68.90	66.88	68.62	65.44
4	Indianapolis, IN MSA	73.47	72.61	69.70	67.52	68.01	65.50
5	Lansing-East Lansing, MI MSA	71.78	70.13	68.23	66.29	68.60	65.85
6	Little Rock-North Little Rock, AR MSA	76.06	74.89	72.32	70.36	72.23	69.00
7	Miami, FL PMSA	55.42	52.33	54.42	50.27	58.23	53.41
8	Newark, NJ PMSA	60.92	58.57	53.54	51.09	62.61	58.04
9	Orange County, CA PMSA	55.84	52.46	54.02	48.85	50.22	46.11
10	Phoenix-Mesa, AZ MSA	70.04	68.56	67.45	64.85	70.43	66.84
11	Seattle-Bellevue-Everett, WA PMSA	66.71	64.24	57.18	54.89	60.13	56.38
12	Syracuse, NY MSA	71.39	69.74	70.20	68.89	62.60	59.75
	High-intensity sites (with field)	66.21	64.07	62.85	60.13	62.14	58.54
	Sites Grouped by Population Size (RDD Site Sample)						
	3 million+	59.62	56.59	59.12	55.34	54.32	50.48
	2-3 million	62.36	59.66	60.05	56.52	55.54	51.75
	1-2 million	65.42	63.07	63.66	60.59	59.18	55.77
	<1 million large MSA	68.33	66.15	66.76	64.17	60.38	57.48
	Small MSA (<200,000)	73.75	71.50	69.68	67.48	61.66	59.32
	Non-MSA	74.28	72.64	73.01	71.03	67.25	64.93
	Total (RDD site + supplemental + field)	**67.04**	**64.82**	**65.10**	**62.46**	**61.95**	**58.56**

percent. Despite considerable efforts to convert refusals (discussed below), respondent resistance to survey participation increased in all areas of the country.

c. Comparison of Rounds Two and Three Response Rates by Sample Group

Unlike Round One, which did not have an overlap sample, the second and third rounds of the CTS had similar sample designs. A comparison of response rates by sample group shows that the decline in the Round Three response rate for the new sample (old and new working banks combined) was greater than for the overlap sample (Table IV.9). For the overlap sample (prior round completes and non-completes combined), the weighted RDD FIU response rate declined from 65.5 to 60.8 percent, whereas the weighted new sample response rate decreased from 59.1 to 49.8 percent. The relatively smaller decline in the overlap sample was due to the impact of completed prior round interviews, a group whose response rate decreased relatively less than most other sample groups. This was due in part to the impact of sample weights, as overlap completed interviews were the only Round Three RDD sample group whose response rate increased as a result of weighting. As noted above, prior survey participation and receipt of a monetary incentive also may have increased the likelihood of their participation in Round Three.

The weighted field response rate changed very little (from 73.3 to 72.6 percent) between Rounds Two and Three, indicating that the people in low income areas are still very responsive to personal visits.

d. Impact of Weighting on Response Rates

The decrease in the weighted FIU response rate from 62.5 to 58.6 percent was greater than the decline in the unweighted response rate of from 63.2 to 60.7 percent. We investigated the

TABLE IV.9

COMPARISON OF ROUND CTS TWO AND ROUND THREE FIU RESPONSE RATES BY SAMPLE GROUP

	Unweighted		Weighted	
	Round Two	Round Three	Round Two	Round Three
RDD				
Overlap Sample				
Complete	80.5	73.6	81.6	75.5
Noncomplete	40.1	38.9	39.6	32.3
Total Overlap	66.7	65.7	65.5	60.8
New Sample				
Old Working Banks	56.9	51.2	58.3	49.8
New Working Banks	61.4	55.8	62.4	50.0
Total New Sample	58.0	51.4	59.1	49.8
Total RDD	62.9	60.5	62.3	57.4
Field	79.2	70.2	73.3	72.6
Total RDD and Field	63.2	60.7	62.5	58.6

impact of weighting on response rates and determined that the largest decreases were within the categories of prior round non-completed interviews in the overlap sample and the small sample of new working banks included in the survey.

The nearly six percentage point decrease from the unweighted (55.8 percent) to weighted (50.0 percent) response rate for new working banks was related to the small sample size of this group (see Table IV.2), which can result in variable weighting impacts. For Round Three, new working bank strata with lower sampling rates in several of the high intensity sites had higher nonresponse than new working bank strata with higher sampling rates.

Within the overlap sample, weighting had a greater impact on response rates for prior round noncompleted interviews in Round Three than in Round Two. In Round Two, the unweighted and weighted response rates for this component were both 40 percent, whereas in Round Three the unweighted response rate was virtually the same (39 percent) but the weighted response rate was only 32 percent.

For Round Three, overlap completed interviews were sampled at a higher rate and overlap noncompleted interviews at a lower rate compared with Round Two, when all overlap sample were selected at the same rate. Some categories of overlap non-completes that were sampled at a low rate had very low response rates in high intensity sites. Since these cases had large weights, they had a disproportionate impact on the overlap noncomplete response rate. However, the high response rate for overlap completed interviews reduced the impact of declining response rates from other sampling components on the total RDD response rate.

e. Impact of Other Surveys Using the CTS Household Survey as a Sample Frame

Part of the decline in the overlap response rate between Rounds Two and Three also was due to the impact of other surveys using the CTS as a sample frame. The Round Three CTS Household Survey included a large number of individuals in the overlap complete sample group

who had previously been sampled for two non-CTS surveys. The Health Care for Communities (HCC) survey selected individuals from the Round Two CTS survey, over sampling those who had a higher probability of experiencing mental health problems; sampling was based on scores to the mental health component of the SF-12 and use of mental health resources. The HCC survey, which was more heavily concentrated in low intensity sites, included about 40 minutes of questions dealing with mental health and substance abuse issues. The Community Quality Index (CQI) Survey sampled individuals who had completed interviews in high intensity sites during Round Two for a brief interview in which they were asked to release medical records for a quality of care study.

In Table IV.10, we show cooperation rates[7] for households in the overlap sample that completed Round Two interviews and were eligible to be sampled for either the HCC or CQI survey. The HCC survey had almost no impact on the weighted cooperation rate, since 81.9 percent of HCC participants who were contacted for the third round of the CTS household survey cooperated, compared with 82.3 percent who were not selected for HCC or CQI. On the other hand, the Round Three household survey weighted cooperation rate was only 76.5 percent for individuals sampled for the CQI survey; CQI non-respondents were much less likely to participate in the CTS Round Three survey (66.7 percent cooperation rate), while CQI respondents were about as likely to respond to the CTS Round Three survey (81.3 percent) as individuals who had not been selected for CQI or HCC (82.3 percent).

Persons sampled for the CQI survey participated in a brief (20 minute) telephone interview, after which they were asked to provide access to their medical records, whereas those selected

[7]The cooperation rate used here is the number of households in which at least one FIU was completed divided by the number of confirmed households. Telephone numbers that were ineligible for the survey or for which eligibility could not be determined were excluded.

TABLE IV.10

IMPACT OF COLLATERAL STUDIES ON CTS HOUSEHOLD COOPERATION RATE FOR OVERLAP COMPLETES[a]
(Percents)

	Unweighted	Weighted
No Overlap with HCC or CQI[b]	83.0	82.3
HCC Overlap	79.8	81.9
CQI Overlap	76.5	76.5
CQI complete	81.8	81.3
CQI not complete	65.2	66.7
All Round Two Completes	79.2	80.7

[a]The cooperation rate is the number of households in which at least one FIU was completed divided by the number of confirmed households. Telephone numbers that were ineligible for the survey or for which eligibility could not be determined were excluded.

[b]The CTS was used as a sample frame for the Health Care for Communities (HCC) and the Community Quality Index (CQI) surveys. The HCC and CQI samples were largely mutually exclusive; households in both samples were considered part of the HCC sample.

for HCC were asked to participate in a longer (40 minute) telephone interview, but were not asked to provide access to medical or other personal records. Since the sample frames and data collection methods were similar and the HCC interview was longer, the most likely explanation for the lower CTS cooperation rate for CQI participants was the request for confidential records.

A request for sensitive information in one survey can affect participation rates in a follow-up survey. A recent study by Loosveldt et. al. (2002) found that refusals to sensitive items (for example, income) in an initial survey were positively related to the probability of unit nonresponse in the follow-up survey.

We also estimated the impact of the CQI and HCC surveys on the total weighted Round Three FIU response rate by assuming that households sampled for these surveys would have responded at the same rate as other Round Two completed interviews selected for the overlap sample. The decision to use the CTS as a sample frame for these surveys resulted in an overall decline of 0.6 percentage points in the weighted response rate.

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

The initial FIU interview was conducted with an informant who answered for all sampled FIU members. However, each adult in the FIU was asked to self-respond to a subset of questions (the self-response module), including subjective assessments of health, chronic conditions, tobacco use, satisfaction with care and health plan, aspects of the physician–patient interaction, and consumer information. Although the length of the self response module has increased with each round, we have been able to sustain completion rates of 94 percent across the three rounds (Table IV.11). In certain circumstances, such as when an adult FIU member was too ill to respond, temporarily unavailable, or unwilling to respond after several interviewing efforts had been made, the family informant was allowed to complete the self-response module for that FIU member. The use of proxies has declined with each round, from 2.3 percent of the

TABLE IV.11

RESPONSE RATES FOR THE CTS ADULT SELF-RESPONSE MODULE, BY ROUND
(Percents)

	Round One	Round Two	Round Three
Completed Module	94.5	94.0	94.3
Proxy Accepted			
Illness	0.5	0.5	0.6
Away and unavailable	0.6	0.7	0.1
Language barrier	0.1	0.1	0.0
Other or unspecified reason	1.1	0.3	0.2
Refused or Unable to Complete for Other Reasons	3.2	4.4	4.8
Total	100.0	100.0	100.0
Number of Adults	**49,807**	**48,724**	**49,603**

self-response modules in Round One, to 1.6 percent in Round Two, and 0.9 percent in Round Three.

D. EFFORTS TO INCREASE RESPONSE TO THE SURVEY

During data collection, we used a variety of efforts to increase response, including:

- Making at least 12 calls to determine residency and 50 or more calls to complete interviews with residential households
- Offering Spanish-speaking interviewers to respondents who preferred to conduct the interview in that language
- Making multiple rounds of refusal conversion calls
- Offering monetary incentives
- Leaving messages on mechanical answering devices

1. Follow-Up Calls for the RDD Sample

Telephone numbers in the RDD sample were controlled by the CATI scheduler, which randomly assigned sampled telephone numbers to interviewers; nonscheduled calls were based on optimal calling patterns, dispersed over different times of the day and different days of the week. (As described in Chapter III, the survey introduction for the initial call varied according to whether the telephone number was linked to a household that had been interviewed in Round Two and according to whether a letter had been mailed prior to the call.) Firm appointments were scheduled within a 20-minute window; other appointments were scheduled within a 60-minute time period, based on information provided by the interviewer. Separate queues were set up for Spanish-speaking interviews and for refusal conversions (discussed below). The limit on the number of calls to determine residential and other outcomes for different disposition codes

was varied to provide distributions needed to estimate residency rates under a modeling procedure discussed in Appendix E[8]:

1. Ring no answer

   70% - 12 calls
   20% - 20 calls
   10% - 30 calls

2. Mechanical answering service

   70% - 20 calls
   20% - 30 calls
   10% - 50 calls

3. Effort ended with contact but without confirmed residence (probable residence not coded by interviewer)

   70% - 40 calls
   20% - 50 calls
   10% - 60 calls

4. Effort ended with contact but without confirmed residence (probable residence coded by interviewer)

   70% - 40 calls
   20% - 50 calls
   10% - 60 calls

2. Follow-Up Calls for the Field Sample

Dwelling units selected for the field sample were screened by interviewers to identify households that had not had telephone service for a period of two weeks or more since the beginning of the RDD data collection period. Field interviewers made up to six visits to complete the household interview. Refusal rates were low, and we did not make refusal

[8]The modeling procedure was not used in computing residency rates for the Round Three survey; however, we staggered limits on calls to meet an assumption needed to test it.

conversion calls for the field sample. However, considerable efforts were made to obtain access to locked apartment buildings, which comprised a significant portion of sampled dwellings in some interviewing areas. These efforts included letters and calls by survey managers to supplement efforts by field interviewers.

3. Interviews Conducted in Spanish

We prepared a Spanish version of the CATI instrument and trained bilingual telephone interviewers to conduct interviews with family informants or adults for whom self-response modules were required and who preferred to conduct the interview in Spanish. In addition, two of the field interviewers (one in Miami and one in Orange County) spoke Spanish; in other sites, field interviewers attempted to use family members to translate the screener questions, when necessary. From the first to third round of the CTS, the percentage of family interviews completed in Spanish has increased from 3.8 to 4.9 percent, an increase of 29 percent (Table IV.12). This finding is not surprising, given the nationwide increase in the Hispanic population noted in reports on the 2000 Census.

Spanish-speaking interviewers were critical in the Miami site, where just over 30 percent of the interviews were conducted in Spanish in each round, and in three other high-intensity sites: (1) Newark (5 to 6 percent in each round), (2) Orange County (about 13 percent in each round), and (3) Phoenix (about 7 percent in the first two rounds but nearly 16 percent in Round Three).

In addition, interviews in Spanish comprised at least five percent of the completed Round Three FIU interviews for 11 low-intensity sites.

4. Refusal Conversions

Based on our experience in prior rounds of the CTS Household Survey, we anticipated a high volume of refusals and trained a pool of our best interviewers to convert refusals. Refusal

TABLE IV.12

SPANISH-SPEAKING FIU INTERVIEWS, BY SITE AND OVERALL

	Round One	Round Two	Round Three
	FIU Interviews Completed in Spanish (Percent)	FIU Interviews Completed in Spanish (Percent)	FIU Interviews Completed in Spanish (Percent)
High-Intensity Sites			
Boston MA	2.93	4.38	3.43
Cleveland OH	0.60	0.44	0.42
Greenville SC	0.49	0.33	1.54
Indianapolis IN	0.08	0.27	1.80
Lansing MI	0.17	0.34	1.11
Little Rock AR	0.08	0.82	1.09
Miami FL	31.62	31.12	32.24
Newark NJ	5.07	5.02	5.79
Orange County CA	13.47	12.53	12.72
Phoenix AZ	6.92	7.48	15.59
Seattle WA	0.62	1.00	1.47
Syracuse NY	0.32	0.52	0.49
Low-Intensity Sites			
Atlanta GA	1.35	2.27	2.18
Augusta-Aiken GA/SC	0.34	0.35	1.54
Baltimore MD	0.00	1.03	2.14
Bridgeport CT	2.82	2.23	4.29
Chicago IL	3.41	8.03	8.16
Columbus OH	0.00	0.36	1.22
Denver CO	4.81	4.74	7.22
Detroit MI	0.65	0.72	0.65
Greensboro NC	0.74	0.80	1.41
Houston TX	11.07	9.03	11.76
Huntington-Ashland WV/KY/OH	0.00	0.00	0.00
Killeen TX	2.35	2.04	2.16
Knoxville TN	0.00	0.00	0.00
Las Vegas NV/AZ	5.62	5.56	3.86
Los Angeles CA	19.92	20.68	18.34
Middlesex NJ	2.25	3.47	5.84
Milwaukee WI	0.64	1.85	0.98
Minneapolis-St. Paul MN/WI	0.60	0.32	0.62
Modesto CA	12.75	11.76	10.42
Nassau NY	4.99	3.46	3.42
New York City NY	13.36	14.38	13.27
Philadelphia PA/NJ	3.56	2.06	2.39
Pittsburgh PA	0.00	0.00	0.00
Portland-Salem OR/WA	3.58	1.48	3.65
Riverside CA	9.87	14.95	12.69

TABLE IV.12 (*continued*)

	Round One	Round Two	Round Three
	FIU Interviews Completed in Spanish (Percent)	FIU Interviews Completed in Spanish (Percent)	FIU Interviews Completed in Spanish (Percent)
Low-Intensity Sites (continued)			
Rochester NY	0.85	0.56	0.72
San Antonio TX	5.35	6.79	8.86
San Francisco CA	5.34	8.20	4.71
Santa Rosa CA	4.91	3.31	4.17
Shreveport LA	0.00	0.00	0.00
St. Louis MO/IL	0.00	0.30	0.00
Tampa FL	0.00	3.97	2.52
Tulsa OK	0.00	0.30	1.20
Washington DC/MD	3.87	4.90	4.03
West Palm Beach FL	5.14	3.32	5.40
Worcester MA	0.97	3.17	2.33
Dothan AL	0.00	0.00	0.30
Terre Haute IN	0.00	0.00	0.00
Wilmington NC	0.00	0.00	1.48
West Central Alabama	0.00	0.00	0.00
Central Arkansas	0.26	2.11	1.22
Northern Georgia	2.20	2.30	4.67
Northeast Illinois	0.34	0.00	0.33
Northeast Indiana	0.00	0.00	0.00
Eastern Main	0.00	0.00	0.00
Eastern North Carolina	1.32	2.40	0.89
Northern Utah	2.12	0.73	2.32
Northwest Washington	1.21	2.78	2.27
Supplemental Sample	**3.48**	**3.91**	**4.75**
Total Overall	**3.80**	**4.25**	**4.87**

converters used information about the reason and intensity of the prior refusals in planning their calls. We attempted several rounds of refusal conversions, allowing a minimum of four weeks between refusal conversion attempts to minimize the likelihood of antagonizing respondents. The refusal pool included respondents who hung up the telephone before the interviewer completed the introduction (HUDIs), those who said they preferred not to be interviewed (refusals), those who terminated the call after the screener was completed (breakoffs), and those with electronic privacy managers.[9] Overall, at least one refusal occurred in 21,839 households, or 55.4 percent of the 39,413 households for which residency was determined. (See Table IV.13 for number of households with refusals and Table IV.2 for the total number of completed and residential nonresponding households.)

Refusal conversion efforts were necessary to achieve a high response rate, since at least one FIU interview was completed in 32.1 percent of the households that refused the initial call (Table IV.13).[10] Conversion rates were more successful among Round Two completes in the overlap sample (41.1 percent were converted), than with Round Two noninterviews (19.3 percent) or new and residual sample (31.1 percent). Most of the refusal conversions occurred after one refusal (20.9 percent), with 7.6 percent occurring after two refusals, and 3.6 percent after three or more.

Refusal conversion efforts are designed to reduce non-response and the risk of biases from excluding households reluctant to participate in the survey. The impact of refusal conversion

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

[10]Since refusal conversion rates are performance measures, we report unweighted conversion rates.

TABLE IV.13

ROUND THREE HOUSEHOLD LEVEL UNWEIGHTED REFUSAL CONVERSION RATES, BY SAMPLE TYPE
(Percents)

Refusal Conversion Attempts[a]	Overlap Round Two Completes	Overlap Round Two Noninterviews	New and Sample[c]	Total Sample
Converted After One Refusal	28.0	12.3	19.4	20.9
Converted After Two Refusals	9.4	5.2	7.3	7.6
Converted After Three or More Refusals[b]	3.8	1.8	4.2	3.6
Total Converted	41.1	19.3	31.0	32.1
Not Converted	58.9	80.7	69.0	67.9
Number of Households with Refusals	**7,626**	**4,505**	**9,708**	**21,839**

[a]Refusals were defined as respondents who actively refused or hung up the telephone during the survey presentation.

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

[c]Old and new working banks combined.

efforts on unweighted response rates for Rounds Two and Three RDD sample is shown in Table IV.14. We chose unweighted response rates here to demonstrate the effectiveness of an operational procedure before sample weights were applied. Initial household level response rates prior to refusal conversions were higher on Round Two (48.3 percent) than on Round Three (43.6 percent), with the difference between surveys diminishing with conversion efforts. For Round Two, the difference between the initial and final household level unweighted response rate was 17.6 percentage points and for Round Three it was 20.5 percentage points, indicating that the impact of refusal conversions increased between rounds.

A related question is the impact of refusal conversions on survey estimates, that is, does the substantial effort required by several rounds of refusal conversions change estimates and presumably reduce bias? We are considering an assessment of this key question in a future methodological report.

5. Monetary Incentives

Except for a decision to offer monetary incentives to individuals in Round Three rather than to families as in prior rounds, the amount and form of incentives used in Round Three were based on experiments conducted in prior rounds. For Round One, we tested the impact of different levels of promised incentives ($0, $15, $25, and $35) on response rates and level of effort, and selected $25, which was used for all three rounds (Strouse 1997). We also tested larger promised incentives of $50 and $100 on small samples in Round Two as part of refusal conversion efforts, but did not observe significant gains from the larger incentives and did not use them again.

For Round Three, we decided to offer $25 to each adult responding to the survey rather than

TABLE IV.14

ROUNDS TWO AND THREE RDD (UNWEIGHTED) CUMULATIVE HOUSEHOLD LEVEL RESPONSE RATES, BY NUMBER OF REFUSAL CONVERSIONS
(Percents)

	Round Two Household Level	Incremental Change	Round Three Household Level	Incremental Change
Assumes No Refusal Conversion[a]	48.3	—	43.6	—
Assumes One Refusal Conversion[b]	58.5	10.2	56.5	12.9
Assumes up to Two Refusal Conversions[c]	63.7	5.2	61.4	4.9
Actual Survey Results[d]	65.9	2.2	64.1	2.7

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

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

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

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

to each family. We shifted the target of the incentive from the family to the individual primarily because the Round Two CTS household survey was used as a sample frame for other surveys that compensated individuals; consequently, we assumed that respondents would expect comparable incentives from the next round of the CTS household survey. In addition, the length of the self response module increased for Round Three and was about two-thirds as long as the core interview completed by the family informant. Since data obtained from the self response module was critical to many analyses, we did not want to risk losing observations as the length of this module increased.

For Round Two, we tested the impact of prepaid incentives on households that had refused. Although research by Singer et. al. (1999) showed that prepaid (cash) incentives yield higher response rates than promised incentives, we did not attempt to replicate cash incentives. The CTS payment amount ($25) is too high for cash incentives because of the financial risk of mailing such large amounts of cash. Testing smaller cash incentives was not practical since many survey respondents had participated in prior rounds and were accustomed to receiving $25 or would be sampled for related surveys offering $25 incentives. We were also concerned that sending a check prior to the initial call would remove the prepaid incentive as an option for families that subsequently refuse.

During Round Two, we limited prepaid incentives to households that had refused and for whom we had the names and addresses of potential respondents. We tested the impact of offering $50 versus $25 in an experiment where eligible households were randomized between the two treatments. The $50 incentive resulted in a slightly higher completion rate (44.0 percent) than the $25 incentive (38.4 percent), but the difference was not large enough to warrant such a large incentive and was discontinued (Technical Publication Number 34, Chapter IV).

For Round Three, we mailed $25 checks to households that had refused (most had refused more than once) or were very difficult to contact and for which we had current names and addresses. Checks were mailed to the person responding for the primary FIU (completed interviews in the overlap sample) or the person linked to the address (noncompleted interviews in the overlap sample or new sample).

The results of prepayment efforts used on Rounds Two and Three are shown in Table IV.15. Prepayment efforts were more successful with households that had participated in a prior round of CTS and were more effective in Round Two than in Round Three. For Round Two, 41.0 percent of households mailed prepaid incentives completed interviews, compared with 20.7 percent in Round Three. The completion rate was higher in Round Two than Round Three for completed interviews in the overlap sample (42.9 versus 28.0 percent) and for new sample (38.5 versus 14.2 percent).

However, there were three differences between procedures used in the two rounds that could have affected completion rates. First, the Round Two sample was limited to refusals, whereas the Round Three sample included both households that refused and those that were difficult to contact, which also included a larger number of ineligible cases. Second, the Round Two effort included some cases that were offered $50; as noted above, the completion rate for the larger incentive was 44.0 percent versus 38.4 percent for the $25 incentive. Third, less time was allowed to follow up nonrespondents in Round Three (particularly for the new sample) because the survey effort was truncated after September 11.

TABLE IV.15

DISPOSITION OF PRIMARY FIUS IN HOUSEHOLDS OFFERED PREPAID INCENTIVES BY ROUND AND SAMPLE TYPE

(Percents)

	Round Two[b]				Round Three[b]		
Disposition	Overlap Round One Complete	Overlap Round One Noncomplete	New Sample[c]	Total	Overlap Round Two Complete	New Sample[c]	Total
Complete	42.9	27.5	38.5	41.0	28.0	14.2	20.7
Nonresponse	56.3	68.9	60.6	58.0	64.3	81.3	73.3
Ineligible[a]	0.8	3.5	0.9	1.0	7.7	4.5	6.0
Number of FIUs	3,132	287	899	4,318	3,286	3,713	6,999

[a]Nonresidential, non-working telephone, cellular telephone, non-civilian household.

[b]The Round Two sample was limited to refusals; the Round Three sample included both refusals and other nonrespondents.

[c]Old and new working banks combined.

6. Messages on Mechanical Answering Devices

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

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

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

E. QUALITY ASSURANCE

1. RDD Sample

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

Interviewer conduct during interviews was evaluated primarily by having supervisors

monitor actual calls, supplemented by review of interviewers' notes maintained in the CATI system. (All calls and notes recorded about monitored calls are maintained by the CATI system.) Supervisors monitored approximately 10 percent of the RDD interviews, increasing the monitoring level for new interviewers and those experiencing problems. The monitoring system enables supervisors to listen to interviews without either the interviewers' or respondents' knowledge. It also allows supervisors to view interviewers' screens while an interview is in progress. Interviewers are informed they will be monitored but do not know when observations will take place. Supervisors concentrate on identifying behavioral problems involving inaccurate presentation of information about the study; errors in reading questions; biased probes; inappropriate use of feedback in responding to questions; and any other unacceptable behavior, such as interrupting the respondent or offering a personal opinion about specific questions or about the survey. The supervisor reviews results with the interviewer after the interviewer completes her or his shift.

2. Field Sample

Eligible households for field interviews (interrupted or no telephone service) were interviewed by cellular telephone and were subject to the same monitoring procedures that were used for the RDD sample, described above. In addition, eligibility for the field component was verified by the telephone center interviewer for all completed interviews.

We verified reported telephone service for ten percent of households with listed addresses that had a disposition code of "no interruption in telephone service" (see Table IV.4) by checking against reverse address directories. During Round One, we attempted to obtain verification of telephone status of households with this disposition code that had unlisted telephone numbers by mailing postage paid post cards to a random sample of 10 percent. However, the return rate was poor and was discontinued for subsequent rounds.

F. DATA EDITING, CODING, AND CLEANING

One of the most important advantages of computer-assisted surveys is that errors can be identified and corrected during the interview by building logic, range, and consistency checks into the program. The CATI program (CASES) also permits interviewers to back up and change answers to previously answered questions without violating instrument logic.

For Round Three, a single CATI instrument was written for the RDD and field components of the survey. Separate Spanish versions of the programs were written, but their structures were the same as those of the corresponding English versions. A cleaning program was written that enforced questionnaire logic. An interview could not be certified as clean until all appropriate questions had either been answered or assigned an acceptable nonresponse value, and until the data record for each interview was consistent with the instrument program logic.

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

Other open-ended items included personal contact information, insurance plan names, and employer names. Personal identifying information remained confidential and was maintained in a separate file used only to assign respondent payments and subsequent interviews. Information on insurance plan names and employer names had been used in prior rounds to conduct a followback survey to link data provided by insurers with the household file. However, after

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

conducting a pilot test, we decided not to conduct a Round Three Insurer Followback. In addition, the survey included text responses to closed-ended questions, with options for answers that did not correspond to precoded categories. Files including text responses to these questions were delivered to HSC.

G. REFORMATTING DATA FILES AND FILE DELIVERY

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

V. WEIGHTING AND ESTIMATION

A. OVERVIEW

In this chapter, we discuss weighting and estimation procedures. The CTS Household Survey sample design was complex, using stratification, clustering, and oversampling to produce various national and community level estimates. The use of unweighted data is likely to produce seriously biased estimates because the unweighted samples are distributed differently than the populations they represent. Weights were designed to restore proportionality to the sample and were adjusted to compensate for nonresponse at the household, FIU, and person levels. This difference in proportionality occurred for the following reasons:

- ***Design decision***. Fixed sample sizes for sites, restricting the high-intensity sites to MSAs with populations of 200,000 or more, and subsampling children resulted in different sampling rates for population subgroups.
- ***Incomplete sample frame coverage***. The RDD frame excluded telephone banks of 100 numbers containing no published household numbers; the field sample, excluded areas with high telephone penetration and was restricted to MSAs with populations of 200,000 or more.
- ***Differing chances of selection***. Some households had differing chances of selection because of the number of telephones they owned or interruptions in telephone service. Telephone numbers (RDD sample) selected for Round Two were sampled at different rates for Round Three, depending on the final disposition of the case in the prior round.
- ***Nonresponse***. Survey response rates differed among sites and population subgroups.

Although the correct use of weights in analyzing CTS Household Survey data substantially reduces the bias of estimates resulting from the sample design and survey nonresponse, the weights do not address the potential for bias resulting from item nonresponse or response errors. The procedures used to impute missing data for individual variables are discussed in the Household Survey Round Three Public Use File (technical publication forthcoming). Estimates

of sampling error that do not account for the use of weights and the complex nature of the sample are likely to be severely understated. Specialized software is required to properly estimate standard errors of estimates from this survey; procedures for using different statistical software packages will be discussed in "Comparison of Statistical Software Packages for Variance Estimation in the CTS Surveys" (Technical Publication 40, May 2003).

1. Weights Provided for Public Use

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

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

1. ***Site Sample.*** Weights for national estimates that use data from the site sample
2. ***Supplemental Sample.*** Weights for national estimates that use the supplemental sample
3. ***Combined Sample.*** Weights used for national estimates that combine data from the site sample and supplemental sample

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

4. ***Augmented Site Sample (Site-Specific Estimates).*** Weights for site-specific estimates that use data from the site sample, augmented with observations from the supplemental sample located within the boundaries of the 60 sites

5. ***Augmented Site Sample (National Estimates*).** Weights for national estimates that use data from the site sample, augmented with observations from the supplemental sample located within the boundaries of the sites

TABLE V.1

NAMES OF ROUND THREE CTS HOUSEHOLD SURVEY WEIGHTS

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

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

In many surveys, nonresponse, poststratification, and other adjustments can introduce variation in the sampling weights. In some situations, the combination of these adjustments produces disproportionately large weights. These large weights can decrease the accuracy and precision of point estimates. We chose to reduce the sampling error caused by extremely large weights by trimming them and distributing the excess among other weights. Although the difference between estimates using the trimmed or untrimmed weights is quite small (the extent of trimming was not great), the trimmed weights result in better precision, with little or no additional bias.

The combined weights combine individual-level and FIU weights from the site and supplemental samples to produce national estimates. These weights are based on the relative

variances of the two samples and enable researchers to more easily take advantage of the increased precision of the combined samples.

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

2. Constructing Weights

Each weight is the product of several factors:

- An initial weight, the inverse of the probability of selection, to correct for differences in probabilities of selection
- Nonresponse adjustment factors, to correct for differential nonresponse at the individual, FIU, and household levels
- Factors to adjust for interruptions in telephone service
- Poststratification adjustments of 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 sample weights, and combined weights
- A variance-based factor for the combined weights that allows the site and supplemental samples to be used together for national estimates

3. Sampling Error Estimation

Because sample-based estimates of population characteristics are not based on the full population, some element of uncertainty is always associated with these estimates. This element of uncertainty, known as *sampling error,* is an indicator of the precision of an estimate. Sampling error is generally measured in terms of the standard error or the sampling variance, which is the square of the standard error.[2]

The complexities of the CTS Household Survey design preclude the use of statistical software packages for variance estimation that do not account for such a design in their algorithms. The variance estimates from these statistical packages may severely underestimate the sampling variance in the Household Survey. Therefore, the CTS data require the use of survey data analysis software or specially developed programs designed to accommodate the sample design and the statistic being estimated.

The sampling variance in the Household Survey is a function of the sampling design and the population parameter being estimated and is referred to as a *design-based sampling variance*. The CTS database contains fully adjusted sampling weights for site-specific estimates and national estimates of FIUs and persons, as well as the information on sample design parameters (that is, strata and clusters) necessary to estimate the sampling variance for a statistic.

[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 members of the population, whereas a sampling variance is a measure of the variation of the *estimate* of a population parameter (for example, a population mean or proportion) over repeated samples. The population variance is different from the sampling variance in the sense that the population variance is a constant, independent of the sample design, whereas the sampling variance decreases as the sample size increases. The sampling variance is zero when the full population is observed, as in a census.

Most common statistical estimates and analysis tools (such as percentages, percentiles, and linear and logistic regression) can be implemented using Taylor series approximation methods. Survey data software, such as SUDAAN (Shah et al. 1997), uses the Taylor series linearization procedure and can handle the multistage design, joint inclusion probabilities, and variance components in the Household Survey design. (See Technical Publication 40 for a discussion of how to use different statistical packages to analyze CTS Household Survey data.)

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

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 issues pertaining to the construction of the three types of RDD sample weights.

1. General Weighting Approach

As explained in Chapter II, sampling took place in several stages. In the first stage, we selected the 60 sites (with probability proportional to size) and then randomly selected the high-

intensity sites from among the 60. For the RDD sample, we selected telephone numbers, identified households, defined FIUs within households, and collected data on FIUs and individuals within FIUs (all eligible adults 18 years of age and older and one randomly selected child). Each of these stages was considered in weighting. The steps necessary for calculating FIU- and person-level weights are listed here and described in the sections that follow.

- Calculate site probability of selection (only for weights using the site sample to make national estimates)
- Account for the distribution of cases in high- and low-intensity sites (only for weights using the site sample to make national estimates)
- Calculate probability of selection of telephone numbers
- Adjust for the telephone number resolution rate (determination of whether the telephone number was a working residential number)
- Adjust for the household screener rate (determination of the household's eligibility using household enumeration questions)
- Adjust for household nonresponse among eligible households
- Adjust for multiple telephones and telephone service interruption within a household
- Poststratify household weights to external estimates of telephone and nontelephone households
- Adjust for secondary FIU nonresponse within responding households
- Calculate the probability of selection for the randomly selected child
- Adjust for high person level item nonresponse within responding FIUs

a. Telephone Number Initial Weight

The telephone number was the second stage of selection for the site sample,[3] and the first stage of selection for the supplemental sample. The telephone sampling weight accounted for the probability of selection of telephone numbers within each site, stratum, and overlap sampling category.[4] The probability of selection accounts for the fact that most sampled telephone numbers in Round Three could have been selected for the first time in Round Three or been selected for the first time in Rounds One or Two.

We calculated the probability of selection of a Round Three telephone number (in stratum *h*, with Round Two disposition *d*) as follows:[5]

$$(1) \quad Q_{hd} = (Q1_h \cdot Q2_{hd}) + (Q3_h \cdot Q4_{hd}),$$

where $Q1_h$ is the cumulative probability of selection in the prior round (Round Two); $Q2_{hd}$ is the probability of selection in Round Three given that the telephone number was selected for release in Round Two; $Q3_h$ is the probability of *not* being selected for release in Round Two ($1 - Q1_h$); and $Q4_{hd}$ is the probability of selection of new cases in Round Three.

---

[3] The site was the first stage of selection for the site sample (see Metcalf, et al., 1996).

[4] There are six overlap sampling categories:(1) Round Two complete; (2) Round Two refusal; (3) Round Two other nonresponse; (4) Round Two non-household; (5) Round Two no answer; and (6) Round Two answering device. There are two new telephone number sampling categories: (1) old working banks ("residual") and (2) new working banks ("new").

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

The first step was to calculate these probabilities of selection for each telephone number. By the third round of data collection, the calculation of these probabilities becomes complex. Different methods were used depending on whether the telephone number was part of the overlap sample (that is, Round Two sample members subsampled for Round Three); part of the residual sample (sampled for the first time in Round Three, but theoretically could have been selected in Round Two); or part of the new sample (sampled for the first time in Round Three, and had no chance of selection in Round Two). We will address each one separately below.

Probability of Selection in Round Three for Overlap Sample. Because these cases were actually selected in Round Two, we had previously calculated the value of $Q1_h$, the probability of selection in Round Two. $Q1_h$ already incorporated whether the case could have come into the sample for the first time in Round Two or as part of the Round One sample. We then have to calculate $Q2_{hd}$, the subsampling rate for Round Three. For overlap cases, the Round Two disposition, *d*, has six values (described previously). $Q2_{hd}$ is calculated within stratum *h* and Round Two disposition *d* as:

(2) $$Q2_{hd} = P(\text{overlap case subsampled in R3 given R2 status } d, \text{stratum } h) = \frac{n_{hd}}{N_{hd}} \cdot \frac{nrel_{hd}}{n_{hd} - nbad_{hd}},$$

$d=1,2,\ldots,6$ *(overlap),*

where N_{hd} is the number of Round Two sample telephone numbers; n_{hd} is the number of these telephone numbers initially selected in Round Three; $nbad_{hd}$ is the number of these telephone numbers found to be nonworking or business numbers before being released (using Genesys-ID or –ID-Plus); and $nrel_{hd}$ is the number of these telephone numbers released for interviewing.

Because $Q3_h$ is simply 1-$Q1_h$, we are left to estimate only $Q4_{hd}$ for the overlap cases, which is the probability of selection in Round Three for those not part of the Round Two sample. As discussed above, we must calculate the probability of selection for all possible ways these

telephone numbers could have gotten into the sample. Overlap cases could have come into the sample for the first time in Round Three as residual sample. We calculated their alternative probability of selection ($Q4_{hd}$) by applying residual cases' Round Three probabilities of selection (disregarding probabilities of selection in prior rounds) to overlap cases by strata.

Probability of Selection in Round Three for New Sample. A telephone number *bank* is defined as the first 8 digits of a 10-digit telephone number; a bank has 100 possible 10-digit telephone numbers associated with it. If at least 1 of these 100 possible telephone numbers was listed in a telephone directory as a residential number, then the bank was designated as a *working bank*. New sample did not have a chance of selection in Round Two because these telephone numbers were not in a working bank at the time of the prior round. Because of that, $Q1_h$ is set to 0, $Q2_{hd}$ is not applicable, and $Q3_h$ is equal to 1. That leaves only $Q4_{hd}$, the probability of selection in Round Three, given the telephone number was not part of the Round Two sample. $Q4_{hd}$ is calculated, within stratum *h*, as:

$$(3)\quad Q4_{hd} = Q4_{h\text{-}new} = P(\text{new case selected in R3 given not in R2, stratum } h) = \frac{n_{h\text{-}new}}{N_{h\text{-}new}} \cdot \frac{nrel_{h\text{-}new}}{n_{h\text{-}new} - nbad_{h\text{-}new}},$$

where $N_{h\text{-}new}$ is the number of new working telephone banks times 100; $n_{h\text{-}new}$ is the number of these telephone numbers initially selected in Round Three; $nbad_{h\text{-}new}$ is the number of these telephone numbers found to be nonworking or business numbers prior to release (using Genesys-ID or –ID-Plus); and $nrel_{h\text{-}new}$ is the number of these telephone numbers released for interviewing.

Probability of Selection in Round Three for Residual Sample. Calculating the probabilities of selection for the residual sample was more complex than for the overlap or new samples. Calculating the actual probabilities of selection in Round Three (disregarding probabilities of selection in prior rounds) was straightforward. However, calculating the alternative way in

which residual sample telephone numbers could been selected (as overlap cases) was difficult because sampling rates for overlap sample are based on the prior round's disposition, which do not exist for residual sample cases.

We decided to impute hypothetical prior round dispositions based on Round Three dispositions. By cross-tabulating groupings of Round Two and Round Three dispositions for telephone numbers in the overlap sample, we were able to estimate the proportion of telephone numbers with a Round Three disposition of complete that had been: (a) a Round Two complete; (b) a Round Two refusal, no answer, or answering device; or (c) a Round Two other nonresponse or non-household. Similarly, we calculated these three proportions for each of the other five Round Three disposition groups (refusal, other nonresponse, non-household, no answer, answering device). If we assign these proportions the notation R_{ij}, where *i* designates the Round Three status group and *j* designates the Round Two status group, then the probability of selection for a residual case with Round Three status *i* can be shown as:

$$(4)\quad Q_{hid} = Q_{hi-residual} = \sum_{j=1}^{3}\left[\overline{Q1}_{hij}\,\overline{Q2}_{hij} + (1-\overline{Q1}_{hij})Q4_{hd}\right]\cdot R_{ij}$$

where i = 1 (R3 complete), 2 (R3 refusal), 3 (R3 other nonresponse), 4 (R3 non-household), 5 (R3 no answer), 6 (R3 answering device); and j = 1 (R2 complete), 2 (R2 refusal or no contact), 3 (R2 other nonresponse or non-household). $\overline{Q1}_{hij}$ is the mean of *Q1* (the Round Two probability of selection) among overlap cases in stratum *h* with Round Three status *i* and Round Two status *j*, and $\overline{Q2}_{hij}$ is the mean of *Q2* (the Round Three subsampling rate) among these same overlap cases. The value of $Q4_{hd}$ is calculated, within stratum *h*, as:

(5)

$$Q4_{hd} = Q4_{h\text{-}residual} = P(\text{residual case selected in R3 given not in R2, stratum } h) = \frac{n_{h\text{-}residual}}{N_{h\text{-}residual}} \cdot \frac{nrel_{h\text{-}residual}}{n_{h\text{-}residual} - nbad_{h\text{-}residual}},$$

where $N_{h\text{-}residual}$ is the number of working telephone banks (banks not new to Round Three) times 100; $n_{h\text{-}residual}$ is the number of these telephone numbers initially selected in Round Three as part of the residual sample; $nbad_{h\text{-}residual}$ is the number of these telephone numbers found to be nonworking or business numbers prior to release (using Genesys-ID or –ID-Plus); and $nrel_{h\text{-}residual}$ is the number of these telephone numbers released for interviewing.

Sampling weights. Probability formulas differed slightly depending on the type of estimate (national or site-specific) and sample type (site, supplemental, or augmented) for which the weight was designed (described in more detail in Sections V.B.2 through V.B.5). Once the probability of selection is calculated, the sampling weight is simply the reciprocal of that probability of selection:

(6) $SW(phone_{hd}) = \frac{1}{Q_{hd}} or \frac{1}{Q_{hid}}$ for the residual sample.[6]

b. Adjustment to Telephone Weight for Undetermined Residency

For the telephone number weight, an adjustment was made for undetermined eligibility status (inability to determine whether a sampled telephone number was a working residential number). For the telephone numbers with undetermined residency, we estimated the percentage with working residential telephone numbers from a study conducted for the National

[6]The estimated Round Three probability of selection for the residual sample depends on Round Three disposition *i*.

Immunization Survey (Shapiro et al. 1995) and discussed in Section C of Chapter IV. Based on this study, the status of these telephone numbers was randomly assigned as follows:[7]

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

Corresponding percentages of the weights for these unresolved telephone numbers were included in the subsequent weighting steps as households. Telephone numbers known not to be working residential numbers were deleted in this step. The remaining telephone numbers were those known to be working and associated with residential households.

c. Nonresponse Adjustment to Household Weight for Screener Nonresponse

After adjusting for undetermined residency, we formed weighting cells to adjust for two kinds of household-level nonresponse: (1) nonresponse to survey questions used to determine whether the household was eligible, and (2) nonresponse to the survey by eligible households; that is, residences in which there is at least one eligible adult.[8]

We formed primary weighting cells by crossing site, sampling strata, and Round Two disposition; for the supplemental sample, the cells were defined by stratum and Round Two

[7]Sample weights were computed before we analyzed different procedures for estimating residency for undetermined cases (see Appendix E). Although we decided to discontinue use of the business office method for resolving undetermined cases, the difference in the residency rate between the business office and CASRO method (which we decided to use for computing response rates) was too small to justify re-computing the weights.

[8]See Section II.F.1 in Chapter II for the complete definitions of eligible and ineligible households and Section II.B.4 for the definition of individuals excluded from the survey.

disposition. The Round Two disposition categories used to form cells were (1) Round Two complete; (2) Round Two noncomplete (combining all noncomplete sampling categories); (3) residual (not sampled in Round Two, but could have been); and (4) new sample (not sampled in Round Two because its telephone bank was nonworking at that time). Based on generally accepted guidelines, we decided that each cell should contain at least 20 respondents, and that the adjustment factor in each cell should be less than two. Cells failing these criteria were combined with similar cells.

The first adjustments accounted for whether a residential household was eligible for the survey. For Round Three households with incomplete information on household eligibility, we created the following household eligibility nonresponse adjustment factor:

$$(7)\quad A'_{nr}(hhold_c) = \frac{\sum_{hh \in c} SW(phone_{hd})}{\sum_{det\ hh \in c} SW(phone_{hd})},$$

for households in stratum h with disposition d, which are in cell c, where the numerator is summed over all telephone numbers in cell c known to be households (plus telephone numbers imputed to be households using the external empirical estimates described in the previous section), and the denominator is summed over households in cell c with a known survey eligibility status.

A telephone number weight adjusted for determination of household eligibility was then calculated for these cases:

(8) $W1(phone_{hd}) = SW(phone_{hd}) \cdot A'_{nr}(hhold_c)$, if eligibility of household determined

$W1(phone_{hd}) = 0$, otherwise.

After this adjustment, households with undetermined eligibility status and households known to be ineligible for the survey were removed from the weighting process.[9]

d. Interview Nonresponse Adjustment to Household Weight

We then adjusted these weights for survey nonresponse among eligible households. A responding household was one in which at least one eligible FIU responded to the survey. We performed a weighting class adjustment for households using the same cells as defined for the household eligibility adjustment. We created a household survey nonresponse adjustment factor as follows:

$$(9) \quad A''_{nr}(survey_c) = \frac{\sum_{elig\ hh \in c} W1(phone_{hd})}{\sum_{resp\ hh \in c} W1(phone_{hd})},$$

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

$$(10) \quad W2(phone_{hd}) = W1(phone_{hd}) \cdot A''_{nr}(survey_c), \text{ if household responded}$$

$$W2(phone_{hd}) = 0, \text{ otherwise.}$$

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

e. Poststratification and Other Adjustments to Household Weight

We then adjusted for multiple telephones in the household and for interruptions in telephone service.[10] Because some households have multiple nonbusiness telephone numbers,[11] a household multiplicity factor was used to adjust for the number of telephone numbers in the household. This factor, which is simply the inverse of the total number of these telephones in the household, was applied to the nonresponse-adjusted household weight:

(11) $W3(hhold_{hdi}) = W2(phone_{hd})/(number\ of\ phones\ in\ household\ i).$

One of the last steps in creating the household-level weight was to poststratify the sum of the weights to external estimates of current population totals. We created two sets of weights for the RDD sample: (1) ones that sum to telephone households and (2) ones that sum to all households. We used estimates from the March 2000 Supplement to the Current Population Survey (U.S. Census Bureau, http://www.bls.census.gov/cps/cpsmain.htm) of the number of telephone and nontelephone households nationally (by whether or not in an MSA). For counts of telephone and nontelephone households in each site, we used a combination of estimates from the March 2000 Supplement to the Current Population Survey, the Census 2000 Supplemental Survey, and the Census 2000 Short Form data. The poststratification adjustment factor for telephone households is:

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

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

(12) $$A_{ps\text{-}tel}(\text{metro status}) = \frac{TELHH_{(non)metro}}{\sum_{resp\ hh_i \in (non)metro} W3(hhold_{hdi})}$$ for the national weights, and

(13) $$A_{ps\text{-}tel}(\text{site}) = \frac{TELHH_{site}}{\sum_{resp\ hh_i \in site} W3(hhold_{hdi})}$$ for the site-specific weights,

where *TELHH* is the estimated number of telephone households in 2000 (for metro U.S., non-metro U.S., or by site), and the denominator is the sum of the nonresponse-adjusted weights for all responding households in the corresponding category (metro or nonmetro for national weights, site for site-specific weights). The household-level weight poststratified to telephone households is:

(14) $$WT_{tel}(hhold_{hdi}) = W3(hhold_{hdi}) \cdot A_{ps\text{-}tel}\ .$$

To create the weights summing to all households, we used information on telephone service interruption to inflate the RDD sample weights for telephone households in order to account for nontelephone households.[12] Even though all cases in the RDD telephone sample had working telephones when interviewed, they were asked whether they had had any interruption in telephone service during the year preceding the interview.[13] We used cases with interruptions in telephone service to represent nontelephone households and those with no reported interruptions to represent telephone households. Then, we adjusted weights to the number of months of interrupted service. The interruption-adjusted weight is:

[12]These weights were used for the supplemental sample and low-intensity site-specific weights when combining with the rest of the RDD and field samples.

[13]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...?").

(15) $$WT_{interruption}(hhold_{hdi}) = \frac{W3(hhold_{hdi})}{proportion\ \ of\ \ \ year\ \ household\ i\ with\ \ phone}.$$

The poststratification adjustment factor for total households is:

(16) $$A_{ps\text{-}all}(phone\ status\ g, metro\ status) = \frac{TOTHH_{\,g\text{-}metro\ status}}{\sum_{resp\ \ hhold_i\ \ with\ phone\ status\ g \in metro\ status} WT_{interruption}(hhold_{hdi})}$$

for the national weights, and

(17) $$A_{ps\text{-}all}(phone\ status\ g, site) = \frac{TOTHH_{\,g\text{-}site}}{\sum_{resp\ \ hhold_i\ \ with\ phone\ status\ g\ in\ site} WT_{interruption}(hhold_{hdi})}$$ for site-specific

weights,

where *TOTHH* is the estimated number of all households in 2000 (for metro U.S., non-metro U.S., or by site) by phone status, and the denominator is the sum of the nonresponse- and phone status-adjusted weights for all responding households in the corresponding category. *Phone status g* is equal to one (interruption in telephone service) or is equal to two (no known interruption in telephone service), $TOTHH_1 = TOTHH - TELHH$ and $TOTHH_2 = TELHH$.

The household-level weight poststratified to all households is:

(18) $$WT_{all}(hhold_{ghdi}) = WT_{interruption}(hhold_{hdi}) \cdot A_{ps\text{-}all}(phone\ status\ g).$$

f. Interview Nonresponse Weight Adjustment for FIUs

The probability of selection of each FIU was equal to the probability of selection for its household (that is, all FIUs in a selected household were selected for the interview). We therefore used the final household weight as the starting point for developing the FIU weight. The FIU weights accounted for FIU interview nonresponse within responding households.

Within responding households, FIU eligibility was based on information provide by the household informant.

We started with an FIU-level file containing all FIUs enumerated within responding households and assigned to each FIU its final household weight. Using the same cells as defined for the telephone- and household-level adjustments, we created an FIU survey nonresponse adjustment factor for FIUs in responding households *i* (stratum *h*, Round Two disposition *d*):[14]

$$(19) \qquad A_{nr}(FIU_c) = \frac{\sum_{elig\ fiu \in c} WT(hhold_{hdi})}{\sum_{resp\ fiu \in c} WT(hhold_{hdi})},$$

where the numerator is summed over all eligible FIUs in cell *c*, and the denominator is summed over responding eligible FIUs in cell *c*.

An FIU weight adjusted for survey nonresponse was then calculated for these cases:

$$(20) \qquad W4(FIU_{hdi}) = WT(hhold_{hdi}) \cdot A_{nr}(FIU_c), \text{ if FIU responded}$$

$$W4(FIU_{hdi}) = 0, \text{ otherwise.}$$

g. Initial Person Weight

The probability of selection for each adult member of an eligible responding FIU was equal to the probability of selection of the FIU (that is, all adults in each responding FIU were selected for the interview). We therefore used the final FIU weight to develop the person weight for adults. However, because only one child was selected at random per FIU, the within-FIU probability of selection for a child was equal to the inverse of the number of children in the FIU.

[14] To simplify notation, we use $WT(hhold_{hdi})$ here to refer to both national and site-specific household weights, weighted up to all households (WT_{all}) or weighted up to just telephone households (WT_{tel}). Parallel adjustments are made for all versions of these household weights.

The overall probability of selection for person *k* in FIU *j* in household *i* in stratum *h* can be expressed as:

$$(21) \qquad P(person_{hijk}) = \frac{P(FIU_{hij})}{(\delta \cdot numkids_{hij}) + (1-\delta)},$$

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

$$(22) \qquad W5(person_{hdijk}) = W4(FIU_{hdi}) \cdot [(\delta \cdot numkids_{hij}) + (1-\delta)],$$

for all persons *k* in FIU *j*, household *i*, stratum *h*, with disposition *d*.

All eligible individuals in responding FIUs were assigned this weight regardless of whether we had complete data on an individual. Most of the survey data was obtained from the FIU informant about all family members; however, responses to subjective questions were obtained from a self-response module completed by each adult. Therefore, for some individuals we had data provided by the FIU informant but were missing data from that person's self response module.

h. Nonresponse Adjustment to Person Weight

The next adjustment to the person weight accounted for high levels of missing data among individuals selected for the survey.[15] There were 30 person records deleted because of high

[15]An editing program was used to determine whether a person record contained too many missing items to be usable. The editing rule was that all person records with 75 percent or more missing data for variables from Sections B through G of the questionnaire were considered to be nonrespondents.

levels of missing information. This step in the weighting process adjusted for unit nonresponse at the person level, using the same weighting cells as defined for previous adjustments. We created a person-level survey nonresponse adjustment factor as follows:

$$(23) \qquad A_{nr}(missing_c) = \frac{\sum_{elig\ person \in c} W5(person_{hijkd})}{\sum_{resp\ person \in c} W5(person_{hijkd})},$$

for person *k* (in FIU *j*, household *i*, stratum *h*, disposition *d*) in cell *c*, where the numerator is summed over all eligible and selected individuals in cell *c*, and the denominator is summed over individuals with complete responses. A person weight adjusted for survey nonresponse was then calculated for these cases:

(24) $W6(person_{hijkd}) = W5(person_{hijkd}) \cdot A_{nr}(missing_c)$, if person met the editing rule for individuals

$W6(person_{hijkd}) = 0$, otherwise.

2. Calculating the Base Weight for National Estimates Using the Site Sample

The previous section described the general weighting approach used for the RDD sample, including the initial sampling weight for each telephone number and various adjustments to account for eligibility determination, nonresponse, and household-level poststratification. This section shows how the general approach is applied to the weight used when making national estimates using the site sample.

Weights used for making national estimates using the site sample must account for the probability of selection of the site, as well as for the distribution of cases in the high-intensity and low-intensity sites. (The selection of the 60 sites is discussed in detail in Metcalf et al. [1996].)

The sample size of RDD telephone numbers was about four times larger 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 used the expected number of selected telephone numbers in each site, $E(n_{sh})$, rather than the actual number of selected telephone numbers, n_{sh}. For site s in stratum h, where the site is an MSA with 200,000 or more people, the expected number of selected telephone numbers is:

$$
\begin{aligned}
(25) \quad E(n_{sh}) &= [n_{lo} \cdot 4 \cdot P(\text{high intensity})] + [n_{lo} \cdot P(\text{low intensity})] \\
&= [n_{lo} \cdot 4 \cdot 12/48] + [n_{lo} \cdot 36/48] \\
&= n_{lo} \cdot (4/4 + 3/4) \\
&= n_{lo} \cdot 1.75 \ ,
\end{aligned}
$$

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

When making national estimates based on the site sample, the combined site and telephone number probability of selection can then be defined as:

$$
(26) \quad QN_{hds}(telephone) = PSUPROB_s \cdot \frac{E(n_{sh})}{n_{sh}} \cdot Q_{hd} \ ,^{16}
$$

where $PSUPROB_s$ is the probability of selection of site s[17] and n_{sh} is the actual number of telephone numbers selected in the site sample in stratum h in site s (set equal to n_{lo} for low-

[16] For residual sample cases, the last term is Q_{hid} instead of Q_{hd}.

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

intensity sites and equal to $4 \times n_{lo}$ for high-intensity sites, for the actual calculation). Q_{hd} is the probability of selection of telephone numbers defined in Section 1.a.

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

3. Calculating the Base Weight to Make National Estimates Using the Supplemental Sample

When using only the supplemental sample to make national estimates, the probability of selection of the telephone number is Q_{hd}, defined in Section 1.a above. The general weighting approach described in Section 1 can be directly applied to this situation without further adjustment.

4. Using the Augmented Site Sample to Make National Estimates

When calculating the probability of selection for the augmented site sample, we must account for the fact that the telephone numbers could have entered the sample through either the site sample or the supplemental sample. For the augmented site sample weight, we assigned to each telephone number the "alternative site" and "alternative stratum" it would have been assigned had the case been selected in the other sample. Thus, we determined the stratum into which each site sample telephone number would have fallen, had it been selected through the supplemental sample instead, and determined the site (and substrata, if applicable) into which each supplemental sample telephone number would have fallen, had it been selected through the site sample.

We then calculated the mean value of the probability of selection for site sample cases by site, substratum, and sample type (overlap by Round Two status, residual, or new sample) and merged these mean probabilities onto the supplemental cases according to their alternative site,

substratum, and sample type. Similarly, we calculated the mean value of the probability of selection for supplemental sample cases by stratum and sample type and merged these mean probabilities onto the site sample cases according to their alternative stratum and sample type.

The probability of selection was then calculated as:

$$(27)\quad Q_{hd}(aug) = Q_{hd} + (1 - Q_{hd})\left(altQ_{hd}\right)$$

where $altQ_{hd}$ is the alternative probability of selection.

In other words, for each telephone number in the site sample (in stratum *h* with disposition *d*), we calculated the actual probability of selection through the site sample (Q_{hd}) and an alternative probability of selection if the telephone number were to have come into the sample through the supplemental sample instead ($altQ_{hd}$). The overall probability of selection is a sum of the two probabilities, accounting for the fact that a household could not have actually come into the sample both ways (1-Q_{hd}). Similar calculations were made for telephone numbers that came in through the supplemental sample.

When using the augmented sample to make *national* estimates, we also had to modify the value of the *expected* sample size in each site as follows:

$$(28)\quad E'(n_{sh}) = (n_{lo} \times (4/4 + 3/4)) + ((n_{lo} \times 10) \times usprop),$$

where the first term of the sum is the same as that described in equation 25 above,[18] and the second term of the sum represents the expected sample size contributed by the supplemental sample for a particular site. In this second term, 10 is the size of the supplemental sample, as a

[18]This part of the formula accounts for the fact that a telephone number in a large MSA had a ¼ chance of being in a high-intensity site in the site sample (with four times the sample size of a low-intensity site), and a ¾ chance of being in a low-intensity site in the site sample.

multiple of the size of a low-intensity site,[19] and *usprop* is the proportion of the U.S. population (excluding Alaska and Hawaii) that fell within each site in July 1992. This part of the formula accounts for the fact that a telephone number also had a chance of being in the supplemental sample in proportion to its population (with ten times the sample size of a low-intensity site).

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

(29) $n_{sh} = (n_{lo}) + ((n_{lo} \cdot 10) \cdot usprop)$ for low-intensity sites,

and

(30) $n_{sh} = (n_{lo} \cdot 4) + ((n_{lo} \cdot 10) \cdot usprop)$ for high-intensity sites.

The probability of selection of the telephone number for the augmented site sample (national estimates) can then be defined (for site *s* in stratum *h* with disposition *d*) as:

(31) $$QN_{hds}(aug - telephone) = PSUPROB_s \cdot \frac{E'(n_{sh})}{n_{sh}} \cdot Q_{hd}(aug).$$

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

When combining the site sample and the supplemental sample to make the site-specific estimates, the probability of selection of the telephone number is $Q_{hd}(aug)$, defined in Section 4 above.

[19]This is in accordance with the relative sample sizes of the supplemental sample and a low-intensity site in the sample allocation design (see Chapter II).

C. WEIGHTS FOR THE FIELD SAMPLE

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

We produced two sets of weights for the field survey data. Although neither set is intended to be used alone in policy analysis, these two sets of weights and the weights representing the RDD sample were used to create integrated weights for making inferences about the entire U.S. population (excluding Alaska and Hawaii). Field sample weights for households, FIUs, and individuals were constructed for (1) individual sites in which the field survey was conducted, and (2) all MSAs with 1992 populations of 200,000 or more. We refer to the second set of weights as *national* weights. As with the RDD sample, each weight was the product of several factors that reflected differences in probabilities of selection and nonresponse. The set of weights (household, FIU, and person level) also included poststratification adjustments so that the sample matched external estimates of the relevant population.

1. Steps in the Weighting Process

The first weighting factor for a unit (listed housing unit [LHU], household, FIU, or individual) for any of the weights was the inverse of that unit's probability of selection.[20] This

[20]We use the Census definition of a housing unit, that is, a structure that is occupied or intended for occupancy by person(s) living separately from other person(s) in the building and must meet one of the following criteria: (a) It has complete kitchen facilities for the exclusive use of that unit regardless of whether the kitchen is used or not, or (b) The housing unit has a

factor differed for site-specific estimates and national estimates. The weights also account for various types of nonresponse at the household, FIU, or individual level, and are ratio-adjusted to estimated population totals (poststratification).

a. Initial Weights

The initial weight was the inverse of the overall probability of selection of a unit. For a listed housing unit LHU_i in listing area LA_c in secondary sampling unit SSU_b and primary sampling unit PSU_a, the preliminary supplemental sample weight, *SWN*, is:

(32) $\mathrm{SWN}(\mathrm{LHU}_{i\in abc}) = 1/\mathrm{P}(\mathrm{LHU}_{i\in abc})$, where:

(33) $$P(LHU_{i\in abc}) = P(PSU_a)\cdot P(SSU_b \mid PSU_a)\cdot P(LA_c \mid SSU_b)\cdot P(LHU_i \mid LA_c).$$

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 $\mathrm{P}(\mathrm{LHUi}\mid\mathrm{LAc})$ accounts for the fact that only a subsample of listed housing units was selected for interviewing in some listing areas.

For site-specific estimates, the same formula can be modified by omitting the term for the site selection probability $P(PSU_a)$. Thus, for site-level estimates for site *a*:

(34) $$SWS(LHU_{i\in abc}) = 1/P_a(LHU_{i\in abc}),$$

(35) $$P_a(LHU_{i\in abc}) = P(SSU_b \mid PSU_a)\cdot P(LA_c \mid SSU_b)\cdot P(LHU_i \mid LA_c)$$

(continued)
separate entrance directly from the outside of the structure or through a common or public hall, lobby or vestibule.

Further adjustments to the field sample weights were carried out similarly to those for the RDD sample weights.

b. Adjustment to Field Sample Weight for Undetermined Residency

For the field weight, household level adjustments were made for whether: (1) a sampled address was coded as an inhabited residence; (2) telephone status of a residence was determined; (3) survey eligibility was determined (defined the same way as for the RDD sample); and (4) the eligible household responded. These adjustments were done within weighting classes defined as each site.

To adjust for whether a sampled address was known to be an occupied housing unit, we created the following adjustment factor:[21]

$$(36) \qquad A'_{nr}(hhold_c) = \frac{\sum_{LHU_i \in c} SW(LHU_i)}{\sum_{det\ LHU_i \in c} SW(LHU_i)},$$

where the numerator is summed over all addresses in cell (site) *c*, and the denominator is summed over addresses in cell *c* with a known residency status. An address weight, adjusted for determination of address eligibility, was then calculated for these cases:

$$(37) \qquad W1(LHU_i) = SW(LHU_i) \cdot A'_{nr}(hhold_c)\text{, if eligibility of address determined}$$

$$W1(LHU_i) = 0\text{, otherwise.}$$

[21]To simplify notation, we switch from $SWN(LHU_{i \in abc})$ for national weights and $SWS(LHU_{i \in abc})$ for site-specific weights to $SW(LHU_i)$. Parallel adjustments are made for both versions of these weights.

After this adjustment, addresses with undetermined eligibility status and addresses known to be vacant or nonresidential were removed from the weighting process. The remaining addresses were those known to be households.

c. Adjustment to Field Sample Weight for Undetermined Telephone Status

Field sample households were screened to identify households that (1) did not have telephone service at the time we contacted them, or (2) had an interruption in telephone service of two weeks or more in the last 12 months. Households meeting either criteria were eligible for the field survey.

To adjust for whether a household's telephone status was known (that is, whether the telephone screening questions were completed), we created the following adjustment factor:

$$(38) \qquad A''_{nr}(nonphone_c) = \frac{\sum_{households \in c} W1(LHU_i)}{\sum_{det\ phone \in c} W1(LHU_i)},$$

where the numerator is summed over all known residential addresses in cell (site) *c*, and the denominator is summed over addresses in cell *c* with a known telephone status. A household weight adjusted for determination of telephone status was then calculated for these cases:

$$(39) \qquad W2(hhold_i) = W1(LHU_i) \cdot A''_{nr}(nonphone_c)\text{, if telephone eligibility of household determined}$$

$$W2(hhold_i) = 0\text{, otherwise.}$$

After this adjustment, households with undetermined telephone status and households with telephone service at the time of the interview and no interruption in service of two weeks or more during the 12 months prior to the interview were removed from the weighting process. The

remaining households were those known to be households without telephone service or meeting our telephone interruption criterion.

d. Survey Eligibility Nonresponse Adjustment to Household Weight in the Field Sample

After adjusting for undetermined telephone status, we used a weighting cell adjustment to account for households that: (1) did not complete the survey enumeration questions to determine whether there was at least one civilian adult in the household, and (2) eligible households completing the enumeration questions that did not complete the survey.[22] For Round Three households with incomplete enumeration questions, we created the following household eligibility nonresponse adjustment factor:

$$(40) \qquad A'''_{nr}(hhold_c) = \frac{\sum_{nonphnhh \in c} W2(hhold_i)}{\sum_{det\,nonphn\,hh \in c} W2(hhold_i)},$$

where the numerator is summed over all nonphone households in cell (site) *c*, and the denominator is summed over nonphone households in cell *c* with known survey eligibility status. A household weight adjusted for determination of survey eligibility was then calculated for these cases:

$$(41) \qquad W3(hhold_i) = W2(hhold_i) \cdot A'''_{nr}(hhold_c)\text{, if survey eligibility of household determined}$$

$$W3(hhold_i) = 0\text{, otherwise.}$$

After this adjustment, households with undetermined eligibility status and households known to be ineligible for the survey were removed from the weighting process.

[22]See Section F.1 in Chapter II for the definition of eligible and ineligible households and Section II.B.4 for a definition of individuals excluded from the survey.

e. Interview Nonresponse Adjustment to Household Weight in Field Sample

We then adjusted these weights for survey nonresponse among eligible households. A responding household was one in which at least one eligible FIU responded to the survey.

We performed a weighting class adjustment for households using the same cells as defined for the household eligibility adjustment (see previous section). We created a household survey nonresponse adjustment factor as follows:

$$(42) \qquad A''''_{nr}(survey_c) = \frac{\sum_{elig\ hh \in c} W3(hhold_i)}{\sum_{resp\ hh \in c} W3(hhold_i)},$$

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

$$(43) \qquad W4(hhold_i) = W3(hhold_i) \cdot A''''_{nr}(survey_c), \text{ if the household responded}$$

$$W4(hhold_i) = 0, \text{ otherwise.}$$

f. Poststratification and Other Adjustments to Household Weight in Field Sample

Next, we poststratified the sum of the weights to estimated population totals. We used estimates of the number of nontelephone households in large MSAs from the March 2000 Supplement to the Current Population Survey (U.S. Census Bureau, http://www.bls.census.gov/cps/cpsmain.htm). For site-specific counts of nontelephone households in each of he twelve high-intensity sites, we used a combination of estimates from the March 2000 Supplement to the Current Population Survey, the Census 2000 Supplemental

Survey, and the Census 2000 Short Form data. The poststratification adjustment factor for non-telephone households is:

(44) $$A_{ps\text{-}nontel}(\text{large metro areas}) = \frac{NONTELHH}{\sum_{resp\ hh_i} W4(hhold_i)}$$ for national estimates, and

(45) $$A_{ps\text{-}nontel}(site) = \frac{NONTELHH_{site}}{\sum_{resp\ hh_i \in site} W4(hhold_i)}$$ for site-specific estimates,

where *NONTELHH* is the estimated number of non-telephone households in large metro areas of the U.S. in 2000. The household-level weight poststratified to non-telephone households is:

(46) $$WT_{nontel}(hhold_i) = W4(hhold_i) \cdot A_{ps\text{-}nontel} .$$

g. Nonresponse Weight Adjustment for FIUs in Field Sample

As with the RDD weighting adjustments, we developed the FIU weight for the field sample from the final household weight. The FIU weights accounted for secondary FIU nonresponse to the survey within responding households.[23] All FIUs in responding households were assumed to have known eligibility status.

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

[23]Secondary FIU nonresponse was less of an issue in the field sample than in the RDD sample because the household informant was allowed to respond on behalf of other FIUs in the household.

$$(47) \qquad A_{nr}(survey_c) = \frac{\sum_{elig\ fiu \in c} WT_{nontel}(hhold_i)}{\sum_{resp\ fiu \in c} WT_{nontel}(hhold_i)},$$

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

$$(48) \qquad W5(FIU_i) = WT_{nontel}(hhold_i) \cdot A_{nr}(survey_c), \text{ if FIU responded}$$
$$W5(FIU_i) = 0, \text{ otherwise.}$$

i. Initial Person Weight

We used the final FIU weight to develop the person weight for adults in the field sample. However, because only one child was selected at random per FIU, the within-FIU probability of selection for a child was equal to the inverse of the number of children in the FIU. The overall probability of selection for person *k* in FIU *j* in household *i* can be expressed as:

$$(49) \qquad P(person_{ijk}) = \frac{P(FIU_{ij})}{(\delta \cdot numkids_{ij}) + (1-\delta)},$$

where $numkids_{ij}$ is the number of children in FIU_{ij}, and δ is equal to zero for adults and is equal to one for children.

The initial person-level weight for all people was calculated as follows:

$$(50) \qquad W6(person_{ijk}) = W5(FIU_i) \cdot [(\delta \cdot numkids_{ij}) + (1-\delta)].$$

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

j. Nonresponse Adjustment to Person Weight

Using the same editing program and rule described previously for the RDD sample, there were no person records in the field sample with high levels of missing information. Therefore, there was no person-level nonresponse adjustment.

D. INTEGRATED WEIGHTS FOR THE HOUSEHOLD SURVEY

The integrated weights combined the field and RDD survey data from the site-based sample to make national and site-specific estimates. For areas represented by both the RDD and field components, the integrated weights accounted for the likelihood of being chosen in each of the two components. For areas not represented by the field component, the RDD survey data alone were weighted up to represent all households and individuals in those households, 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):[24]

1. Poststratify the RDD and field components to our best estimates of the telephone and nontelephone populations, respectively (household level nonresponse-adjusted weights)
2. Create household telephone service interruption adjustment factors (IAFs) for both components (see Section D.1)
3. Apply IAFs to the weights for the two household components
4. Combine the RDD and field telephone components into one data file
5. Poststratify the combined RDD and field components again at the household level (recalibrate)
6. Apply the recalibrated IAFs to the weights for the two FIU components
7. Apply the recalibrated IAFs to the weights for the two person-level components

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

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

For RDD households in the 48 large MSAs, we began with the weights that represented the telephone portion of the population *(WT_{tel})*. For the field households, we began with the weight that represented the nontelephone portion of the population. Large MSA households in the RDD component that had intermittent telephone service and households in the field component that had any telephone service during the year preceding the survey were adjusted for dual selection probabilities (they had a chance of being selected into both the RDD and field components), while accounting for the length of interruption. (This adjustment is described in more detail below.) Table V.2 illustrates how the RDD and field components were combined for national estimates. Note that these steps were carried out twice for the national weights: once based on the site sample only, and once based on the augmented site sample.

TABLE V.2

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

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

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

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

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 factor complicating the combination of the RDD and field samples was the inclusion of households with interrupted telephone service during the preceding year in both sample components. The integrated weights assumed that (1) households with no interruption in service could have been sampled only for the telephone survey; (2) those with no telephone service could have been sampled only for the field survey; and (3) the remainder could have been sampled for both surveys. For the RDD site sample, 3.3 percent of households completing interviews had an interruption in telephone service of two weeks or longer during the year preceding the survey, with most of these households in areas eligible for the field component. For the field sample, 70.6 percent of households were without telephone service for at least two weeks during the year preceding the survey and could have been sampled for the RDD survey.

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

TABLE V.3

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

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

Instead, we constructed integrated weights that synthetically accounted for multiplicity by using a weighting adjustment that we termed the *telephone interruption adjustment factor* (the IAF). This factor accounted for both length of telephone interruption and multiplicity and was applied only to households in the "integration sites" (that is, sites represented by both the RDD and field components). For national estimates, integration sites included all 48 large MSA sites. For site-specific estimates, they included the 12 high-intensity sites only. For the RDD component, households with no telephone interruption would have been ineligible for the field component and so had an IAF set equal to one. For the field component, households with no telephone service would have had no chance of selection into the RDD component and also had an IAF equal to one. For households in the field component with some telephone availability and for households in the RDD component with some telephone interruption, we multiplied the value of IAF by the households' weights poststratified to the populations represented by their components (telephone or nontelephone). We calculated IAF_m as:

$$(51)\quad IAF_m = \frac{1/RelP_m}{1/MEDIAN(RelP)} . k \qquad m = (1,2,...,12),$$

where:

$$(52) \quad RelP_m = [PRatio \cdot \frac{(12 - m)}{12}] + 1,$$

and

$$(53) \quad PRatio = \frac{(\text{unwgted hholds in RDD sample / telephone hholds in population})}{(\text{unwgted hholds in field sample / nontelephone hholds in population})},$$

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 households with an interruption in telephone service remained the same; $RelP_m$ is the relative combined likelihood of selection into either component, estimated on the basis of the number of months with telephone service;[25] and $PRatio$ is the probability of selection into the RDD component, relative to selection into the field component; and the "population" refers to either large metropolitan areas in the U.S. or to a high-intensity site.

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

$$(54) \quad WTINT_m = WT_{tel} \cdot IAF_m, \text{ for RDD households in integration sites}$$

$$(55) \quad WTINT_m = WT_{nontel} \cdot IAF_m, \text{ for field households}$$

$$(56) \quad WTINT_m = WT_{all}, \text{ for RDD households outside of integration sites,}$$

[25]In equation (52), the first term (in square brackets) represents the likelihood of selection into the RDD component, and the second term (the number 1) reflects the likelihood of selection into the field component.

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

2. Poststratification of Person-Level Integrated Weights

For national estimates, person-level weights were poststratified by sex and age group, then by sex and whether or not Hispanic, then by sex and race (black or non-black), and then by level of education.[26] For high-intensity sites, site-specific weights were poststratified by sex and age group, then by whether or not Hispanic, then by race (whether or not black), and the estimated site population.[27] Weights for low-intensity site-specific estimates were poststratified to site totals only. After person-level weights were trimmed, weights were poststratified again by the same demographic variables, as well as by the distribution of telephone and nontelephone households prior to trimming (discussed below). The re-poststratification was done within site for site-specific weights.

E. TRIMMING PERSON WEIGHTS

In analyses of survey data, even a few extremely large weights can reduce the accuracy of point estimates and inflate the sampling variance. To reduce the sampling variance, excessively large weights are trimmed, and the amount trimmed is distributed among the untrimmed weights to preserve the original sum of the weights. However, trimming of sampling weights can introduce bias into some point estimates. The objective in trimming weights is to reduce the impact of excessively large weights while minimizing the introduction of bias.

---

[26]Age, sex, Hispanic, and race distributions and totals were from the 2000 Census. The education distribution was from the March 2000 Current Population Survey (excluding Alaska and Hawaii).

[27]Age, sex, race, ethnicity, and total population, by site, were based on figures from the 2000 Census.

For site-specific and national estimates, we trimmed the person-level and family-level integrated weights and then assessed the effect of the trimming. We evaluated the extent of trimming and the inflation factor for the untrimmed weights necessary to preserve the original sum of the weights and then estimated the effect of the trimming on the sampling variance. We used a weight-trimming algorithm that compares each weight with the square root of the average value of the squared weight used to identify the trimming cutpoint and the weights to be trimmed. 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. Unequal weighting (a result of differential selection rates and response rates) has the potential to decrease precision because variation in the weights affects the variance of weighted estimates. Person-level weights were trimmed to reduce this design effect; however, the extent of trimming was limited to minimize the risk of introducing bias into the sample estimates.

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

$$(57) \qquad T_c = (k \sum_{i \epsilon c} WT_i^2 \; / \; n_c)^{1/2},$$

where n_c is the number of observations in the trimming class, k is an arbitrary number (generally assigned a value of 10), and the summation is over the observations in the trimming class. Any weight exceeding the cut-off point, T_c, is assigned the value of T_c, and excess is distributed

among the untrimmed weights, thereby ensuring that the sum of the weights after trimming is the same as the sum of the weights before trimming.

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

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

The weights designed to produce site-specific estimates were evaluated separately for adults and children in each high-intensity site. Because only one child was randomly selected in each FIU and the sample size of children was smaller than that of adults, weights for children had greater variation and were larger on average than for adults. The weights for trimming were identified by using the NAEP procedure, as well as by visual inspection of outlier weights 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.

We used a similar method to trim the weights designed to produce national estimates by using the NAEP procedure and assessing the impact of the trimming on the design effect from unequal weights. For the site sample, the weight-trimming classes were defined by the three site-selection strata (large MSAs, small MSAs, and non-MSAs), geographic region (four

regions), and adult versus child. For the supplemental sample, the weight-trimming classes were defined by the sample strata (metropolitan areas in each of four geographic regions and the nonmetropolitan areas of the United States) and adult versus child.

FIU-level weights for site-specific and national estimates were also trimmed. We used the same trimming classes and procedures as were used for the two groups (adults and children) of person-level weights. Fewer than 0.5 percent of the cases were trimmed for any of the groups of weights discussed above.

F. WEIGHTS FOR COMBINING THE SITE AND SUPPLEMENTAL SAMPLE SURVEYS

The goal of the supplemental sample is to increase the precision available from the site sample by combining the two samples. To simplify the combined sample analyses, we explored procedures to determine whether a single combined sample weight (or a set of combined sample weights) could be constructed that would achieve variance estimates near to the minimum variance. The following sections describe the procedure to achieve minimum variance estimates from the combined samples, and the results for computing the combined sample weights.

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

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

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

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

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 λ will result in an unbiased estimate of the survey estimate, but not necessarily an estimate with the minimum sampling variance. The value associated with minimum variance, λ, can be computed as:

$$(60) \quad \lambda = [1/V(Y(Site))] / [1 / V(Y(Site)) + 1/ V(Y(Supp))]$$
$$= V(Y(Supp)) / [V(Y(Site)) + V(Y(Supp))].$$

In this case, the minimum variance is:

$$(61) \quad V(Y) = [V(Y(Site)) \bullet V(Y(Supp))] / [V(Y(Site)) + V(Y(Supp))].$$

To compute the combined sample estimate with minimum variance, a survey estimate is derived by first computing the estimate for each survey component, and then computing a value of λ using the estimated variance from each survey component. The combined sample point estimate is computed using the point estimate from each survey component and this value of λ (as in equation (58)). The sampling variance is estimated using the sampling variance estimate from each component survey and the computed value of λ (as in equation (59)). Although this process produces the minimum variance estimates, it is computer intensive. In addition, because of differing values of λ among levels of a categorical variable, it results in some inconsistencies among estimates of percentages and proportions. For example, proportional distributions, such as the proportion of the population by insurance type, sometimes did not sum to 100 percent because the component proportions had different values of λ. In addition, this two-step process for computing estimates would likely pose analytic problems for regression analyses and more

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

The concept was that a value (or values) of λ was needed that would result in the best estimate and smallest variance for a variety of analysis variables and key populations. Because any value would result in an unbiased estimate, the key statistic for the analysis was the change in the sampling variance relative to the minimum variance. We also evaluated the change in the survey estimate relative to the survey estimate with minimum variance. For that analysis, we identified 15 analysis variables (10 categorical and 5 continuous) and nine populations (the full population and eight subpopulations). For dichotomous variables (for example, a yes/no variable), the sampling variances for both response options were equal and therefore redundant.

In Round Three, the mean value of the λs was 0.829, with a median of 0.857; the distribution of the λs was slightly skewed. The value of λ was affected by design effects in the site sample (that is, by the average number of people in a site) and by the correlation among responses within a site (that is, the intracluster correlation). As expected, because of the number of persons in each site, the mean of the λs for estimates for the full population was the lowest (mean, 0.761; median, 0.763). For three key subpopulations (children, blacks, and Hispanics), the mean value of λ was between 0.82 and 0.85, and the mean of the median values for the three subpopulations was 0.865. The mean of the median values (0.865) was used as the λ for combining the weights for three reasons. First, it was close to the median value for all λs (0.857). Second, the sample sizes of the three subpopulations were relatively small, and it was desirable to minimize the variance estimates for point estimates for these subpopulations. Third, the optimal λ for the full population would result in less-than-optimal variances for

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

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

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

and for individuals in the supplemental sample as:

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

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

G. SAMPLING ERROR ESTIMATION

1. Background

The CTS Household Survey sample design is complex and therefore requires specialized techniques for estimation of sampling variances. Procedures in standard statistical packages, such as SAS and SPSS, compute variances using formulas under the assumption that the data are from a simple random sample from an infinite population. Although the simple random sample variance may approximate the sampling variance in some surveys it is likely to substantially underestimate the sampling variance with a design as complex as the CTS Household Survey. 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:

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

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

For the CTS Household Survey, the sampling variance estimate, called the *design-based sampling variance,* is a function of the sampling design and the population parameter being estimated. The design-based variance assumes the use of fully adjusted sampling weights, which are derived from the sampling design, with adjustments to compensate for nonresponse and for ratio-adjusting the sampling totals to external totals (for example, to data on population totals by age and race/ethnicity generated by the Bureau of the Census from the Current Population Survey).

The data files for the CTS Household Survey contain a set of fully adjusted sampling weights and information on analysis parameters (that is, stratification and analysis clusters) necessary for the estimation of the sampling variance for a statistic. Because of the stratification and unequal sampling rates, it was necessary to account for the sampling weights and the sampling design features in order to compute unbiased estimates of population parameters and their associated sampling variances. The estimation of the sampling variance required the use of

special survey data analysis software or specially developed programs designed to accommodate the population parameter being estimated and the sampling design.

Survey estimators fall into two general classes: (1) linear estimators, and (2) nonlinear estimators. Linear estimators are weighted totals of the individuals with an attribute, or means and proportions, if the denominators are known (for example, when the denominator is a poststratum total or a sum of poststrata totals). Nonlinear estimators include proportions and means (when the denominators are unknown and are estimated from the survey), ratios, and correlation and regression coefficients. In general, the variances of nonlinear statistics cannot be expressed in a closed form. Woodruff (1971) suggested a procedure in which a nonlinear estimator is linearized by a Taylor series approximation. The sampling variance equation is then used on this linear form (called a *linearized variate*) to produce a variance approximation for the original nonlinear estimator.

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

Other software packages use the Taylor series approximations (for example, Stata and SAS SurveySelect), 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. Alternatives to using SUDAAN are discussed in Technical Publication Number 40 (May, 2003).

2. Variance Estimation

The CTS Household Survey contains a series of weights that are designed for site-specific and national estimates. The site-specific weights are designed for estimates that include units (either FIUs or individuals) from the site sample and units selected in the supplemental sample that were within the boundary of a site. The weights available for national estimates include the national site sample weights, the supplemental weights, the augmented site sample weights, and the combined weights that incorporate the site and supplemental samples. All four national weights were poststratified to the same population totals to ensure comparability; however, the four national samples may not produce precisely the same point estimates. The following discussion provides the variance estimation protocols for each of these weights. (The forthcoming Household Survey User's Guide will provide instructions for deriving appropriate variance estimate for different samples.)

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

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

b. Weights for National Estimates Based on the Site Sample

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

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

For the MSA sites, the samples were selected without replacement and with unequal probability. To account for the finite population correction, we computed the probability of

selection of any pair of selected sites jointly into the sample. These joint inclusion probabilities and a site's probability of selection were used to compute the finite population correction factor using the Yates-Grundy-Sen variance estimation equation (Wolter 1985). The SUDAAN software package permits direct variance estimates based on this equation.

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

- Nine analysis strata, one corresponding to each of the nine sites selected with certainty
- Nine analysis strata formed among the 39 noncertainty sites in the stratum of large MSAs (to facilitate the computation of the joint selection probabilities)
- One stratum for small MSAs
- One stratum for nonmetropolitan areas

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

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

c. Weights for National Estimates Based on the Supplemental Sample

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

d. Weights for National Estimates Based on the Combined Sample

The maximum precision for national survey estimates is obtained by combining the site sample and the supplemental sample. For computing survey estimates, combined across the two sample components, *Est(Y)*, separate estimates can be computed for each sample component and combined using equation (58). The sampling variance of this estimate, *V(Y)*, is computed using equation (59). Section G of this chapter describes the value of λ we derived to simplify processing without substantial loss in precision. The combined weights incorporated this value.

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

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

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

REFERENCES

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

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

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

Brick, J.M., J. Montaquila, and F. Scheuren. "Estimating Residency Rates for Undetermined Numbers in RDD." *Public Opinion Quarterly*, Vol.66, Spring 2002, No. 1, pp. 18-40.

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

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

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

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

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

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

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

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

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

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

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

Loosvelt, Geert, J. Pickery and J. Billiet, *Item Nonresponse as a Predictor of Unit Nonresponse in a Panel Survey*, Journal of Official Statistics, Vol 18, No,. 4, 2002, pp. 545-557

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

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

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

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

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

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

Rand Health, *Consumer and Health Quality Information: Need, Availability, Utility*. California Care Foundation, October 2001.

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

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

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

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

Singer, E., J Van Howeyk, N Gebler, T Raghunathan, and K McGonagle, *The Effect of Incentives on Response Rates in Interviewer-Mediated Surveys*, Journal of Official Statistics, Vol. 15, No. 2, 217-230, 1999.

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

Strouse, R., J.W. Hall, F. Potter, B.L. Carlson, P. Cunningham, J. Pascale, and R. Stapulonis. "Report on Survey Methods for the Community Tracking Study's 1996-1997 Round One Household Survey—Final." Report submitted to the Center for Studying Health System Change. Princeton, NJ: Mathematica Policy Research, Inc., November 1998 (HSC Technical Publication #15).

Strouse, R., B. Carlson, and J. Hall. "Report on Survey Methods In the Community Tracking Study's 1998-1999 Round Two Household Survey-Final." Report submitted to the Center for Studying Health System Change. Princeton, N.J. Mathematica Policy Research Inc., November 2001 (HSC Technical Publication #34).

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

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

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

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

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

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

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

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