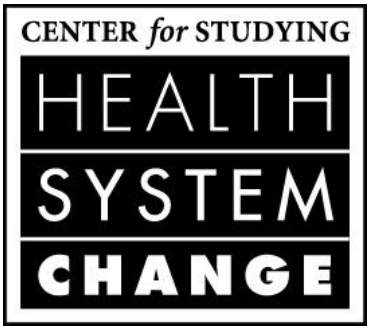
Followback Survey Methodology Report

(Round Two)



600 Maryland Ave., SW
Suite 550
Washington, D.C. 20024

Technical Publication No.

35

September 2002

MPR Reference No.: 8553

Community Tracking Study Health Insurance Followback Survey Technical Report on Data Collection Operations and Final Data Files

Final Report

September 2002

Michael Sinclair, Senior Statistician
Richard Strouse, Vice President
David Uglow, Associate Director Systems and Programming
Barbara Carlson, Senior Statistician
Kathy Sonnenfeld, Survey Director

Submitted to:

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

Project Officer: Peter Cunningham

Submitted by:

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

Project Directors: Richard Strouse,
Michael Sinclair

CONTENTS

CONTENTS *(continued)*

Chapter **Page**

I. OVERVIEW OF DATA COLLECTION PROCEDURES FOR ROUND 2 OF THE FOLLOWBACK SURVEY

A. OBJECTIVES OF THE COMMUNITY TRACKING STUDY

The Community Tracking Study (CTS), which is funded by the Robert Wood Johnson Foundation (RWJF), is designed to provide a sound information base for decision making by health care leaders. It does so by collecting information on how the health system is evolving in 60 nationally representative communities across the United States and on the effects of those changes on people. The CTS, which has been under way since 1996, is a longitudinal project that relies on periodic site visits and surveys of households, physicians, and employers. The CTS addresses two broad questions that are important to public and private health decision makers:

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

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

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

The CTS examines 12 of the 60 communities in depth by conducting site visits and using survey samples large enough to draw conclusions about change in each community. The 12

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

communities comprise a randomly selected subset of sites that are metropolitan areas with more than 200,000 people (as of July 1992). We refer to them as *high-intensity sites*.

B. ANALYTIC COMPONENTS OF THE COMMUNITY TRACKING STUDY

The CTS has qualitative and quantitative components. The qualitative component consists of case studies in the 12 high-intensity sites, which are conducted every two years. Survey data from the 12 high-intensity sites and from 48 additional sites, listed in Table I.1, complement this information.

The CTS also includes independent surveys of households, physicians, and employers in all 60 sites, thereby enabling researchers to explore relationships among purchasers, providers, and consumers of health care.[2] An Insurance Followback Survey, which is linked to the Household Survey, was conducted for the first two rounds of the survey. In this survey, the privately financed health insurance policies covering respondents to the survey of households are "followed back" to the organization that administers the policy. The purpose of the Insurance Followback Survey is to obtain information about the private policies that is more detailed and more accurate than Household Survey respondents are able to provide. For the first three rounds of the study, data were collected on a two-year cycle; however, surveys will be conducted every three years beginning with round 4 in 2003. The round 1 surveys of households and physicians, completed during 1996 and 1997, and the Insurance Followback Survey, completed in 1997 and 1998, are the baseline. Data collection for the round 2 surveys of households and physicians began in 1998 and was completed in 1999. Round 2 Insurance Followback Survey data collection was conducted during 1998 and 2000. The round 3 household and physician surveys

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

TABLE I.1

SITES SELECTED FOR THE COMMUNITY TRACKING STUDY

PSU	Site Label	Site State
1	Boston, MA	MA
2	Cleveland-Lorraine-Elyria, OH	OH
3	Greenville-Spartanburg-Anderson, SC	SC
4	Indianapolis, IN	IN
5	Lansing-East Lansing, MI	MI
6	Little Rock-North Little Rock, AR	AR
7	Miami, FL	FL
8	Newark, NJ	NJ
9	Orange County, CA	CA
10	Phoenix-Mesa, AZ	AZ
11	Seattle-Bellevue-Everett, WA	WA
12	Syracuse, NY	NY
13	Atlanta, GA	GA
14	Augusta-Aiken, GA-SC	GA-SC
15	Baltimore, MD	MD
16	Bridgeport-Danbury-Stamford, CT	CT
17	Chicago-Kenosha-Kankakee, IL-WI	IL-WI
18	Columbus, OH	OH
19	Denver-Boulder-Greeley, CO	CO
20	Detroit, MI	MI
21	Greensboro-Winston Salem-High Point, NC	NC
22	Houston-Galveston-Brazoria, TX	TX
23	Huntington-Ashland, WV-KY-OH	WV-KY-OH
24	Killeen-Temple, TX	TX
25	Knoxville, TN	TN
26	Las Vegas, NV-AZ	NV-AZ
27	Los Angeles-Long Beach, CA	CA
28	Middlesex-Trenton, NJ	NJ
29	Milwaukee-Racine, WI	WI
30	Minneapolis-St. Paul, MN-WI	MN-WI
31	Modesto, CA	CA
32	Nassau-Suffolk, NY	NY
33	New York City, NY	NY
34	Philadelphia, PA-NJ	PA-NJ
35	Pittsburgh, PA	PA
36	Portland-Salem, OR-WA	OR-WA
37	Riverside-San Bernadino, CA	CA
38	Rochester, NY	NY
39	San Antonio, TX	TX
40	San Francisco, CA	CA
41	Santa Rosa, CA	CA
42	Shreveport-Bossier City, LA	LA
43	St. Louis, MO-IL	MO-IL
44	Tampa-St. Petersburg-Clearwater, FL	FL
45	Tulsa, OK	OK
46	Washington-Hagerstown, DC-MD-VA-WV	DC-MD-VA-WV
47	West Palm Beach-Boca Raton, FL	FL
48	Worcester-Fitchburg, MA	A
49	Dothan, AL	L
50	TerreHaute, IN	N
51	Wilmington, NC	C
52	West Central Alabama, AL	L
53	Central Arkansas (excluding Little Rock), AK	R
54	Northern Georgia (excluding Atlanta), GA	A
55	Northern Illinois (excluding metro Chicago), IL	L
56	Northeast Indiana, IN	N
57	Northern and Central Maine, ME	E
58	East Central North Carolina, NC	C
59	Northern and Eastern Utah (excluding Salt Lake area), UT	T
60	Western Washington (excluding Seattle), WA	A

were conducted during 2000 and 2001; we did not conduct a third round of the Insurance Followback. Documentation of CTS data collection activities is available at www.hschange.org.

C. THE INSURANCE FOLLOWBACK SURVEY

Purpose. This report describes the results of the second round of the Insurance Followback Survey, in which privately financed health insurance policies covering Household Survey respondents are "followed back" to the organization that administers the policy. The Insurance Followback was designed to obtain information on the characteristics of insurance products that Household Survey respondents are unable to provide themselves. Various studies have shown that people have difficulty accurately reporting even very basic attributes of their insurance plans, such as whether they belong to an HMO, are required to sign up with a primary care provider, or require referrals to obtain maximum in-network coverage (Nelson 2000, Cunningham 2001).

Content. For the first two rounds of the CTS, the Insurance Followback Survey obtained basic attributes of insurance plans (e.g., product type, primary care provider sign up requirement, use of referrals to obtain maximum in-network coverage), typical method of payment for primary care, specialty and hospital services for the product, whether an HMO is a for-profit or not, and estimates of physician and hospital network size. (The round 2 survey attempted to obtain more detailed information on in-network coverage than was obtained for round 1.) In addition, the survey attempted to obtain other types of information that employers could more easily have provided for individual contracts (e.g., estimates of copayment amounts and co-insurance rates, and deductibles). Since we were not conducting an employer followback for those rounds, we asked insurers to provide estimates of typical cost sharing arrangements for products at the site level.

Methods. In planning the Insurance Followback, we reviewed related efforts conducted for the National Medical Expenditures Survey (NMES) and its successor, the Medical Expenditures Panel Survey (MEPS). Emmons and Hill (1991) describe the design of the Health Insurance Plans Survey for the 1987 NMES, which verified health insurance status and collected supplementary information about the private health insurance coverage of household survey respondents. Participants in the household survey were asked to provide signed permission forms for employer based or individually purchased plans, and employers or insurers were contacted by a combination of mail, telephone, and personal visits to obtain interviews and copies of insurance booklets, which were abstracted. The 1996 MEPs also included a Health Insurance Plan Abstraction study, but policy booklets were obtained directly from household survey respondents, who were offered a monetary incentive.

We decided against following back to the employer or obtaining policy booklets for several reasons. First, many of the features of health insurance policies we wanted to collect were more likely to be available from insurers than employers. Second, it was not feasible to obtain policy booklets from CTS household survey respondents; unlike the MEPS, which is conducted in person, our household survey is conducted by telephone. Third, a survey of employers was more expensive than a survey of insurers because data collection for the latter could be grouped by plan.

There were two components to Insurance Followback data collection. First, we had to link Household Survey reports of insurance policies to insurance products. Then, we had to obtain information about those products from insurers (or other knowledgeable sources of information, such as third party administrators, self insured employers, or unions). For round 1, Household Survey respondents described the private health insurance polices under which they received health care services. Based on the names of health insurance plans and employers provided by

respondents, we contacted health plans and other organizations. We conducted brief telephone interviews with insurers to obtain organizational information, a list of products offered in local markets, and attributes of those products. Then, we faxed forms with product and employer names to all but the smallest plans (data were obtained entirely by telephone from insurers linked to fewer than five policies). Respondents receiving faxed forms were asked to verify that the listed employers had contracts with the organization during the data collection period and to identify the products linked to the listed plan and employer names. The procedures used in round 1 are described in the HSC technical publication No 30.

We believed that a more fully automated process would increase the percentage of policies linked to insurers and provide more control over information provided by insurers. For round 2, insurance databases and product data obtained in round 1 were used to develop a product file to prompt household survey respondents during their interviews with insurer plans and products offered in their state. If Household Survey respondents could not link their policies with products in the file, we contacted their employers, using the same product file as a memory aide to identify insurance products. Insurers were then contacted by telephone to provide information on organizational characteristics, and product attributes for the products in the database. Round 2 data collection was fully automated using computer assisted telephone interviewing methods. We also used product data available from selected web sites and plan booklets to fill in missing product attributes for a few national insurers.

For round 1, we were able to link 52.5 percent of the eligible private policies reported in the household survey to a unique product. For 19.4 percent, we were able to link to an insurer or other entity, such as an employer, with information about the policy, but not to a specific product. For those policies, we used statistical matching procedures to assign a product to the

policy from among multiple products identified by the insurer. We could not link 28.1 percent of the policies to an insuring entity; they were accounted for in the survey weighting procedures.

For round 2, we were able to link 51.3 percent of the estimated 20,199 eligible policies with a unique product, 9.4 percent to an insurer, but not to a unique product, and 39.3 percent of the policies could not be linked.

We were disappointed that the more automated process did not result in higher linkage rates; the percentage of directly linked policies remained about the same and the percentage of statistically matched polices declined. A key factor explaining our inability to increase the direct linkage rate was that product data obtained from databases and the round 1 followback were often dated and sometimes included duplicate records. Many insurers merged or were acquired by other firms and product names often changed. Second, insurer nonresponse was much higher in round 2 than in round 1. Consequently, we were unable to obtain product attribute data for many linked policies. A third problem was relatively high nonresponse to selected questions. Some insurers had difficulty answering questions for product attributes at the market level because they varied at the contract level and were reluctant to estimate for the market.

We were unable to statistically link as many policies to insurers (soft-matches) in round 2 because we did not have complete data on potential products offered by an insurer within a CTS market for as many insurers as in round 1. In large part, this reflected a difference in methodology. For round 1, we relied on self-reported products and assumed that questionnaires returned by insurers represented a product inventory. For round 2, we did not attempt to obtain an inventory of all products as we assumed the database built from round 1 and round 2 responses was complete. After comparing reported plans to other data sources, we realized this was not the case, and attempted statistical linkages in round 2 only when we could verify that products were listed for all product lines offered by an insurer in a site.

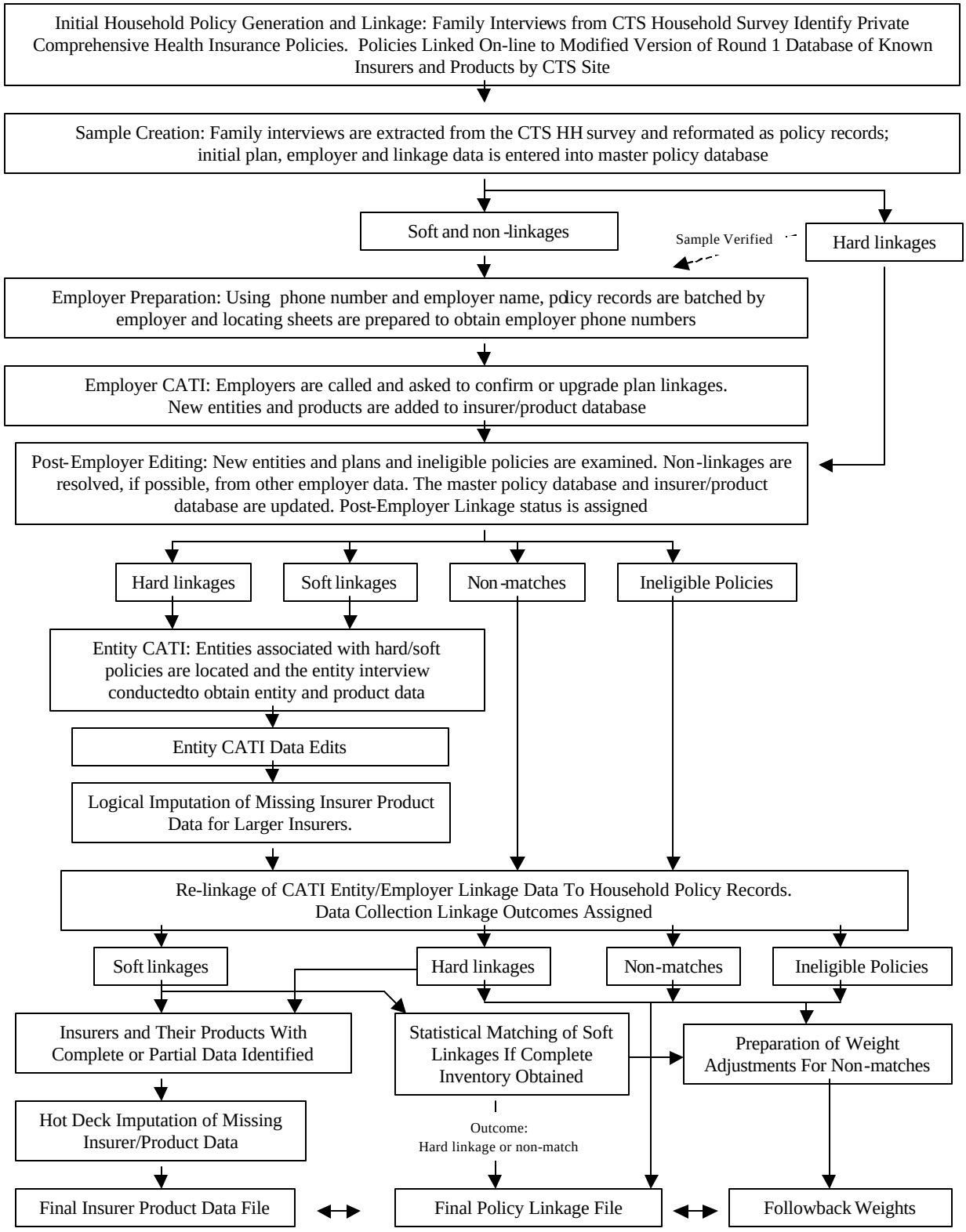
D. ORGANIZATION OF THE ROUND 2 FOLLOWBACK SURVEY

The process used to link plans reported by household survey respondents to appropriate insuring entities and to obtain product attribute data for these policies was complex. An overview of the steps required to collect and process the data for the Insurance Followback Survey is presented in Figure I.1 and summarized below.

- **Initial Household Policy Generation and Linkage**: Family interviews from the CTS Household Survey identified up to three private health insurance *policies* providing comprehensive health care services to family members. When asked to describe their health plans, most survey respondents mentioned insurers, HMOs, PPOs, and other health plan providers, but some described employers, unions, or third party administrators (TPA's). To aid respondent's recall, the CATI program accessed an *insurance product database* with information on the names of insurers and the products they offered in the respondent's state. The database was created from the round 1 insurance product file (a file listing insurers and their products), supplemented in some cases with more current information from web sites.

- **Sample Creation:** We extracted employer insurance plans from household survey families with private health insurance and reformatted the information into *policy-based records* (one record per plan, with up to three plans per family). Policies were then classified into one of three categories: *hard-match* – policy name matched insurer database listing; *soft-match* – insurer reported in the household survey matched a known insurer on the database but the insurance product was unknown or did not match; and *non-match* --all remaining policies. A *master policy database* was created to store policy information on the plan name, employer name, and linkage status to the insurer product database.

FIGURE 1.1 OVERVIEW OF THE INSURANCE FOLLOWBACK PROCEDURES



- **Employer Survey Preparation:** Using the master policy database, we identified employers associated with soft-matched or non-matched policies. We also selected a sample of hard-matched policies to be verified during the employer survey (approximately 15 percent of the hard matched policies). We batched policy records by employer name and site, and prepared locating sheets, which were used to obtain the employer phone numbers. We then used an automated program to re-batch policies by employer. Since some employers were linked to multiple insurance policies, we designated one policy as the "master" CATI case, and the remainder as "subordinate cases; the record ID number of the master case served as the employer identification number. Policy records from employers that were expected to be linked to large numbers of policies were held back until the household survey was completed to avoid calling employers several times.

- **Employer Interview:** A CATI program was created to interview employers. The purpose of the interview was to (1) verify whether the plan name reported by the household survey respondent was offered and was a comprehensive medical plan and (2) to correct any errors in insurer and product nomenclature. If different insurance plans were linked to an employer, we began the interview with the "master" CATI case, and an external program added insurance plans for the other policies.

- **Post-Employer Editing**: After the CATI employer interviews were completed, we conducted a manual review of ineligible policies, policies linked to multiple insurers (i.e., the employer worked with multiple insurers and a unique insurer or insurance plans could not be determined) and non-matches. In some cases, non-matches could be resolved from information provided by other interviewed employers. We also updated the insurance product database with any new products identified during the employer survey. The review process resulted in a post-employer linkage status of hard, soft, non-match, or ineligible policies.

- **Insurer Survey Preparation:** We examined the policy links to insurers after post-employer editing, and recoded policies when 1) the insurer database erroneously showed two or more entities representing a single insurer location, and 2) policies were linked to the wrong location of a large national insurer. Following these edits, we printed contact sheets for insurers and other entities[3] linked to policies; these were used for locating and for managing interviewing.

- **Insurer CATI Survey:** Insurers were contacted to obtain product attributes and background information for all insurance products offered in a site that were linked to policies reported on the household survey. Respondents for insurers were often, although not always, local or regional marketing staff. In many cases a marketing manager could provide data on insurance products offered in several markets and the CATI program was designed to so that several sites could be grouped together.

[3]The overwhelming majority of firms providing information on insurance product attributes were insurance companies that owned HMO, POS, PPO, or indemnity products. However, some were third party administrators, unions, or self insured employers.

- **CATI Editing:** Data were edited for consistency and open-ended responses were re-coded.
- **Logical Imputation of Missing Product Data:** Several insurers were linked to large numbers of policies for which we did not obtain data from the entity interviews. In these cases, we reviewed insurer web sites, plan booklets, and contacted insures' staff in national headquarters to obtain missing product attribute data.
- **Re-Linkage to Household Survey Data:** We re-linked the household survey policy data to the master policy database to resolve errors. In some cases, the data collection process resulted in duplicate products and the creation of new products that had no matching policies. We reconciled these errors to produce an updated master policy database.
- **Data File Preparation:** We identified insurers and products associated with hard or soft-matched policies and prepared a data file containing the product attribute data for these products.
- **Hot deck imputation:** We conducted a weighted sequential hot-deck imputation on the missing product attribute data for a selected set of items to produce a final product-based data analysis file.
- **Statistical Matching:** For the soft-matches, we conducted statistical matching procedures to assign products to policies that were linked to individual entities but could not be matched to a unique insurance product. Matching procedures were only conducted for insurers in sites for which we could verify that a complete list of the products available to the policyholder was obtained.
- **Linkage Data File:** We prepared a linkage data file that provided the final linkage status (hard-matched, soft-matched, non-matched, or ineligible) for all 21,701 policies identified from the household survey. Based on the family identification number and product number, this file links the product file based on entity, site and product identification numbers with the household survey data.
- **Survey Weights:** We prepared survey weights using the original person weights from the household survey and a propensity modeling procedure to adjust for non-matched policies.

We conducted the data collection from December 1998 through October 2000. File preparation, including editing, imputation, and statistical matching, was completed in October 2001.

In the remainder of this report, we describe survey instrumentation, data management procedures, survey operations, data editing and file processing methods, and imputation and

weighting methods. The survey instrument used to contact employers is included in Appendix A, advance letters and refusal conversion letters are shown in Appendix B. The entity instrument used to interview insurers and other organizations about insurance products linked to household survey policies is included in Appendix C. Editing specifications and item nonresponse rates prior to imputations are shown in Appendix D. More detailed discussion of imputation and statistical matching methods, preparation of survey weights and Sudaan specifications, and data imputations specifications are included in Appendices E-G, respectively. Documentation of files created from the Followback Survey—the round 2 weight file, product file, and policy linkage file—are included in Appendices, H, I, and J, respectively.

II. INSTRUMENTATION

This chapter presents an overview of the contents of both the Employer and Entity CATI instruments (see Appendix A for the Employer and Appendix C for the Entity instrument).

A. EMPLOYER INSTRUMENT

The employer instrument was designed to obtain accurate descriptions of insurers and insurance products for plans named by family informants that were not linked to known products by the insurer database during the household interview. In addition, we surveyed a sampled of employers to validate insurers and insurance products that household survey informants were able to link to the database. The survey included an introduction, verification of insurers and insurance plans, and information about any other health insurance plans offered to employees.

Section 1: Introduction. The introduction described the purpose of the call and included scripts to handle incorrect information about the employer, refusals, and employers that do not offer health plans.

Section 2: Verification of Insurers and Insurance Plans. Depending on the results of the effort to link information provided by the family informant to the insurer database, the instrument could follow one of three paths:

- Path 1: Both the product and entity name matched (hard-match)
- Path 2: Entity name matched, but either the product name was not provided or did not match (soft-match)
- Path 3: Neither the product nor entity name provided by the informant matched the insurer database (non-match).

In ***path 1***, both the entity name (such as Blue Cross and Blue Shield) and a product name (such as Blue Care Preferred) matched the insurer database; a sample of these cases was called to validate the information. Either the employer confirms that this information is correct or the employer provides other information (e.g., they don't offer their employees that plan), or corrects the information.

In ***path 2***, the entity name (such as Blues Cross and Blue Shield) was provided and matched, but the FIU was unable to provide a product name. The employer either provided the correct product or no product. If they offered only one product, then this case became a hard-match; otherwise it remained a soft-match.

In ***path 3***, neither the product nor entity name provided by the family informant was matched to the insurer database. The respondent's text response was given to the employer, and the interviewer probed to achieve a hard or soft-match.

If the employer responses resulted in potential matches to multiple entities and/or plans, the instrument recorded information on each. If the employer offered products that did not match those listed by the family informant, the interviewer could code multiple plans using the insurer database and probe for more information regarding products not listed in the database.

The instrument also verifies that each plan is a comprehensive health plan. If it is not, (e.g. Medicare or retirement supplement, military health plan, specialty plan), policies linked to those plans were coded as ineligible.

Section 3: Other Health Plans and Closing. In addition to thanking the employer for completing the survey, the instrument asked if they offer any other health plans to their employees; additional information is gathered about the other health plans.

B. ENTITY INSTRUMENT

The entity survey was a computer-assisted instrument designed to obtain information about product attributes, network size and physician payment arrangments, and organizational information from insurance companies and other entities (unions, self-insured employers, and third party administrators) that were identified on the household survey or employer interview. The structure of the survey is summarized below; advance letters mailed to respondents are shown in Appendix B and the CATI instrument and training manual are included in Appendix C. A summary of the questions and differences between rounds 1 and 2 (in pdf) are shown in the HSC Technical Publication No. 33.

The survey was organized in four sections:

SECTION A: Entity Information

SECTION B: Product Attributes

1. Product type and network model type
2. Availability for individual purchase
3. Out-of-network coverage
4. In-network coverage
5. Requirement to sign up with primary care physician, group of doctors, or clinic
6. Types of providers who can serve as primary care physicians
7. Consumer cost sharing (copayment, coinsurance, and deductible)

SECTION C: Network Size and Physician Payment Arrangements

1. Physician and hospital network size
2. Payment methods for primary care providers, specialists, and hospital services
3. Separate provision or management of mental health and/or substance abuse services

SECTION D: Organizational Information

1. For-profit/non-profit
2. National/multi-state

Section A determined the correct person to answer the questions and identified the type of organization – Blue Cross Blue Shield, licensed insurer or HMO, management care provider organization such as a PPO or IPA (not licensed to sell insurance), a third party administrator, employer union or trust plan administrator, or something else (specified by the interviewer and generally coded back into one of the other categories).

Section B obtained information about the attributes of each product linked to the entity at the site level. The CATI instrument was structured so that a single interview could be conducted with a respondent answering for products offered in several sites. For example, a respondent answering for insurer products offered in five markets would be asked the product attribute questions for each product in the five sites for which she was responsible.

Section C requested information on network size and physician payment arrangements for the products enumerated in Section B.

Section D obtained the organization's tax status, relationship to a national or multi-state organization, name, and location of the parent company. We also obtained a contact who could answer questions about mental health and/or substance benefits to assist a collateral organization interested in these data.

III. DATA MANAGEMENT

A. HOUSEHOLD SURVEY PRODUCT DATABASE DEVELOPMENT AND MAINTENANCE

In planning for the Round 2 CTS Household Study, we believed we could increase the percentage of policies reported by household survey respondents that could be matched to insurance plans if interviewers could access accurate lists of insurance plans and the entities linked to them (insurers, TPAs, unions) and use them as prompts to aide the respondent's recall. Using data from round 1 we constructed an Insurer Database,which could be accessed by interviewers during the round 2 household interview. The database, which included entities and insurance plans from round 1 and the states served by them, were assigned numeric codes and entered into a series of database tables during the interview. The process for accessing entity and plan names in the database during the household interview is described below:

1. The interviewer asked the respondent for up to three private insurance plans.

2. To improve spelling consistency, the program displayed the names of twelve of the larger insurers offering plans in the state on the screen.

3. The interviewer recorded whether a document was used by the respondent to aid recall, and if so, whether it was an insurance card, a claims form, or an insurance policy.

4. Based on a complex matching algorithm, the program searched for entities matching the text string which served the respondent's state.

5. For all entities retrieved, a list of all plans offered nationally by that entity were displayed and read to the respondent.

6. At this point, the respondent could a) confirm one of the choices; b) decide that the name of the insurance company had been given or recorded incorrectly, and the name re-entered for another search as in step 4, or c) verify that the name was indeed recorded correctly, but that none of the plan names offered matched their policy. In that case the respondent was asked if the insurance plan had been

obtained in a state other than their state of residence. If that was the case, the "new" state was recorded, and a new search undertaken as in step 4, using this state.

The screens used by family informants to the household survey to identify health insurance plans are shown below.

>b1a< Are READ NAMES covered by a health insurance plan from (your/any of your/either of your) current or former employers or unions. [CPS]

IF YES: Who is covered?

INTERVIEWER: DO NOT INCLUDE MILITARY COVERAGE.

PROBES:

(1) Do not include plans that only provide extra cash while in the hospital or plans that pay for only one type of service, such as dental care, vision care, nursing home care, or accidents.

(2) Include health insurance plans provided by colleges and universities to students.

CODE ALL THAT APPLY

[fill NAME] 1
[fill NAME] 2
[fill NAME] 3
[fill NAME] 4
[fill NAME] 5
[fill NAME] 6
[fill NAME] 7
[fill NAME] 8

NONE/NO ONE/NO OTHER RESPONSES ..... n
NEED TO DELETE A RESPONSE x

DON'T KNOW ..... d

REFUSED ..... r

===>

PROGRAM ACCESSES INSURER DATABASE MATCHING PROGRAM

>zb211< What is the complete name of [the; the SECOND; the THIRD] plan?

PROBE:IF R. HAS DIFFICULTY RECALLING NAME, ASK: Do you have an insurance card or something else with the (first) plan name on it?

DISPLAY: Read-Only List Of 12 entity names

DON'T KNOW
[fill "this plan" in subsequent questions] ..... d [GO TO b231]

REFUSED
[fill "this plan" in subsequent questions] ..... r [GO TO b231]
===> GO TO zb221

>zb221< INTERVIEWER: CODE WHETHER DOCUMENT USED. [NO ERASE]

INSURANCE CARD	1
CLAIMS FORM	2
INSURANCE POLICY	3
NO DOCUMENT USED	0
INSURANCE COMPANY NAME INCORRECT, BACKUP AND CORRECT	9

===>

If the respondent could identify both an entity and plan in the Insurer Database, the policy was considered to be a "hard-match." If only the entity could be matched, it was considered to be a "soft-match." If neither entity nor plan existed on the Insurer Database, it was considered to be "non-match," even if the recorded text answer of the respondent clearly identified an entity and plan. Where the respondents said they could not identify an entity, or refused to do so, the appropriate codes were stored in the CATI database.

B. EMPLOYER SURVEY DATA PREPARATION

The next step was to build a larger database structure, known as the Master Policy Database (MPDB), around the Insurer Database. The first elements of this database, which uploaded data from the CTS Round 2 Household Study, consisted of several steps:

1. The household survey data for each family insurance unit (FIU)[4] was reformatted to create a single record of data for each of up to three private insurance policies per FIU.

2. Sequential record identifiers, or plan Ids ("PLIDs") were assigned to each of these records.

3. If the policy represented an employer- or union-issued plan (where Household Study variable PB25=1), the record was uploaded to MPDB.

4. At the time of upload, the text strings representing the entity and plan name and the employer name were edited to improve comparability and usability, and scored as to whether either were usable for further processing.

5. A code was assigned to the variable SMPS (sample status) determining whether the record represented a "hard-match," "soft-match" or "non-match" according to the criteria described above.

Since we planned to begin round 2 insurance followback data collection before the round 2 household survey was completed, the data output from the household to the followback survey could not undergo comprehensive data cleaning. Data needed for the followback survey were reviewed and entity/plan and employer text strings were manually examined and corrected, if necessary.

[4]Individuals in surveyed households 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 and other health measures. The FIU includes an adult member, his or her spouse, if any, any dependent children up to 17 years, or 18 to 22 years of age if a full time student. 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.

Entities named in text string responses for the "no match" policies – those cases where the respondent's answer did not match at least an entity in the original Insurer Database-- were not included in the database because there was insufficient information at this stage even for a soft-match.

The next step in database processing was preparation for the Employer CATI Survey. A Employer Coder module displayed records sorted by employer within site. The coder chose a single policy record linked to an employer , and then all other policy records that were a "close match" were displayed to the coder. The coder would decide which records shared the same employer and code them into an employer batch, with one record arbitrarily designated the "master" record and the remainder "subordinate" records. Subordinate records were linked to each of their related master records in the MPDB("Master_PLID"); this variable served as an identifier for employers within sites. Records whose employer text strings did not produce any "close matches" with other records, or were not batched by a coder with other records, were classified as "uniques."

This process produced a nearly unduplicated universe of employers. Locaters then used a "Phone Log" module to obtain a telephone number for each employer from directory assistance and enter the number into the database. This provided a second check on duplicate employers yet to be batched, and also allowed for additional records generated by the household survey to be batched into employer groups. These records also were printed on contact sheets for use by the Employer CATI Survey interviewers.

The MPDB was used to output the sample for the Employer CATI Survey. As new records became available from the household survey, they were processed and uploaded to the database. The resulting policy records were screened such that only records with usable employer and plan text strings were output to the Employer CATI Survey. Records were transferred from the

MPDB to the Employer CATI Survey via file transfer. At the same time, an interviewer contact sheet was printed for each master or unique record, which included the name, city, county, state, telephone number and time zone of the employer, the record ID, and match status (SMPS) of the case for which the sheet was printed, and grids to record call histories.

Records from large employers were held back to minimize the likelihood of multiple employer calls. When household survey data collection was completed, these records were loaded to the Employer CATI sample and another MPDB coding operation was performed. Coders viewed the policy batches for a large employer within a site on-line and grouped records that had similar entity and plan responses. Only one record per employer was uploaded to the Employer CATI Survey.

The Employer CATI Survey included records for all employers linked to an entity or plan that was not hard-matched to the insurance database. The survey also included a sample (14.5 percent) of hard-matched records to verify the accuracy of the information provided by the household survey informant. No calls were made to employers if the household survey respondent did not identify an entity or insurer on the household survey.

The MPDB supported data collection by accessing information on all plans linked to an employer and by providing employer survey respondents with the same health insurance plan lists shown to household survey respondents. This enabled the employer to upgrade soft-matches to hard-matches. For example, the household survey informant may have indicated that he or she had health insurance from Blue Cross/Blue Shield but didn't know the name of the product. The employer could have identified the product or products offered to employees, thereby creating a hard-match to the insurance database or reducing the number of products that would have to be statistically matched. Upgrades were uploaded to the MPDB.

C. ENTITY SURVEY DATA PREPARATION

The Employer CATI Survey data were then loaded into an MS Access Coding Database for further editing and cleaning. This process included the following steps:

1. Patterns of responses were examined to differentiate: (a) cases with either a hard-match status (coded to reflect a single entity and plan) or an unambiguous soft-match status (coded to reflect a single entity but where the plan could not be determined); (b) cases where the interview indicated that the plan reported by the family respondent was not a valid health plan (such as a dental plan, etc.), and (c) cases that required further editing.

2. Cases in categories (a) and (b) were uploaded to the MPDB, and the match status of the cases (LINKSTAT) updated.

3. Cases in category (c) were referred to a coder, who used the MS Access Coding Database which was linked to the MPDB to perform the following actions:

    a. If the employer reported entities and plans which were not already in the Insurer Database, new numeric codes were assigned for each and the policy record was appropriately coded.

    b. If the interviewer recorded either multiple possible entities or multiple possible plans under which the policy could be coded, the case was resolved by the coder or a match status code (LINKSTAT) was assigned to reflect multiple possible links.

After editing and coding, all remaining records were uploaded to the MPDB. The next step involved editing the MPDB records to prepare for the Entity CATI Survey. We examined links to insurers after post-employer editing, and recoded policies when 1) the insurer database erroneously showed two or more entities representing what was in reality a single insurer, and 2) policies were coded to the wrong site of a large national insurer. The entity and product codes were updated for edited policy records.

These edits of the MPDB and the Insurer Database produced the entity universe with which we began the Entity CATI Survey. Contact sheets were produced for each of the entities which continued to show links to Household Study policies. These sheets were used first for updating

or locating entity telephone numbers, and then for scheduling the Entity CATI Survey interviews.

The MPDB was used to output the initial sample for the Entity CATI interviews. It also served as an on-line resource to structure these interviews. The Entity CATI interview was organized as a hierarchal data collection instrument: for each entity, information was collected for plans within sites. As each Entity CATI interview was released for interviewing, the instrument accessed the MPDB to determine which sites and plans should be included in the interview based on links to household survey policies.

As the interviews proceeded, the Insurer Database was manually updated to incorporate new entity and plan names and states served by an entity; policy links in the MPDB were updated to reflect these changes. If information from the Entity CATI interview resulted in a new entity being added to the Insurer Database, a contact sheet was printed for that entity, contacts at that entity were identified, and it was added to the Entity CATI sample.

At the conclusion of the Entity CATI interviews, the data were cleaned to remove extraneous responses (responses that were corrected by interviewers during the interview) and transferred to SAS cleaning and editing, as described in the next section.

D. ENTITY SURVEY DATA CLEANING

We edited the CATI survey data using guidelines included in Appendix D. These procedures were designed to achieve the following objectives:

- To ensure that that various reported values meet specified norms.
- To translate CATI program codes into standard analytical values and to reformat variables as needed for analysis.
- To review verbatim responses for possible coding of the survey outcome.
- To prepare a set of "created variables" from the edited CATI responses
- To resolve differences between the CATI outcome codes and data availability and to merge the final product data collection outcome to the master policy database to determine the post-CATI linkage status.

The CATI survey procedures did not create an outcome code for each product . Instead they identified an outcome code for each entity which was coded as complete, if the entity completed any portion of the survey. As a result, we prepared an outcome code for each product to denote whether any of the product attributes were obtained. This code, RCDSP, was set to a value of 1 if the entity provided data for at least some of the product attributes in modules 2, 3, 4 or 5.

The CATI entity product outcome determined the linkage status of each policy prior to the logical editing and statistical matching process. If an entity did not provide data on the product for an eligible policy, the interim CATI linkage status was set to a non-match.

The final step in the entity data cleaning process was to reconcile differences between the household survey data file and the master policy database. We updated the master policy database based on the CATI outcome for the associated product to assign a post-CATI linkage status. We then re-merged the master policy database to the household survey data file to identify any discrepancies.

IV. SURVEY OPERATIONS

A. THE HOUSEHOLD SURVEY

1. Eligibility

The household survey was the sampling frame for the followback survey. The family informant for each FIU provided plan and employer names for up to three health insurance policies providing health care services to one or more family members under the age of 65. Ineligible plans include specialty plans that provide only one type of service, such as accident, vision, dental, or nursing home coverage, Medicare, Medigap polices providing supplemental coverage to Medicare recipients, Medicaid, Military health plans, including CHAMPUS, CHAMP-VA, TRICARE, AND VA.

2. Household Survey Linkages and Outcomes

During the round 2 household survey, interviewers accessed known health insurance plans through the Insurer Database to assist family informants in recalling their health insurance plans (see Chapter III.A). Of the 21,701 policies identified in the household survey, hard-matches were established with 36.1 percent, soft-matches for 20.6 percent, and non-matches for 43.3 percent (see Table IV.1). To contact employers for more information, we needed information on the employer's name; this information was available for 71.8 percent of the soft-matches and 81.5 percent of the non-matches.

Household survey interviewers also asked family informants to access their health insurance cards or other documentation to help aid recall of plan names. Only 17.1 percent of informants with private health insurance coverage provided some form of documentation (see Table IV.2).

TABLE IV.1

HOUSEHOLD SURVEY OUTCOMES

Data-Base Matching Outcome	Policies		
	Count	Percent (Subgroup)	Percent (Overall)
Hard-Matched to Insurer Database	7,831		36.1
Soft-Matched to Entity			
Employer name	3,209	71.8	
No employer name	1,263	28.2	
Subtotal soft-matched to entity	4,472	100.0	20.6
Non-Match			
Some plan information and employer name	7,662	81.5	
No plan name but employer name	1,295	13.8	
No employer name but some plan Information	441[a]	4.7	
Subtotal non-match	9,398	100.0	43.3
Total Policies	**21,701**		**100.0**

[a]Includes 61 cases later determined by manual review not to be a health plan.

TABLE IV.2

DOCUMENTS USED TO NAME PLAN

	First Plan (b221)		Second Plan (b222)		Third Plan (b223)	
Outcome	Count	Percent	Count	Percent	Count	Percent
No Coverage	12,794		29,410		31,948	
No Document Used	15,962	82.9	2,300	87.2	89	89.9
Insurance Card	3,042	15.8	290	11.0	7	7.1
Claims Form	113	0.6	26	1.0	1	1.0
Insurance Policy	136	0.7	21	0.8	2	2.0
		-		-		-
With Policy	19,253	100.0	2,637	100.0	99	100.0
Total Records	**32,047**		**32,047**		**32,047**	

B. THE EMPLOYER SURVEY

1. Data Collection and Survey Management

The next phase of data collection consisted of a CATI employer survey to resolve soft and non-matched policies from the household survey and to validate a sample of hard-matches. The survey was conducted from December 1998 to May 2000. Thirty interviewers experienced in calling businesses worked on the survey. Two one-day training sessions were conducted focusing on use of the CATI instrument and access of the Health Care Plan Database (HCPDB). Trainees conducted practice interviews with one another to gain proficiency in using the program.

Polices shown in Table IV.1 were allocated to six classes, with the first three classes excluded and the last three included in the employer survey:

1. No employer name provided by soft-matched (1,263) and non-matched policies (441)
2. Employer name but no plan name (1,295 policies)
3. Hard-matches not selected for verification (6,696)
4. Hard-matches selected for verification (1,135)
5. Soft-matches with employer name (3,209)
6. Non-matches with some plan information and employer name (7,662).

A sample of hard-matched policies (class 4) were validated to assess the accuracy of household survey linkages used the insurer database. Soft-matched policies (class 5) and non-matched policies (class 6) were included to obtain insurer and product names that could be identified by insurers in the entity survey. Employers were located through directory assistance and Internet "yellow page" services. Calls were made to employers during tracing to clarify variations in spelling of employer names. Record of contact sheets were generated for each located employer.

The sheet included the name, city, county, state, telephone number and time zone of the employer, the record ID and match status of the case for which the sheet was printed, and grids to record call histories. The sample was released as employers were located, except for large employers which were held until the end of the household survey to avoid multiple calls. Interviews were scheduled throughout business hours with time zone adjustments. There was no limit on the number of call attempts before efforts were ended; instead, contact sheets were reviewed by interviewing supervisors on a weekly basis to decide whether further efforts were likely to be successful.

To protect confidentiality, we did not identify the respondent during the interview; instead, we indicated that one of the employees stated they had plan "X", and then asked the employer to verify this name or make corrections. Policies linked to the same employer were grouped to avoid contacting employers more than once.

After the employer survey was completed, we classified policies into one of four linkage outcomes: hard-match, soft-match, non-match, and "not-a-health plan" (see questions b11, b20 and b30 of the survey). Subsequently, after the analytical files were finalized we determined that the interviewer could assign a not-a-health-plan code for outcomes in which the employer either had not heard of the plan or did not offer a plan from that company to their employees. A total of 1,188 of the 1,343 policies coded as not-a-health plan should have been treated as non-matches. These cases were treated as ineligible plans during imputations and statistical weighting (see Chapter V).

2. Survey Results

Table IV.3 shows the impact of the employer survey on household survey linkages. Key findings are summarized below:

TABLE IV.3

SUCCESS OF EMPLOYER RESOLUTION OF HOUSEHOLD POLICIES

Household Survey		Employer and Post-Employer Edit	
Outcome	Policies	OUTCOME	Policies (Percent)
Hard-Matched	7,831		
Verified	1,135	Hard-Matched	1,048 (92.3%)
		Other	87 (6.7%)
Not Verified	6,696	Hard-Matched	6,696
Soft-Matched	4,472		
With Employer name	3,209	Hard-Matched	1,476 (46.0%)
		Soft-Matched	1,484 (46.2%)
		Non-Matched	34 (1.1%)
		Not- a- Health Plan	215 (6.7%)
W/o Employer name	1,263	Soft-Matched	1,263
Non-Matched	9,398		
With Employer name	7,662	Hard-Matched	4,137 (54.0%)
		Soft-Matched	1,899 (24.8%)
		Non-Matched	625 (8.2%)
		Not a Health Plan	1,001 (13.1%)
Insufficient Data	1,736	Non-Matched	1,675 (96.5%)
		Not-a-Health Plan	61 (3.5%)
Summary			
Total Hard-Matched	7,831	Total Hard-Matched	13,357 (61.6%)
Total Soft-Matched	4,472	Total Soft-Matched	4,666 (21.5%)
Total Non-Match	9,398	Total Non-Plans	1,343 (6.2%)
		Total Non-Match	2,335 (10.8%)

- Of the 1,135 hard-matches that were verified, 92.3 percent were found to be correctly linked to insurers and insurance products. Consequently, we concluded that the use of the insurer database aided respondent recall with minimal error.
- Of the 3,209 soft-matched policies included in the employer survey, 46.0 percent were upgraded to a hard-match, 46.2 percent remained in a soft-match state, and only 1.1 percent were not matched. A total of 1,263 soft-matched policies were not included in the employer survey because of insufficient information on employer name and location.
- Of the 7,662 non-matched policies included in the employer survey, 54 percent were upgraded to a hard-match, and 24.8 percent were upgraded to a soft-match. A total of 1,736 had insufficient information on employer or plan to be included in the employer survey.
- Overall, the employer survey increased the hard-match rate from 36.1 percent to 61.6 percent prior to contacting insurers. Soft-match rates increased slightly from 20.6 to 21.5 percent after the employer survey. Non-matches were reduced from 43.3 percent to 10.8 percent. Interviewers coded 1,343 (6.2 percent) policies as non-a-health plan.

Table IV.4 shows the outcomes of the survey. The response rate among employers was 93.8 percent. No differences were found in employer cooperation between the high-intensity and low intensity sites (Table IV.4a). Employer cooperation rates were slightly higher for larger employers, those linked with two or more policies. Nearly all (97.9 percent) of the larger employers cooperated compared with 92.8 percent for employers with one policy linkage. While these cooperation rates are extremely high, the information requested from employers was limited to verification of offered plans. We did not request information about plan characteristics; also, we were generally able to obtain this information from junior staff in Human Resource departments, which simplified the data collection task.

TABLE IV.4

EMPLOYER SURVEY RESPONSE RATE

CATI Outcome	Employers
Cooperated	8,799
Eligible Refusal	55
Ineligible	21
Not Reached	531
Total Attempted	9,406
Response Rate Percent	93.8%[5]

TABLE IV.4A

COOPERATION RATES BY NUMBER OF LINKED POLICIES AND TYPE OF SITE

Type of Site	Number of Policies	Count of Employers	Employers Completed	Response Rate
High Intensity Site	1	3,421	3,182	93.0%
	2+	735	719	97.8%
Subtotal		4,156	3,901	93.9%
Low Intensity Site	1	4,603	4,264	92.6%
	2+	647	634	98.0%
Subtotal		5,250	4,898	93.3%
All Sites	1	8,024	7,446	92.8%
	2+	1,382	1,353	97.9%
Total		**9,406**	**8,799**	**93.5%**

[5]Assumes based on 8,875 employers with a known eligibility status that 99.8 percent of employers were eligible. Using this as an eligibility rate we estimate that 9,384 of the 9,406 employers attempted were eligible to yield a response rate of 8,799 divided by 9,384 =0.938 or 93.8 percent. The policies associated with these employer calls totaled to 12,006 (see Table IV.2 which includes 1,135 verified hard linkages, 3,209 soft linkages and 7,662 non-matches) yielding a policy-to-employer ratio of 1.28.

C. THE ENTITY SURVEY

1. Data Collection And Survey Management

The final phase of data collection was a survey of insurers and other organizations (Entity Survey) that could provide information on the characteristics of insurance products identified in the household and employer surveys. The entity survey was conducted from March to October 2000. Sixteen interviewers experienced interviewing health professionals or employers were trained. Supervisors were selected from those who had worked on the round 1 followback or round 2 employer survey. A three-day interviewer training session was held in June 2000.

Many of the insurers to be contacted for round 2 had participated in round 1, so contact information often was available. This information was updated through interviewer calls and web site review Advance letters were mailed to insurers and other entities; those that refused were mailed another letter to encourage cooperation (see Appendix B).

2. Survey Results

Overall, the cooperation rate to the Entity Survey was quite low. Among the 1,122 entities contacted, we obtained a response rate of only 56.4 percent (Table IV.5). Respondents included a mix of local, regional, and national representatives of insurers, third party administrators and a self insured employer; the most common respondent was a local or regional marketing representative for an insurance company. As a result, the hard-matched rate from the employer phase declined from 61.6 percent to only 41.9 percent; non-matches increased from 10.8 percent to 36.7 percent. To compensate for the low response rate, key information was obtained from various sources (see Chapter V).

The Entity Survey response rate included a contact as a completed interview if the respondent provided data for any of the products requested. Based on a grouping of entity

TABLE IV.5

ENTITY SURVEY RESPONSE RATES

Outcome	Entity Survey		
	Estimated Companies	Entities Contacted	Associated Policies
Cooperated[6]	431	633	12,390
Eligible refusal	404	489	5,633
Ineligible	0	0	0
Not reached	0	0	0
Total attempted	835	1,122	18,023
Response rate percent	51.6%	56.4%	68.7%

[6]An attempted interview is counted as a complete if product attribute data was provided for any of the products requested for the corresponding linked polices.

respondents by company name, we estimate that these respondents are associated with 835 unique companies, of which we obtained at least one product for 51.6 percent (Table IV.5). Of these 835 unique companies, 30.3 percent provided data for all products in all requested sites. An additional 4.0 percent provided data for at least one product in all sites, but did not respond for some products. A total of 17.4 percent provided data for some, but not for all products. The remaining 48.4 percent failed to complete any portion of the survey.

During entity interviewing, we attempted to collect product attribute data for 18,023 hard-matched and soft-matched policies. Given the low entity cooperation rate, only 50.4 percent were hard-matched, 18.3 percent were soft-matched, and 31.3 percent were not matched (Table IV.6).

Among the 13,357 policies that were hard-matched after the employer survey, we were unable to obtain product attribute data from insurers and other entities for 4,269 (32.0 percent). Of these, we were subsequently able to obtain data from published sources via logical editing for 1,243 (discussed in Chapter V).

Among the 4,666 soft-matched policies at the close of the employer survey, we were unable to obtain product attributes for 1,364 (29.2 percent). Of the 3,302 soft-matched policies, we subsequently determined that we did not have a complete inventory of the products offered for 1,622, which were reclassified to a non-match status. For the 1,364 policies with missing attribute data, we were able to obtain a complete inventory of the products and their attributes from published sources for 252.

Due to non-response on the entity survey, the hard-match rate, which had been 61.6 percent at the close of the employer survey operations, declined to 47.6 percent and the non-match rate increased from 10.8 percent to 37.3 percent. The soft-match rate also declined from 21.5 percent

TABLE IV.6

ENTITY AND POST SURVEY OPERATIONS POLICY OUTCOMES

Employer and Post-Employer Edit Outcome Status		Entity Survey Results		Logical Editing	Soft-Match Review	Final Status	
Status	Policies	Outcome	Policies Attempted	Conducted (Yes/No)	Outcome	Outcome	Policies
Hard-Match		Cooperated	9,088	No	→	Hard-Matched	9,088
	13,357						
		Did not Cooperate	4,269	Yes	→	Hard-Matched	1,243
		Non-Match					
				No	→	Non-Match	3,026
Soft-Match		Cooperated	3,302	No	Match-able	Soft-Match	1,680
	4,666						
					Insufficient Data[7]	Non-Match	1,622
		Did no Cooperate	1,364	Yes	All Match-able	Soft-Match	252
		Non-Match					
				No	N.A.	Non-Match	1,112
Non-Match (Eligibility Unknown)	2,335	→				Non-Match	1,040
Ineligible	1,343	→				Identified Ineligible	1,343
Total Policies	21,701		18,023				21,701
Hard-Match			9,088 (50.4%)				10,331 (47.6%)
Soft-Matches			3,302 (18.3%)				1,932 (8.9%)
Non-Match			5,633 (31.3%)				8,095 (37.3%)
Ineligible			0				1,343 (6.2%)

[7]See Chapter 5, Section C.1 for discussion. For these plans we did not have the full inventory of products offered in the site.

to 15.2 percent. Subsequent lack of complete product inventory data on the soft-matches reduced the final soft-match final rate to 8.9 percent.

The overall match rate (hard-matches and resolved soft-matches) is between 56.5 and 60.7 percent (Table IV.7). The lower rate is based on the assumption that all of the policies coded as not-a-health-plan were in fact non-matches and the higher rate assumes that none of them were eligible. Based on subsequent review, we confirmed that at least 155 were non-comprehensive health plans; excluding only these policies as ineligible yields a linkage rate of 57.1 percent.

Subsequent analysis (not shown) indicated that the design of the survey instrument may have contributed to the low entity survey response rate. The CATI questionnaire was designed to cycle across sites and within sites by product if the entity respondent was responsible for more than one site. Among respondents providing data for at least one product, 65 percent did not provide data for all sites. This indicated that the added burden of providing product data on multiple sites greatly reduced the overall response rate. In contrast, providing data for all products within a site was not as problematic.

TABLE IV.7

FINAL OUTCOMES

Outcome	Policies	Percent
Matched (Hard and Resolved Soft-Matches)	12,263	56.5%[8]
Non-Matched	8,095	37.3%
Known Eligible	5,795	26.7%
Unknown Status	2,300	10.6%
Ineligible	1,343	6.2%
Total	21,701	100.0%
Total Known Eligible Policies	18,058	
Total Policies with Known Eligibility Status	19,401	
Estimated Eligibility Rate (18,058 / 19,401)	93.1%	
Total Estimated Eligible Plans (21,701 times 0.931)	20,199	

[8]The linkage rate among estimated eligible plans (12,263/20,199) is 60.7%.

V. IMPUTATION AND WEIGHTING METHODS

For round 2, the Household Survey identified a total of 21,701 private health insurance polices, of which 18,058 were confirmed as eligible plans, 1,342 were considered to be ineligible plans, and 2,300 had unknown eligibility,[9] yielding an estimated eligibility rate of 93.1 percent. For the 20,358 policies that were eligible or had unknown eligibility, we were able to link 10,368 to a unique insurance product, defined as a unique combination of the CTS site (PSU), the insurer company or entity (defined by the entity's identification number), and the insurance plan it offers. For 1,895 policies, we were able to link the policy to an insurance entity that provided data on their full product line, but not to the specific product. For those policies, we used statistical matching procedures to assign a product to the policy from products offered by the entity in that site. We could not link 8,095 of the policies, either because we could not determine the insuring entity or because we did not have data on an identified entity's products that were offered in a site; they were accounted for in the survey weighting procedures.

In the following sections, we describe the procedures used to replace missing product attribute data. The first step (section A) was to replace selected missing product data for large insurers with information obtained from web sites, plan booklets, and interviews with senior officials. In section B, we describe a weighted sequential hotdeck imputation procedure to replace missing values for a specified list of product attributes (additional details on these procedures are presented in Appendix E.) In Section C, we describe the statistical matching procedures used to link soft-matched policies with one of the products offered by the entity in the

[9]These include two types of policies: (1) those with a non-match status at the end of the household survey for which we attempted to interview the employer, but either the employer did not respond or the employer had not heard of the plan, and (2) nonmatched policies without an employer name for with an employer interview could not be attempted.

policy's site. Both the product attribute imputations and the statistical matching procedures followed the same methodology used for round 1. Lastly, in section D, we discuss the weight adjustment used to account for the non-matched policies.

A. LOGICAL EDITING OF MISSING PRODUCT INFORMATION FROM WEB-SITES

A large number of the hard and soft-matched polices with missing product data were linked with a small number of larger insurers that had low cooperation rates (see Table V.1). Consequently, we concentrated efforts to obtain administrative data for missing product attributes on the largest insurers. Based on a review of sites with completed data, we limited this effort to companies that we expected would have little or no variation on a small number of key product attributes. Of the 17 companies we reviewed, we decided to focus on four that had low cooperation rates and relatively little product attribute variation across sites.

The logical imputation effort used data from web sites, follow-up calls, and plan booklets, obtaining product attributes from web sites for two insurers, from plan booklets for one, and from a call to a national official for the fourth. Overall, this effort improved the hard-match rate from 41.9 percent to 47.6 percent, with the resolution of 1,243 hard-matches, and the soft-match rate from 15.2 percent to 16.4 percent, with the addition of product data for 252 policies. Of the final set of 2,946 products, we obtained product attributes through logical editing for 452.

B. PRODUCT IMPUTATION

Product data could be missing because the respondent did not provide the data or because of unresolved inconsistencies. Table V.2 lists the 14 data items selected for imputation and the number of responses imputed for each. While the survey instrument contained several other items that had missing responses, we decided not to impute an item if it met any of the following criteria:

TABLE V.1

TOP INSURER HARD AND SOFT-MATCHES PRE AND POST LOGICAL IMPUTATION

Insurer	Hard-Match Total		Hard-Match Entity Provided Data		Hard-Match Logically Imputed	Soft-Matches Total	Soft-Matches Entity Provided Some Data	Soft-Matches Logically Imputed (All Products Obtained)
	Count	Percent	Count	Cooperation Rate	Count	Count	Count	Count
1	3,894	29.2%	3,213	82.5%	-	1,306	1,239	-
2	1,485	11.1%	644	43.4%	815	406	201	204
3	1,075	8.0%	858	79.8%	183	210	197	13
4	997	7.5%	643	64.5%	150	278	225	31
5	518	3.9%	278	53.7%	-	105	103	-
6	232	1.7%	217	93.5%	-	63	63	-
7	212	1.6%	107	50.5%	95	78	57	4
8	198	1.5%	183	92.4%	-	90	85	-
9	186	1.4%	162	87.1%	-	70	62	-
10	167	1.3%	153	91.6%	-	23	22	-
11	143	1.1%	137	95.8%	-	17	17	-
12	120	0.9%	103	85.8%	-	61	58	-
13	96	0.7%	71	74.0%	-	17	15	-
14	92	0.7%	62	67.4%	-	29	24	-
15	91	0.7%	89	97.8%	-	36	31	-
16	91	0.7%	58	63.7%	-	34	22	-
17	88	0.7%	69	78.4%	-	23	23	-
	9,685	**72.5%**	**7,047**	**72.8%**	**1,243**	2,846	**2,444**	**252**
	13,357	**100.0%**	**9,088**	**68.0%**	**1,243**	4,666	**3,302**	**252**

TABLE V.2

PRODUCT DATA IMPUTED FOR EACH QUESTIONNAIRE ITEM

(AMONG 2,946 PRODUCTS)

Data Item(s) In Order Imputed		Abbreviated Description	Number Imputed
1.	B2A_1 B2A_2 B2A_3 Imputed as vector	Model Type _1 = Staff model _2 = Group model _3 = Network or IPA _4 = Something else	220 (same for all three)
2.	B8	Maximum in-network coverage	17
3.	B10	PCP Sign-up Required	7
4.	B91A10	Any in-network coverage	294
5.	B12_2 B12_3 Imputed as vector	Which types of physicians can serve as PCPs PCP Type 2=OB/GYNs PCP Type 3=Other Specialists	164 167
6.	B13	Co-payment or co-insurance	203
7.	B14	Deductible Amount	483
8.	B13AMT	Co-payment amount	431
9.	B13PER	Co-insurance amount	245
10.	B13OUT	Out of network co-payment or co-insurance	71
11.	B14OUT	Out-of-network deductible requirement	156
12.	B14OD	Out-of-network deductible amount	240
13.	B13OP	Coinsurance for out-of-network	176
14.	C4_2	PCP Payment	621

[10]Items B8, B10 and B91A were imputed together as a vector (see Appendix E).

- If the missing rate was too high to support inference (generally above 40 percent).
- If the variable was applicable to a small number of products (less than 20 percent).
- If after attempting imputation we found the ratio of donors-to-recipients to be less than 2.5-to-1 in the final set of classing characteristics.
- If a variable's skip pattern was dependent on another variable that we decided not to impute.

We used a weighted sequential hot deck procedure to impute the missing responses. For the item to be imputed, the procedure selects a "donor" respondent from cases with nonmissing values and substitutes the donor's response for the missing value. The procedure can be controlled by restricting the donor pool to products that have the same or similar responses to variables that are related either to the data item being imputed or to the item's likelihood of being missing. The potential donors and the recipients are then sorted according to their sampling weights and the donors are selected based in part on their weight values. Over repeated imputations, the expected values of the distributional characteristics using both imputed and actual data will equal those using only the reported data. We used the square root of the number of policies successfully linked to each product as the sampling weight for each product in the execution of the weighted sequential hot-link imputation procedures.

C. STATISTICAL MATCHING OF SOFT-MATCHES

A soft-match is defined as a policy that is linked to an entity, but not to one of the entity's products; a total of 4,666 policies were soft-matched to a set of products after employer interviewing (Table IV.6). For the soft-matches, the objective was to select a product that best represents the policy from products the insurer made available to the household respondent at the CTS site. The CATI data collection methodology did not inventory all products offered by an insurer within a site. As a result, we had to determine whether a complete set of products were obtained for an insurer or other entity within a site before statistically matching soft-matches.

For soft-matches for which we had a full product inventory (1,932 [11] of the 4,666 soft-matches) we selected a product that best "fit" the policy reported by the family informant.

1. Availability of Full Set of Product Offerings

We examined the mix of product lines (HMO, POS, PPO, indemnity) for each entity and site combination with one or more related soft-matches. If the entity provided us with data for one or more products within each product lines we considered our choice set to be complete. Of the 3,554 soft-matched policies for which we had some data on their products after logical imputation, 1,615 offered one or more products within each product line. For the remaining 1,939 soft-matched policies, we attempted to obtain administrative data on the number of product lines offered in the site. In most cases, we were able to obtain information about product line offerings at the state rather than site level. If the product lines offered at the time of our review matched the product lines obtained during the survey for the site, we consider the choice set complete.

The use of administrative data to determine whether the choice set was complete did not substantially increase the number of soft-matches that could be resolved. We were able to find a web site containing information on product line offerings for 196 entities representing 1,060 policies. Thirty-five entities (237 policies) were dropped because the company had changed ownership since the survey. Only 66 of the remaining 161 entities, representing 317 policies, listed product lines on web sites that matched the survey data. Overall, we were only able to verify that the choice set was complete for these 317 policies to bring the total number that could be statistically matched to 1,932. A summary of the process is shown in Table V.3.

[11]Of the 1,932 policies that were statistically matched, 37 were linked to an insurer offering only one product in the site; consequently, 1,895 soft-matches were resolved using the statistical matching procedures.

TABLE V.3

SUMMARY OF SOFT-MATCHED POLICIES

Situation	Total Policies by Situation	Policy Subtotals	Total Policies Matched
No Entity Data Collected	1,112		0
Data Collected (All)	3,554		1,932
With All 4 Product Lines Reported	1,615		1,615
With Less than 4 Reported Lines	1,939		317
Researched		1,240	
Found related Web Data		1,060	
Company Same Ownership		823	
Matched product lines		317	
Total	**4,666**		

2. Selecting the Matches

By examining hard-matched data, we were able to identify relationships between characteristics reported by entities (typically insurers) and family informants from the household survey. We modeled these relationships and applied the models to the soft-matched cases to select a final policy from among the alternatives recorded. Based on procedures used in round 1, these methods produced an estimated 64 percent exact match rate and 67 percent match rate with a product of the same type.

Statistical matching required several steps summarized below:

- First, we prepared a set of policy level variables from the Household Survey to summarize the policy holder's demographic, socioeconomic, family structure, and health plan characteristics that may relate to product choice.
- Second, we prepared a weighting class adjustment for the hard-matched policies so that the weighted distribution of the site and self-reported HMO membership for these cases would mimic that of the hard and soft-matched policies combined. This was done because the hard-matched policies potentially represented a skewed sample of the policy holders represented by the hard and soft-matched policies.
- Third, based on the round 1 methodology, we used the set of nine entity-reported product attributes to describe the differences among the products offered.
- Fourth, we developed a logistic regression model, to predict each of the nine product attributes using the household reported policy level information from the hard-matched policies.
- In round 1, we also prepared "mock" or test data sets to assess the accuracy of different matching procedures. We assumed that a similar process would have yielded the same results for round 2 and did not prepare another set of test data sets.
- In round 1, the test data files led to the development of a second logistic regression model to predict the probability of a correct linkage based on the difference (or "gaps") between the predicted values for the product attributes and the product attributes for the possible choices (step 5 below). For round 2, we applied these same model coefficients to the "gaps" to assign a probability of a match to each product in the choice set (step 6).
- Fifth, we used the policy-level information from the soft-matches in the nine attribute models developed in step four to obtain predicted values for the nine product attributes. For each of the nine attributes, we then used the predicted values to compute a measure of the difference between the predicted values and actual attribute values among the product choices.

- Finally, we used both the measured differences between the predicted values and the actual attribute values among the soft-matched choices and the modeling results from round 1 (see step 4) to compute an estimated probability of a match for each choice. The product with the highest probability of a match was selected as the final match. If the highest probability of a match was the same for two or more products, we chose one at random.[12]

D. REWEIGHTING TO ADJUST FOR NON-MATCHES

A total of 8,095 policies were not linked to a product in the Followback Survey.[13] We used a weight adjustment based on the inverse of the modeled probability of a link for these non-matched policies. A stepwise regression approach was used to build this model. We started with a set of variables that might be related to the likelihood of a successful followback linkage, including the same set of variables used to build the comparable model for round 1. In addition, we added a set of indicator variables (one for each of the five largest entities) to account for the effect of the logical editing procedures. If one of these entities was mentioned by the family informant, the indicator variable for that entity was set to one.

Using a stepwise regression model at the policy level, weighted by the national family-level weight for the household survey, we arrived at a model to be used for predicting a successful linkage. The predicted probability of a linkage resulting from this model was used to adjust the appropriate person-level weights from the household survey. This adjustment factor was merged onto the person-level file by policy. The followback weights for respondents whose policies

[12]This occurred in only 25 of the 1,932 linkages with two product choices each.

[13]Family informants to the household survey reported 1,343 policies that were determined by the employer survey and manual review to be something other than comprehensive health plans. These policies were included with the hard and soft matches for weighting as ineligible plans, so that the appropriate proportion of the non-matched cases were assumed to be ineligible. After the survey weights were prepared, we conducted another review of these policies and concluded that it was more likely that the actual eligibility status was unknown rather than ineligible for most of these plans (1,188).

were either hard-matches, soft-matches, or not-a-health-plan were set equal to their final CTS Household Survey person-weights (either national or site-specific weights, based on the augmented site sample), multiplied by the inverse of the probability of a link from the model. These weights were set to zero for persons with non-matched policies.

The next steps involved poststratifying (via raking) and trimming outliers for these adjusted person-level weights. We used a combination of both weighted and unweighted least squares raking procedures to align the weighted distribution among the followback matched cases to the weight distribution among all the cases in the followback sample with private insurance. After trimming the outlier weights, we re-raked as necessary, so that all the distribution matched again.

REFERENCES

Community Tracking Study Followback Instrument (Round Two), 2002, Technical Publication No. 33.

Emmons, C., and Hill, C. 1991. Questionnaires and Data Collection Methods in the Health Insurance Plans Survey, Methods 5, AHCPR Publication Number 94-0016.

Nelson DE, Thompson BL, Davenport NJ, Penaloza LJ. 2000. What people really know about their health insurance: a comparison of information obtained from individuals and their insurers. Am J Public Health 90(6): 924-8.

Cunningham PJ, Denk C, Sinclair M. 2001. Do consumers know how their health plan works? Health Affairs 20(2):159-66.

Sinclair, M., Strouse, R., Denk, C., Potter, F., Carlson, B., and Pascale, J., Community Tracking Study Health Insurance Followback Survey Technical Report on Data Collection Operations and Final Data Files, HSC Technical Publication No. 30

MPR Reference No.: 8553

Community Tracking Study Health Insurance Followback Survey Technical Report on Data Collection Operations and Final Data Files

Appendices A-J: Instruments, Data Editing Specifications and Imputations, and Weights

Final Report

September 2002

Michael Sinclair, Senior Statistician
Richard Strouse, Vice President
David Uglow, Associate Director Systems and Programming
Barbara Carlson, Senior Statistician
Kathy Sonnenfeld, Survey Director

Submitted to:

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

Project Officer: Peter Cunningham

Submitted by:

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

Project Directors: Richard Strouse, Michael Sinclair

APPENDIX A

EMPLOYER SURVEY QUESTIONNAIRE

I. THE CATI QUESTIONNAIRE

>a0< Hello, is this [FILL enam]? Could I speak to the person in charge of employee health benefits?

<1> CONTINUE [goto a1]
<2> WRONG NUMBER [goto a3]
<3> DOES NOT OFFER HEALTH BENEFITS [goto a3]
<4> CALL-BACK [goto z1]
<5> REFUSES [goto a2]
<6> ALREADY CALLED THIS SOURCE AND PRIOR DATA APPLIES [goto z1]

>a1< Hello, my name is _____ and I'm calling on behalf of the Robert Wood Johnson Foundation. We are conducting a nationwide study to track the rapid changes that are going on in the health care system in particular communities. In our family survey we recently spoke to someone in your area who said they have health care coverage through your company. I'd just like to ask a few questions to confirm that coverage. Everything you or that employee told us will be kept confidential.

<1> CONTINUE WITH INTERVIEW [goto b01]
<2> MORE INFO [goto a2]
<3> DOES NOT OFFER HEALTH BENEFITS [goto a3]
<4> CALL-BACK [goto Z1]
<5> REFUSES [goto a2]
<6> ALREADY CALLED THIS SOURCE AND PRIOR DATA APPLIES [goto z1]

>a2< MORE INFO: DO NOT READ

* Refer to posted materials for more information about the study.

* If permission from a higher corporate office is required, explain limited purpose of call. If necessary, record corporate information on contact sheet.

* Use contact sheet to record new information about respondents.

<1> CONTINUE WITH INTERVIEW [goto b02]
<2> NEW RESPONDENT TO PHONE [goto a1]
<3> DOES NOT OFFER HEALTH BENEFITS [goto a3]
<4> CALL-BACK [goto z1]]
<5> REFUSES [goto z1]
<6> ALREADY CALLED THIS SOURCE AND PRIOR DATA APPLIES [goto z1]

>a3< DOES NOT OFFER HEALTH BENEFITS, OTHER WRONG NUMBERS:
READ AS APPROPRIATE

* Confirm phone number; is there any other employer in the area with similar name? Get contact info.

* Is there a union or professional association providing health benefits to employees? Get contact info.

* Has this company offered health benefits in the past that might be continuing for retirees or employees that have left you (COBRA)?

<1> CONTINUE WITH INTERVIEW [goto b02]
<2> NEW RESPONDENT TO PHONE [goto a1]
<3> DOES NOT OFFER HEALTH BENEFITS [goto z1]
<4> CALL-BACK [goto z1]
<5> REFUSES [goto z1]
<6> ALREADY CALLED THIS SOURCE AND PRIOR DATA APPLIES [goto z1]
<7> WRONG NUMBER, NO NEW CONTACT [goto z1]
<8> WRONG NUMBER, NEW CONTACT SUPPLIED [goto z1]

B. SECTION B – GENERAL NOTES

You will now try to validate the information of the household interview. You will also start using the database as described above. The company representative might not be able to identify one single plan from the information provided. You will have the ability to enter three different plans for each entity.

PLAN CONFIRMATION

>b02< For reasons of confidentiality we will not tell you the name of the person we spoke to in our survey. But we can tell you that they live in the [FILL psu] area, and said they had health care coverage through employment with [FILL enam].
<g> Continue

>b03< [if smps=1 goto b11] This is a validation of a hard hit from household interview.
[if smps=2 goto b20] This is to convert a product-not-known entity match from household interview to specific products. [if smps=3 goto b30] This is to follow-up a no-match from household interview.

>b11< call cvr1 into dtxt (text string from Insurer Database)The person we spoke with said they were enrolled in [FILL dtxt] . I'd just like to confirm that [FILL dtxt] is a comprehensive health plan that you offer to your employees? A comprehensive health plan is one that covers most medical and hospital services?

PROBE: We want to exclude plans that provide only one type of service, such as dental and vision plans, or provide only supplemental cash, or is another kind of plan such as a disability and pension plan.

PROBE FOR INCORRECT PLAN: Do you offer any plans from that company?
USE <2> FOR YES.

<1> YES [goto c1]
<2> GET A DIFFERENT PRODUCT FROM THIS ENTITY [goto b201]
<3> GET A NEW ENTITY [goto b301]
<4> NO, NOT A HEALTH PLAN [goto b60]
<d> DON'T KNOW [goto b99]

GET THE RIGHT PLAN WITH KNOWN ENTITY:

There are two versions of > b20< depending on how usable the plan name was the household respondent provided. The variable dtxt is the entity name coded in the Household Study by using the Insurer Database; pnme is the original text string supplied by the Household respondent. Thus whether pnme is usable determines whether b20 or B20 is used. Probe and answer categories are the same in both versions.

Name was useful:

>b20< The person we spoke with seemed sure of the company but not specific plan offered by that company. They mentioned it might be called [FILL pnme] and they indicated that the plan was offered by [FILL dtxt]. They mentioned it might be called " [FILL dtxt."] First, does this sound like a comprehensive health plan you offer to your employees? A comprehensive health plan is one that covers most medical and hospital services?

Name was <u>not</u> useful:

>b20< The person we spoke with seemed sure of the company but not specific plan offered by that company. First, does this sound like a comprehensive health plan you offer to your employees? A comprehensive health plan is one that covers most medical and hospital services?

PROBE: We want to exclude plans that provide only one type of service, such as dental and vision plans, or provide only supplemental cash, or is another kind of plan such as a disability and pension plan.

<1> Yes
<0> No, not a health plan [goto b60]
<3> Need to search different entity name [goto b301]
<d> Don't know [goto b99]
<r> Refused [goto b99]

>b201< I'm going to read a list of plans offered by that company in the area. Please tell me if one or more of them matches a plan you offer to employees that might be the one that person meant.

<g> Continue – [Accesses the Insurer Database]
<d> Don't Know [goto b99]
<r> Refused [goto b99]

Example of the database screen, showing all the plans offered nationwide by that entity:

I'm going to read a list of plans offered by that company. Tell me if one
of them is the name of (FILL: call number 1=your; 2=the SECOND; 3=the THIRD) plan.

(READ FROM THE LIST OF PRODUCTS.
USE ARROW KEYS TO PAGE.)

100 *	Aetna	EPO
101 *	Aetna	FFS
102 *	Aetna	HMO
103 *	Aetna	POS
104 *	Aetna	PPO
105 *	Aetna	Quality POS
106 *	Aetna	PRODUCT NOT SPECIFIED

7 matches found

<1> Confirm highlighted entry
<0> Insurance company name does not match
<9> Insurance company name incorrect; backup and correct (aetna)

===>

Items >b23< to >b29< allow the interviewer to record additional plans for the entity in question.

Similar to >b20< two versions of >b30< exist depending on the usability of the Household respondent's identification of their health plan (pnme).

Usable name:

>b30< The person we spoke with mentioned a plan called [fill pnme] First, does this sound like a comprehensive health plan you offer to your employees? A comprehensive health plan is one that covers most medical and hospital services?

PROBE: We want to exclude plans that provide only one type of service, such as dental and vision plans, or provide only supplemental cash, or is another kind of plan such as a disability and pension plan.

<1> Yes
<0> No, not a health plan [goto b60]
<d> Don't Know [goto b99]
<r> Refused [goto b99]

No usable name provided:

>b30b< The person we spoke with was not sure of the health plan they were enrolled in. I have a database of many plans offered in the area. Please tell me about any comprehensive health plan you offer to your employees, that is, any health plan that covers most medical and hospital services.

PROBE: We want to exclude plans that provide only one type of service, such as dental and vision plans, or provide only supplemental cash, or is another kind of plan such as a disability and pension plan.

<g> Continue [goto b31]
<0> No applicable health plan [goto b60]
<d> Don't Know [goto b99]

>b301< I'm going to read a list of plans offered by that company in the area. Please tell me if one or more of them matches a plan you offer to employees that might be the one that person meant.

<g> Continue – [Accesses the Insurer Database]
<d> Don't Know [goto b99]
<r> Refused [goto b99]

Example of the database screen, showing all the plans offered nationwide by that entity:

I'm going to read a list of plans offered by that company. Tell me if one
of them is the name of (FILL: call number 1=your; 2=the SECOND; 3=the THIRD) plan.

(READ FROM THE LIST OF PRODUCTS.
USE ARROW KEYS TO PAGE.)

100 *	Aetna	EPO
101 *	Aetna	FFS
102 *	Aetna	HMO
103 *	Aetna	POS
104 *	Aetna	PPO
105 *	Aetna	Quality POS
106 *	Aetna	PRODUCT NOT SPECIFIED

7 matches found

<1> Confirm highlighted entry
<0> Insurance company name does not match
<9> Insurance company name incorrect; backup and correct (aetna)

===>

Items >b33< to >b38< allow the interviewer to record additional plans for the entity in question.

NEW ENTITIES

>b50< RECORD THE NAME OF THE NEW ENTITY [72 char]

>b511-B513<

RECORD THE NAME OF THE NEW PRODUCT [72 char]. Up to three new products.

>b52< SOLICIT CONTACT INFORMATION FOR ENTITY:

contact person
address
phone number
fax number
(leave blank for DK)
[goto c1]

NOT A COMPREHENSIVE MEDICAL PLAN

>b60< What type of plan is it?

<1> Medicare or retirement supplement ("Medigap")
<4> Military Health Plan (active duty and dependents, e.g., VA, CHAMPUS, CHAMP-VA, TRICARE STANDARD and PRIME)
<5> Specialty plan that provides only one type of service, such as dental and vision plans, or provides only supplemental cash
<6> OTHER INCLUDING MEDICAID OR OTHER PUBLIC [SPECIFY]
<d> DON'T KNOW
[goto z1]

NEW PRODUCTS

>b981-b983<

What is the name of the new product [72 char]. Up to three new products.

NEW CONTACTS

>b99< Is there someone else who can answer this question for me?

IF YES, GET NEW CONTACT INFORMATION.

<1> YES, CALL BACK [z1 with status new contact]
<2> YES, TRANSFER NOW [record new contact, goto a1]
<2> NO [z1 with final status no info available]

C. SECTION C:

There are only 2 questions in section c that ask for general information about the company's health insurance program.

OTHER VALIDATION QUESTIONS

>c1< [if sample status ne validation goto z1]

>c2< For how long have you offered the plan(s) you indicated to your employees?
RECORD YEARS

>c3< Do you offer any other health plans to your employees? (I won't ask what they are.)

<1> YES
<2> NO

CLOSEOUT

>zz< Closing text for interviewers; Instructions for filling out contact sheets; Instructions for coding case status.

II. FAQ

What is this about?/Will the data be confidential?

All the information you provide will be kept strictly confidential. Our reports will describe different types of health plans and policy holders as groups; at no time will individual health plans or employers be identified by name.
How do I participate, and how much time with this take?

The interview will take less than 5 minutes and we can schedule an appointment for anytime that's convenient for you.

Who is sponsoring the survey?

The survey is being sponsored by the Robert Wood Johnson Foundation, a non-profit organization based in Princeton, New Jersey, whose sole mission is to improve health care. Some of the other projects sponsored by the foundation include:

- Medicaid Managed Care Program: Aimed at helping states, managed care organizations, providers, and consumers take advantage of the unique opportunities presented by managed care to meet the needs of Medicaid recipients.

- Service Credit Banking in Managed Care: Intended to help HMOs and other prepaid delivery systems respond to growing numbers of enrollees in need of informal care by developing and implementing volunteer caregiver programs for their elderly member.

- Addressing Tobacco in Managed Care: Designed to help managed care providers help people avoid harm caused by tobacco and promote exemplary tobacco intervention practices.

How was I selected?

This is the second part of a large survey about health plans. One of your employees told us in part I of the survey that he or she has health insurance through your company. For reasons of confidentiality we cannot reveal who this employee was.

Corporate Headquarters should be responding to these questions!

This is a very short survey and our experience tells us that these questions are best answered by the person responsible for health benefits at the location where the employee that initially took part in our survey works.

We have a very complicated benefits program!

This survey only deals with one health plan you may be offering to your employees. We already got most of the information from one employee and we are only calling to clear up some details.

Who can I call to get more information about the survey?

For more information about the study or to schedule an interview appointment, please call Dr. Charles Denk toll-free at (877)840-4770. Thank you in advance for your help.

APPENDIX B

ENTITY SURVEY ADVANCE AND REFUSAL LETTERS

Dear Respondent:

The U.S. health care system is undergoing change at an unprecedented pace, and new forms of managed care are emerging as it serves a growing portion of the population. However, little systematic information is available to understand the nature and extent of health system change and its impact on the local marketplace. In response to this information gap, the Robert Wood Johnson Foundation is sponsoring the "Community Tracking Study"--a major multi-year study to track changes in the health care system at the community level. The study involves gathering information from community residents, insurers, health plans, physicians, and other organizations that make up the health care system in 60 randomly-selected communities across the country. Data are being gathered on a recurring basis, permitting tracking of health system change in these communities. Some of the individual surveys for this larger study have already been completed, and we are now asking for your participation in the survey of health plans and insurance companies. Some of the questions you may have about this survey are answered below:

How was my organization selected to be part of the survey?

The first phase of the Community Tracking Study--a survey of residents in these 60 communities--has just been completed. In the residential survey we gathered basic identifying information about residents' health care plans (such as the plan name and the name of the employer providing the coverage), and we are now conducting a survey of the health plans and insurance companies cited by these residents.

Why are you doing this survey?

In the residential survey we also gathered basic information on the general characteristics about the plan, such as the type of plan (HMO, PPO, etc.), and whether a primary care physician is required. Because individual policyholders frequently do not know about or understand the details of their coverage, we'd like to validate the health plan information obtained from these community residents and gather supplemental information about those plans. Please note we are not seeking information on individual enrollees (e.g., claims experience), only general descriptive information about health care products your company offers.

What are you offering me in return for my participation?

When we've completed the study, we'd like to send your organization a summary report that can help your staff understand how your local market compares to others in the industry. The report will include aggregate statistics on the characteristics of residents' plans, such as the percentage of all plans that include out-of-network coverage, the percentage that require referrals for specialists, and typical copays and deductibles.

Will the data be confidential?

Yes. All the information you provide will be kept strictly confidential. Our reports and analyses will group individual enrollees by type of health plan (e.g., HMO, POS, PPO, indemnity); at no time will individual health plans or insurers be identified by name.

How do I participate, and how much time with this take?

An interviewer from Mathematica Policy Research, an independent survey research organization, will be contacting you soon by phone. The interview will take only about 15-20 minutes. We can schedule an appointment for anytime that's convenient for you, and we can break up the interview into several shorter sessions.

Who is sponsoring the survey?

The survey is being sponsored by the Robert Wood Johnson Foundation, a non-profit organization based in Princeton, New Jersey, whose sole mission is to improve health care. Some of the other projects sponsored by the foundation include:

- Medicaid Managed Care Program: Aimed at helping states, managed care organizations, providers, and consumers take advantage of the unique opportunities presented by managed care to meet the needs of Medicaid recipients.

- Service Credit Banking in Managed Care: Intended to help HMOs and other prepaid delivery systems respond to growing numbers of enrollees in need of informal care by developing and implementing volunteer caregiver programs for their elderly members.

- Addressing Tobacco in Managed Care: Designed to help managed care providers help people avoid harm caused by tobacco and promote exemplary tobacco intervention practices.

Who can I call to get more information about the survey?

For more information about the study or to schedule an interview appointment, please call Ellen Siegel at 800-840-4770. Thank you in advance for your help.

Sincerely,

Sally Waltman

Refusal Conversion Letter

Dear [fill respondent name]:

Recently an interviewer from Mathematica Policy Research called on behalf of The Robert Wood Johnson Foundation to ask your organization to participate in a major nationwide health care study (attached is a letter, faxed earlier, explaining the study). The foundation regards this study as one of the most important research. projects on health care ever undertaken. While there is an abundance of anecdotal information on health care today, little systematic information is available on the nature of health care products in specific markets and how those products are changing over time. We need your participation to make the study a success.

We understand you may have concerns about the confidentiality of the information you provide. The Robert Wood Johnson Foundation and Mathematica Policy Research take their pledge of confidentiality very seriously. All information gathered from individuals and organizations will be aggregated and reported only in terms of broad categories; no individual, organization or employer will ever be identified by name. Furthermore, no marketing lists will be produced from this study.

In exchange for your participation we'd like to send you a report, which will include aggregate statistics on the characteristics of community residents' plans. The report will focus on managed care attributes and network characteristics, such as the percentage of persons whose plans include out-of-network coverage, the percentage that require referrals for specialists, and typical copays and deductibles in specific markets. We believe this report will be extremely valuable to your company for the following reasons:

- It will be based on a representative sample of community residents and the health plans in which they are actually enrolled;
- It will provide a detailed picture of product lines and plan features in local markets across the country;
- Future reports will track changes in these product lines and features over time; and
- Only organizations that participate in the survey will receive the report.

Finally, to reimburse your organization for the staff time involved in completing the survey, we will send a check commensurate with the survey task.

A staff member from the foundation or Mathematica Policy Research will be calling you soon to see if you have any other questions or concerns about the study. You may also call the foundation at 800-719-9419 and ask for Maureen Michael. If you prefer to make an appointment to complete the survey, please call Ellen Siegel of Mathematica Policy Research at 800-840-4770.

Thank you in advance for your help. Your participation is crucial for this innovative study to provide the most up-to-date, accurate portrait of local health care markets and how they're changing over time.

Sincerely,

Steven A. Schroeder, M.D.
President
The Robert Wood Johnson Foundation

APPENDIX C

ENTITY SURVEY INSTRUMENT

For the Entity Survey Instrument please refer to the document submitted in September 2001. (See attached report).

MPR Reference No.: 8553-500

COMMUNITY TRACKING STUDY — FOLLOWBACK ENTITY SURVEY
Round 2
Instrument

September 2001

Submitted by:

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

Submitted to:

Center for Studying Health System Change
600 Maryland Ave., SW
Suite 550
Washington, DC 20024-2512
(202) 484-5261

CONTENTS

I. INTERVIEWER TRAINING

A. PROJECT OVERVIEW

The Follow-Back Survey is an integral part of the Community Tracking Study (CTS), which is conducted by Mathematica Policy Research, Inc. (MPR) for the Center for Studying Health System Change. The Follow-Back Study follows the CTS Household Survey in which information is collected from household family members on the kind of private insurance coverage they have. This information is "followed back" to the organization that administers the health insurance plan named by the household respondent. In the Follow-Back Survey health plans and health insurance entities are interviewed to collect information on the characteristics of specific health insurance products that are linked to household respondents.

B. BACKGROUND

The purpose of the Community Tracking Study is to develop an information base designed to track and analyze change in the nation's health care market and to inform public and private decision-makers about these changes. The study has three overall objectives.

- ***Tracking Changes in Health Systems.*** The study's first objective is to document changes in the health system through intensive study of selected communities. The major changes that have been reported in the health system include consolidation of the market at all levels (medical groups, hospitals, insurers, and health plans); vertical integration of providers (for example, hospitals and physicians) and of insurers and providers; increased risk sharing by providers; growth of large, national, for-profit health care enterprises; and the adoption of new techniques for managing clinical care (clinical information system quality improvement techniques, utilization management, and so forth). This research is conducted directly by HSC and through other contractors.

- ***Tracking Changes in Access, Service Delivery, Cost and Perceived Quality.*** The second objective of the study is to monitor the effects of health system change on people by

tracking indicators of these effects, including favorable or unfavorable changes in access to care, service use and delivery, and quality and cost of care.[1]

- ***Understanding the Effects of Health System Change on People.*** The third objective of the study is to understand how differences in health systems are related to differences in access, service delivery, cost and perceived quality. This objective will be achieved by analyzing – qualitatively and quantitatively – the relationship between health systems and access, delivery, cost, and perceived quality.[2]

C. THE ENTITY SURVEY

1. The Entity Survey Instrument – Basics

The instrument uses a CATI program consists of essentially six modules or sets of questions. These modules are arranged to allow repetition of questions across multiple sites and multiple products. The six modules are:

Module 1: ENTITY AND SITE INFORMATION

Module 2: PRODUCT OFFERINGS

Module 3: CORE ATTRIBUTES OF PRODUCTS

Module 4: CO-PAYMENTS AND DEDUCTABLES

Module 5: NETWORK AND PROVIDER RELATIONS

Module 6: CLOSING

[1]These issues are covered in the CTS Household Survey.

[2]The Followback Survey contributes mainly to this objective, by obtaining information from the health plan organizations about the specific policies that cover individual people.

The ***first module*** consists of ***questions a1– a3***, and asks basic screening questions about the ***entity***; determines that the respondent is the correct person to answer for all of the sites included in the sample for that entity; and identifies which type of organization the entity is.

ENTITY TYPES:

Licensed Insurer or Health Maintenance Organization (HMO) – Traditional insurance companies sell policies that cover specified services for a fixed premium, usually with a policy-holder contribution for each service. An HMO is an organization that combines the delivery and financing of health care in a single organization that provides comprehensive health services to a defined population of patients or "enrollees" for a set fee.

Managed Care Organization, not licensed to sell insurance – Includes Preferred Provider Organization (PPO) or Independent Practice Association (IPA) that provide health care services but are not licensed to sell insurance to the patients it serves.

Third Party Administrator (TPA) – An organization which administers health care plans (e.g., claims processing), but does not bear any risk for health care costs as an insurer, and does not employ any physicians.

Other Provider Organization – An organization that has selective provider arrangements through a network of providers who are typically paid according to a negotiated fee schedule.

Employer, union or trust plan administrator - An organization that provides health care for its employees or members.

The ***second module*** consists of ***questions b1a – b3***, and asks questions about each of the ***health care products*** **that may be offered** or administered by the entity. If the product is an HMO or a POS, the respondent is also asked what *model* best describes the organization.

HEALTH CARE PRODUCTS:

HMO – A Health Maintenance Organization, in which enrollees may see only the providers within the HMO network. Services provided outside the network are generally not covered

POS – A Point of Service plan, in which enrollees may use the HMO's network of providers for a set co-payment, or may self-refer to providers outside of the network for a larger co-payment or other additional cost. May be referred to as "open-ended or open-access HMO", or as "triple option" or "dual option" plan

PPO – A Preferred Provider Organization, in which enrollees do not need a referral, but have financial incentives to use a "preferred" set of providers, usually through difference in coinsurance or deductibles

INDEMNITY – Traditional fee-for-service arrangement in which enrollees pay for the cost of the service, and do not have financial incentive to use a select set of providers

MODEL TYPES:

STAFF MODEL – An HMO or POS that employ salaried physicians who only serve the HMO's enrollees in the HMO's facilities

GROUP MODEL – An HMO or POS that contracts with a single physician group to provide care to the HMO's enrollees

NETWORK MODEL – An HMO or POS that contracts with two or more group practices or with individual physicians to provide care to the HMO's enrollees

The questions in the second module are repeated for each product in that site before moving on to module three for the same site.

The ***third module*** consists of ***questions b5 – b12*** and asks about ***core attributes*** for each product including:

- whether or not there is a ***list of doctors***,
- whether enrollees have ***any coverage*** for visits ***to out-of-network*** doctors,
- whether ***primary care physicians*** are ***required***;
- whether enrollees may ***"self-refer"*** to ***in-network*** specialty doctors.

COVERAGE TERMINOLOGY:

OUT-OF-NETWORK – visits to providers of major medical services NOT included in the list of doctors specified as belonging to the health plan's network

IN-NETWORK – visits to providers who are included in the list of doctors specified as belonging to the health plan's network of doctors

PRIMARY CARE PHYSICIANS – Physicians who act as the first point of contact for patients seeking care. May include generalists, internists, pediatricians, or family practitioners; and sometimes OB/GYNs and selected specialists

SPECIALIST – Physicians whose services are not considered primary care. May include surgeons, dermatologists, etc. A referral may be required before the patients can see a specialist physician

SELF-REFERRAL – Visits to an in-network specialist without obtaining a referral. Different products offer differing levels of coverage, or none at all

The questions in Module 3 are repeated for each product within the site. Module 2 and Module 3 are completed as a block for each site. In other words,. once Module 3 is completed for all products within a site, the program returns to Module 2 for any additional sites.

The ***fourth module*** consists of ***questions b13-b14*** which ask about any ***co-payments, co-insurance*** **or** ***deductibles*** associated with in-network and out-of-network office visits. It is important to note that the co-payment or co-insurance is the amount that the enrollee pays.

PAYMENT TERMS

CO-PAYMENT – a fixed dollar amount that the enrollee is responsible for paying at each visit

CO-INSURANCE – a fixed percentage of the cost of each visit that the enrollee is responsible for

DEDUCTIBLE - A fixed sum of money that the enrollee is responsible for paying in a given year before coverage begins

The questions in Module 4 are repeated for each product and for each site before moving on to the fifth module.

The ***fifth or Network and Payment module*** consists of ***questions c1r – c7b*** and asks about the primary care physicians and specialist physicians associated with the health plan, as well as the payment mechanism used by the entity for various provider services.

PROVIDER REIMBURSEMENT TERMINOLOGY:

FEE-FOR-SERVICE – A system of payment in which a separate fee is charged for each specific medical service performed

FIXED FEE SCHEDULE – Also called discounted fee for service, or relative value units. A variant fee-for-service in which a pre-negotiated discounted fee is charged for each specific medical service performed

SALARIED – Reimbursement in which the health plan pays the provider a set salary for all covered services the enrollees may require

CAPITATION – Reimbursement to a provider or provider group through the payment of a fixed, periodic payment (usually monthly) for a defined set of services delivered to a set population of patients (literally per head). Under this type of payment system, financial risk for the patients' utilization is borne by the providers

Module 5 is repeated for each product and each site before moving on to questions in the sixth module.

The ***sixth module*** or ***closing section*** consists of ***questions d1 – d3*** which ask for the entities ***tax*** and ***organizational*** status and affiliations.

2. The Entity Survey Instrument – Navigational Specifics

The Follow-Back instrument was programmed for computer-assisted-telephone-interviewing (CATI) because of variation in sites covered and products offered. Some entities may have been reported by a number of different household respondents across a number of different sites. During the entity interview, you will ask about characteristics of the health insurance products that are offered in each of the sites that has been sampled. Each case will be pre-loaded with a number of sites (for example, Little Rock Arkansas, Western Washington state) and a number of products within those sites (for example an HMO, a PPO, etc.), based on reports from the Household and Employer Surveys.

a. Moving Through Modules by Product and Site

Most entities have more than one site. Therefore, the CATI program is designed to move through the product and site combinations on a site bases and then within the site on a product basis. For example, if a company has two products (A and B) both of which are offered in two sites (1 and 2), we begin with site 1, product A, then move to site 1, product B. Then we move on to site 2, product A and finally onto site 2, product B. While this site-by-product flow is the same throughout the CATI program, the process starts and finishes separately in some modules. See Figure 2.1 for details. For example in Module 2 questions about all of the products (A and B) for the first site (site 1) are asked. Then the program moves onto Module 3 and asks all of the questions for all of the products (A and B) for the first site (site 1) before moving back to Module 2 for the next site (site 2). Once all the questions in Module 3 have been asked for all of the sites, the program will move onto Module 4. Module 4 is completed on all the products and sites before moving onto Module 5 and likewise, Module 5 is complete on all products and sites before moving onto Module 6 (however in c7b, in Module 5 is only conducted on a site basis).

FIGURE 2.1

NAVIGATION THROUGH SIX MODULES BY PRODUCT AND SITE

<table>
<tr><td rowspan="2">Module 1
Entity and Site Information

a1 – a3</td><td colspan="2">For each site:
Module 2 then Module 3 for each site before moving onto Module 4</td><td>For each site:</td><td colspan="2">For each site:</td><td rowspan="2">Module 6
Closing

d1 – d3</td></tr>
<tr><td>For each product:</td><td>For each product:</td><td>For each product:</td><td>For each product:</td><td></td></tr>
<tr><td></td><td>Module 2
Product Offerings

b1a – b3</td><td>Module 3
Core Attributes

b5 – b12</td><td>Module 4
Co-Payments & Deductibles

b13 – b14</td><td>Module 5
Network and Provider Payment
c1r – c7a</td><td>Module 5
Network and Provider Payment
c7b</td><td></td></tr>
</table>

b. Deleting, Combining, and Adding Products in Module 2

Product differentiation begins in Module 2. The CATI program allows you to record whether a listed product is offered or not offered for a specific site. It also allows you to record a product that has been renamed, and whether a product should actually be combined with another product, or if it should be deleted from the list. If the product is not offered in one site, but is offered in other sites, you should select the "not offered" category for the first site. The product name will come up again when you ask about additional sites. If the product is not offered in ***any site*** for that entity, you should select the "delete" category. Please note that selecting the "delete" or "combine" categories will remove products entirely for all sites, so use these options only when you are certain that the product is not offered by ***any site***. This should be done at the end of the Module 2 after all sites have been interviewed. Figure 2.2 lists code responses to item B1a.

FIGURE 2.2

KEY TO HANDLING PRODUCTS AT ITEM B1A---

"DID YOU OFFER [PRODUCT] IN [SITE]?"

Code	Label	Use to indicate….
<1>	Offered	[Product] was offered in [site] during the Summer 1998, either through employers or individual purchasers
<2>	Offered, rename	Same as <1>, but change the name according to the Respondent
<3>	Not Offered	[Product] was **not** offered in [site] during Summer 1998
<4>	Combine	[Product] is the same as another on our list, except for name and minor variations, such as benefit limits, eye or dental coverage; do not combine with another if it differs in ways covered in **core attributes, co-payments and deductibles,** or **provider payment modules**
<5>	Delete	Respondent does not recognize [product] as associated with entity

c. Product Splits Within the Core Attribute Section in Module 3

The questions in Module 3 ask about core attributes for each product, including issues such as out-of-network coverage. According to this study's definition, a "product" must have uniform answers for items b5, b6, b8, and b10 for all contacts and enrollees within that product for a specified site. After each question, you will be asked if the response is true for all enrollees in the site. If it is not true for all enrollees, it cannot be considered one product. This event should be rare, but it if happens you have two alternatives:

- The CATI program will instruct you to split out a separate product. The original product will retain the value you already recorded for the item. An additional product will be added. You will be prompted to name the new product. The program will then copy all product data recorded to that point, and then return you to the original "parent" product that launched the split. On-screen instructions will allow you to skip all previously answered questions. Later you will return to finish other items for the newly created "child" product. The instrument will let you know when you are switching from one product to another.

- "Un-combine" this and another product that were previously combined at b1a. You may wish to wait to un-combine until after you have finished the current item block for the current product. Immediately, however, change your answer to the current item verifying that all contracts are the same to "yes".

d. Copying Product Data Across Sites

Once you have completed entering product information for the first site, the CATI program will ask you in Module 2 if certain answers are the same for this product later. If the information is the same, answers from Modules 2 and 3 will be copied for the current site, and you will go on to the next product. If not the same, you will repeat the questions for the products in the new site.

- Item b1d will prompt immediately after item b1a – [product] offered in [site]. The program will first check whether the product has been offered in another site; if not, this sequence will be skipped.

- You will then be asked which site to copy from. You may only copy from a site where the product is offered, and the program checks which sites are eligible.

- Some items in the inventory block will be copied, but b2a and b3a will be asked anyway, as they are site specific. Most data from the core attributes block will also be copied, without further confirmation.

- You cannot copy data from one product to another within a site, or from a different product in another site.

- Once you have entered or copied data for a product, you will not be able to copy again. Do not attempt any irregular movement within this copy sequence.

- Products that are copied can later be edited using myedit function

.

e. The MyEdit Function

In addition to splitting and combining products within questions, the program has an editing feature that allows you to change the status of any product or site. You access the ***MyEdit function*** by pressing the F6 key. This will take you to a series of editing screens. The first screen is a menu with items for: (1) editing, deleting or combining products; (2) adding a product; (3) editing specific site or product data; or (4) leaving the editing program and returning to the question you last encountered. Select the number for the function you want to use. This will move you to the next edit screen. For example, if you select (1), *editing, deleting or combining products* you will move to a screen which lists all of the products in the first column. The delete status for each of those products is shown in the second column, and the combine status for each product is shown in the third column. Use the arrow keys to move the cursor to the product you want to edit, and the item you want to change. All products that have not been deleted will show a default "n" in column two. Change the "n" to an "x" to delete this product. All products not combined will have a default "o" in column three. Change the "o" to the number of the product you want to combine this product with. When you are done editing from this screen, type "d", as directed by the screen instructions. This will take you back to the main *MyEdit function* menu. If you select (2) *adding a product* you will move to an edit screen that

asked for the name of the product, and then appends that product to all sites that the entity operates so that product attribute questions can be asked. If you select (3) *edit specific site/product data,* one of three screens can be selected allowing you to input the site, product and items you wish to edit. If you select (4) *editing complete,* you will leave the editing function and return to the last substantive question you encountered. **NOTE:** It is important to press the Control F key after leaving the *MyEdit function*. This will move the program through all answered questions, changing the data to reflect your editing action and take you to the last unanswered question.

SCREENER--INTRODUCTION

>ia0< Hello, is this [entity]? Could I speak to the Director of Marketing?

PROBE: to someone in charge of group health insurance products and contracts?
<1> CONTINUE [goto ia1a]
<2> WRONG NUMBER [goto ia3a]
<3> DOES NOT MARKET OR ADMINISTER HEALTH PLANS [goto ia3a]
<4> CALL BACK [goto cb][3]
<5> REFUSES [goto ia2a]
<6> NEED TO CODE FINAL STATUS (** FOR FINAL STATUS CODES ONLY) [goto fdis][4]

>ia1a< *Use for entities that are being interviewed for the first time in Round 2*
Hello. My name is _____, calling on behalf of the Robert Wood Johnson Foundation. We are conducting a nationwide study of health plans and organizations, and we'd like your organization to participate in a brief survey. The purpose of the study is to track the local-level rapid changes that are going on in the health care industry. In our household survey we recently spoke to people in your area who said they have health care coverage through your organization. We know how busy you are, and we would like to send you our final report in appreciation for your help with the study.

or

Use for entities that were interviewed in Round 1
Hello. My name is_____, calling on behalf of the Robert Wood Johnson foundation. I want to thank (you and) your organization for participating in our survey of health plans for the Community Tracking Study. A few months ago, we sent (you/your organization) reports on the results of the first round of our survey and background on the project. We hope you found the reports and other information available on our research web site (www.hschange.com) helpful.
Since one of the most important objectives of our study is tracking change, we would like to update information (you/your organization) gave us two years ago. We have made many changes to our survey and it shouldn't take much time to complete the telephone interview. Once again, we will provide you with customized reports for your (region/community), as well as updates on other reports we are producing on changes in the health care system.

<1> CONTINUE WITH INTERVIEW [goto a1]
<2> MORE INFO [goto ia2a]
<3> DOES NOT MARKET OR ADMINISTER HEALTH PLANS [goto ia3a]
<4> CALL-BACK [goto cb]
<5> REFUSES [goto ia2a]

[3] See page 50 for data element >cb<.
[4] See page 51 for data element >fdis<.

>ia2a< MORE INFO: DO NOT READ

* Refer to posted materials for more information about the study.

* If permission from a higher corporate office is required, record corporate information on contact sheet.

* Use contact sheet to record new information about respondents.

<1> CONTINUE WITH INTERVIEW [goto a1]
<2> NEW RESPONDENT TO PHONE [goto ia1a]
<3> DOES NOT MARKET OR ADMINISTER HEALTH PLANS [goto ia3a]
<4> CALL-BACK [goto cb]
<5> REFUSES [goto cb]

>ia3a< DOES NOT MARKET OR ADMINISTER HEALTH PLANS, OTHER WRONG NUMBERS: READ AS NEEDED.

Confirm phone number; is there any other organization in the area with similar name? Get contact info.

If they are eligible (according to a1) and can answer our questions about health plan features then they are qualified.

FOR PROVIDER ORGANIZATIONS, STATE AGENCIES, OTHER PLAN SPONSORS: Is your organization affiliated with another organization that does provide or administer basic medical health care plans? Get contact info.

<1> CONTINUE WITH INTERVIEW [goto a1]
<2> NEW RESPONDENT TO PHONE [goto ia1a]
<3> DOES NOT MARKET OR ADMINISTER HEALTH PLANS [goto cb]
<4> CALL-BACK [goto cb]
<5> REFUSES [goto cb]
<7> WRONG NUMBER, OR NEW CONTACT SUPPLIED [goto cb]

MODULE 1: ENTITY AND SITE INFORMATION

>a1< NAME: [Fill Entity Name]
ID: [Fill Entity Identification Number]

YOU SHOULD NOW BE TALKING TO A "REAL" RESPONDENT

Before we begin, I want to confirm that your organization did offer or administer basic medical health care plans at any time since the summer of 1998?

PROBE: Exclude specialty-only health plans (such as cancer only), workers' compensation, supplemental and pharmacy only plans, military facilities, free clinics, and individual providers' offices.

<1> YES
<2> NO [goto ia3a]

>a2_pre< [If the number of sites equals 1, then goto a2a]

>a2< In our household survey, we interviewed people from the following geographic areas who said they were enrolled in one of your health plans since the summer of 1998:
[SITE(S)]

Are you able to answer questions about your health plan products in all of these areas, or are there some areas you do not cover or cannot answer about?

<1-n[5]> SITE NUMBER TO EXCLUDE [goto PROBE]
<g> ALL/REST OKAY [goto a3a]
<d> DISCONTINUE, NEW CONTACT [goto cb]

>PROBE< WHAT'S THE PROBLEM WITH SITE: [SITE]

Is there a different person or office that handles accounts in [SITE], or is it that your organization doesn't offer products in that area?

<1> DIFFERENT PERSON [goto a2t]
<3> NO PRODUCTS OFFERED [if there are no other sites to exclude from a2, goto a3a]
<4> RECIPROCATING ENTITY (RECORD ON PROBLEM SHEET) [goto a3a].

[5] Last site for the entity.

>a2t< Could you give me the name/address/telephone number of a contact person for . . . [SITE]

RECORD CONTACT INFO ON REFERRAL FORM

<g> CONTINUE [goto a3a]

>a2a< In our household survey, we interviewed people from [SITE(S)]

who said they were enrolled in one of your health plans.

Are you able to answer questions about your health plan products in that area?

<1> YES, CONTINUE [goto a3a]
<0> NO, GET NEW CONTACT INFO [goto cb]

>a3a< ASK IF NOT ALREADY KNOWN, OR CODE
Are you a Blue Cross/Blue Shield Plan?

<1> YES [goto b_intro]
<2> NO

>a3< Please tell me which of the following categories best describes your organization:

<2> A licensed insurer or HMO
<3> A managed care provider organization, such as a PPO or IPA (not licensed to sell insurance)
<4> A Third Party Administrator (TPA)
<6> An employer, union or trust plan administrator (including a government employee plan)
<8> Or something else [SPECIFY][6]

INTERVIEWER: USE STATUS OF CORPORATE PARENT IF APPLICABLE

[6] See page 49 for data element >SPECIFY<

MODULE 2: PRODUCT OFFERINGS

b-intro In this part of the interview we'd like to establish what types of health insurance products you offer ([if the number of sites greater than 1] in each geographic area you serve).

By "product," I mean groups of plans or contracts that are similar in how they handle network coverage and referrals. If specific plans or contracts are similar in these ways but differ on copays, deductibles, co-insurance rates, or supplemental benefits such as prescription drugs or dental care, we can consider them the same product. Examples are open-ended HMOs, PPOs without a primary care physician, and traditional indemnity plans.

<g> CONTINUE

begin with the first/next SITE for the entity

>b1a_pre< From our survey of households, we have compiled the following list of products people said they were enrolled in:

[PRODUCT(S)]

Now, I'm going to ask if you offer each of these product(s) in [SITE]. Also tell me if any name is incorrect, or if any of these are different names for the same basic product.

<g> CONTINUE

>editb1< *If a product has been deleted, combined with another product or is not offered the program will skip questions*[7]
[if p_delete equals <X> or p_combine is greater than or equals <1> goto b1_end]

>b1a< The (first/next) product is [PRODUCT]
Did you offer that in [SITE]?

<1> OFFERED [goto b2]
<2> OFFERED, BUT RENAME PRODUCT [goto newname]
<3> NOT OFFERED [goto b1_end]
<4> COMBINE WITH ANOTHER PRODUCT [goto combine]
<5> DELETE FOR ALL SITES [goto b1_end]

>newname< What do you want to call this product instead of [PRODUCT]?
[goto b1d]

[7] See page 45 for data elements >p_delete<, and >p_combine<.

>combine< What product is this the same as?
[goto b1_end]

>b1d< Is that product essentially the same as you previously described?
<1> YES
<2> NO [goto b2]

>b1d_site< READ IF NOT OBVIOUS:
Which site has the same data for [PRODUCT]?

DATA COPIED INTO TEMP SPACE. PRESS {ENTER} TO CONTINUE.

>b2< Do you think of [PRODUCT] as a(n) . . .
<1> HMO (Health Maintenance Organization)
<2> POS (Point of Service Plan)
<3> PPO (Preferred Provider Organization) [goto b3]
<4> Indemnity Plan (Traditional FFS) [goto b3]
<5> or something else? [SPECIFY] [goto b3][8]

>b2a< Which of the following describes the medical providers available in [SITE]?
PROBE: Exclude dental, mental and vision providers.

INTERVIEWERS: ASK FOR INDIVIDUAL COMPONENTS OF "MIXED MODEL." SELECT ALL THAT APPLY
<1> Staff model
<2> Group model (plan contracts with a single group)
<3> Network and/or IPA (contracts with multiple individual and/or group providers) or
<4> something else [SPECIFY][6]
<d> DON'T KNOW
<r> REFUSED
<n> NO MORE CODES
<x> DELETE A CODE [goto xb2a] (*used to correct errors for multiple response questions*)
[go to b3]

>xb2a< THIS SCREEN IS TO DELETE A RESPONSE.
<1> Staff model
<2> Group model (plan contracts with a single group)
<3> Network and/or IPA (contracts with multiple individual and/or group providers) or
<n> NO CODES TO DELETE
[goto b2a]

[8] See page 49 for data element >SPECIFY<.

>b3< Is [PRODUCT] ever sold to individuals in [SITE]?
<1> YES
<2> NO
<d> DON'T KNOW
<r> REFUSED

>b1_end< *go back to b1a to the next product, if there are no more products, then goto b1c.*

>b1c< Do you offer any other products in [SITE]?
<1> YES [pname]
<2> NO [goto b5s_pre]

>pname< What is the name of the product?
Product name is stored.

>pre_sure< YOU HAVE TO DO B1a FOR THESE PRODUCTS / SITES NOW
<g> to continue

>sure< *This item should prompt the interviewer to jump forward. A jump forward occurs when the interviewer has added a new product or in some way changed the path. After the 'jf' the path will correct and the data will be correct as well.*

ENTER g TO JUMP FORWARD. YOU MAY ENCOUNTER NEW QUESTIONS BEFORE RETURNING TO ORIGINAL ITEM!
[If b_intro is null goto b_intro else go to the next blank question]

MODULE 3 - CORE ATTRIBUTES OF PRODUCTS

>b5s_pre< *begin with the first/next PRODUCT for the SITE*

>editb5< *If a product has been deleted, combined with another product or is not offered the program will skip questions*[9]
[if p_delete equals <X> or p_combine is greater than or equals <1> or offer_flag equals <0> goto b9_end_real]
[if SPLIT equals <0> goto b5s_end]
[if SPLIT less than <0> goto b5s]
[if SPLIT greater than <0> goto child_beg]
[goto b5s_end]

>b5s< We just finished adding a new product. We'll discuss that product again later.
Now, let's return to the product you called [PRODUCT]
That was a(n) [PRODUCT TYPE].
<g> CONTINUE
[if NET eq <0> goto b9_end]
[if SPLIT eq <-8> goto b8b_end]
[if SPLIT eq <-10> goto b10b_end]
[if SPLIT eq <-6> goto b6b_end]
[if SPLIT eq <-5> goto b5nt]
[goto b9_end]

>child_beg< Now let's continue with the product you called [PRODUCT].
We know some information about this product from a previous product, but we'll review those items.
<g> CONTINUE

>b5s_end< [if b1d equals <1>][goto b9_end]

>b5_pre< [if b2 is less than or equals <3> goto b5nt]
[if SPLIT is greater than or equals <5> goto b5_split]
[goto b5]

>b5_split< When we added it, we agreed that [PRODUCT] did [if b5 equals <2>]NOT[endif] have a network.

<1> ACCEPT [goto b5nt]
<2> CHANGE IT ANYWAY [goto b5]

>b5s_end< [if b1d equals <1> goto b9_end][goto b5_pre]

[9] See page 44 for data elements >p_delete<, >p_combine<, and >offer_flag<, and page 50 for data element >SPLIT<.

>b5_pre< [if b2 is less than or equals <3> goto b5nt][if SPLIT is greater than or equals <5> goto b5_split][goto b5]

>b5_split< When we added it, we agreed that [PRODUCT] did ([if b5 equals <2> NOT] have a network

<1> ACCEPT [goto b5nt]
<2> CHANGE IT ANYWAY

>b5< Is there a directory or list of doctors associated with [PRODUCT] in [SITE]?

<1> YES
<2> NO

>b5a< Does that apply to all contracts and enrollees under [PRODUCT] in [SITE]?

<1> YES [got to b5nt]
<2> NO

>b5b< We need to treat contracts where there is ([if b5 equals <1>] not) a network of participating providers as a separate product.

<1> CONTINUE
<2> IGNORE, GO ON
<d> DON'T KNOW
<r> REFUSED

>b5b_end< [if b5b eq <1> [goto append][10]

>b5nt< [if b2 is less than or equals <3> or b5 equals <1>, store <1> in NET, otherwise store <0> in NET][11]
[if NET equals <0> goto b9_end]
[if SPLIT is greater than or equals <6> goto b6_split][12]
[goto b6]

>b6_split< When we added it, we agreed that [PRODUCT] did ([if b6 equals <2>] not) cover out-of-network office visits without referrals.

<1> ACCEPT [goto b6b_end]
<2> CHANGE IT ANYWAY

[10] See page 47 for data element >append<.
[11] See page 47 for data element >NET<.
[12] See page 47 for data element >SPLIT<.

>b6< Under [PRODUCT] in [SITE], if enrollees do not have a referral and goto **out-of-network** doctors, does the plan cover **any** of the costs for these visits?

PROBE: Exclude emergency care and non-major medical services such as dental and vision care.

PROBE: By "out-of-network" we mean providers of major medical services **NOT** associated with [PRODUCT].

<1> YES
<2> NO
<7> THERE IS NO NETWORK IN THIS SENSE
<d> DON'T KNOW[13]
<r> REFUSED[11]

>b6_end< [if b6 equals <7>][store <0> in NET] [goto b9_end][else goto b6a].
The program skips over network questions if b6 response is "there is no network".

>b6a< Does that apply to all contracts and enrollees under [PRODUCT] in [SITE]?

<1> YES [goto b6b_end]
<2> NO
<d> DON'T KNOW[14]
<r> REFUSED[12]

>b6b< We need to treat contracts where out-of-network visits without referrals are ([if b6 equals <1>] **not**) covered as a separate product.

<1> CONTINUE
<2> IGNORE, GO ON
<d> DON'T KNOW
<r> REFUSED

>b6b_end< [if b6b equals <1> goto append]
[if SPLIT is greater than or equals <8> goto b10_split]
[goto b10]

>b10_split< When we added it, we agreed that [PRODUCT] did ([if b10 equals <2>]**NOT**) require a PCP.

<1> ACCEPT
<2> CHANGE IT ANYWAY [goto b10]

[13] In final CATI program, <d> and <r> are set to goto the next question, b6a. If this program is to be used again a more appropriate skip for <d> and <r> responses would be goto b10.

[14] In final CATI program, <d> and <r> are set to goto the next question. If this program is to be used again a more appropriate skip at b6a for <d> and <r> responses should be b6b_end.

>b10_sp_end< [if b10_split equals <1>]
[if b10 equals <1> goto b12a_pre]
[goto b8_split_beg]

>b10< Does [PRODUCT] in [SITE] require members to have a primary care doctor, group of doctors, or clinic to receive maximum coverage for all routine care?
<1> YES
<2> NO
<d> DON'T KNOW[15]
<r> REFUSED[13]

>b10a< Does that apply to all contracts and enrollees under [PRODUCT] in [SITE]?
<1> YES [got to b10b_end]
<2> NO
<d> DON'T KNOW[13]
<r> REFUSED[13]

>b10b< We need to treat contracts where primary care providers are ([if b10 equals <1>] **not**) required as a separate product.
<1> CONTINUE
<2> IGNORE, GO ON
<d> DON'T KNOW
<r> REFUSED

>b10b_end< [if b10b equals <1> goto append]
>b12a_pre< [if b10 does not equal<1> goto b8_split_beg]

>b12< ([if SPLIT is greater than or equals <5> and b12 is not null]
I need to confirm….) *This question stem is asked anytime a product has split*
Which types of providers can serve as primary care physicians for enrollees in this product?

INTERVIEWERS: CODE ALL THAT APPLY
<1> Generalists, such as internist, pediatrician, or family practice
<2> OB/GYNs
<3> Other specialists
<n> NO MORE CODES
<x> DELETE A CODE [goto xb12]
<d> DON'T KNOW
<r> REFUSED
[goto b8_split_beg]

[15] In final CATI program, <d> and <r> are set to goto the next question. If this program is to be used again a more appropriate skip at b10a and b10b for <d> and <r> responses would be b10b_end.

>xb12< THIS SCREEN IS TO DELETE A RESPONSE.

<1> Generalists, such as internist, pediatrician, or family practice
<2> OB/GYNs
<3> Other specialists
<n> NO CODES TO DELETE
[goto b12]

>b8_split_beg< [if SPLIT equals <8> goto b8_split][goto b8]
>b8_split< When we added it, we agreed that [PRODUCT] did ([if b8 equals <2>] NOT ever[else] sometimes[endif]) require referrals for in-network office visits.

<1> ACCEPT [goto b8b_end]
<2> CHANGE IT ANYWAY

>b8< We are interested in whether referrals are required for specialty care, and how they affect coverage, under [PRODUCT] in [SITE]. For these questions, please consider only major medical services, but not emergency care and other services such as dental, vision, and mental health care.
Under [PRODUCT] in [SITE], is a referral or authorization **ever** required to obtain maximum coverage for an initial visit to an in-network specialist?

PROBE: If specialists can arrange authorization on-the-spot or after the visit, consider this a requirement to get a referral.

<1> YES
<2> NO
<d> DON'T KNOW[16]
<r> REFUSED[14]

>b8a< Does that coverage rule apply to all contracts and enrollees under [PRODUCT] in [SITE]?
<1> YES [goto b8b_end]
<2> NO
<d> DON'T KNOW[14]
<r> REFUSED[14]

>b8b< We need to treat contracts where referral is/is not required as a separate product.

<1> CONTINUE
<2> IGNORE, GO ON
<d> DON'T KNOW
<r> REFUSED

[16] In final CATI program, <d> and <r> are set to goto the next question. If this program is to be used again a more appropriate skip at b8 and b8b for <d> and <r> responses would be b8b_end.

>b8b_end< [if b8b equals <1> goto append]
[if b8 does not equal<1> goto b9_end]

>b91a< For the next few questions "self referral" refers to visits where a patient sees an in-network specialist without obtaining a referral or authorization, even through this is required to obtain maximum coverage.

([if SPLIT is greater than or equals <5> and b91a is not null]
I need to confirm…..) *This question stem is asked anytime a product has split.*

Does [PRODUCT] provide at least some coverage for self-referrals to **any** types of in-network specialists?

<1> YES
<2> NO
<d> DON'T KNOW
<r> REFUSED

>b91a_end< [if b91a does not equal<1> go to b9_end]

>b91b< ([if SPLIT is greater than or equals <5> and b91b is not null]
I need to confirm…..) *This question stem is asked anytime a product has split.*

Does this coverage for self-referral apply to most types of in-network specialists?

<1> YES [goto b92]
<2> NO
<d> DON'T KNOW
<r> REFUSED

>b91c< ([if SPLIT is greater than or equals <5> and b91c is not null]
I need to confirm…..) *This question stem is asked anytime a product has split.*

Does this coverage for self-referral apply to **most** visits to in-network OB/GYNs?

INTERVIEWER: COVERAGE OF ONE ANNUAL VISIT DOES NOT COUNT.

<1> YES
<2> NO
<d> DON'T KNOW
<r> REFUSED

>b91d< ([if SPLIT is greater than or equals <5> and b91d is not null]
I need to confirm…..) *This question stem is asked anytime a product has split.*

Does this coverage for self-referral apply to **any other** types of in-network specialists?

<1> YES
<2> NO
<d> DON'T KNOW
<r> REFUSED

>b92< ([if SPLIT is greater than or equals <5> and b92 is not null]
I need to confirm…..) *This question stem is asked anytime a product has split.*

When [PRODUCT] covers in-network self referrals, is the level of coverage the same as with a physician referral, or is it less than the coverage with a physician referral?

PROBE: Lesser coverage means that the copayment or coinsurance rate that the enrollee pays is **higher**.

<1> SAME
<2> LESS
<3> VOLUNTEER: Varies
<d> DON'T KNOW
<r> REFUSED

>b92t< [if (b92 equals <3> and (b91b equals <1> or b91c <1>)) go to b92b]
[goto b93_pre]

>b92b< ([if SPLIT is greater than or equals <5> and b92b is not null]
I need to confirm…..) *This question stem is asked anytime a product has split.*

What about in-network self-referrals to OB/GYNs - is the level of coverage the same as with a physician referral, or less than with a physician referral?

PROBE: Lesser coverage means that the co-payment or coinsurance rate that the enrollee pays is **higher**.

<1> SAME
<2> REDUCED
<d> DON'T KNOW
<r> REFUSED

>b93_pre< [if b6 equals <1> and (b92 equals <2> or b92 equals <3>) goto b93]
[goto b9_end]

>b93< ([if SPLIT is greater than or equals <5> and b93 is not null]
I need to confirm…..) *This question stem is asked anytime a product has split.*

Under [PRODUCT] when the level of coverage for in-network self-referrals is reduced, is that level of coverage better than for out-of-network self-referrals, or the same?

<1> SAME
<2> LESS
<3> VOLUNTEER: varies
<d> DON'T KNOW
<r> REFUSED

>b93t< [if b93 equals <3> and (b91b equals <1> or b91c equals <1>) goto b93b
[goto b9_end]

>b93b< ([if SPLIT is greater than or equals <5> and b93b is not null]
I need to confirm…..). *This question stem is asked anytime a product has split.*

What about in-network self-referrals to OB/GYNs--Is that level of coverage better than for out-of-network self-referrals, or the same?

<1> BETTER
<2> SAME
<d> DON'T KNOW
<r> REFUSED

>b9_end<

WE ARE DONE WITH INVENTORY/SPLIT SECTION FOR THIS PRODUCT.
PRESS ENTER TO CONTINUE

>b9_end_real< *goto b5s_pre and ask Module 3 questions for the next PRODUCT on the list, if there are no more PRODUCTS for that SITE got to b9s_end At this point a negative value in SPLIT is changed to a positive value.*

>b9s_end< *goto b1a and ask Module 2 questions for the next SITE on the list if there are no more SITES goto b13_intro.*

MODULE 4: CO-PAYMENTS AND DEDUCTIBLES

b13_intro Now I would like to ask some questions about co-payments and deductibles for each product you offer in each of the sites we have discussed.

<g> CONTINUE

>b13ros< *begin with the first/next SITE for the entity*
begin with the first/next PRODUCT for the SITE

>editb13< *if a product has been deleted, combined with another product or is not offered the program will skip questions on copayment and deductibles.*
[if p_delete equals <X> or p_combine is greater than or equals <1> or offer_flag equals <0> goto bp_end][17]

>b13< Does [PRODUCT] in [SITE] have a fixed co-payment per visit, or percentage co-insurance payment for [if NET equals <1>]in-network [endif] office visits?[18]

PROBE: Whichever is most common for enrollees in this product?

<1> CO-PAYMENT [goto b13amt]
<2> CO-INSURANCE RATE [goto b13per]
<0> NONE
<d> DON'T KNOW[19]
<r> REFUSED[17]
[goto b13out_pre]

>b13amt< What is the typical co-payment amount per office visit for [PRODUCT] in [SITE]?

PROBE: The lowest co-payment that typically applies for [if NET equals <1>] in-network [endif] office visits with referrals. Exclude "well" visits if these are different.

<1-2000> dollars
<d> DON'T KNOW [20]
<r> REFUSED[18]

>b13amt_end<
[if b13amt is less than or equals <50> goto b13out_pre, else goto b13ch1]

---

[17] See page 45 for data elements >p_delete<, >p_combine<, and >offer_flag<.

[18] See page 47 for data elemert >NET<.

[19] In final CATI program, <d> and <r> are treated as negative numbers, therefore at >b13ch2_end<, <d> and <r> meet the criteria to goto b13out_pre.

[20] In final CATI program, <d> and <r> are treated as negative numbers, therefore at >b13amt_end< , <d> and <r> meet the criteria to goto b13out_pre.

>b13ch1<

I just heard you say that the co-payment for [in-network] office visits is [b13AMOUNT], is that correct?

PROBE: Is that typical for all contracts under this product?

PROBE: The co-payment is the amount the enrollee pays for the visit.

PROBE**:** The lowest co-payment that typically applies for [if NET equals <1>] in-network [endif] office visits with referrals. Exclude "well" visits if these are different.

<1> CORRECT [goto b13out_pre]
<2> NO, GO BACK AND CHANGE [goto b13amt]
<d> DON'T KNOW[21]
<r> REFUSED[19]

>b13per< What is the typical coinsurance percentage for office visits under [PRODUCT] in [SITE]?

PROBE: The coinsurance rate is the percentage for which the enrollee is responsible.

PROBE: The lowest coinsurance that typically applies for [if NET equals <1>] in-network [endif] office visits, with referrals. Exclude "well" visits if these are different.

<1-100> percent
<d> DON'T KNOW[22]
<r> REFUSED[20]

>b13ch2_end<

[if b13per is less than or equals <50> goto b13out_pre, else, goto b13ch2]

[21] In final CATI program, <d> and <r> are set to goto the next question, b13per instead of b14. If this program is to be used again the skip for <d> and <r> response should be changed to goto b14.

[22] In final CATI program, <d> and <r> are treated as negative numbers, therefore at >b13ch2_end<, <d> and <r> meet the criteria to goto b13out_pre.

>b13ch2< I just heard you say that the coinsurance rate for in-network office visits is **[b13PERCENT]**, is that correct?

PROBE: Is that typical for all contracts under this product?

PROBE: The coinsurance rate is the percentage for which the enrollee is responsible.

PROBE: The lowest coinsurance rate that typically applies for ([if NET equals<1>] in-network) office visits with referrals. Exclude "well" visits if these are different.

<1> CORRECT
<2> NO, GO BACK AND CHANGE [goto b13per]
<d> DON'T KNOW[23]
<r> REFUSED[21]

>b13out_pre< [if b6 equals <2> or NET equals <0> goto b14]

>b13out< For out-of-network office visits without a referral does [PRODUCT] in [SITE] have a fixed co-payment per visit, or percentage coinsurance payment?

PROBE: Whichever is most common for enrollees in this product?

<1> CO-PAYMENT [goto b13od]
<2> COINSURANCE [goto b13op]
<0> NONE
<d> DON'T KNOW
<r> REFUSED
[goto b14]

>b13od< What is the typical co-payment amount for out-of-network office visits under [PRODUCT] in [SITE]?

PROBE: The co-payment that typically applies for office visits without referrals, outside of any network.

<1-2000> dollars
<d> DON'T KNOW[24]
<r> REFUSED[22]

>b13od_end< [if b13od is less than or equals <50> goto b14]

[23] In final CATI program, <d> and <r> are set to goto the next question, b13out_per and onto b13_out instead of b14. If this program is to be used again the skip for <d> and <r> response should be changed to b14.

[24] In final CATI program, <d> and <r> are treated as negative numbers, therefore at >b13od_end<, <d> and <r> meet the criteria to goto b14.

>b13ch3< I just heard you say that the co-payment for out-of-network office visits is **[b13odAMOUNT]**, is that correct?

PROBE: Is that typical for all contracts under this product.
PROBE: The co-payment is the amount the enrollee pays for the visit.
PROBE: The co-payment that typically applies for office visits, without referrals, outside of any network.

<1> CORRECT [goto b14]
<2> NO, GO BACK AND CHANGE [goto b13od]
<d> DON'T KNOW[25]
<r> REFUSED[23]

>b13op< What is the typical coinsurance percentage for out-of-network office visits under [PRODUCT] in [SITE]?

PROBE: The coinsurance rate is the percentage for which the enrollee is responsible.

PROBE: The coinsurance that typically applies for office visits, without referrals, outside of any network?

<1-100> percent
<d> DON'T KNOW[26]
<r> REFUSED[24]

>b13op_end<

[if b13op is less than or equals to <50> goto b14]

>b13ch4< I just heard you say that the coinsurance for out-of-network office visits is **[b13opAMOUNT]**, is that correct?

PROBE: Is that typical for all contracts under this product.
PROBE: The coinsurance is the amount the enrollee pays for the visit.
PROBE: The coinsurance that typically applies for office visits, without referrals, outside of any network.
<1> CORRECT
<2> NO, GO BACK AND CHANGE [goto b13op]
<d> DON'T KNOW
<r> REFUSED

[25] In final CATI program, <d> and <r> are set to goto the next question, b13op instead of b14. If this program is to be used again the skip for <d> and <r> response should be changed to b14.

[26] In final CATI program, <d> and <r> are treated as negative numbers, therefore at >b13op_end< <d> and <r> meet the criteria to goto b14.

>b14< Under [PRODUCT] in [SITE], what is the dollar amount of the individual deductible that applies to [if NET equals <1>] in-network [endif]office visits?

PROBE: Your best estimate is fine. Please tell me what is **typical** for this product in [SITE]?

INTERVIEWER: Enter "0" if none.

<0-100000> dollars[27]
<d> DON'T KNOW[28]
<r> REFUSED[26]

>b14_end< [if (b14 is greater than <0> and b14 is less than <50>) or b14 is greater than <5000> goto b14ch1][goto b14out_pre]

>b14ch1< I just heard you say that the deductible for [if NET equals <1>] in-network [endif] office visits is **[b14AMOUNT]**, is that correct?

PROBE: Is that typical for all contracts under this product?

PROBE: The deductible is the annual amount the enrollee pays.

<1> CORRECT [goto b14out_pre]
<2> NO, GO BACK AND CHANGE [goto b14]
<d> DON'T KNOW
<r> REFUSED

>b14out_pre<
[if NET equals <0> or b6 does not equal <1> goto bp_end]

>b14out< Is there a **separate** deductible for [PRODUCT] in [SITE] that applies to out-of-network office visits?

<1> YES
<2> NO [goto bp_end]
<d> DON'T KNOW[29]
<r> REFUSED[27]

---

[27] The dollar range at b14 is <0-10000>. If this program is to be used again the dollar range should be revised.

[28] In final CATI program, <d> and <r> are treated as negative numbers, therefore at >b14_end<, <d> and <r> meet the criteria to goto b14out_pre.

[29] In final CATI program, <d> and <r> are set to goto the next question, b14od instead of bp_end. If this program is to be used again the skip for <d> and <r> response should be changed to bp_end.

>b14od< What is the dollar amount of the **individual deductible** for out-of-network office visits?

PROBE: Your best estimate is fine. Please tell me what is typical for this product in [SITE]?

<0-50-5000> dollars[30]
<d> DON'T KNOW[31]
<r> REFUSED[29]

>b14od_end<
[if (b14od is greater than <0> and b14od is less than <50>) or b14od is greater than <5000> goto b14ch2][gotobp_end]

>b14ch2< I just heard you say that the deductible for out-of-network office visits is **[b14od AMOUNT],** is that correct?

PROBE: Is that typical for all contracts under this product?

PROBE: The deductible is the annual amount the enrollee pays.

<1> CORRECT [goto bp_end]
<2> NO, GO BACK AND CHANGE [goto b14od]
<d> DON'T KNOW
<r> REFUSED

>bp_end< *goto b13 and ask Module 4 questions for the next PRODUCT on the list, if there are no more PRODUCTS for that SITE got to bs_end*

>bs_end< *goto b13 and ask Module 4 questions for the next SITE on the list if there are no more SITES goto c_beg.*

[30] The dollar range at b14od the dollar range is <0-50-5000>. If this program is to be used again the dollar range should be revised.

[31] In final CATI program, <d> and <r> are treated as negative numbers, therefore at >b14od_end< <d> and <r> meet the criteria to goto b14ch2. But the criteria should be set for <d> and <r> to goto bp_end.

MODULE 5: NETWORK AND PROVIDER RELATIONS

>c_beg< I now have a few questions about the providers associated with each product.

<g> CONTINUE

begin with the first/next SITE for the entity
begin with the first/next PRODUCT for the SITE

>editc1r< *if a product has been deleted, combined with another product or is not offered the program will skip questions on network and provider relations.*
[if p_delete equals <X> or p_combine is greater than or equals <1> or offer_flag equals <0> goto cp_end][if NET equals <0> goto c4_1]

>c1r< Approximately what percentage of all primary care and specialist physicians in [SITE] are associated with [PRODUCT]? Would you say? . . .

INTERVIEWER: DO NOT **PROBE DK.**
<1> fewer than 25 percent
<2> at least 25 percent but less than 50 percent
<3> at least 50 percent but less than 75 percent
<4> 75 percent or more
<d> DON'T KNOW
<r> REFUSED

>c2r< Under [PRODUCT] are enrollees limited to a single hospital system for general acute care services in [SITE]?

PROBE: A single hospital system would be one or more hospitals under the same ownership or management.

<1> YES
<2> NO
<d> DON'T KNOW
<r> REFUSED

>c4_1< Next, I have some questions about payment arrangements for primary care physicians, specialists, and hospitals for each product in [SITE]. Since this may vary somewhat depending on the provider, I just want to know what is **typical** for the providers who serve a majority of enrollees in each product.

<g> CONTINUE

>c4_2< In [PRODUCT] in [SITE], what is the typical method of payment that your organization uses for primary care services? Is it? . . .

PROBE: By that I mean how your organization pays individual providers, medical groups, or other entities for primary care services in [SITE].

PROBE: Capitation is a fixed payment per member per month for a class of services.

INTERVIEWER: Probe carefully between <1> and <2>.

<1> Fee-for-service, for example, usual and customary rates
<2> Fixed fee schedule, including discounted FFS or relative value units
<3> Salaried by your organization, or . . .
<4> Capitation (includes combined "professional" or "global" capitation)
<5> OTHER (SPECIFY) [SPECIFY] END WITH // [32]
<d> DON'T KNOW
<r> REFUSED

>c4a_pre< [if c4_2 does not equal <4> goto c5]

>c4a< What **other** services are included in this capitated payment?

INTERVIEWER: CODE ALL THAT APPLY

<1> Referrals to specialists
<2> Hospitalizations
<3> Other services
<n> NO MORE CODES/NONE OF THESE
<x> DELETE A CODE [goto xc4a]
<d> DON'T KNOW
<r> REFUSED
[goto c5_pre]

>xc4a< THIS SCREEN IS TO DELETE A RESPONSE.

<1> Referrals to specialists
<2> Hospitalizations
<3> Other services
<n> NO CODES TO DELETE
[goto c4a]

>c5_pre< [if c4a equal <1> goto c6_pre]

[32] See page 49 for data element >SPECIFY<.

>c5< In [PRODUCT] in [SITE] what is the typical method of payment that **your** organization uses for specialty services. Is it? . . .

PROBE: By that I mean how **your** organization pays individual providers, medical groups, or other entities for specialty services in [SITE].

PROBE: Capitation is a fixed payment per member per month for a class of services.

<1> Fee-for-service, for example, usual and customary rates
<2> Fixed fee schedule, including discounted FFS or relative value units
<3> Salaried by your organization, or . . .
<4> Capitation
<5> OTHER (SPECIFY) [SPECIFY] END WITH // [33]
<d> DON'T KNOW
<r> REFUSED

>c6_pre< [if c4a equals <2> goto c7t][34]

>c6< In [PRODUCT] in [SITE] what is the typical method of payment for hospital services? Is it? . . .

PROBE: By that I mean how **your** organization pays for hospital services in [SITE]. Exclude physician services delivered during the hospital stay.

PROBE: Capitation is a fixed payment per member per month for a class of services.

<1> Per diem
<2> According to DRG or per stay
<3> Capitation
<4> Billed charges, or discounted billed charges, or
<5> Something else (SPECIFY) [specify] END WITH //
<7> NOT APPLICABLE; HOSPITALS OWNED BY ORGANIZATION
<d> DON'T KNOW
<r> REFUSED

[33] See page 49 for data element >SPECIFY<.

[34] The final CATI program reads, if c4a equals <2> goto c7t. The program should have read, if c4a equals <2> goto c7.

>c7< Does [PRODUCT] in [SITE] ever include any mental health and/or substance abuse services?

PROBE: Includes "chemical dependency."

PROBE: Includes mental health or substance abuse services you provide by subcontract only if your organization administers that benefit.

<1> YES
<2> NO
<d> DON'T KNOW
<r> REFUSED

>c7t< [if c7 does not equal <1> goto cp_end]

>c7a< Are mental health and/or substance abuse services ever provided or managed separately by a specialty managed behavioral health organization?

<1> YES
<2> NO
<d> DON'T KNOW
<r> REFUSED

>cp_end< *goto c_beg and ask Module 5 questions for the next PRODUCT on the list, if there are no more PRODUCTS for that SITE got to c7b.*

>c7a_any< [if p_delete does not equal <X> or p_combine equals <0> or offer_flag equals <1>][if c7a equals <1>][store <1> in c7a_any][endif]
[if c7a_any equals <0> goto cs_end]

>c7b< Please tell me the name of the specialty managed behavioral health organization you use in [SITE]?

SELECT ALL THAT APPLY

<1> American Psych System
<2> Healthcare Value Mgt.
<3> MAGELLAN Behavioral Health
<4> Managed Health Network
<5> MAPSI Mid Atlantic Psych Services
<6> Private Health Care Systems (PHCS)
<7> Pro Behavioral Health Plan
<8> Sagamore
<9> Something else [goto c7c]
<d> DON'T KNOW
<r> REFUSED
<n> NO MORE CODES
<x> DELETE A CODE [goto xc7b]
[goto cs_end]

>xc7b< THIS SCREEN IS TO DELETE A RESPONSE.
<1> American Psych System
<2> Healthcare Value Mgt.
<3> MAGELLAN Behavioral Health
<4> Managed Health Network
<5> MAPSI Mid Atlantic Psych Services
<6> Private Health Care Systems (PHCS)
<7> Pro Behavioral Health Plan
<8> Sagamore
<9> Something else: SPECIFY WHICH ORGANIZATION TO DELETE [specify] [35]
<n> NO CODES TO DELETE
[goto c7b]

>c7c< [if c7cName is not null goto c7c2]

RECORD NAME IF NOT LISTED
NAME: >c7cName<

In what city and state is this specialty behavior health company located?
CITY:
STATE:

[35] See page 49 for data element >SPECIFY<.

>c7c2< [if c7c2Name is not null goto c7c3][36]
RECORD NAME IF NOT LISTED

NAME: >c7c2Name<

In what city and state is this specialty behavior health company located?

CITY:
STATE:
[goto tc7b]

>c7c3< RECORD ADDITIONAL OTHERS ON PROBLEM SHEET. PRESS ENTER TO CONTINUE

>cs_end< *Goto c_beg and ask Module 5 questions for the next SITE on the list, if there are no more SITEs got to d1_pre.*

[36] c7c2 may be an unnecessary data element

MODULE 6: CLOSING

>d1_pre< [if a3a does not equal <1> and a3 does not equal <2> and <3> and <5> goto d2]
>d1< I just have a few final questions about your organization . . .

What is your organization's tax status? Is it? . . .

INTERVIEWER: CODE ORGANIZATIONS WITH A 501(c)3 or 501(c)4 TAX STATUS AS NON-PROFIT
<1> for-profit, privately held
<2> for-profit, publicly held, or . . .
<3> nonprofit
<4> OTHER (SPECIFY) [SPECIFY] END WITH // [37]
<d> DON'T KNOW
<r> REFUSED

>d2< Is your organization a division or subsidiary of another health plan organization?

<1> YES [goto d2a]
<2> NO
<d> DON'T KNOW
<r> REFUSED
[goto d3]

>d2a< Is this parent company a national or multi-state organization?
<1> YES [goto d2b]
<2> NO
<d> DON'T KNOW
<r> REFUSED
[goto c7a_ent]

>d2b< What is the name of that parent company?
NAME:
<s> scroll [nat_codes][38]
<o> other [goto d2c]
<d> DON'T KNOW [goto c7a_ent]
<r> REFUSED [goto c7a_ent]
or ENTER CODE a between <01 – 66>[39]

[if code is greater than or equal to <01> and less than or equal to <66>, goto c7a_ent]
[goto d2b]

---

[37] See page 49 for data element >SPECIFY<.
[38] See page 44 for data element >nat_codes< -- a list of National Parent Company Names.
[39] See page 43, the codes (>nat_codes<) and their associated National Parent Company Names are listed

>d2c< In what city and state is this parent company located?

NAME:
CITY:
STATE:
[goto c7a_ent]

>d3< Is your organization a national or multi-state organization?

<1> YES
<2> NO
<d> DON'T KNOW
<r> REFUSED

>c7a_ent< [If no sites were deleted and c7a equals <1>, store <1> in c7a_ent].
[if c7a_ent equals <0> goto d5]

>d4< **IF ANY PRODUCT COVERS MENATL HEALTH SERVICES:**
Finally, may I have the name and phone number of the person within your organization who could answer questions about mental health and/or substance about benefits?

PROBE: I'd like the name of someone within your organization, not at the managed behavioral health organization.
NAME:
PHONE NUMBER:
ORGANIZATION:

>d5< Finally, in order to send you our report on this study, may I have your name, title and mailing address.
NAME:
TITLE:
ORGANIZATION:
STREET ADDRESS OR PO BOX:
CITY:
STATE:
ZIP CODE:
[goto done_pre][40]

[40] See page 50 for data element >done_pre<

>nat_codes< *List of National Parent Company Names*

01 Admar Corp. Med Network
02 Aetna Life Insurance Co.
03 Aetna Services Inc. (Aetna Health Plans -- managed care)
04 Allstate Life Insurance
05 AMERICAID, Inc.
06 American HMO
07 American Medical Security, Inc.
08 AmeriChoice Corp.
09 AmeriHealth, Inc.
10 Anthem Health Plans
11 Apex Health Care, Inc.
12 Beech Street Corp.
13 Blue Cross and Blue Shield System
14 CAPP Care
15 CIGNA Health Plans, Inc.
16 Community Health Plan, Inc.
17 Connecticut General Life Insurance Co.
18 Coventry Corp.
19 FHP, Inc.
20 Fortis Benefits
21 Foundation Health Corp.
22 Great Western Life and Accident
23 Group Health Cooperative of Puget Sound
24 Guardian Life Insurance Co.
25 Harvard/Pilgrim Health Care
26 Health Insurance Plan of Greater New York
27 Health Management Associates
28 Health Systems International, Inc.
29 HealthCare COMPARE Corp./The AFFORDABLE Medical Networks
30 HealthSource, Inc.
31 Henry Ford Health Care Corp.
32 Home Life Financial
33 Humana, Inc.
34 John Alden Life
35 John Deere Health Care, Inc.
36 John Hancock Life
37 Kaiser Foundation Health Plans, Inc.
38 Managed Health Network, Inc.
39 Maxicare Health Plans, Inc.
40 Medica
41 MedView Services Inc.
42 Mid-Atlantic Medical Services, Inc.
43 MultiPlan Inc.
44 Mutual of Omaha (managed care division)
45 Mutual of Omaha Insurance Co.
46 National Preferred Provider Network, Inc.
47 New York Life
48 NYLCare Health Plans, Inc.
49 Oxford Health Plans, Inc.
50 PacifiCare Health Systems, Inc.
51 PHS, Inc.
52 Physician Corp. of America
53 Preferred Health Network
54 Principal Financial
55 Principal Health Care, Inc.
56 Principal Mutual
57 Private Healthcare Systems
58 Provident Life and Accident Insurance Co.
59 Prudential Health Care Plans, Inc. (managed care division)
60 Prudential Insurance Co. of America
61 Sisters of Providence
62 United American HealthCare Corp.
63 United HealthCare Corp.
64 US Healthcare, Inc.
65 USA Health Network
66 WellCare Management Group, Inc.

DATA ELEMENTS : ADD, COMBINE, EDIT, OR SPLIT PRODUCTS

MyEdit Function[41]

>myedit<

<1> Edit Delete and Combine Flags	[goto flags]
<2> Add a product	[goto add]
<3> Edit specific Site/Product data	[goto edit]
<4> Editing complete	[goto sure]

IF YOU HAVE CHANGED DELETE AND COMBINE STATUS FOR PRODUCTS (OPTION 1)

YOU MAY WANT TO CHANGE OFFER STATUS BY SITE NOW (OPTION 3)

>flags<

ENTER 'd" ON ANY FIELD (except combine) WHEN DONE EDITING

Column1	*Column2*	*Column3*
Product Roster	*Delete Status*	*Combine Status*
	Store in >p_delete<	*Store in>p_combine<*
	(enter x here)	*(enter another product number here)*

[goto myedit]

>p_delete< *A variable per product, x = product was deleted*

>p_combine< *A variable per product. If greater than 0 the product has been combined with another product*.

>offer_flag< *Indicator set to 1 if product offered, or 0 if product not offered*

>add< What is the name of the product you want added to ALL sites
new product name is stored
[goto myedit]

>edit< ENTER THE SITE YOU WANT TO EDIT
list site roster

Enter SITE number [goto p_edit]

[41] See page 12 for a detailed description of the MyEdit Function.

>p_edit< What product did you want to edit in [SITE]?

Enter PRODUCT number

>item< What item do you want to edit in SITE for PRODUCT?

<1> Offer/combine/delete products (item b1a)
<2> Core/split attributes (item b5)
<3> Copays (item b13)
<4> Network and payment (item c1r)

>warn< This ends the editing function.
You will be entering the questionnaire for [SITE]'s [PRODUCT] product at item
[if item equals <1> b1a]
[if item equals <2> b5]
[if item equals <3> b13]
[if item equals <4> c1r]

If you made a mistake then back up now to change it.
[goto gotoprod]

>gotoprod<

[store <> in myedit]
begin list of Sites and Products
[if PRODUCT does not equal p_edit goto prod_no]
[if item equals <1> goto editb1]
[if item equals <2> goto editb5]
[if item equals <3> goto editb13]
[if item equals <4> goto editc1r]
[goto myedit]

>prod_no<

end list of PRODUCTS

>site_no<

end list of Sites

SPLITTING

The following data elements are used to split a product from either, b5b_end, b6b_end, b10b_end or b8b_end.

>append< YOU ARE ABOUT TO SPLIT AN ADDITIONAL PRODUCT FROM THIS PRODUCT
<g> CONTINUE with newname2

>newname2<
SITE: [SITE]
PARENT PRODUCT: [PRODUCT]

What do you want to call this product? – return to b5s, set SPLIT= negative value of number corresponding to the item where the split occurred. For example, if a split occurred at b5, SPLIT becomes <–5>.

>app< *If the number of products is less than 18, add the product to the product list and go to the next unanswered question, otherwise goto maxprod.*

>maxprod<

You have exceeded the maximum number of products allowed in this interview.
You can NOT add this product.
Please back up and change your answers so you don't add this product.
If you have any questions please contact your supervisor.
Go back to b5b or b6b or b10b or b8b

>SPLIT< Since *splitting a product into multiple products can occur at four different points in the interview b5b or b6b or b10b or b8b, the value of the >SPLIT< indicates where the split occurred. For example, if >SPLIT< equals <–5>, the split occurred at b5b. The value of >SPLIT< will determine what series of questions will be asked next. SPLIT will remain a positive number (i.e., at b5 SPLIT equals 5) until it is reset to a negative number at either b5b, b6b, b10b or b8b. SPLIT is reset to a positive value at b9_end_real if a split occurred.*

>NET< *The value of >NET< is determined by the response to questions b2, b5 and b6 and the value will determine what series of questions will be asked next.*

Example of Program Flow for Splits.

As discussed in section I.C.2.c, the CATI program was designed to allow the interviewer to split a product into two if the attribute responses for items b5,b6,b8 and b10 were not the same within the site. This process helped to ensure that product definitions were attribute based to account for within-site contract level variation. Overall, this situation occurred infrequently with a maximum of 54 (1.8%) of the 2,946 product site combinations resulting from splits. As an example, suppose that an insurer offered an "POS Gold" plan in a specified site. If upon calling the insurer we found out that for most contracts, if enrollees went to an out-of-network doctor without a referral, the plan would not cover any of the costs (b6=No). On the other hand, for other contracts, the POS Gold plan does cover some costs in this situation (b6=Yes). As a result, what we originally thought was one product became two.

In the above example, the CATI program was designed to allow the interviewer to identify a new name for the second product and then to collect information separately about both. The program after identifying the added product returns to the item where the split occurred and proceeds to collect the rest of the attributes on the original product. When the original product data is complete, the program begins a questioning cycle on the added product. For the items up until the split occurred, the questioning cycle for the added product is designed to verify the information collected on the original product also applies to the added one. When the interviewer reaches the split item, the questioning for the added product begins anew.

Given the questionnaire flow in a split situation is different depending on where the split occurred, in an effort to avoid confusion and to keep this document to a manageable format, we designed it to describe the general questionnaire path in the absence of such situations. To provide the reader with a general understanding of the skip patterns in a split situation, we provide the following example.

Using our example above, suppose that for the original product, POS Gold, the respondent provides a response of "Yes" to b6, and then answers "No" to b6a, indicating this is not true for all contracts in that site. The interviewer then confirms this fact in b6b (b6b=1) which as indicated in b6b_end takes the interviewer to the append screen. At the append screen, the interviewer obtains a new product name for the added product (newname2), for example, "POS Gold Freedom." At this point the program returns to b5s with SPLIT equal to <-6>, since the split occurred at question b6. At b5s with SPLIT=<-6> we go to b6_end which takes the interviewer to b10 to finish the rest of the questions on the original product. Once the original product attributes are obtained, at b9_end the program, via b9_end_real, returns to b5s_pre to begin the collection of the product attributes for the added product. **Also at this point, the value of SPLIT is changed from <-6> to <6>.**

For the added product, at editb5, with SPLIT now equal to <6> we go to child_beg and from b5s_end and b5_pre, the interviewer asks b5_split and b6_split to verify the original product responses apply to the added product. At b6b_end, since SPLIT=<6> the program returns to b10, the normal question path. At modules 4 and 5, the appended product is treated as a separate product.

Storing verbatim responses

>SPECIFY< *A opened ended response will be stored for the particular question. Open ends responses are allowable at a3, b2, b2a, c4_2, c, c6, xc7b and d1.*

DATA ELEMENTS : INTERIM OR FINAL STATUS CASE

Screen that comes up at the end if the respondent could not answer for one or more products.

>done_pre<

Hold for one minute while I check all my questions.....

[goto sure][42]
[If probe equals <1> goto new_resp, otherwise goto comp][43]

>new_resp<
Earlier you said you could not tell me about products in:
[SITE]

Do you think you could answer for any of those sites now?

<1> YES, WILL DO MORE SITES [goto a2]
<2> NO
[goto comp]

>cb< CODING CURRENT DISPOSITION

<20> CALLBACK--FIRM APPOINTMENT [goto cb20]
<21> CALLBACK--NO APPOINTMENT [goto xit1]
<22> NO ANSWER/BUSY [goto xit1]
<23> ANSWERING MACHINE--NO HUMAN CONTACT [goto xit1]
<24> SEARCHING--WRONG#--# CHANGED [goto xit1]
<25> INITIAL REFUSAL [goto xit1]
<26> INCOMPLETE--BREAKOFF [goto xit1]
<27> WAITING FOR LETTER [goto xit1]
<28> RESPONDENT WILL CALL MPR [goto xit1]
<40> OUT OF BUSINESS [goto fdis]
<42> NOT A HEALTH PLAN--NO NEW CONTACT [goto cls1]
<38> SUPERVISOR REVIEW [goto fdis]

>cb20< ENTER DATE & TIME FOR A FIRM APPOINTMENT [goto note]

[42] See page 20 for data element >sure<.
[43] See page 16 for data element >PROBE<.

>fdis< CODING FINAL DISPOSITION
<11> RETIRED—MERGER—ALL LINKS GO ELSEWHERE [goto xit2]
<12> MAX CALLS—END OF EFFORT [goto xit2]
<13> UN-LOCATABLE [goto xit2]
<15> FINAL REFUSAL [goto xit2]
<16> INCOMPLETE/BREAKOFF [goto xit2]
<17> CHRONIC NO ANSWER/BUSY [goto xit2]
<19> OTHER [goto sp2]
<40> OUT OF BUSINESS [goto xit2]
<41> INELIGIBLE—NOT A HEALTH PLAN [goto xit2]

>sp2< INTERVIEWER: SPECIFY OTHER FINAL DISPOSITION REASON [goto xit2]

>xit1< YOU ARE ABOUT TO EXIT THIS CASE.
ENTER DISPOSITON ON CONTACT SHEET AND TYPE <g> TO EXIT [goto note].

>c1s1< That's all the questions I have.
Your establishment is not selected to participate in the study.

INTERVIEWER: GIVE REASON FOR INELIGIBILITY IF ASKED.

>xit2< THIS CASE WILL BE SENT TO CLEAN.
ENTER DISPOSITON ON CONTACT SHEET AND TYPE <g> TO EXIT.

>note< INTERVIEWER: ARE THERE ANY NOTES ABOUT THIS CASES ?
<1> YES [specify]
<2> NO

>zend< *CATI sets time, date and exits*

>comp< CASE IS ABOUT TO COMPLETE.
TYPE <g> TO CONTINUE.

>cmp1<

Thank you very much. We appreciate your participation. Those are all the questions I have, though we may contact you again.

Thank you and goodbye.

APPENDIX D

ENTITY CATI EDITING SPECIFICATIONS AND PRE-IMPUTATION PRODUCT ATTRIBUTE ITEM NONRESPONSE RATES

This appendix presents the editing rules that were implemented on the insurer entity CATI data prior to the logical imputation procedures. The logical imputation procedures also resulted in some editing of the reported data for those plans in which post-CATI data was obtained from other sources.

I. STARTING PROCEDURES

The steps below should be performed before conducting the edits.

B91A-B93B Series Changes.

The method for applying the data to the edited CATI file is as follows, for cases where b1a={1,2} and the site was not deleted, nor product deleted or combine, is as follows: for each item in the b91a-b93b series, if the item is null OR IS EQUAL TO "6", find that record's value of b1d_site, and set the item to the value of the item for the same entity and product, and site equals the damaged (record's value of b1d_site).

An additional series of specs:

1. If B12_all = 001, set B12_1=1 and **set B12_2 and B12_3 to -5** (we can't know if the answer to B12_2 and B12_3 should have been YES or NO)

2. If B12_all = 012, set B12_1=1 and B12_2=2 and **set B12_3 to –5** (we can't know if the answer to B12_3 should have been YES or NO)

3. If B12_all = 101, set B12_2 and B12_3 =-5 (we can't know if the answer to B12_2 and B12_3 should have been YES or NO)
4.
5. If B12_all = 122, set B12_3=-5 (we can't know if the answer to B12_3 should have been YES or NO)

C7B Variables.

Incorporate the new C7B items from David Uglow’s file

Verbatims (Other specify answers)

Incorporate verbatim responses to other-specifies (To be provided).

Then, print out the verbatim responses for the following situations:

1. if A3=8 (something else) for what describes the organization
2. if B2=5 (something else) for the type of product
3. B2A_4=4 (or 1 after edit) something else – model type
4. C4_2=5 (Other) typical method of payment uses for primary care services
5. C5=5 (Other) typical method of payment uses for specialty services
6. C6=5 (Something else) typical method of payment uses for hospital services
7. D1=4 (Other) What is your tax status.

Note: the verbatim listing should include the id0, s_no and p_no identification numbers or fin_ent, prod_id and psu.

If the verbatim is not resolved (re-coded set the value of the variable to missing as indicated.

Order of Processing

NOTE: THE PROCEDURES ON THE FOLLOWING PAGES SHOULD BE IMPLEMENTED IN THE ORDER PRESENTED.

In addition, we need to conduct the editing in two stages. First, implement the edit specifications that are underlined for each variable and create a file with only these edits imposed for review. Second, implement the remaining code (creating a new program file with the initial edits and the second set) to create a final edited file. Place the final SAS code and the edited file in the appropriate q: drive directory.

CATI POST-EDIT PRE-LOGICAL IMPUTATION PRODUCT ATTRIBUTE MISSING RATES

Table D1 that follows presents the edit specifications for each attribute.

Table Notes

After conducting the logical imputation procedures, we considered the product information to be sufficiently complete for inclusion in the final analytical data file, if the record had post logical imputation, a non-missing for B2, entity reported product type. Of the original set of 7,643 CATI product records, we identified 2,946 products {defined by a combination of the entity identification number (FIN_ENT), product identification number (PROD_ID) and PSU membership} that meet this criteria. These 2,946 product records contain 452 products for which the product type was obtained via the logical editing process (see Chapter V, Section A as reflected on page D6).

The tables that follow present the CATI post-edit, pre-logical imputation item nonresponse/missing rates, equal to the number of records with missing data divided by 2,946. As a result, these missing rates, reflect the status of the data at the time the edits were performed from the CATI interviewing procedures prior to the logical imputation process for the set of records that were ultimately included in the data file. These rates are conservative in the sense that skipped responses are treated as non-missing. Some editing of the reported data also occurred during the logical imputation procedures, that is not reflected in these tables. During the logical editing procedures, reported data was modified, if the information reported was inconsistent with the post-CATI logical imputation data. Of the 2,946 product records, 1,976 ultimately were linked to a household policy (representing 12,263 policies); 970 we obtained data on a product that could not be successfully linked. These 970 product records resulted from

the data collection procedures that were designed to collect information on all known products in the database and any new products identified by the insurer.

More detailed information on the missing rates, as reflected in the post-logical imputed data, for the 1,976 policy linked products is presented in Appendix E. Item nonresponse counts for the 2,946 products post-logical edits among the items imputed are presented in Appendix G.

TABLE D.1

FOLLOWBACK CORE ATTRIBUTE EDITING SPECIFICATIONS
(CORE ITEMS ONLY)

Item #	Description	Edit Specifications	Number Missing of 2,946	% Missing
A3a	BC/BS ?	Answered by all. If A3a other than (1,2) set to missing	82	2.8
A3	Entity Type – Describe Organization	Dependent on A3a. If A3a=missing then set A3=missing. If A3a=1 then set A3=-1. If A3=8 then send verbatims to HSC for review **After verbatim recode, If A3=8 then set A3=missing.** Otherwise, if A3 other than (2,3,4,6) set to missing	133	4.5
B2	Product Type 1=HMO 2=POS 3=PPO 4=Indemnity 5=Other Specify	Answered by all. *If B2=5 send verbatims to HSC for review.* ***After verbatim recode, If B2=5 then set B2 to missing.*** Otherwise, if B2 other than 1,2,3,4 then set B2 to missing.	452	15.3
B2A_1 - B2A_4	Model Type _1 = Staff model _2 = Group model _3 = Network or IPA _4 = Something else	**Dependent on B2. Respondent can answer all that apply.** If B2=missing, then set B2A_1-B2A_4 to missing. If B2=3,4,or 5 set B2A_1, B2A_2, B2A_3, and B2A_4=-1. **If B2=1,2 and B2A not equal to 'd' then:** **If B2A_1 not equal to 1 then set B2A_1=0.** If B2A_2 not equal to 2 then set B2A_2=0. If B2A_3 not equal to 3 then set B2A_3=0. If B2A_4 not equal to 4 then set B2A_4=0. +Recode values of 2, 3 and 4 in B2A_2-B2A_4 to 1. THEN RECODE 0s to 2. If B2A_4=(1 post-edit) send verbatims to HSC **Eliminate B2A_4 from final dataset.**	B2A_1-B2A_3 All: 463	15.7
B3	Individual Purchase (not imputed)	Answered of all. If B3 other than 1,2 set B3 to missing.	498	16.9
B5	Directory of Physicians? (Network)	Dependent on B2. If B2=missing then set B5=missing. If B2=1,2, or 3 then set B5=-1 If B2=4,5 then if B5 other than 1,2 set B5 to missing.	453	15.4
B6	Out of Network Coverage	Dependent on B2 and B5. If B2=missing or B5=missing then set B6=missing. If B2=4,5 and B5= 2 set B6=-1 If B2=1,2 or 3 OR B5=1 then if B6 other than 1,2,7 set B6 to missing.	462	15.7

TABLE D.1 *(continued)*

Item #	Description	Edit Specifications	Number Missing of 2,946	% Missing
NET	Created Variable	Dependent on B2, B5, and B6 Set initially to missing then, If B2=4,5 and B5= 2 set NET=2 If B2=4,5 and B5=1 and B6=7 set NET=2 If B2=1,2, or 3 and B6=7 then set NET=2 **If B2=4,5 and B5=1 and (B6=1,2) then set NET=1** If B2=1,2, or 3 and (B6=1,2) then set NET=1	462, based on missing values for B2, B5, and B6	15.7
B10	PCP Sign-up	Dependent on NET. If NET=missing then set B10=missing. If NET=2 set B10=-1 If NET=1, then if B10 other than 1,2 set B10 to missing.	469	15.9
B8	B8 Maximum in-network coverage	Similar to B10. If NET=missing then set B8=missing. If NET=2 set B8=-1 If NET=1, then if B8 other than 1,2 set B8 to missing.	479	16.3
B12_1-B12_3	PCP Types. _1=Generalists _2=OB/GYNs _3=Other Specialists	**Dependent on B10. Respondent can answer all that apply.** **If B10=missing then set B12_1, B12_2 and B12_3=missing. Also set B12=missing.** **If B10 =2 or –1 (after edit) then set B12_1, B12_2 and B12_3= -1** **If B10 =1, and [B12='n' or B12=' '] then:** **a. set B12_1=1.** **b. If B12_2 is not –5 or 0 then set B12_2=1** **c. If B12_3 is not –5 or 0 then set B12_3=1** **d. If B12_2=-5 then set B12_2=missing** **e. If B12_3=-5 then set B12_3=missing** **If B10 =1, and B12='d' then set B12_2 and B12_3=missing, set B12_1=1 (2 cases):** ***Finally recode values of 0 to 2. * all but 2 missing B12_2 and B12_3 cases will be the result of CATI program problems.	In order B12_1, B12_2 and B12_3 469 547 550	 15.9 18.6 18.7

TABLE D.1 *(continued)*

Item #	Description	Edit Specifications	Number Missing of 2,946	% Missing
B91A	Any in-network coverage	Dependent on B8 (and B2, B5,B6) If B8=missing then set B91A=missing. If B8=2 or -1 then set B91A=-1. If B8=1 and B91A other than 1,2 then set B91A to missing.	756	25.7
B91B	In network coverage for most specialists	Dependent on B91A (B8,B2, B5 and B6) If B91A=missing then set B91B=missing. If B91A=2 or –1 then set B91B=-1. If B91A=1 and B91B other than 1,2 then set B91B to missing.	757	25.7
B91C	In network coverage for OB/GYN	Dependent on B91B. If B91B=missing then set B91C=missing. If B91B=1 or –1 then set B91C=-1. If B91B=2 and B91C other than 1,2 then set B91C to missing	757	25.7
B91D	In network coverage for non-OB/GYN	Dependent on B91B. If B91B=missing then set B91D=missing. If B91B=1 or –1 then set B91D=-1. If B91B=2 and B91D other than 1,2 then set B91D to missing.	761	25.8
B92	In network coverage compared to referral	Dependent on B91A. If B91A=missing then set B92=missing. If B91A=2 or –1 then set B92=-1. If B91A=1 and B92 other than 1,2,3 then set B92 to missing.	760	25.8
B92B	In network coverage compared to referral for OB/GYN.	Dependent on B92, B91B and B91C. If B92=missing then set B92B=missing. If B92=1,2 or –1 then set B92B=-1 If B92=3 and combination of B91B/B91C is (2,2) then set B92B=-1 If B92=3 and any combination of B91B/B91C is (1,-1) or (2,1) then if B92B other than 1,2 set B92B to missing.	760	25.8

TABLE D.1 *(continued)*

Item #	Description	Edit Specifications	Number Missing of 2,946	% Missing
B93	Any in-network coverage compared to out-of-network coverage	Dependent on B6 and B92 If B6=missing or B92=missing then set B93=missing. If B6=2,7 or –1, or (if B6=1 and B92=-1 or 1) then set B93=-1 If B6=1 and B92=2,3 and B93 other than 1,2,3 then set B93 to missing.	644	21.9
B93B	In network coverage compared to out-of-network coverage for OB/GYN.	Dependent on B93, B91B and B91C. If B93,B91B or B91C=missing then set B93B=missing. If B93=1,2 or –1 then set B93B=-1 If B93=3 and combination of B91B/B91C is (2,2) then set B93B=-1 If B93=3 and any combination of B91B/B91C is (1,-1) or (2,1) then if B93B other than 1,2 set B93B to missing.	645	21.9

TABLE D.1 *(continued)*

Item #	Description	Edit Specifications	Number Missing of 2,946	% Missing
B13	Co-payment or co-insurance	Answered of all, if B13 other than 1,2,0 then set B13 to missing. **Exception: If B13AMT>0 and B13=. Then set B13=1.** If B13PER>0 and B13=missing, then set B13=2.	583	19.8
B13PER	Co-insurance	Dependent on B13. If B13=0,1 then set B13PER=-1 If B13=2 then if: **a. B13PER other than numeric >0 set B13PER to missing.** b. if B13PER > 50 then set B13PER =(100 minus the reported value).	599 (16 with B13=2)	20.3
B13AMT	Co-payment amount	Dependent on B13. If B13=0,2 then set B13AMT = -1 If B13=1 then if: B13AMT other than numeric, or if *B13AMT >50* then set B13AMT to missing.	622 (39 with B13=1)	21.1
B13OUT	Out of network co-payment or co-insurance	Dependent on B6. If B6=missing, then set B13OUT=missing. If B6=-1,2,7 then set B13OUT=-1. Otherwise, if B13OUT other than 0,1,2 then set B13OUT to missing.	525	17.8
B13OP	Coinsurance for out-of-network	If B13OUT=missing then set B13OP=missing. If B13OUT=-1,0 or 1 then set B13OP=-1. If B13OUT=2, then do: a. B13OP>50 then set b13OP=(100 minus the reported value). b. If, B13OP is not numeric, set B13OP=missing.	570	19.4
B13OD	Co-payment for out-of network	If B13OUT=missing then set B13OD=missing. If B13OUT=-1,0 or 2 then set B13OD=-1. If b13OUT=1, then if B13OD is not numeric, or if B13OD>50 then set B13OD=missing.	532	18.1

TABLE D.1 *(continued)*

Item #	Description	Edit Specifications	Number Missing of 2,946	% Missing
B14OUT	Out-of-network deductible requirement	Dependent on B6. If B6=missing, then set B14OUT=missing. If B6= -1,2,7 then set B14OUT=-1. Otherwise, if B14OUT other than 1,2 then set B14OUT to missing.	575	19.5
B14OD (& B14CH2)	Out-of-network deductible amount	Dependent on B14OUT. If B14OUT=missing, then set B14OD=missing. If B14OUT=2 or –1 then set B14OD=-1. If B14OUT=1, then if a. B14OD other than numeric, or b. 0<B14OD<50 or c. B14OD>5000, then set B14OD to missing. If B14OD=missing, set B14CH2=missing	659	22.4
C1R	Physician Network Size	Dependent on NET. If NET=2 then set C1R=-1. If NET=missing, then set C1R=missing. Otherwise, If C1R other than 1,2,3,4 then set C1R to missing.	1,212	41.1
C2R	Hospital Network Size	Dependent on NET. If NET=2 then set C2R=-1. If NET=missing, then set C2R=missing. Otherwise, If C2R other than 1,2 then set C2R to missing.	658	22.3
C4_2	PCP Payment	Answered of all. If C4_2=3,4 and NET=2 then set C4_2= missing. If C4_2=5, send verbatims to HSC **If C4_2=5 after verbatim review then set C4_2=missing.** Else, if C4_2 other than 1,2,3,4 then set C4_2 to missing.	703	23.9

TABLE D.1 *(continued)*

Item #	Description	Edit Specifications	Number Missing of 2,946	% Missing
C4A_1-C4A_3	C4A_1=1 Referrals to specialists C4A=2 Hospitalizations C4A=3 Other services	Dependent on C4_2. Respondent can answer all that apply. If C4_2=missing then set C4A_1, C4A_2 and C4A_3 to missing. If C4_2 =1,2,3 then set C4A_1, C4A_2 and C4A_3=-1. If C4_2=4 and C4='n' and C4A_1-C4A_3 are all missing then set C4A_1-C4A_3=0. Note: If C4_2=4 and C4='d' then all of C4A_1-C4A_3 will be missing. If C4_2=4 and C4='n' and C4A_1-C4A_3 not all missing then do: If C4A_1 not equal to 1 then set C4A_1=0. If C4A_2 not equal to 2 then set C4A_1=0. If C4A_3 not equal to 3 then set C4A_1=0. Then recode values of 2 and 3 in C4A_2 and C4A_3 to 1. Finally recode values of 0 to 2.	All 779	26.4

TABLE D.1 *(continued)*

Item #	Description	Edit Specifications	Number Missing of 2,946	% Missing
C5	Specialist Payment	Dependent on C4_2 and C4A_1 If C4_2=missing or C4a_1=missing then set C5=missing If C4a_1=1 then set C5=-1. **If NET=0 and C5=3,4 then set C5=missing.** If C5=5, then send verbatims to HSC for review. **After verbatim review if C5=5 set C5=missing.** Otherwise, if C5 other than 1,2,3,4 then set C5 to missing.	784	26.6
C6	Hospital Payment	Dependent on C4_2 and C4A_2 If C4_2=missing or C4a_2=missing then set C6=missing. If **C4a_2=1** then set C6=-1. If NET=0 and **C6**=3 ,7 then set C6=missing. **If C6=5, send verbatims to HSC for review** **If C6=5 after verbatim review then set C6=missing.** Otherwise, if C6 other than 1,2,3,4,7 then set C6 to missing.	993	33.7
C7	Mental Health Benefit	Answered of all. If C7 other than 1,2 then set C7 to missing. If C7a=(1,2) and C7=missing then set C7=1.	753	25.6
C7a	Specialty mental health organization	Dependent on C7. If C7=missing then set C7a=missing. If C7=2 then set C7a=-1. If C7=1 and C7a other than 1,2 then set C7a to missing.	794	27.0
C7b	Name of mental health organization	**Don't Edit**	N.A.	N.A.

TABLE D.1 *(continued)*

Item #	Description	Edit Specifications	Number Missing of 2,946	% Missing
D1	Tax Status	Dependent A3A and A3. If A3a=2 and A3=(-1,4,6,8) then set D1=-1. If A3a =missing or A3=missing then set D1=missing. If A3a=1 or A3=2,3 then: If D1=4, send verbatims to HSC **IF D1=4 after verbatim review then set D1=missing.** Else, D1 other than 1,2,3 then set D1 to missing.	309	10.5
D2	Subsidiary	Answered of all. If D2 other than 1,2 then set D2 to missing.	178	6.0
D2a	Parent company or service area	Dependent on D2. If D2=2 then set D2a=-1 If D2=missing, then set D2a=missing. Otherwise, if D2a other than 1,2 then set D2a to missing.	178	6.0
D3	Entity service area	Dependent on D2. If D2=1 then set D3=-1 If D2=missing, then set D3=missing. Otherwise, if D3 other than 1,2 then set D3=.	178	6.0

APPENDIX E

PRODUCT IMPUTATION AND STATISTICAL MATCHING METHODS

In this appendix we describe the stochastic product imputation and policy weighting methodology presented in Chapter V of the report in more detail. Preparation of survey weights and Sudaan specifications are discussed in Appendix F and data imputation specifications in Appendix G. We refer to the methodology as stochastic imputation because we used a probability-based method to assign values for policies with missing data (referred to as recipients) from other policies with completed information (referred to as donors).

A. PRODUCT IMPUTATIONS

Based on consultations with staff at The Center for Studying Health System Change (HSC), Social and Scientific Systems (SSS) conducted a series of stochastic data imputation steps based on specifications developed by MPR and HSC to assign values to selected items that had missing values after the data editing process and logical imputation procedures had been completed. Section A presents the missing rates for the data items considered for imputation, and which of these items we selected to be imputed. Section B outlines the general imputation methodology and section C presents details on the implementation for each variable. Throughout this appendix we define a product to be a unique combination of the CTS site (PSU), the insurer company or entity (defined by the entity's identification number, FIN_ENT), and the insurance product it offers (as indexed with each PSU and FIN_ENT by PROD_ID). Each of these products is linked to one or more policies as reported in the household survey.

1. Selecting Questionnaire Items for Imputation

As in round 1 of the survey, we selected a subset of data items to be imputed based on the planned use of the variables and their missing rates. A list of the data items is provided in Appendix C. First, we excluded variables that were used only to control the flow of the

instrument (e.g., B5s_pre, B5_split). Missing rates were then examined for remaining analytical variables to select those to be imputed.

Table E.1 includes items initially considered for imputation. For each item, the table shows post-edit, post logical imputation, and the percentage of nonskipped or applicable outcomes that were missing. In developing these counts, we considered a outcome to be *applicable* if:

- A response to the question was not dependent on other responses, or
- The response was dependent on one or more other questionnaire items, and
  - The outcomes to the appropriate other items were nonmissing and indicated that a respondent should have answered the question, **or**
  - One or more of the outcomes to the appropriate other items were also missing, so that it was not possible to determine whether the question should have been answered.

An outcome is considered "skipped" only if other outcomes affecting whether the item should have been answered are nonmissing and indicate that the item should not have been asked. Note that these definitions count an outcome as missing if other outcomes that affect whether the item should have been asked are also missing. In many cases, after the other outcomes were imputed, the subsequent missing data could be set to a "skipped" response.

In all cases, the applicable and skipped definitions are based on "prior" outcomes.

Table E.1 presents the missing rates based on 1) the percentage of products and 2) the percentage of policies (with each product weighted by its associated number of linked policies). The data presented is based on the final set of 1,976 products that had one or more resolved policy linkages (totaling to 12,263 policies). Post-edit, pre-logical imputation, and missing rates (defined as the number of products with missing data divided by the total number of products) for the full set of products (based on 2,946 products provided in the final analytical file[44]) are presented in Appendix D.

From the list of potential items to be imputed, we eliminated variables through a two-step process. At the first step, we eliminated items whose missing rates appeared too high to support inference. The second stage, we decided not to impute variables if the ratio of donors to recipients was less than 2.5-to-1 for the final set of classing characteristics. Table E.1 indicates which of the variables in the initial set were finally imputed using stochastic imputation and the number of records that we imputed. Appendix G presents the total number of imputed product values based on the full set 2,946 product records. Reasons for not imputing each item are listed below.

- As in round 1, we decided to limit the data items to be imputed to specific product attributes and did not impute data for the entity based characteristics (A3A, A3, D1, D2, D2A, and D3)

[44]See Table notes in Appendix D. We identified 2,946 products to had sufficient data from the CATI interviewing procedures to be used in product data analysis. Of these, we were able to link 1,976 to one or more policies. We note as in Appendix D that we collected data on 970 products without linkages given that the data collection procedures were designed to collect information on all known products in the database and any new products identified.

TABLE E.1

MISSING RATES FOR ENTITY QUESTIONNAIRE ITEMS

Variable/ Item	Abbreviated Description	Products (1,976)					Linkages (12, 263)			
		Stochastic Imputation	Number Skipped	Number Missing/ Imputed	Percent Applicable [45]	Missing Rate (Percent)	Number Skipped	Number Missing	Percent Applicable	Missing Rate (Percent)
A3a	A Blue Cross or Blue Shield Plan	No	-	-	100.0%	0.0%	-	-	100.0%	0.0%
A3	Entity Type – describe organization	No	277	35	86.0%	2.1%	4,204	108	65.7%	1.3%
B2	Reported Product Type	No	-	-	100.0%	0.0%	-	-	100.0%	0.0%
B2A_1	_1 = Staff model	Yes	1,115	135	43.6%	15.7%	5,299	899	56.8%	12.9%
B2A_2	_2 = Group model	Yes	1,115	135	43.6%	15.7%	5,299	899	56.8%	12.9%
B2A_3	_3 = Network or IPA	Yes	1,115	135	43.6%	15.7%	5,299	899	56.8%	12.9%
B3	Individual Purchase	No	-	267	100.0%	13.5%	-	1,628	100.0%	13.3%
B5	Directory of physicians? (Network)	No	1,733	-	12.3%	0.0%	11,320	-	7.7%	0.0%
B6	Out of Network Coverage	No	177	-	91.0%	0.0%	647	-	94.7%	0.0%
NET	Created Variable, Network Status	N.A.	-	-	100.0%	0.0%	-	-	100.0%	0.0%
B10	PCP Sign-up	Yes	182	3	90.8%	0.2%	656	3	94.7%	0.0%
B8	Maximum in-network coverage	Yes	182	8	90.8%	0.4%	656	27	94.7%	0.2%
B12_1	PCP Type 1=Generalists	Yes[46]	1,155	3	41.5%	0.4%	5,548	3	54.8%	0.0%
B12_2	PCP Type 2=OB/GYNs	Yes	1,155	78	41.5%	9.5%	5,548	378	54.8%	5.6%
B12_3	PCP Type 3=Other Specialists	Yes	1,155	80	41.5%	9.7%	5,548	383	54.8%	5.7%
B91A	Any in-network coverage	Yes	1,136	194	42.5%	23.1%	5,597	995	54.4%	14.9%
B91B	In network coverage for most specialists	No	1,252	195	36.6%	26.9%	6,935	1,083	43.4%	20.3%
B91C	In network coverage for OB/GYN	No	1,531	195	22.5%	43.8%	8,493	1,083	30.7%	28.7%
B91D	In network coverage for non-OB/GYN	No	1,531	198	22.5%	44.5%	8,493	1,125	30.7%	29.8%
B92	In network coverage compared to referral	No	1,252	257	36.6%	35.5%	6,935	1,810	43.4%	34.0%
B92B	In network OB/GYN coverage compared to referral.	No	1,700	268	14.0%	97.1%	10,365	1,840	15.5%	96.9%
B93	Any in-network coverage compared to out-of-network coverage	No	1,707	146	13.6%	54.3%	11,217	633	8.5%	60.5%
B93B	In network coverage compared to out-of-network coverage for OB/GYN.	No	1,826	147	7.6%	98.0%	11,621	638	5.2%	99.4%
B13	Co-payment or co-insurance	Yes	-	117	100.0%	5.9%	-	792	100.0%	6.5%
B13PER	Co-insurance	Yes	1,557	142	21.2%	33.9%	10,306	874	16.0%	44.7%
B13AMT	Co-payment amount	Yes	364	277	81.6%	17.2%	1,423	2,193	88.4%	20.2%
B13OUT	Out of network co-payment or co-insurance	Yes	688	56	65.2%	4.3%	5,852	411	52.3%	6.4%
B13OP	Coinsurance for out-of-network	Yes	912	125	53.8%	11.7%	6,999	920	42.9%	17.5%
B13OD	Co-payment for out-of network	No	1,756	107	11.1%	48.6%	9,838	542	19.8%	22.4%
B14	Deductible Amount	Yes	-	307	100.0%	15.5%	-	1,984	100.0%	16.2%
B14OUT	Out-of-network deductible requirement	Yes	688	104	65.2%	8.1%	5,852	630	52.3%	9.8%
B14OD	Out-of-network deductible amount	Yes	1,037	163	47.5%	17.4%	7,508	1,190	38.8%	25.0%
C1R	Physician Network Size	No	182	705	90.8%	39.3%	656	3,144	94.7%	27.1%
C2R	Hospital Network Size	No	182	301	90.8%	16.8%	656	1,755	94.7%	15.1%
C4_2	PCP Payment	Yes	-	371	100.0%	18.8%	-	2,043	100.0%	16.7%
C4A_1	C4A_1=1 Referrals to specialists	No	1,300	439	34.2%	64.9%	7,632	2,499	37.8%	54.0%
C4A_2	C4A=2 Hospitalizations	No	1,300	439	34.2%	64.9%	7,632	2,499	37.8%	54.0%
C4A_3	C4A=3 Other services	No	1,300	439	34.2%	64.9%	7,632	2,499	37.8%	54.0%
C5	Typical payment method for specialty Services	No	178	443	91.0%	24.6%	1,672	2,538	86.4%	24.0%
C6	Typical payment method for hospital Services	No	87	596	95.6%	31.6%	835	3,548	93.2%	31.0%
C7	Ever include mental health services	No	-	250	100.0%	12.7%	-	1,497	100.0%	12.2%
C7a	Mental health services managed separately?	No	74	417	96.3%	21.9%	146	2,589	98.8%	21.4%
D1	Tax Status	No	325	182	83.6%	11.0%	521	1,276	95.8%	10.9%
D2	Subsidiary	No	-	13	100.0%	0.7%	-	149	100.0%	1.2%
D2a	Parent company or service area	No	1,434	13	27.4%	2.4%	8,848	149	27.8%	4.4%
D3	Entity service area	No	529	13	73.2%	0.9%	3,266	149	73.4%	1.7%

[45] Defined as the number of non-skipped responses divided by the number of products (2,946).

[46] The three missing cases were set to a value of 1 (Yes, generalists can serve as PCPs) without the use of stochastic imputation.

- After logical imputation and editing, all products had non-missing data for item B2, reported product line, and as imputation was not needed. Similarly, no imputation was necessary for items B6, coverage of out-of-network physicians, and B5, existence of a network.

- B3, C7, and C7a (individual purchase and mental health coverage) were not imputed because the questions had analytical use.

- Items B91C and B91D (in network coverage for OB/GYN, and non-OB/GYN, respectively), and item C1R (physician network size) were not imputed due to high nonresponse (approximately 40% or higher).

- We decided not to impute items B92B, B93 and B93B (differences in coverage for in-network specialists with and without a referral) because these questions applied to less than 20 percent of products. Item B92, while having a higher applicability rate, had a high missing rate and was excluded from imputation.

- We did not impute B91B or B13OD because we would not obtain the minimum 2.5 to 1 donor-to-recipient ratio.

- After reviewing the data on C2R, hospital network size, we were unable to identify good predictors for this item and decided not to impute it.

- Initially, we attempted to impute items C4A (services included in capitated payments), C5 (typical payment method for specialty services), and C6 (typical payment for hospital services) Unfortunately as with B91B, the donor to-recipient ratio of 2.5-to-1 could not be obtained for the C4A items, so we decided not to impute. Since C4A was not imputed, imputation also was not conducted on C5 and C6.

2. Methodology: Sequential Hot Deck Imputation

Sequential hot deck imputation procedures are designed to use responses from another respondent for assignment to a respondent with missing data. Respondents with nonmissing responses for an item are referred to as "donors," and those with missing data are "recipients." This type of imputation procedure selects a donor for each recipient whose response to a question has a value that is closest to the recipient's unknown, but most likely or expected, response. One of the strengths of this approach is that both categorical variables and continuous variables can be used in selecting a donor.

In sequential hot deck imputation, the set of potential donors is restricted to those that have the same responses as the recipient to a group of data items or variables, called "classing variables." Traditionally, classing variables are chosen so that each donor pool has a sufficient count of donors. Donors and recipients having the same values to the classing variables are then sorted by a set of "sorting variables," which may be continuous or categorical. The sort is conducted in a card-like deck fashion so that donors and recipients with similar values are in proximity to each other. The donor who immediately precedes the recipient is then selected to provide the replacement value. This sequential card-like deck sort and selection process gives the method its name.

We required the resulting weighted imputed product data to have the same distributional properties on each imputed item as the weighted distribution reflected in the cases with nonmissing data. To best meet this requirement, we conducted a weighted sequential hot deck imputation process (Cox 1980).[47] Repeated applications of this method produce, on average, weighted estimates (using reported and imputed data) that match the weighted estimates using only cases with nonmissing data.

Weighted and standard sequential hot deck procedures differ in their selection processes. A weighted sequential hot deck imputation process uses a selection process that is similar to the methods used in probability proportionate to size sampling. First, the donors and recipients with the same responses to a classing variable are grouped, and each group is sorted on the sorting variables. Next, the selection procedure uses a random mechanism that interweaves respondents and nonrespondents to divide the donors into subgroups, where the number of subgroups is equal

[47]Cox, Brenda G., "The Weighted Sequential Hot Deck Imputation Procedure," Proceedings of the American Statistical Association Survey Research Section, 1980, pages 721-726.

to the number of recipients in the group. Within a subgroup, the donors are selected based on the relative sizes of their survey weights. With this approach, in repeated applications of the selection process, the weighted imputed data will have the same distributional properties as the weighted data for the non-missing cases. The other benefit of this approach is that it limits the number of times an individual donor can be used. In contrast, in traditional sequential hot deck imputation schemes, a series of missing cases occurring together could receive the same donor. We used the square root of the number of policies successfully linked to each product as the sampling weight for each product in the execution of the weighted sequential hot-deck imputation procedures.

3. Implementation

Our first task was to select a group of data items and variables as the classing and sorting variables. We based our choices on two criteria. First, we wanted to include variables that predicted the missing status of the data item. A desirable property of missing responses is that once the predictors are considered, the responses will be missing-at-random (MAR)[48] (and as such missing status does not depend on the outcome). Second, we wanted to pick predictors of the item to be imputed, so that a donor with the same value as the recipient's expected value. In reality, the set of predictors of the data item value and the set of predictors of missing status often identified the same variables, but using these sets is expected to improve the accuracy of the imputation process. In particular, it is possible for a variable to be a good predictor of the outcome but to be not related to the missing status. By including such a variable in the imputation process we reduce the variability in the imputed values by increasing in accuracy of

[48]See Rubin, D.B., (1976) "Inference and Missing Data," *Biometrika,* 63, 581-592 and Rubin D.B, (1987) *Multiple Imputation for Nonresponse in Surveys*, John Wiley, New York.

the individual responses. We conducted a variety of cross-tabulations to identify these relationships, using the item being imputed, as well as other items and external variables, such as site. We also conducted a Chi-square Automatic Interaction Detector (CHAID) analysis[49] to help rank the predictive power of the predictors and to isolate optimal classing sets. We reviewed the recommended classing set to develop final specifications (see Appendix G).

With the classing and sorting variables identified, we next determine how to use these items to maximize the accuracy and stability of the imputed values. We defined stability based on the donor-to-recipient ratio in the classing set and accuracy based on the level of detail used to build the classing set (i.e., the number of classing variables used and their corresponding categories). Since some of the combinations of possible values among classing variables contained few, if any, possible donors, we decided to use the classing variables in a stepwise manner to improve the accuracy of the prediction. (This approach also was used in round 1). As a general rule, we considered the donor pool to be too limited to conduct the imputations if the number of donors represented less than 33 percent of all the respondents in a cell. For the first step in the imputation process, we used the imputation cells based on the most restrictive/detailed set of classing variables to impute cases in cells that meet the required 3-to-1 donor-to-recipient ratio. For each subsequent step in the process, we imputed the remaining cases by either deleting the classing variable that appeared to have the weakest relationship to the item being imputed or collapsing the categories to achieve the desired 3-to-1 donor-to-recipient ratio in the cells used. For a few variables, we allowed the donor-to-recipient ratio to drop to 2.5 in the final classing set. In general, we performed only a small number of steps of this type; most of the imputations were completed in one or two steps using the most restrictive cell definitions. In defining the

[49]Using Answer Tree 2.0. SPSS Inc. 233 Wacker Drive, Chicago IL.

donors, we required all donors to have a reported nonimputed value for the item imputed. However, to increase the size of the donor pool, we allowed cases with an imputed value for a classing variable to serve as donors. We also implemented the imputation procedures on each variable, so that survey questions that were not dependent on prior responses were imputed first, followed by those the related dependent items.

In addition to the stepwise process, we conducted variable imputations jointly to reduce the likelihood of generating a series of imputed item responses or imputed and reported question responses that were not *observed in the data.* For example, in the imputation of items B8, B10, and B91A, we imputed these values together (as a vector taking some or all of these values from the same donor depending on the missing pattern) in four stages based on four observed missing patterns among these items (see Appendix G). For example, if B8 and B91A were missing, and B10 was reported, we imputed B8 and B91A together from the same donor using B10 in the classing set. The final imputation specifications are presented in Appendix G, which indicate the classing sets used and how many variable values were imputed at each imputation step and for each missing pattern when the variables were imputed jointly.

B. STATISTICAL MATCHING OF SOFT-MATCHES

1. Overview

A second form of missing data in the Followback Survey resulted from the linkage process between products identified by health insurance entities (or by the employer associated with a policy) and persons covered by a product or identified by the CTS Household Survey respondent. Persons within a family insurance unit (FIU) associated with a specific health insurance product were defined as a "policy" unit. An FIU could have more than one policy and a person in an FIU could be associated with more than one policy.

Of the 21,701 eligible policies, 4,666 were initially soft-linked after the employer survey (See Table IV.3 above). A soft link indicates that we could identify the entity (insurer, third party administrator, union, or employer) with which the policy unit was associated, but we were unable to determine which of the offered products covered the policy unit. For a portion of these soft-matches we used a statistical matching procedure to select the "best" linkage from among the potential products the household respondent could have selected. Of the 4,666 soft-linked products, 1,112 could not be resolved because of insufficient product data for that insurer in the respondent's site. For the remaining 3,554 soft-matches, we had a complete product choice set for 1,932 of these linkages. For these 1,932 policies, we began by linking all the associated products the entity offered in the site to each of these policies. This step generated 8,689 product–policy linkages consisting of one or more product choices for each soft-linked policy . The statistical matching procedures then selected one of these products to serve as the final match. Note that 37 of the soft-matches had only one choice available and as such, became hard-matches.

The statistical matching procedures conducted in round 2 of the survey were basically identical to the methods used in round 1 based on the modeling-based procedure suggested by Singh et al.[50] This modeling process consisted of five implementation steps. First, we combined the Household Survey person- and family-level data into a policy-based record file that provided us with the characteristics of the policy-holder and the household reported attributes for each policy. Recall that a policy is defined as a unique relationship between a private health plan and

[50]Singh A.C., H.J. Mantel, M.D. Kinack, and G. Rowe. "Statistical Matching: Use of Auxiliary Information as an Alternative to the Conditional Independence Assumption." *Survey Methodology*, vol. 19, 1993, pp. 59-79.

the set of household members it covers Second, we decided to use the same set of nine insurer reported product attributes in the modeling process. Based on round 1, we found that these nine attributes could be predicted from the household reported policy holder and policy characteristics and shown related to self-reported product line (health maintenance organization (HMO), point-of-service (POS), preferred provider organization (PPO), and fee-for-service (FFS). Third, we developed logistic regression models to predict each of these nine attributes based on policy-level CTS household data prepared in step 1. Fourth, using the model results, we prepared predicted values for the nine attributes for the soft-matched policies based on their household reported information. As the final step, we selected one of the product's in each soft-matched choice set as the final match. We selected the link by comparing the predicted values for the insurer product attributes with the insurer reported attributes based on the methodology developed in round 1. The round 1 validation procedures showed a 64 percent overall exact match rate and 67 percent match rate with a product of the same type.

2. The Product Attributes Used For Matching

We selected the same nine round 1 product attributes for the round 2 matching process to keep the same matching procedures. The nine attributes are listed below.

1. HMO Line status From Line From Item B2.
2. Out of Network Coverage of Primary Care Physician Visits, B6
3. Coverage of Self-referral to in-network specialists (B8 and B91A combined)
4. Primary Care Physician Sign-up Required, B10
5. Coinsurance or Copayment Requirement, B13

6. Physician Payment Methods, Three indicators from C4_2, Fee-for-service (FFS)
7. Physician Payment Methods, Three indicators from C4_2, Discounted FFS'
8. Physician Payment Methods, Three indicators from C4_2, Capitated Payment.
9. Network Status (created variable from B2, B5 (does plan have directory of physicians), and B6)

While we used the same set of attributes, we needed to adjust for changes in question wording. In round 2, the question on the coverage of costs to visits to in-network specialists, item B8, changed from a single question to two, items B8 and B91A. Table E.2 presents the round 1 question and the corresponding questions in round 2 and their equivalent value combinations. To keep the number of matching items the same as in round 1, we decided to recode the round 2 values from the two new questions into a single item for use in the matching process. Also, to verify that we were still using the best set of attributes in the modeling process, we conducted a series of discriminant analysis procedures to predict product line classification (from question B2) from the available attributes.

We conducted two unweighted stepwise discriminate analysis procedures to identify the entity-reported variables that together best predicted the self-reported product type. We conducted the discriminant analysis procedure on the 2,946 products, using the self-reported product type (four categories) as the dependent variable. For this task, we converted the continuous data items with the deductible (item B14) into series of range indicators. Since the level of coinsurance or copayment added little, if any, predictive power, compared with the binary variable showing whether a copayment or coninsurance was required, we used only the copayment/coinsurance status outcome in the models.

We also transformed the response categories for primary care physicians payment methods (item C4) into three categorical indicators. We conducted the first model using the round 1 and new items including B91B (whether the in-network coverage of self-referrals applies to most types of specialists), B91C if this coverage (applies to OB/GYN specialists), and B92 (if the coverage is the same or less compared to coverage with a referral). For comparative purposes,

TABLE E.2

COMPARISON OF ROUND 1 VS. ROUND 2 QUESTIONING

STRATEGY FOR DETERMINING COVERAGE OF IN – NETWORK SPECIALISTS WITHOUT A REFERRAL

Round 1	**Round 2**
Format	
B8: IF NET=1: If enrollees do not have a referral and go to in-network specialists, does the plan cover any of the costs for these visits?	**B8:** IF NET=1: Is a referral or authorization ever required to obtain maximum coverage for an initial visit to an in-network specialist?
	B91a: IF b8=YES: Does product provide at least some coverage for self-referrals to any types of in-network specialists?
Values	
Yes	(B8=Yes and B91A=Yes) or B8=No.
No	B8=Yes and B91A=No.

the second model included the same items included in the round 1 analysis. Table E.3 presents the partial *R*-square values for the variable selected at each step in the stepwise discriminant analysis procedure.[51] The results indicated the that the items used in round 1 provided nearly the same level of predictive power as the new model (a difference of only .03 in the average canonical correlation). The relative strength of each of the items used in round 1 also appeared to be similar; however, coinsurance/copayment status appeared to stronger in round 2 than in round 1 and capitated payment status, (from item C4_2) was weaker for round 2 in identifying product line. As in round 1 the deductible level (item B14) showed some predictive ability; however, we could not accurately predict the deductible level from the CTS policy-level information; hence, we excluded this variable from the matching list.

3. Modeling the Product Attributes for the Hard-Linked Cases

The modeling procedures for round 2 mimicked those used for round 1. We prepared a series of weighted logistic regression models to predict each of the nine matching product variables, using the Household Survey policy-level variables for the hard-linked policies. For the hard-linked policies (10,331 products), we weighted the hard-linked data (10,331 products) to account for variation in the hard linkage rates between the sites and for differences in the prevalence of HMO membership by site. We developed the models using a combination of stepwise and non-

[51] In Table E.3, the product attributes are listed in order of their inclusion in the model. The partial *R*-square values reflect the squared partial correlation for predicting the self-reported product type from the product attributes, controlling for the effects of the attributes listed previously in the table.

TABLE E.3

RELATIVE PREDICTIVE POWER OF THE PRODUCT ATTRIBUTES TO DESCRIBE SELF-REPORTED PRODUCT TYPE

Variable[a]	Description	Order Entered	Partial *R*-Square Round 2	Partial *R*-Square Round 1
	Model Using New Round 2 B9 Series Variables (Average Squared Canonical Correlation =.5889)			
vb6	B6 Cover Out Net 1Y/0N	1	0.6718	
vb10	B10 Require PCP Dx 1Y/0N	2	0.6528	
vb13	B13 Coinsurance/copayment 1=Dollar 0=%	3	0.2781	
vb8	B8 Self Referral to In-Network 1=Yes	4	0.1068	
vb91b	Does this coverage apply to most in network	5	0.1165	
vc4_2_1	C4 Fee For Service	6	0.0768	
vb14d_3	B14 $ 100-$300	7	0.0357	
vb91c	Does this coverage apply to most GOB?GYN	8	0.0352	
vb14d_1	B14 $0	9	0.0176	
vc4_2_2	C4 Fixed Fee Schedule	10	0.0165	
	Model Using Same Items as in Round 1 (Average Squared Canonical Correlation =.5553)			
VB6	B6 Cover Out Net Docs 1Y/0N	1	0.6718	0.7338
VB10	B10 Require PCP 1Y/0N	2	0.6528	0.6580
VB13	B13 Coinsurance/copayment	3	0.2781	0.1738
VB8	B8/B91A Self Refer in Net 1Y/0N	4	0.1068	0.0852
VC4_2_1	C4 Payment Method PCP=FFS	5	0.0894	0.2105
VB14D_1	B14 $0	6	0.0393	0.0437
VC4_2_2	C4 Disc FFS	7	0.0254	0.0637
VB14D_3	B14 $ 100-$300	8	0.0183	0.0069
VB14D_2	B14 $ 0-$100		Not Entered	0.0015

[a]Listed in order of entry.

stepwise procedures. In these models, we set the significance level for the model selection process liberally at 0.15 to ensure that all potential predictors were included in the model. In most cases, we started with a full model. After reviewing the output from the full model and the stepwise procedures, we eliminated variables that were not significant when a chi-square test of significance was performed. Table E.4 presents the list of the top 10 variables (as ranked by the Chi-square test of significance) predictors in each model. For comparison purposes, Table E.4 also provides model fitting diagnostics from the final models in rounds 1 and 2. In general, the models provided about the same level of accuracy in the predictions as in round 1, with HMO status, and B6 showing a slight improvement in the model fit, but a reduction in the fit for B8, and C4_2.

4. Selecting the Linkages

For each of the 1,932 soft-linked cases, the initial matching procedures appended from two to nine potential products to each policy to yield 8,689 potential soft-linked products. Table E.5 shows a frequency distribution of the number of potential products linked to each policy. Contrary to round 1, most policies (86.5 percent) had more than two choices.

We selected one of the soft-matched products as the final product for a policy using a three-step process. For the first step, we computed predicted values for each of the nine product attributes for the 1,932 soft-linked policies, using the coefficients from the models presented in Table E.4. The predicted values were computed using the general formula given in equation (1):

$$(1) \quad \hat{A}_j = \frac{\exp\left(\sum_{i=1}^{k} b_i \times x_{ij}\right)}{1+\exp\left(\sum_{i=1}^{k} b_i \times x_{ij}\right)}$$

TABLE E.4

	Hmo Status	Out Of Network Coverage Of Primary Care Physician Visits	Coverage Of Self-Referral To In-Network Specialists	Primary Care Physician Sign-Up Required	Co-Insurance Or Co-Payment Requirement	Primary Care Physician Payment Methods Is Fee-For-Service (FFS)	Primary Care Physician Payment Methods Is Discounted Fee-For-Service (FFS)	Primary Care Physician Payment Methods Is Capitation	Network St atus
Variable	HMO	B6	B8	B10	B13	C4_1	C4_2	C4_4	NET
				Round 1					
R-Square	29.7%	16.5%	9.0%	30.5%	14.4%	11.3%	10.2%	21.7%	12.4%
HL Test (p-value)	0.0036	0.0001	0.0035	0.0001	0.0083	0.0001	0.0001	0.0102	0.0014
Pearson (p-value)	0.0001	0.0092	0.0001	0.0001	0.0019	0.03	0.0002	0.0016	0.0001
				Round 2					
R-Square	33.9%	22.2%	5.5%	29.2%	12.4%	6.2%	9.3%	16.8%	10.9%
HL Test (p-value)	0.0207	0.037	0.0098	0.023	0.6876	<0.0001	<0.0001	0.0025	0.0027
Pearson (p-value)	0.0291	0.2362	0.4623	0.3994	0.0019	0.0188	0.2928	0.7952	0.0001
Differences									
R-Square R2 vs R1	4.2%	5.7%	-3.5%	-1.3%	-2.0%	-5.2%	-0.9%	-4.9%	-1.5%
				Top 10 Predictors					
1	Plan Is An HMO (Like B2=1)	Plan Is An HMO (Like B2=1)	Plan Pays W/O Referral (Like B6=1) Has Missing Value	Plan Is An HMO (Like B2=1)	Plan Pays W/O Referral (Like B6=1) Has Missing Value	Plan Is An HMO (Like B2=1)	Plan Is An HMO (Like B2=1)	Plan Is An HMO (Like B2=1)	Plan Pays W/O Referral (Like B6=1) Has Missing Value
2	Plan Pays W/O Referral (Like B6=1)	Plan Pays W/O Referral (Like B6=1)	Plan Is An HMO (Like B2=1)	Plan Requires Referral (Like B8)	Plan Is An HMO (Like B2=1)	Plan With List Of Phys (Like NET/B5)	Common Place Of Care Is HMO	Member Of Sites 13-48	Plan Requires Referral (Like B8)
3	Plan Requires Referral (Like B8)	Plan Pays W/O Referral (Like B6=1) Has Missing Value	Employee Of State Gov.	Member Of Sites 52-60	Employee Of State Gov.	Member Sites 52+	Sites 13-48	Plan Requires Referral (Like B8)	Employee Of State Gov.
4	Common Place Of Care Is HMO	Common Place Of Care Is HMO	Plan Pays W/O Referral (Like B6=1)	Plan Pays W/O Referral (Like B6=1) Has Missing Value	Plan Requires Referral (Like B8)	Plan Requires Referral (Like B8)	Plan Pays W/O Refer(Like B6=1)	Plan Pays W/O Referral (Like B6=1)	Plan With List Of Phys (Like NET/B5
5	Policy Holder Years In HMO Is Missing	Plan Requires Referral (Like B8)	Years Enrolled In HMO	Policy Holder Years In HMO Is Missing	Plan With List Of Phys (Like NET)	Employee Of State Gov.	Government Emp.	Common Place Of Care Is HMO	Member Of Sites 13-48
6	Member Of Sites 52-60	Member Of Sites 52-61	Policy Holder Completed High School	Plan Has List Of Phys (Like NET/B5)	Employee Of Local Gov.	Employee Of Federal Gov.	Common Place Of Care Is Go To Other Health Center	Policy Holder Years In HMO Is Missing	Yrs Enrolled In HMO
7	Plan Requires Sign Up (Like B10) Has Missing Value	Employee Of State Gov.	Previous Insurance Is Missing	Common Place Of Care Is HMO	Member Of Sites 13-48	Firm Size=1-9 Emp.	Employee Of Local Gov.	Member Sites 1-12 High Intensity	Prior Insurance Is None
8	Government Employee	Previous Insurance Is Missing	Government Employee.	Plan Requires Sign Up (Like B10) Has Missing Value	Common Place Of Care Is Hmo	Age 60 & Above	3 Or More Kids In Family	Member Of Sites 52-61	Member Sites 1-12 High Intensity
9	Common Place Of Care Is Go To Other Health Center	Race Is Asian Pacific	Common Place Of Care Is Hospital	See Same Staff At Usual Pl Ace Of Care	Race Is Asian Pacific	Usually See Nurse	Age 60 & Above	Prefer More Choice In Plans	2 Kids In Family
10	Prefer More Choice In Plans	Policy Holder Years In HMO Is Missing	Smoking Above 20 Cigarettes	Firm Size=10-99 Employees.	Firm Size=1-9 Emp.	1 Kid In Family	Member Of Sites 52-60	Common Place Of Care Is Doctor Office	Plan Requires Referral (Like B8)

TABLE E.5

DISTRIBUTION OF THE NUMBER OF POTENTIAL LINKS
ASSOCIATED WITH EACH SOFT-LINKED POLICY

Number of Potential Links	Records (Number)	Cases (Number)
1	37	37 (Re-statused as hard links)
2	128	64
3	477	159
4	3,156	789
5	2,700	540
6	1,488	248
7	455	65
8	176	22
9	72	8
Total	**8,689**	**1,932**

where b_i denotes the coefficient associated with a Household Survey policy-level variable, i as estimated from the logistic regression procedures, and $x_{i,j}$ denotes the value of that characteristic for policy j. The procedures in equation (1) produced a predicted value for each of the nine attributes that represented the estimated probability that the policy had the associated trait.

We then compared the predicted values of the nine attributes with the actual values among the linked products. For each possible link, we computed the absolute difference between the predicted and actual value. This computation produced nine "gap" measures for each potential product link. Because the predicted value was the estimated probability of having the trait, the gap measures had the form of either (1) the absolute difference between a value of zero (not having the trait) and the predicted probability, or (2) the absolute difference between a value of one (having the trait) and the predicted probability. We then applied the model coefficients from the model prepared in the validation step in round 1 to convert the gap measures into an estimated probability of a match. The coefficients from round 1 model are presented in Table E.6. Finally, we selected the match with the highest probability as the final link. In 25 of the 64 choice sets with 2 products each, the products had the same probabilities of a match, and therefore, we selected the final product at random.

For round 2, we decided not to validate the linkage selection procedures by preparing a new set of "mock" files to examine various linkage selection procedures. Since the round 1 models and corresponding linkage methodology showed a good estimated accuracy rate and we expected the data relationships for round 2 to be similar; consequently conducting a new validation study was not cost-effective.

TABLE E.6

RESULTS OF THE LOGISTIC REGRESSION ANALYSIS
TO PREDICT LINK STATUS FROM ROUND ONE

Gap Measure	Round 1 Model Estimated Coefficient
Intercept	1.5326
HMO Status	1.0844
B6	2.2476
B8	1.1188
B10	1.9837
B13	0.3831
C4 (FFS)	0.4442
C4 (Disc FFS)	0.8406
C4 (Capitation)	0.2802
Network Status	0.5086

APPENDIX F

PREPARATION OF SURVEY WEIGHTS AND SUDAAN SPECIFICATIONS

I. WEIGHTING ADJUSTMENT FOR FOLLOWBACK SURVEY NONMATCHES

In this section, we discuss how we accounted for policies that we were able to link between the Household and Followback surveys. In part A, we describe the weighting approach used to account for these non-linkages. In part B, we discuss methods for selecting one plan when a person was covered by multiple plans. We then describe different aspects of the modeling process used to determine the weighting adjustment factor, including the selection of independent variables (part C) and the modeling results (part D). In part E, we describe the weighting adjustment and in part F the Followback person level weight to be used for national estimates.

A. Weighting Approach

Some policies that Household Survey respondents described did not have a corresponding record in the Followback Survey.[52] We decided to adjust for these non-linkages in the weights, rather than perform probabilistic matching with the Followback data, as was conducted for the soft matches. The weighting adjustment is based on the inverse of the modeled probability of a link.

B. Dealing with Persons Covered by Multiple Plans

Because we were going to adjust person-level weights for non-linked policies, we had to select one plan for persons who were covered by multiple plans. We developed the following hierarchy to choose among multiple plans: (1) status as a policyholder took precedence over whether the policy was linked to a product, (2) coverage by a policy that was a hard link took

[52]Weights were not applied to policies linked to households outside the 60 sites, since these households were not included in the Followback survey.

precedence over a policy that was a soft link or a non-link, and (3) soft links took precedence over non-links. If policyholder status and link status were insufficient to narrow the choices to one policy, we chose the policy the respondent had mentioned first.

C. Selection of Independent Variables

For the modeling, we had to determine which variables in the Household Survey would be good predictors of a link. For this purpose, we considered links to be Household Survey-reported policies that were hard links, soft links, and ineligible plans. All other policies within the 60 sites were considered non-links. We started with the set of variables that were used in the round 1 modeling; that is, those variables that were read into the round 1 stepwise regression modeling. In addition, we added a new set of variables that indicated whether the name of any of five insurance entities was mentioned by the Household Survey respondent (Aetna, Blue Cross Blue Shield, Cigna, United HealthCare, or Humana).

We then developed a model that we believed would best predict a link for national estimates. The policy was the analytic unit for these models. The weight was based on the final family insurance unit (FIU)-level weight used for the augmented site sample national estimates, and then normalized so that the sum of the weights was equal to the unweighted sample size. The dependent variable for these logistic regression models was always the dichotomous link variable.

D. Summary of Modeling Results

Using a stepwise approach (slentry = slstay = .15), we arrived at a single model for adjusting weights for national estimates (based on the augmented site sample). Information about this model can be found in Table F.9, including the independent variables, a brief description of each, their coefficients, and their levels of statistical significance.

TABLE F.9

LOGISTIC REGRESSION, WEIGHTED BY NATIONAL WEIGHT
BASED ON AUGMENTED SAMPLE

Variable	Parameter Estimate	Pr > Chi-Square	Variable Description
INTERCPT	-0.7195	0.0001	Intercept
PLAN_NUM	-0.0833	0.1196	Plan number 1-3
PB33	0.0487	0.2938	Plan requires signing up with primary doctor
PB34	0.1994	0.0001	Plan requires referral for specialists
PB35	-0.0025	0.9607	Plan has a list of doctors
PB36	0.1967	0.0001	Plan is an HMO
IPRSIG	-0.4868	0.0001	PB33 is imputed
IPRREF	-0.1567	0.0143	PB34 is imputed
IPRLST	-0.3881	0.0001	PB35 is imputed
IPRHMO	-0.3227	0.0001	PB36 is imputed
_PRVMOR	-0.9357	0.0001	PRVMOR (employer offers more than one plan) missing (logical skip)
_PRVMORD	-0.2505	0.0003	PRVMOR (employer offers more than one plan) is coded other missing
PEMPGOVT	0.133	0.0089	Person is government employee (PHETYPE=2,3,4)
CPLACE2	0.2726	0.0009	Usual source of care (CPLACE) is an HMO
CPLACE0	-0.1777	0.0011	Usual source of care (CPLACE) is none
_CPLACE	-0.2833	0.0301	Usual source of care (CPLACE) is missing
LARGFIRM	0.1232	0.0029	Employer size is 1000 or more (PHEMPSZ>=6)
HSGRAD	0.1473	0.0483	High school graduate (or missing value)
AETNA	3.8503	0.0001	Entity name contains "Aetna"
BCBS	3.1214	0.0001	Entity name contains "BC/" (Blue Cross Blue Shield)
CIGNA	3.3331	0.0001	Entity name contains "CIGNA"
UNITED	2.2607	0.0001	Entity name contains "United HealthCare"
HUMANA	1.9822	0.0001	Entity name contains "Humana"
SITE1	1.6968	0.0001	(Site indicator variables)
SITE2	0.6353	0.0001	
SITE3	-0.4608	0.0001	
SITE4	0.1724	0.0723	
SITE6	-0.4663	0.0001	
SITE8	0.2495	0.0255	
SITE9	0.7912	0.0001	
SITE10	0.5595	0.0001	
SITE11	0.5282	0.0001	
SITE12	-0.3717	0.0006	
SITE13	0.5193	0.0069	
SITE15	-0.9571	0.0001	
SITE16	-0.3934	0.0259	
SITE18	-0.2999	0.137	
SITE19	0.4687	0.0068	
SITE20	0.8423	0.0001	
SITE21	0.3682	0.045	
SITE23	-0.5444	0.0208	

TABLE F.9 *(continued)*

Variable	Parameter Estimate	Pr > Chi-Square	Variable Description
SITE24	0.5576	0.0093	
SITE25	-1.3779	0.0001	
SITE26	-0.3319	0.0359	
SITE27	-0.292	0.0184	
SITE29	0.6994	0.0001	
SITE30	0.5954	0.0006	
SITE31	0.2682	0.1462	
SITE32	0.4202	0.0169	
SITE33	0.6668	0.0001	
SITE34	-1.3331	0.0001	
SITE38	0.7661	0.0174	
SITE39	0.4821	0.0317	
SITE40	-0.3489	0.0624	
SITE44	-0.4053	0.0417	
SITE45	0.6619	0.0009	
SITE46	0.659	0.0001	
SITE47	-0.4484	0.0292	
SITE48	1.6056	0.0001	
SITE49	-1.8532	0.0001	
SITE52	-2.4702	0.0001	
SITE55	0.1815	0.1028	
SITE56	-1.048	0.0001	
SITE59	-0.5818	0.0001	

	DF	Value	Value/DF	Chi-Square
Pearson Goodness-of-Fit Statistic	10205	12621.1	1.2368	0.0001

RSquare = 0.3512

Hosmer and Lemeshow Goodness-of-fit Statistic = 68.828 with 8 DF (p=0.0001)

E. Weighting Adjustments

We used the predicted probability of a link that the model produced to adjust the appropriate person-level weight from the Household Survey. This adjustment factor was merged onto the person-level file, by policy. If the person was covered by more than one policy, the person-level file already had an indicator for his or her selected policy.

People who resided outside the boundaries of the 60 sites or who did not have private health insurance or who were age 65 or older were out of scope for the Followback Survey. These individuals were assigned their final CTS Household Survey person-level weights as their "Followback weights" (WTPER1 for site-specific estimates based on the augmented sample, and WTPER5 for national estimates based on the augmented site sample). We set the two Followback weights to zero if a person was part of the Followback process, but his or her policy was a non-link.

We set the Followback weights of people whose policies were hard or soft-matches (22,235) or ineligible plans (1,343) equal to their final CTS Household Survey person-level weights, multiplied by the inverse of the probability of a match from the models. That is:

> If P is the predicted probability of a link from the national model, then,
>
> FBWTPER1 = (1/P) × WTPER1, for site-specific estimates based on the augmented site sample
>
> FBWTPER5 = (1/P) × WTPER5 for national estimates based on the augmented site sample

We then poststratified (raked) and trimmed outliers for the person-level weights. We used a constrained least squares raking procedure for the national estimate and a weighted least squares raking procedure for the site-specific Followback weight to achieve the same weighted

proportions.[53,54] Both procedures differ from the traditional iterative proportional fitting procedure in that they use a least squares loss function (also in an iterative fashion) to find a raked set of survey weights that meet the desired constraints while minimizing the squared differences between the pre and post-raked weights. The program also offers the opportunity to control the minimum or maximum size and relative change in the weights (as such is referred to as a constrained raking procedure) that results from the raking procedure. Hence, it offers a greater control over the variation added to the weights from the calibration process. The weighted procedure differs from the other approach by minimizing the relative squared difference, rather than the actual difference between the pre-raked weights and the new weights. As such, the square differences are minimized relative to the starting weights which gives this process its name.

For the national estimates we raked the weight WTPER5 (the final CTS Household Survey person-level weight for national estimates based on the augmented site sample) for the full CTS sample with separate target distributions for the cases with and without private insurance that were age 65 or under and were members of the 60 CTS sites. The raking procedure aligned the full sample on six dimensions: (1) telephone interruption status; (2) age group, by sex; (3) Hispanic, by sex; (4) black, by sex; (5) educational level; and (6) for those with private insurance, by HMO status. After trimming outlier weights (using the methodology that was used for the original Household Survey person-level weights), we re-raked on the same six factors.

---

[53]Claude Deville and Carl-Erik Sarndal (1993) "Generalized Raking Procedures in Survey Sampling." JASA, vol 88, no 423, pages 1013-1020.

[54]Claude Deville and Carl-Erik Sarndal (1992) "Calibration Estimators in Survey Sampling." JASA, vol 87, no 418, pages 376-382.

We also used a raking and trimming procedure for the site-specific Followback weights to achieve the same within-site weighted proportions, using WTPER1 (the final CTS1 Household Survey person-level weight for site-specific estimates based on the augmented site sample. Given the smaller sample sizes, we were able to use the weighted least squares raking procedure which helped to further reduce the impact of the raking/adjustment process. We conducted the raking within sites and carried out the adjustments separately for high-intensity sites and low-intensity sites. We conducted these steps on the full CTS sample in each site as for the national estimates. For each of the high-intensity sites, we used five factors in the raking process: (1) telephone interruption status; (2) age group; (3) Hispanic, by sex, (4) black, by sex; and (5) if privately insured, HMO status. After trimming outlier weights, we conducted the raking again to realign the weighted distribution on these factors. For the low-intensity sites we poststratified on two items: (1) telephone interruption status, and (2) a combination variable with six values (child, adult female, and adult male crossed with HMO status). After trimming the outlier weights, we re-poststratified as for the high intensity sites.

F. Summary of Weighting

The Followback–adjusted person-level weight to be used for national estimates based on the augmented site sample is named FBWTPER5. We set this weight to missing for individuals outside the boundaries of the 60 sites and over the age of 65 (n = 3,539). For those within the boundaries of the 60 sites, age 65 or under, and without private insurance (that is, individuals who were not part of the Followback process), we initially set the Followback weight to what would have been the original person-level weight for this type of estimate (n = 21,470). For persons with private insurance, in the 60 sites and age 65 or under, we adjusted the original weights for the 24,323 matched and ineligible cases (including the 22,235 matched cases and the

2,088 designated as not-a-health plan – ineligible plans) to account for the nonmatches and set the weights for the nonmatches (n = 13,163) to zero. For individuals with positive values for FBWTPER5 (n = 42,254), the design effect due to unequal weighting was 2.72. For those with Followback matches (n = 22,235), the average design effect due to unequal weighting was 2.85.

The Followback–adjusted person-level weight to be used for site-specific estimates based on the augmented site sample is named FBWTPER1. We computed this weight using similar methods as for FBWTPER5, assigning missing values to those outside the boundaries of the 60 sites or over the age of 65 (n = 3,539) and to zero for those with nonmatches (n = 13,163). Table F.10 shows the design effect for each site due to unequal weighting for 22,235 Followback matches.

TABLE F.10

DESIGN EFFECT DUE TO UNEQUAL WEIGHTING FOR FOLLOWBACK
SITE-SPECIFIC WEIGHT (FBWTPER1)
(Among persons for whom FBWTPER1 is greater than zero.)

Site Number	Frequency	Coefficient of Variation	Design Effect Due to Unequal Weighting
1	1,895	63.514	1.40341
2	1,791	70.52	1.49731
3	1,763	87.895	1.77256
4	1,716	75.848	1.5753
5	1,853	88.132	1.77672
6	1,807	87.016	1.75717
7	1,627	81.832	1.66965
8	1,908	69.478	1.48271
9	1,793	72.019	1.51867
10	1,952	74.355	1.55286
11	1,463	69.255	1.47963
12	1,575	80.439	1.64704
13	446	77.191	1.59584
14	411	84.367	1.71178
15	353	94.018	1.88394
16	336	97.551	1.95163
17	555	93.67	1.87741
18	395	78.822	1.6213
19	439	92.206	1.85019
20	524	75.387	1.56832
21	386	62.064	1.38519
22	474	70.192	1.49269
23	379	94.572	1.89438
24	463	85.14	1.72488
25	362	93.421	1.87275
26	353	79.515	1.63226
27	557	76.792	1.5897
28	476	86.614	1.75019
29	382	80.003	1.64005
30	490	65.953	1.43498
31	475	71.593	1.51255
32	540	83.456	1.6965
33	508	76.997	1.59285
34	361	121.156	2.46789
35	493	72.716	1.52876
36	426	77.817	1.60555
37	500	95.055	1.90355
38	646	58.241	1.3392
39	499	79.051	1.62491

TABLE F.10 *(continued)*

Site Number	Frequency	Coefficient of Variation	Design Effect Due to Unequal Weighting
40	291	92.58	1.8571
41	351	80.532	1.64853
42	430	78.828	1.62138
43	467	80.192	1.64308
44	401	82.864	1.68665
45	530	68.233	1.46557
46	585	70.212	1.49298
47	362	95.199	1.90629
48	522	57.303	1.32837
49	344	127.388	2.62277
50	342	92.059	1.84748
51	399	81.842	1.66981
52	281	116.658	2.36091
53	557	78.697	1.61933
54	413	93.27	1.86993
55	404	90.048	1.81086
56	322	107.595	2.15767
57	499	72.704	1.52859
58	432	80.369	1.64592
59	532	109.206	2.1926
60	418	101.648	2.03323

SUDAAN SPECIFICATIONS

Round Two Household Survey With Followback Data

---

1. Site estimates from the Followback survey (augmented site sample)

```
PROC DESCRIPT data="&workdir\\asite" filetype=sas design=wr;
    nest site_str fsu / missunit;
    weight FBWTPER1;
    subgroup siteid;
    levels 60;
    var unmetx putoffx unmetput;
    setenv linesize=159 pagesize=52;
    print nsum wsum mean semean deffmean / style=nchs
        wsumfmt=f10.0 meanfmt=f8.4 semeanfmt=f8.4 deffmeanfmt=f8.4;
    output nsum wsum mean semean deffmean /
        filename="&workdir\\sud1" filetype=ascii replace
        wsumfmt=e25.16 meanfmt=e25.16 semeanfmt=e25.16 deffmeanfmt=e25.16;
    title "Augmented Site Estimates";
```

Weight

FBWTPER1

Nest Variables

SITE_STR (Values: 101-6000)
12 High Intensity Sites: Values: 101-1209
48 Low Intensity Sites: Values: 1300-6000

FSU
Phone sample: Household ID
In person sample: Segment ID

2. National estimates from the 60 site sample

PROC DESCRIPT data="&workdir\\nsite" filetype=sas **design=uneqwor;**
nest pstrataf ppsuf secstraf nfsuf / missunit;
totcnt pstrtotf _zero_ _minus1_ _zero_;
jointprob p1f p2f p3f p4f p5f p6f p7f;
weight FBWTPER5;
subgroup msacat;
levels 3;
var unmetx putoffx unmetput;
setenv linesize=128 pagesize=55;
print nsum wsum mean semean deffmean / style=nchs
wsumfmt=f10.0 meanfmt=f8.4 semeanfmt=f8.4 deffmeanfmt=f8.4;
output nsum wsum mean semean deffmean /
filename="&workdir\\sud4" filetype=ascii replace
wsumfmt=e25.16 meanfmt=e25.16 semeanfmt=e25.16 deffmeanfmt=e25.16;
title "National Estimates";

Weight

FBWTPER5

Nest Variables

PSTRATAF (Values: 1-20)
Categorization of the 60 sites into 20 strata:
a. Certainty selections in PSTRATAF 1-9
b. Unequal probability without replacement in PSTRATAF 10-18, 20
c. With replacement sampling in PSTRATAF 19 (non-metro sites 52-60)

PPSUF (Values: 1-203)
Recode of the sites from 1-52 to 52 unique values between 1 and 203.
For non-metro sites 52-60 (PSTRATAF 19), PPSUF is set to A1".

SECSTRAF (Values: 0-6, 9, 19)
Second stage stratification variable.
Phone sample: Set equal to STRATUM (0-6)
In-person sample: Set equal to @9"
For non-metro sites 52-60 (PSTRATAF 19), SECSTRAF set to A19".

NFSUF
Phone sample: Household ID
In person sample: Segment ID
For households in non-metro sites 52-60 (PSTRATAF 19), NFSUF equals values 191 to 199.

TOTCNT Variables

Provides the population counts at each stage of the sample design; one-to-one (in order) link between the NEST and TOTCNT variables.

PSTRTOTF (Values 1-118)

TOTCNT variable for PSTRATAF - based on the actual strata frame sizes (except in PSTRATAF 19).

```
if (1 le pstrataf le 9) then pstrtotf=1;
else if pstrataf=10 then pstrtotf=18;
  else if pstrataf=11 then pstrtotf=15;
  else if pstrataf=12 then pstrtotf=24;
  else if pstrataf=13 then pstrtotf=17;
  else if pstrataf=14 then pstrtotf=20;
  else if pstrataf=15 then pstrtotf=21;
  else if pstrataf=16 then pstrtotf=14;
  else if pstrataf=17 then pstrtotf=23;
  else if pstrataf=18 then pstrtotf=15;
  else if pstrataf=19 then pstrtotf=1;
  else if pstrataf=20 then pstrtotf=118;
```

For PSTRATAF 19 (non-metro sites 52-60), PSTRTOTF set to "1."

ZERO - SUDAAN keyword which generates a value of "0" for every record.

MINUS1 - SUDAAN keyword which generates a value of "-1" for every record. .

Joint Inclusion Probabilities (JOINTPROB)

The JOINTPROB statement names the variables that give the single and joint inclusion probabilities for each PSU and each pair of PSUs in each first-stage stratum. The probabilities are an *n*-by-*n* matrix, where *n* is the number of PSUs in each stratum. In the household sample, there are 20 stratum and 60 PSUs. Stratum 10, for example, contains 5 PSUs (5 sites) and a 5x5 matrix of joint probabilities (P1F-P5F).

P1F P2F P3F P4F P5F P6F P7F (Values: 0 > P(i) <= 1)
For PSTRATAF 19: P1F=1, P2F-P7F=missing.

APPENDIX G:

DATA IMPUTATION SPECIFICATIONS AND RESULTS

The specifications presented in this appendix represent an abbreviated version of the specifications provided to Social and Scientific Systems (SSS). Sections on the background of the survey and an overview of the methodology as described in Appendix F, have been omitted from the original specifications to reduce redundancy.

Variables To Create For Imputations:

Gatekeeper (create variable GATEK):

Value	Situation
1	IB10=Yes(1) or **{**IB8=Yes(1) and IB91A=**No(2)}**.
0	Otherwise

These variables are not needed until after the component items are imputed, so please wait to compute the values until after IB10, IB8 and IB91A are imputed.

1. Eight B2/B2A categories (create variable B2_8cat)

Value	Situation
1.	IB2=1,HMO, {IB2A_1= 1 (Staff) or IB2A_2=1 (Group model)} and IB2A_3=2.
2.	IB2=1,HMO, IB2A_3= 1 (Network or IPA model) and {IB2A_1=2 and IB2A_2=2}.
3.	IB2=1,HMO, All other combinations of IB2A_1-IB2A_3 (Mixed Model)
4.	IB2=2,POS, {IB2A_1= 1 (Staff) or IB2A_2=1 (Group model)} and IB2A_3=2.
5.	IB2=2,POS, IB2A_3= 1 (Network or IPA model) and {IB2A_1=2 and IB2A_2=2}.
6.	IB2=2,POS, All other combinations of IB2A_1-IB2A_3 (Mixed Model)
7.	IB2=3, PPO
8.	IB2=4, FFS

1. Five B2/B2A categories (create variable B2_5cat)

Value	Situation
1.	IB2=1 or 2 ,HMO/POS, {IB2A_1= 1 (Staff) or IB2A_2=1 (Group model)} and IB2A_3=2}.
2.	IB2=1 or 2,HMO/POS, IB2A_3= 1 (Network or IPA model) and {IB2A_1=2 and IB2A_2=2}.
3.	IB2=1 or 2,HMO/POS, All other combinations of IB2A_1-IB2A_3 (Mixed Model)
4.	IB2=3,PPO
5.	IB2=4, FFS

As for GATEK, please compute the variables B2_8cat and B2_5cat **after** IB2A_1-IB2_3 are imputed. Hence, the GATEK, B2_8cat and B2_5cat variables only need to be computed once during the imputation runs.

Creation of Mean Value Sorting Variables to Be Used in Imputations

As needed, SSS should create mean values (or the percentage of cases having the trait) on a site (PSU) basis for the variables indicated (using the "I" variables) to develop the necessary sorting variables (**weight the means and proportions by the variable SUMPWGT**). **In computing the means, limit the cases used to those with non-missing and non-skipped (-1) values.** If there are no cases in the site that meet the condition set the mean value or percentage to zero (see note below)**.** The mean values should be computed after each variable is imputed (not necessary after each staged step) so that the means are based on the updated information from the prior round (item) of imputation.

.

TABLE G.1

IMPUTATION SPECIFICATIONS BY ITEM
IMPUTE IN ORDER PRESENTED

Questionnaire Item	Variable Name(s)	Initial Computational Steps and the List of Classing Variables to Use for Each Step	Sorting Variables (Site Based)	Notes on 2,946 Records
1. B2A_1 B2A_2 B2A_3 Imputed as vector	IB2A_1 – IB2A_3	The only "missing" situation is when all three of IB2A_1, IB2A_2 and IB2A_3 are missing,. Hence we impute these as vector using classing variables: 1. IB2, IC4_2, and IC5 2. IB2 Note: missing values for IC4_2 and IC5 are allowed as classing categories.	Percentage of HMO products (IB2=1) that use a mixed model type in each site based on values in IB2A_1-IB2A_3. Mixed model - combinations of IB2A IB2A_1 IB2A_2 IB2A_3 1 2 1 2 1 1 1 1 1	220 Missing 9 imputed in Step 1, 211 in Step 2
2. B8, B10 B91A Imputed as vector or individually depending on missing pattern.	IB8 IB10 IB91A	Four missing patterns: A. IB8, IB10 and IB91A all missing. Impute as vector from: 1. IB2, IB6. B. IB8, and IB91A missing. Impute IB8 and IB91A as vector from: 1. IB2, IB6 and IB10 2. IB2, IB6 C. IB10 missing only. Impute IB10 from: 1. IB2, IB6, IB8 and IB91A 2. IB2 and IB6 D. IB91A missing only. Impute IB91A from: 1. IB2, IB6, and IB8 2. IB2 and IB8	Percentage of products with IB6=1, IB8=1, IB10=1 and IB91A=1 of those with non-missing and non-negative (-1) values on these items and IB6 not equal to –1.	Situation A, 5 missing - All step 1 B, 12 missing All step 1 C, 2 missing All step 1 D, 277 missing 222 in step 1 55 in step 2

TABLE G.1 *(continued)*

Questionnaire Item	Variable Name(s)	Initial Computational Steps and the List of Classing Variables to Use for Each Step	Sorting Variables (Site Based)	Notes on 2,946Records
3. B12_2 B12_3 Imputed as vector	IB12_1-IB12_3	If IB12_1-IB12_3 all missing, then : If IB10=2 or –1, set all of IB12_1-IB12_3 to –1. If IB10=1 and IB12_1=., then set IB12_1=1. Two missing patterns remain for IB10=1. A. IB12_2 and IB12_3 both missing. Impute IB12_2 and IB12_3 as vector from: 1. B2_8cat 2. B2_5cat B. Only IB12_3 missing, impute from: 1. B2_8cat and B12_2 2. B2_5cat and B12_2	Percentage of products with (IB12_2=1 or IB12_3=1) (OBGYN and specialists) in each site.	7 cases have IB10=missing which were partially resolved in prior IB10 imputation. Situation A, 164 missing All in step 1 B, 3 missing + any of 7 cases with IB10=1. All in step 1

TABLE G.1 *(continued)*

Questionnaire Item	Variable Name(s)	Initial Computational Steps and the List of Classing Variables to Use for Each Step	Sorting Variables Site-Based	Notes on 2,946 Records
4. B13	IB13	1. B2_8cat and IC4_2 2. B2_8cat 3. B2_5cat	Percentage of cases answering IB13=2 (percent) in each site.	203 missing: 135 in step 1 40 in step 2 28 in step 3
5. B14	IB14	1. B2_5cat and IB13 2. B2_5cat 3. B2_COL B2_COL = 1 if B2=1,2 =2 if B2=3,4	Mean of IB14 by site.	483 missing: 375 in step 1 43 in step 2 65 in step 3
6. B13AMT	IB13AMT	If IB13AMT=., then if IB13=0, or 2 set IB13AMT=-1. 1. IB13, B2_5cat, and IB14 (3 classes (0, 1-300,301+) 2. IB13 and B2_5cat 3. IB13 and B2_COL	For IB13=1, mean of IB13AMT (exc. Values –1) by site.	431 missing: 54 set to skip 299 in step 1 None in step 2 78 in step 3
7. B13PER	IB13PER	If IB13PER=., then if IB13=0, or 1 set IB13PER=-1. 1. IB13, B2_5cat, and IB14 (3 classes (0,1-300,301+) 2. IB13 and B2_5cat 3. IB13 and B2_COL	For IB13=2, mean of IB13PER (exc. Values –1) by site.	245 missing: 155 set to skip 67 in step 1 23 in step 2
8. B13OUT	IB13OUT	1. IB6 and B2_8cat and IB14OUT 2. IB6 and B2_8cat 3. IB6 and B2_5cat 4. IB6 and B2_COL	Percentage IB13OUT=1 (exc. Values –1) in site.	71 missing: 18 in step 1 53 in step 2
9. B14OUT	IB14OUT	1. IB6, B2_8cat and IB13PER, 3 classes (-1,10-20,25+) 2. IB6, and B2_8cat 3. IB6 and B2_5cat 4. IB6 and B2_COL	Percentage IB14OUT=1 (exc. Values –1) in site.	156 missing: 122 in step 1 10 in step 2 24 in step 3
10. B14OD	IB14OD	If IB14OD=., then if IB14OUT=-1,2 set IB14OD=-1. 1. IB14OUT, B2_8cat, and IB13AMT, 3 classes (-1,1-10, 11+) 2. IB14OUT and B2_8cat 3. IB14OUT and B2_5cat 4. IB14OUT and B2_COL	Mean value of IB14OD (exc. values of –1) in site.	240 missing: 35 set to skip 134 in step 1 46 in step 2 3 in step 3 22 in step 4
11. B13OP	IB13OP	If IB13OP=., then if IB13OUT=-1,0 or 1 then set IB13OP=-1. 1. IB13OUT,B2_8cat and IB14 3 classes(0, 1-300, 301+) 2. IB13OUT and B2_8cat 3. IB13OUT and B2_5cat 4. IB13OUT and B2_COL	Mean value of IB13OP (exc. values –1) in site.	176 missing: 16 set to skip 159 in step 1 1 in step 2

TABLE G.1 *(continued)*

Questionnaire Item	Variable Name(s)	Initial Computational Steps and the List of Classing Variables to Use for Each Step	Sorting Variables Site-Based	Notes on 2,946 Records
12. C4_2	IC4_2	1. B2_8cat, Gatekeeper (GATEK), and NET 2. B2_8cat and NET 3. B2_5cat and NET 4. NET	Percentage IC4_2=4 (capitated) in the site.	621 missing: 283 in step 1 103 in step 2 111 in step 3 124 in step 4

APPENDIX H

FOLLOWBACK ROUND 2
PRODUCT FILE DOCUMENTATION

FOLLOWBACK ROUND 2 PRODUCT FILE

The Follow-back product data file (IMPUTE3.SD2) contains the health plan attributes as collected from the insurance entity or from published sources for all the linkable CTS policies (12,263: composed of 10,368 hard linkages and 1,895 soft linkages). This file is linked to the household data via the linkage file (LINKAGEF.SD2) documented separately using a combination of the three variables: FIN_ENT, PROD_ID and PSU which uniquely define a particular entity's product at given location totaling to 2,946 records.

Most of the data for the questionnaire items are stored in a triplicate variable structure with the post-edit/pre-imputation value of the response stored in the variable that is named from the CATI questionnaire item sequence (e.g., B2, B6, B8 etc.). The post imputed and logically edited values are stored in a variable that is preceded by an "I" (e.g., IB2, IB6, IB8). Lastly a flag variable is provided which indicates the source of the value in the "I" variables. The flag variables follow a similar naming format in that an "IF" precedes the questionnaire item. The flag values are as follows:

0 = Reported values from respondent
1 = Value obtained from published sources (or follow-up contacts, post CATI survey)
2 = Value imputed using hot-deck imputation procedures from other respondent values

The following table contains the variable name, description of the variable, and the values for the variable.

File:	IMPUTE3.SD2, SAS v. 6.12 File
Records:	2,946
Identification Variables:	FIN_ENT, PROD_ID, PSU

Item	Imputed Variable	Imputed Flag	Question	Values (exc. missing)
A3	IA3	IFA3 No Hot-deck imputation	Dependent on A3A=2, No Please tell me which of the following categories best describes your organization	2 A licensed insurer or HMO 3 A managed care provider organization, such as a PPO or IPA (not licensed to sell insurance) 4. A Third Party Administrator (TPA) 6. An employer, union or trust plan administrator (including government employee plan 8. Or something else [see A3_OTH] If A3a=1 then A3=-1.
A3A	IA3A	IFA3A No Hot-deck imputation	Are you a Blue Cross/Blue Shield Plan?	1 Yes 2 No
A3_OTH			Text provided by respondent for Other-specify category of A3	Text (for A3=8)
B2	IB2	IFB2	Do you think of [PRODUCT] as a(n) . . .	1. HMO (Health Maintenance Organization) 2. POS Point of Service Plan 3. PPO (Preferred Provider Organization) 4. FFS (Traditional Fee For Service) 5. Or something else? (SPECIFY)
B2_OTH			Text provided by respondent for AOther-please specify A category of B2	Text (if pre-edit B2=5)
B2A_1	IB2A_1	IFB2A_1	Dependent on B2= 1 or 2 Which of the following describes the medical providers available in [SITE]? Model Type = Staff model	1 Yes 2 No If B2=3,4,or 5 then B2A_1, B2A_2, B2A_3, and B2A_4=-1.
B2A_2	IB2A_2	IFB2A_2	Model Type = Group Model	1 Yes 2 No
B2A_3	IB2A_3	IFB2A_3	Model Type = Network Model	1 Yes 2 No
B2A_OTH			Model type = Other	Text (for B2A_4=1 pre-edit)
B3	IB3		Is product ever sold to individuals?	1 Yes 2 No
B5	IB5	IFB5	Is there a directory or list of doctors associated with [PRODUCT] in [SITE]?	1 Yes 2 No If B2=1,2, or 3 then B5=-1
B6	IB6	IFB6	Dependent on B2 and B5 Under [PRODUCT] in [SITE], if enrollees do not have a referral and go to out-of-network doctors, does the plan cover any of the costs for these visits?	1 Yes 2 No If B2=4,5 and B5= 2 then B6=-1
B10	IB10	IFB10	Dependent on B2, B5 and B6 Does [PRODUCT] in [SITE] require members to have a primary care doctor, group of doctors, or clinic to receive maximum coverage for all routine care?	1 Yes 2 No If NET=2 then B10=-1

Item	Imputed Variable	Imputed Flag	Question	Values (exc. missing)
B12_1	IB12_1	IFB12_11	Dependent on B10=1, Yes Which types of providers can serve as a primary care physician: Generalists, such as an internists, pediatricians or family practitioners	1 Yes 2 No If B10 =2 or –1 then B12_1, B12_2 and B12_3= -1
B12_2	IB12_2	IFB12_12	OB/GYNs	1 Yes 2 No
B12_3	IB12_3	IFB12_13	Other specialists	1 Yes 2 No
B8	IB8	IFB8	Dependent on B2,B5 and B6 (see NET=1) Is a referral or authorization ever required to obtain maximum coverage for an initial visit to an in-network specialist?	1 Yes 2 No If NET=2 then B8=-1
B91A	IB91A	IFB91A	Dependent on B8=1, Yes Does product provide for at least some coverage for self-referrals to any types of in-network specialists?	1 Yes 2 No If B8=2 or -1 then B91A=-1.
B91B	IB91B	IFB91B No Hot-deck imputation	Dependent on B91A=1, Yes Does coverage in B91A for self-referral apply to most types of in-network specialists.	1 Yes 2 No If B91A=2 or –1 then B91B=-1.
B91C	IB91C	IFB91C No Hot-deck imputation	Dependent on B91B=2, No Does coverage in B91A for self referral apply to most visits to in-network OB/GYNs	1 Yes 2 No If B91B=1 or –1 then B91C=-1.
B91D	IB91D	IFB91D No Hot-deck imputation	Dependent on B91B=2, No Does coverage in B91A for self-referral apply to any other types of in-network specialists.	1 Yes 2 No If B91B=1 or –1 then B91D=-1.
B92	IB92	IFB92 No Hot-deck imputation	Dependent on B91A=1, Yes When [PRODUCT] covers in-network self referrals, is the level of coverage the same as with a physician referral, or is it less than the coverage with a physician referral?	1 Same 2 Less 3 Volunteer If B91A=2 or –1 then B92=-1.
B92B	IB92B	IFB92B No Hot-deck imputation	Dependent on B92, B91B and B91C What about in-network self-referrals to OB/GYNs - is the level of coverage the same as with a physician referral, or less than with a physician referral?	1 Same 2 Reduced If B92=1,2 or –1 then B92B=-1 Also if B92=3 and combination of B91B/B91C is (2,2) then B92B=-1
B93	IB93	IFB93 No Hot-deck imputation	Dependent on B6 and B92 Under [PRODUCT] when the level of coverage for in-network self-referrals is reduced, is that level of coverage better than for out-of-network self-referrals, or the same?	1 Same 2 Less 3 Volunteer If B6=2,7 or –1, or (if B6=1 and B92=-1 or 1) then B93=-1

Item	Imputed Variable	Imputed Flag	Question	Values (exc. Missing)
B93B	IB93B	IFB93B No Hot-deck imputation	Dependent on B93, B91B and B91C. What about in-network self-referrals to OB/GYNs--Is that level of coverage better than for out-of-network self-referrals, or the same?	1 Better 2 Same If B93=1,2 or –1 then B93B=-1 Also if B93=3 and combination of B91B/B91C is (2,2) then B93B=-1
B13	IB13	IFB13	Under [PRODUCT] in [SITE], what is the co-payment or coinsurance rate [NETWORK PRODUCTS: for in-network office visits]?	0 Neither 1 Co payment (enter dollar amount) 2 Co-insurance rate (enter percentage amount)
B13AMT	IB13AMT	IFB13AMT	What is the typical co-payment amount per office visit for [PRODUCT] in [SITE]?	If B13-1, Dollar Amount -1 for B13=0 or 2
B13PER	IB13PER	IFB13PER	What is the typical coinsurance percentage for office visits under [PRODUCT] in [SITE]?	If B13=2, <1-100> percent -1 for B13=0 or 1.
B13OUT	IB13OUT	IFB13OUT	Dependent on B6=1,Yes For out-of-network office visits without a referral does [PRODUCT] in [SITE] have a fixed co-payment per visit, or percentage coinsurance payment?	1 Co-Payment 2 Coinsurance 0 None If B6=-1,2,7 then B13OUT=-1.
B13OD	IB13OD	IFB13OD No Hot-deck imputation	Dependent on B13OUT=1,Yes What is the typical co-payment amount for out-of-network office visits under [PRODUCT] in [SITE]?	If B13OUT=1, Dollar Amount If B13OUT=-1,0 or 2 then B13OD=-1.
B13OP	IB13OP	IFB13OP	Dependent on B13OUT=2 What is the typical coinsurance percentage for out-of-network office visits under [PRODUCT] in [SITE]?	If B13OUT=2, <1-100> percent If B13OUT=-1,0 or 1 then B13OP=-1.
B14	IB14	IFB14	Under [PRODUCT] in [SITE], what is the dollar amount of the *individual deductible* [NETWORK PRODUCTS: that applies to in-network office visits]?	Dollar Amount
B14OUT	IB14OUT	IFB14OUT	Dependent on B6=1, Yes Is there a separate deductible for [PRODUCT] in [SITE] that applies to out-of-network office visits?	1 Yes 2 No If B6= -1,2,7 then B14OUT=-1.
B14OD	IB14OD	IFB14OD	Dependent on B14OUT=1 What is the dollar amount of the individual deductible for out-of-network office visits?	Dollar Number If B14OUT=2 or –1 then B14OD=-1.
C1R	IC1R	IFC1R No Hot-deck imputation	Dependent on NET=1 Approximately what percentage of all primary care and specialist physicians in [SITE] are associated with [PRODUCT]? Would you say? . . .	1 Fewer than 25% 2 25-50% 3 50-75% 4 75% or more If NET=2 then C1R=-1.
C2R	IC2R	IFC2R No Hot-deck imputation	Dependent on NET=1 Under [PRODUCT] are enrollees limited to a single hospital system for general acute care services in [SITE]?	1 Yes 2 No If NET=2 then C2R=-1.

Item	Imputed Variable	Imputed Flag	Question	Values (exc. Missing)
C4_2	IC4_2	IFC4_2	In [PRODUCT] in [SITE], what is the typical method of payment that your organization uses for primary care services? Is it? . . .	1 Fee-for-service 2 Fixed fee schedule 3 Salaried by your organization 4 Capitation 5 Other
C4_2_OTH			Other Specify Response for C4_2=5 (pre-edit)	Text
C4A_1	IC4A_1	IFC4A_1 No Hot-deck imputation On C4A	Dependent on C4_2=4, Capitation What other services are included in this capitated payment? Referrals to specialists	1 Yes 2 No If C4_2=(1,2,3) then C4A_1-C4A_4=-1.
C4A_2	IC4A_2	IFC4A_2	Hospitalizations	1 Yes 2 No
C4A_3	IC4A_3	IFC4A_3	Other	1 Yes 2 No
C5	IC5	IFC5 No Hot-deck Imputation	Dependent on C4A_1=No In the [PRODUCT] in [SITE], what is the typical method of payment that your organization uses for specialists?	1 Fee For Service (for rates) 2 Discounted Fee-For-Schedule or relative value 3 Salaried by Organization 4 Capitation 5 Other (Specify) If C4a_1=1 then C5=-1.
C5_OTH			Other Specify Response for C5	Text
C6	IC6	IFC6 No Hot-deck imputation	Dependent on C4A_2=No In the [PRODUCT] in [SITE], what is the typical method of payment for hospital services?	1 Per diem 2 According to DRG or per stay 3 Capitation 4 Billed charges 5 Something else (specify) If C4a_2=1 then C6=-1.
C6_OTH			Other Specify Response for C6	Text
C7	IC7	IFC7 No Hot-deck imputation	Does the [PRODUCT] in [SITE] ever include any mental health and/or substance abuse services?	1 Yes 2 No
C7A	IC7A	IFC7A No Hot-deck imputation	Dependent on C7=Yes. Are mental health and/or substance abuse services ever provided or managed separately by a specialty managed behavioral health organization?	1 Yes 2 No If C7=2 then C7a=-1.
C7B			Please tell me the name of the specialty managed behavioral health organization you use in [SITE]? Note: C7B, C7B_ALL and C7C series not edited for consistency.	1 American Psych System 2 Healthcare Value Mgt. 3 MAGELLAN Behavioral Health 4 Managed Health Network 5 MAPSI Mid Atlantic Psych Services 6 Private Health Care Systems (PHCS) 7 Pro Behavioral Health Plan 8 Sagamore 9 Something else [goto c7c]
C7B_ALL			All marked responses in C7B	String of C7B codes

Item	Imputed Variable	Imputed Flag	Question	Values (exc. Missing)
C7C_NAME			In what city and state is this specialty behavior health company located? Name provided:	Text
C7C_CITY			City provided	Text
C7C_ST			State provided	Text
D1	ID1	IFD1	Dependent on A3a and A3 What is your organization's tax status? Is it? . . .	1 For-profit, privately held 2 For-Profit, Publicly held 3 Non profit 4 Other (specify) If A3a=2 and A3=(-1,4,6,8) then D1=-1.
D1_OTH			D1: Other specify response	Text
D2	ID2	IFD2	Is your organization a division or subsidiary of another health plan organization?	1 Yes 2 No
D2A	ID2A	IFD2A	Dependent on D2=Yes Is this parent company a national or multi-state organization?	1 Yes 2 No If D2=2 then D2a=-1
D2B			From D2, What is the name of that parent company? Note: D2B, and D2C series data not edited for consistency.	<s> scroll [parent company list] <o> other [goto d2c] or code entered 01-66 (see D2B code list)
D2C_CITY			For D2B, City of the parent company	Text
D2C_NAME			For D2B, Name of the parent company	Text
D2C_ST			For D2B, State of the parent company	Text
D3	ID3	IFD3	Dependent on D2=No. Is your organization a national or a multi-state organization?	1 Yes 2 No If D2=1 then D3=-1
FIN_ENT			Final Entity ID	Number
NET			Created variable from B2, B5 and B6 Rules: If B2=4,5 and B5= 2 set NET=2 If B2=4,5 and B5=1 and B6=7 set NET=2 If B2=1,2, or 3 and B6=7 then set NET=2 **If B2=4,5 and B5=1 and (B6=1,2) then set NET=1** If B2=1,2, or 3 and (B6=1,2) then set NET=1	1 Yes 2 No
PROD_ID			Product Identification Code	Number
PSU			CTS Site Number 1-60 as coded for CATI Survey	Number
CNAM			Final Entity Name	Text
P_TEXT4			Final Product Name	Text

D2B Code List

01 Admar Corp. Med Network
02 Aetna Life Insurance Co.
03 Aetna Services Inc. (Aetna Health Plans -- managed care)
04 Allstate Life Insurnace
05 AMERICAID, Inc.
06 American HMO
07 American Medical Security, Inc.
08 AmeriChoice Corp.
09 AmeriHealth, Inc.
10 Anthem Health Plans
11 Apex Health Care, Inc.
12 Beech Street Corp.
13 Blue Cross and Blue Shield System
14 CAPP Care
15 CIGNA Health Plans, Inc.
16 Community Health Plan, Inc.
17 Connecticut General Life Insurance Co.
18 Coventry Corp.
19 FHP, Inc.
20 Fortis Benefits
21 Foundation Health Corp.
22 Great Western Life and Accident
23 Group Health Cooperative of Puget Sound
24 Guardian Life Insurance Co.
25 Harvard/Pilgrim Health Care
26 Health Insurance Plan of Greater New York
27 Health Management Associates
28 Health Systems International, Inc.
29 HealthCare COMPARE Corp./The AFFORDABLE Medical Networks
30 HealthSource, Inc.
31 Henry Ford Health Care Corp.
32 Home Life Financial
33 Humana, Inc.
34 John Alden Life
35 John Deere Health Care, Inc.
36 John Hancock Life
37 Kaiser Foundation Health Plans, Inc.
38 Managed Health Network, Inc.
39 Maxicare Health Plans, Inc.
40 Medica
41 MedView Services Inc.
42 Mid-Atlantic Medical Services, Inc.
43 MultiPlan Inc.
44 Mutual of Omaha (managed care division)
45 Mutual of Omaha Insurance Co.
46 National Preferred Provider Network, Inc.
47 New York Life
48 NYLCare Health Plans, Inc.
49 Oxford Health Plans, Inc.
50 PacifiCare Health Systems, Inc.
51 PHS, Inc.
52 Physician Corp. of America
53 Preferred Health Network
54 Principal Financial
55 Principal Health Care, Inc.
56 Principal Mutual
57 Private Healthcare Systems
58 Provident Life and Accident Insurance Co.
59 Prudential Health Care Plans, Inc. (managed care division)
60 Prudential Insurance Co. of America
61 Sisters of Providence
62 United American HealthCare Corp.
63 United HealthCare Corp.
64 US Healthcare, Inc.
65 USA Health Network
66 WellCare Management Group, Inc.

APPENDIX I

FOLLOWBACK ROUND 2 POLICY LINKAGE FILE DOCUMENTATION

FOLLOWBACK ROUND 2 POLICY LINKAGE FILE

The Follow-back policy level identification file (LINKAGEF.SD2) provides the identification variables to link the product data with the person-level data collected in the CTS household survey. Based on the CTS household survey data and subsequent editing of these data, a total of 19,242 family units (FIUs) indicated they had one or more private health insurance plans covering the family members. This generated a total of 21,701 policy records with 19,173 FIUs reporting a first plan (containing 35, 904 person members), 2,443 FIUs reporting a second plan (3,712 person members) and 85 FIUs reporting a 3rd plan (114 person members).

The Followback survey attempted to interview all health insurance entities identified among the 21,701 private health insurance policies; however, linkage was successfully obtained between the FIU-reported plan for 10,368 of the policies. These 10,368 policies are refereed to as hard matches and have a match status indicator (FINSTAT) value of 1. A total of 1,895 policies could only be linked to the entity, but not to a specific product offered by the insuring entity. These are refereed to as soft matches and have a match status indicator value of 2. For these policies, two or more of the entity's products were linked to the policy, from which one was chosen to be the final linkage by statistical matching methods. The remaining 9,438 policies could not be linked to any specific product interview and have a match status indicator value of either 3 or 4 (Policies with a final match status of 4 were determined during the employer survey to not be comprehensive health plans. Eligibility is assumed, for the most part, to be undetermined for the non-matches, FINSTAT=3). Details on the linkage rates and methods used to finalize the linkage status can be found in the memorandum from Michael Sinclair issued on 6/19/2001.

The product data should be linked to this file using a combination of the three variables: FIN_ENT, PROD_ID and PSU which uniquely define a particular entity's product at given location. The CSID number provides a unique family link to the CTS person level records. The variable PLAN_NUM identifies the plan number each person is a member of and corresponds to the PRVINS1-PRVINS3 variables on the CTS person level files. The following table contains the variable name, description of the variable, and the values for the variable.

Note: This file contains a series of indicator variables that reflect certain operational outcomes in developing the final linkage for preparation of the survey weight adjustments that may not be needed on the final file.

File:	LINKAGEF.SD2 v.6.12 SAS File
Records:	21,701 policies
Identification Variables	CSID, PLAN_NUM, FIN_ENT, PROD_ID and PSU.

Notes on Soft-Match Policy Indicators.

The variables in italic text reflect the outcome of various operations conducted on the soft-matched policies to determine if statistical matching was viable. These variables apply only to a subset of the cases that had an original soft match status (reflected by MATCH=2). The variables in hierarchical order include:

1. HAVEDATA, which indicates whether we obtained data from the entity for the soft-matched policy (1=Yes, 0=No).
2. LOOKUP which indicated whether or not we attempted to obtain data on the insurers current product line profile (0=No, 1=Yes, 2=Not necessary, all lines represented).
3. FOUNDD, which indicates if we could find data from web sources on the insurer
4. COMPCHG, which indicates if the company appeared to have been part of a merger or acquisition and as such current data would not be equivalent (and likewise determined to be not matchable).
5. EQUALL, which indicated if the number of lines offered was the same for the companies that appeared to be stable.
6. CANMATCH denotes whether the profile of lines match among those which offered the same number of lines.

VARIABLE	DESCRIPTION	VALUES
B2STRNEW	Numeric string containing a list of the product lines for which we obtained data from the insurer for the site	Combination of values with: Missing = N.A. 1=HMO 2=POS 3=PPO 4=FFS e.g, a value of 1234 denotes the entity offered all four product lines in the site. Values are non-missing only for the soft-matched policies for which a we conducted a comparison of the profile of lines offered currently based on web-data to what was available from the insurer interviews
CANMATCH	Indicates if a soft match cases was determined to be statistically matchable (limited to policies examined).	Numeric, Missing= N.A. 0=Cannot match – profile of lines not the same 1-56= Statistical matching determined to be possible – value denote number of policies in entity/product/site combination.
CNAM	Company Name of Insurer as available after completion of the employer survey	Character String – 50 characters Available for only a subset of policies hard or soft matched at close of employer survey. Missing on 6,473 of non-matched policies.
COMPCHG	Indicates if a soft-matched case's insurer appeared to have changed ownership since the FB2 survey.	Missing = N.A. 0=NO 1=Yes
CSID	Family unit CTS identification number	1000010-55007960 First six digits identifies a unique household
EMPNAM	Employer Name (from CTS Household data)	Character String, 72 characters.
ENTGRPNM	Entity "Group" Name (prepared by Kathy Sonnenfeld for Companies that appeared to be associated with two or more similar names)	Character String – 40 characters Available only for subset of policies hard or soft matched at close of employer survey, based on CNAM. 9,469 polices with missing values.
EQUAL	Indicator that denotes in selection of soft-match linkages during statistical matching, the choices all had equal probabilities of a match (25 policies with 2 choices each).	Missing = N.A. 0=No 1=Yes

VARIABLE	DESCRIPTION	VALUES
EQUALL	Indicates if a soft-matched case's insurer had the same number of lines currently as they provided data for in FB2.	Missing = N.A. 0=No 1=Yes
FINSTAT	Final Matching Status Indicator	1=Hard Match Status (10,368) 2= Soft Match Status (1,895) 3= Non-Match – Eligibility Unknown (includes 1,295 "missed" cases) (8,095) 4= Not a Health Plan (as determined by employer survey and manual review) (1,343)
FIN_ENT	Final Entity Code	4 digits Missing on 9438 non-match and not a health plan policies
FOUNDD	For soft-matched cases, found insurer data on their lines from web-based search	Missing= N.A. 0=No 1=Yes
HAVEDATA	For soft-matched cases, indicates if the entity responded to the FB survey	Missing=N.A. – hard or non-match cases 0=No 1=Yes
HHID	Household Identification number (from CTS FIU-based file)	First six digits of CSID
LINKSTAT	CATI Employer Survey Result of Call Code	50 = not a health plan Other values not informative. In particular, values are semi-entity based, not product specific. A value of 1 indicates the entity supposedly provided data on some products during a CATI interview (but may indicate only that some contact was made).
LOOKUP	For soft-matched cases, indicates if web-based searches were attempted	Missing= N.A. 0=No 1=Yes
MATCH	Post Employer Survey Match Status MATCH FINSTAT Frequency Percent 1 1 10331 47.6 1 3 3026 13.9 2 1 37 0.2 2 2 1895 8.7 2 3 2734 12.6 3 3 2335 10.8 4 4 1343 6.2	1=Hard Match Status 2= Soft Match Status 3= Non-Match 4= Not a Health Plan

VARIABLE	DESCRIPTION	VALUES
MATCHFIN	Secondary Match Status on Soft Cases	2= Soft Match (1,932 cases of which 37 became hard-matches) 3= Non-match (2,734 cases)
MAXPROB	For statistically matched soft cases this indicates the maximum probablity of match determined among the choices Equal to PMATCH1	Between 0 and 1
MINPROB	As for MAXPROB, denotes minimum probability	"
NEWLINES	For soft matched cases researched indicates the number product lines they have currently	Missing= N.A. Numeric values 0-4 (0 not offering anything currently in site).
NLINES	Number of lines entity provided data for the site in FB2	Values 1-4
NUMBRS	Number of Persons in the Policy	No missing, Values 1-7
PENT	Assigned Entity Number from CATI	Numeric – only provides linkages among internal CATI files
PFX	Assigned Plan Number from CATI	Numeric – "
PLANNAM	Plan name as originally supplied by CTS household	Character string, 72 characters.
PLAN_NUM[55]	Reporting sequence number of the plan among all plans reported by the family unit.	NUMBER, 1-3
PLID	CATI ID	CATI Identification number
PMATCH1	Equal to MAXPROB	Numeric between 0 and 1.
POLONLY	Flag to identify non-matched cases that were never attempted (1,295)	Missing =Attempted 1=Not Attempted (1,295)
PROD_ID	Product identification code as obtained from the product interviews	NUMBER, 1-22, Use to Link to Product File.
PSU	Site number from CATI	Missing (1,295) , 1-60 Use to Link to Product File
RCDSP	Final Entity/Product Completion Status	Missing 9,438 cases. 1=Entity responded 2=Logically edited/imputed from booklet or web-based data 8="Newly" identified products (only used for soft-matches)
SITE	Site Number from CTS	0-60, 0= supplemental sample, Links to CTS data files.
SITEID	Family augmented site identifier	CTS Site membership 1-60
STAT	State membership from CATI	Character string, 2 characters. Missing, 1,295 cases.

1Membership in plans on person file are denoted by PRVINS variables, e.g. persons with PRVINS1=1 link to PLAN_NUM=1

For ineligible plans (FINSTAT=4) we have also provided the following variable values from the employer survey

VARIABLE	DESCRIPTION	VALUES
B60	If B11, B20 or B30 is No, not a health plan, this indicates the employers best guess as to the type of plan the employee mentioned having.	1=Medicare or retirement supplement 2=Military health plan 5=Specialty plan 6=Other including Medicaid

APPENDIX J

FOLLOWBACK ROUND 2 WEIGHT FILE DOCUMENTATION

FOLLOWBACK ROUND 2 WEIGHT FILE

The Follow-back weight file (FB2001.SD2) provides the Followback survey weights for each of the 58, 956 person records in the CTS round 2 household survey. This file can be linked to the CTS person file based on CSID and PID, and to the Followback Linkage file based on CSID and PLAN_NUM. This file provides two weights for each person, FBWTPER1, which should be used for site-level analysis and FBWTPER5 for national estimates. The PLAN_NUM variable identifies the plan selected as the "reference" plan for each person. If a person was a member of multiple plans, one plan is selected as the person's reference plan that would be used for analysis purposes per the round 1 specifications.

The Followback survey attempted to interview all health insurance entities identified among the 21,701 private health insurance policies; however, linkage was successfully obtained between the FIU-reported plan for 10,368 of the policies. These 10,368 policies are refereed to as hard matches and have a match status indicator (FINSTAT) value of 1 (which are associated with 18,943 persons). A total of 1,895 policies could only be linked to the entity, but not to a specific product offered by the insuring entity. These are refereed to as soft matches and have a match status indicator value of 2 (3,292 persons). For these policies, two or more of the entity's products were linked to the policy, from which one was chosen to be the final linkage by statistical matching methods. Total persons in a hard or soft linkage status is 22,235. The remaining 9,438 policies could not be linked to any specific product interview and have a match status indicator value of either 3 (8,095 policies, 13,163 persons) or 4 (Policies with a final match status of 4, 1,343 policies, 2,088 persons were determined during the employer survey to not be comprehensive health plans). Eligibility is assumed, for the most part, to be undetermined for the non-matches, FINSTAT=3). Persons that are not part of the Followback survey, including those 65 years of age or older, and those not having private insurance make up the balance of 17,931 persons of the 55,417 eligible CTS person records (3,539 persons are outside the augmented site sample). Details on the linkage rates and methods used to finalize the linkage status can be found in the memorandum from Michael Sinclair issued on 6/19/2001.

File:	FB2001.SD2 v.6.12 SAS File
Records:	58,956 person records.
Identification Variables	CSID, PID (to CTS), CSID, PLAN_NUM to Linkage file

VARIABLE	DESCRIPTION	VALUES
CSID	Family Unit CTS Identification Number	1000010-55007960 First six digits identifies a unique household
FBWTPER1	Followback Site Specific Weight	0 – 115000
FBWTPER5	Followback National Weight	0-181096.37
FINSTAT	Final Match Status	1=Hard Match Status 2= Soft Match Status 3= Non-Match 4= Not a Health Plan missing=Not in FB Sample
FOLLOWB	Followback Analytical Membership	1= Part of Followback Analysis Sample (FINSTAT=1 or 2) 0= Otherwise
PIASNAT	National CTS Person Weight	
PID	CTS Person Id	
PLAN_NUM	Reference Plan Number For PID	1-3
SITEID	Family Augmented Site Identifier	CTS Site membership 1-60
WINTTPP3	CTS Site Level Person Weight	

Note: The values of the Followback weights are set to zero for the non-matched cases (FINSTAT=3). For the cases outside the Followback sample (FINSTAT=missing), the value of the weight is equal to the original CTS weight (subject to some additional trimming adjustments).