

Contract No.: 0600-03-60129
MPR Reference No.: 8978

MATHEMATICA
Policy Research, Inc.

**National Beneficiary
Survey Round 3:
(Volume 1 of 3)
Editing, Coding,
Imputation, and
Weighting Procedures**

December 19, 2008

*Eric Grau
Debra Wright
Yuhong Zheng
Frank Potter*

Submitted to:

Social Security Administration
Office of Disability and Income
Support Programs
ITC Building
500 E. St., SW, 9th Floor
Washington, DC 20254

Project Officer:
Paul O'Leary

Submitted by:

Mathematica Policy Research, Inc.
600 Maryland Ave. SW, Suite 550
Washington, DC 20024-2512
Telephone: (202) 484-9220
Facsimile: (202) 863-1763

Project Director:
Debra Wright

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

ERRATA

(Updated December 20, 2016)

The SF-8 mental component summary (MCS) and physical component summary (PCS) scores provided in the original National Beneficiary Survey (NBS) data files were calculated incorrectly. The original values excluded an intercept constant needed to scale the scores to general population norms. The intercept constant values are -10.11675 for the MCS, and -9.36839 for the PCS.

Because the intercept constants were not applied, the scores provided in the original data files were too high relative to what they should be on the population-based scale. Thus, if comparing NBS respondents to the general population, NBS respondents would appear healthier than they should. However, within the NBS respondent sample, the scores still appropriately represented greater or lesser mental and physical health according to the design of the SF-8.

The MCS and PCS variables included in the current data files have been corrected and are now valid for comparisons to other populations.

This page has been left blank for double-sided copying.

CONTENTS

Chapter		Page
	ACRONYMS	xv
I	INTRODUCTION	1
	A. NBS SAMPLE DESIGN OVERVIEW	1
	B. NBS OBJECTIVES	8
	C. ROUND 3 SURVEY OVERVIEW	9
	D. NBS DATA DOCUMENTATION REPORTS	11
II	DATA EDITING AND CODING	15
	A. DATA EDITING	15
	B. CODING VERBATIM RESPONSES	16
	1. Coding Open-Ended, Other/Specify, and Field-Coded Responses	16
	2. Health Condition Coding	18
	3. Industry and Occupation	22
III	SAMPLING WEIGHTS	25
	A. COMPUTING AND ADJUSTING THE SAMPLING WEIGHTS: A SUMMARY	25
	1. Representative Beneficiary Sample	25
	2. Ticket Participant Cross-Sectional Sample	30
	3. Composite Cross-Sectional Weights	32
	4. Ticket Participant Longitudinal Sample	34
	5. Quality Assurance	35
	B. REPRESENTATIVE BENEFICIARY SAMPLE	36
	1. Initial Weights	36
	2. Nonresponse Adjustment	37
	3. Post-Stratification	50

CONTENTS (continued)

Chapter		Page
	C. TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE	51
	1. Initial Weights	54
	2. Dual Frame Estimation.....	54
	3. Nonresponse Adjustment	58
	4. Trimming.....	84
	5. Post-Stratification.....	86
	D. TICKET PARTICIPANT LONGITUDINAL SAMPLE	86
	1. Initial Weights	87
	2. Dual Frame Estimation.....	88
	3. Nonresponse Adjustment	89
	4. Trimming.....	122
	5. Post-Stratification.....	123
IV	IMPUTATIONS	125
	A. NBS IMPUTATIONS OF SPECIFIC VARIABLES	129
	1. Section L: Race and Ethnicity	130
	2. Section B: Disability Status Variables and Work Indicator	132
	3. Section C: Current Jobs Variables.....	133
	4. Section I: Health Status Variables.....	135
	5. Section K: Sources of Income Other than Employment	140
	6. Section L: Personal and Household Characteristics.....	142
V	ESTIMATING SAMPLING VARIANCE FOR NBS	145
VI	REFERENCES	148

APPENDICES

APPENDIX A: OTHER/SPECIFY AND OPEN-ENDED ITEMS WITH
ADDITIONAL CATEGORIES CREATED DURING
CODING

APPENDIX B: SOC MAJOR AND MINOR OCCUPATION
CLASSIFICATIONS

APPENDIX C: NAICS INDUSTRY CODES

APPENDIX D: PARAMETER ESTIMATES AND STANDARD ERRORS
FOR NONRESPONSE MODELS AND WEIGHTING CLASS
ADJUSTMENTS IN CROSS-SECTIONAL SAMPLES

APPENDIX E: PARAMETER ESTIMATES AND STANDARD ERRORS
FOR NONRESPONSE MODELS IN LONGITUDINAL
SAMPLES

APPENDIX F: SUDAAN AND SAS PARAMETERS USED TO OBTAIN
NATIONAL ESTIMATES

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

TABLES

Table		Page
I.1	NATIONAL BENEFICIARY AND TTW PARTICIPANT SAMPLE SIZES	3
I.2	ROUND 3 SAMPLE SIZES, TARGET COMPLETES, AND ACTUAL COMPLETES	10
II.1	SUPPLEMENTAL CODES FOR OTHER SPECIFY CODING.....	18
II.2	ICD-9 CATEGORY AND SUPPLEMENTAL CODES.....	20
II.3	SUPPLEMENTAL CODES FOR OCCUPATION AND INDUSTRY CODING	24
III.1	SURVEY POPULATION AS OF JUNE 30, 2005, INITIAL AUGMENTED SAMPLE SIZES AND INITIAL WEIGHTS BY SAMPLING STRATA IN THE NATIONAL BENEFICIARY SURVEY	38
III.2	WEIGHTED LOCATION AND RESPONSE RATES FOR REPRESENTATIVE BENEFICIARY SAMPLE, BY SELECTED CHARACTERISTICS	40
III.3	LOCATION LOGISTIC PROPENSITY MODEL: REPRESENTATIVE BENEFICIARY SAMPLE	46
III.4	COOPERATION LOGISTIC PROPENSITY MODEL: REPRESENTATIVE BENEFICIARY SAMPLE.....	47
III.5	SURVEY POPULATION AND INITIAL AUGMENTED AND FINAL CROSS-SECTIONAL SAMPLE SIZES, BY SAMPLING STRATA IN THE PARTICIPANT SURVEY.....	53
III.6	WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 2 STATES, MILESTONE-OUTCOME PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS	61
III.7	WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 2 STATES, OUTCOME-ONLY PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS	64

TABLES (continued)

Table	Page
III.8	WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE PHASE 2 STATES, TRADITIONAL PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS67
III.9	WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 3 STATES, MILESTONE-OUTCOME PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS70
III.10	WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 3 STATES, OUTCOME-ONLY PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS72
III.11	WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 3 STATES, TRADITIONAL PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS74
III.12	VARIABLES USED IN THE LOCATION ADJUSTMENTS: PHASE 2 TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE78
III.13	VARIABLES USED IN THE LOCATION ADJUSTMENTS: PHASE 3 TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE79
III.14	VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODELS: PHASE 2 TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE.....81
III.15	VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODELS: PHASE 3 TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE.....82
III.16	UNADJUSTED AND ADJUSTED R-SQUARED VALUES FOR LOGISTIC PROPENSITY MODELS IN TICKET PARTICIPANT CROSS-SECTIONAL SAMPLES84
III.17	PERCENTAGES OF CONCORDANT AND DISCORDANT PAIRS AND HOSMER-LEMESHOW P-VALUES FOR LOGISTIC PROPENSITY MODELS IN TICKET PARTICIPANT CROSS-SECTIONAL SAMPLES84

TABLES (continued)

Table	Page
III.18	DESIGN EFFECTS DUE TO UNEQUAL WEIGHTS BEFORE AND AFTER TRIMMING, WITHIN TRIMMING STRATA, FOR PHASES AND PAYMENT TYPES IN THE ROUND 3 TICKET PARTICIPANT CROSS-SECTIONAL SAMPLES86
III.19	SURVEY POPULATION AND FINAL LONGITUDINAL SAMPLE SIZES, BY SAMPLING STRATA IN THE PARTICIPANT SURVEY88
III.20	WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 STATES, ROUNDS 1 AND 2 RESPONDENTS, MILESTONE-OUTCOME PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS93
III.21	WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 STATES, ROUNDS 1 AND 2 RESPONDENTS, OUTCOME-ONLY PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS.....95
III.22	WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 STATES, ROUNDS 1 AND 2 RESPONDENTS, TRADITIONAL PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS.....97
III.23	WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 STATES, ROUNDS 1, 2, AND 3 RESPONDENTS, MILESTONE-OUTCOME PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS100
III.24	WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 STATES, ROUNDS 1, 2, AND 3 RESPONDENTS, OUTCOME-ONLY PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS102
III.25	WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 STATES, ROUNDS 1, 2, AND 3 RESPONDENTS, TRADITIONAL PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS.....104
III.26	WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 2 STATES, ROUNDS 2 AND 3 RESPONDENTS, MILESTONE-OUTCOME PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS106

TABLES (continued)

Table	Page
III.27	WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 2 STATES, ROUNDS 2 AND 3 RESPONDENTS, OUTCOME-ONLY PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS.....108
III.28	WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 2 STATES, ROUNDS 2 AND 3 RESPONDENTS, TRADITIONAL PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS.....110
III.29	VARIABLES USED IN THE LOCATION LOGISTIC PROPENSITY MODELS: TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 STATES, ROUNDS 1 AND 2 RESPONDENTS114
III.30	VARIABLES USED IN THE LOCATION LOGISTIC PROPENSITY MODELS: TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 STATES, ROUNDS 1, 2, AND 3 RESPONDENTS115
III.31	VARIABLES USED IN THE LOCATION LOGISTIC PROPENSITY MODELS: TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 2 STATES, ROUNDS 2 AND 3 RESPONDENTS116
III.32	VARIABLES USED IN THE COOPERATION LOGISTIC PROPENSITY MODELS: TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 STATES, ROUNDS 1 AND 2 RESPONDENTS117
III.33	VARIABLES USED IN THE COOPERATION LOGISTIC PROPENSITY MODELS: TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 STATES, ROUNDS 1, 2, AND 3 RESPONDENTS118
III.34	VARIABLES USED IN THE COOPERATION LOGISTIC PROPENSITY MODELS: TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 2 STATES, ROUNDS 2 AND 3 RESPONDENTS119
III.35	UNADJUSTED AND ADJUSTED R-SQUARED VALUES FOR LOGISTIC PROPENSITY MODELS IN THE TICKET PARTICIPANT LONGITUDINAL SAMPLE.....120
III.36	PERCENTAGES OF CONCORDANT AND DISCORDANT PAIRS AND HOSMER-LEMESHOW P-VALUES FOR LOGISTIC PROPENSITY MODELS IN THE TICKET PARTICIPANT LONGITUDINAL SAMPLE.....121

TABLES (continued)

Table	Page
III.37	DESIGN EFFECTS DUE TO UNEQUAL WEIGHTS BEFORE AND AFTER TRIMMING, WITHIN TRIMMING STRATA, FOR EACH TYPE OF LONGITUDINAL WEIGHT AND PAYMENT TYPE123
IV.1	RACE AND ETHNICITY IMPUTATIONS131
IV.2	DISABILITY STATUS IMPUTATIONS133
IV.3	CURRENT JOBS IMPUTATIONS134
IV.4	HEALTH STATUS IMPUTATIONS, QUESTIONNAIRE VARIABLES.....136
IV.5	HEALTH STATUS IMPUTATIONS, CONSTRUCTED VARIABLES.....139
IV.6	IMPUTATIONS ON SOURCES OF INCOME OTHER THAN EMPLOYMENT.....141
IV.7	IMPUTATIONS OF PERSONAL AND HOUSEHOLD CHARACTERISTICS144

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

FIGURES

Figure		Page
I.1	REPRESENTATIVE BENEFICIARY AND TICKET PARTICIPANT SAMPLES AND POPULATIONS AT ROUND 3	7

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

ACRONYMS

AIC :	Akaike's Information Criterion
CAPI:	Computer-assisted personal interviewing
CATI:	Computer-assisted telephone interviewing
CHAID:	Chi-Square Automatic Interaction Detector
EN:	Employment Networks
ICD-9:	International Classification of Diseases, 9th Revision
MPR:	Mathematica Policy Research, Inc.
MSA:	Metropolitan Statistical Area
NAICS:	North American Industry Classification System
NBS:	National Beneficiary Survey
PMSA:	Primary Metropolitan Statistical Area
PSU:	Primary Sampling Unit
SAS:	Statistical software, formerly Statistical Analysis System (SAS is a registered trademark of SAS Institute, Inc., Cary, NC)
SOC:	Standard Occupational Classification
SPSS:	Statistical Package for the Social Sciences (SPSS is a registered trademark of SPSS, Inc., Chicago, IL)
SSA:	Social Security Administration
SSDI:	Social Security Disability Insurance (Title II of the Social Security Act)
SSI:	Supplemental Security Income (Title XVI of the Social Security Act)
SSU:	Secondary Sampling Unit
STATA:	Statistical software (STATA is a registered trademark of StataCorp LP, College Station, TX.)
TTW:	Ticket to Work

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

I. INTRODUCTION

As part of an evaluation of the Ticket to Work and Self-Sufficiency program (TTW), Mathematica Policy Research, Inc. (MPR) conducted the third round of the National Beneficiary Survey (NBS) in 2006. The survey, sponsored by the Social Security Administration's (SSA) Office of Disability and Income Security Programs, collected data from a national sample of SSA disability beneficiaries (hereafter referred to as the Representative Beneficiary Sample) and a sample of TTW participants (hereafter referred to as the Ticket Participant Sample). The Ticket Participant Sample contains cross-sectional and longitudinal components, both of which are discussed in this report. MPR collected data using computer-assisted telephone interviewing (CATI) with computer-assisted personal interviewing (CAPI) follow-ups of CATI nonrespondents and those who preferred or needed an in-person interview to accommodate their disabilities.

A voluntary employment program for people with disabilities, TTW was authorized by the Ticket to Work and Work Incentives Improvement Act of 1999. The legislation was designed to create market-driven services to help disability beneficiaries become economically self-sufficient. Under the program itself, SSA provides beneficiaries with a "Ticket," or coupon, that they may use to obtain employment-support services, including vocational rehabilitation, from an approved provider of their choice (called Employment Networks or ENs).¹

A. NBS SAMPLE DESIGN OVERVIEW

SSA implemented the TTW program in three phases spanning three years, with each phase corresponding to about one-third of the states. The initial NBS survey design called for four

¹ For more information on the Ticket to Work Program, see Thornton et al. 2004.

national cross-sectional surveys (called rounds) of Ticket-eligible SSA disability beneficiaries—one each in 2003, 2004, 2005, and 2006—and cross-sectional surveys of Ticket participants in each of three groups of states (Phase 1, Phase 2, and Phase 3 states)—defined by the year in which the program was rolled out (Bethel and Stapleton 2002).² In addition, the design called for the first TTW participant cohort in each group of Ticket roll-out states to be followed longitudinally until 2006. This design was subsequently revised to accommodate Phase 1 data collection starting in 2004 rather than 2003. In addition, the final round was postponed to address the experiences of TTW participants under the new TTW regulations; implemented in July 2008. The fourth round will include a cross-sectional Representative Beneficiary survey as well as a survey of new Ticket Participants and is planned for 2009. Details of the sample design for round 4 are not yet determined; in a change from the original design, Ticket participants from previous rounds will not be re-interviewed at round 4. Table I.1 gives the original planned sample sizes for all rounds of data collection. Actual sample sizes and number of completes cases are provided in Chapter III.

² The Ticket to Work program, implemented in 2002, was phased in nationwide over three years. In 2002, the first year of the program, SSA distributed Tickets in the following 13 states, known as the “Phase 1” states: Arizona, Colorado, Delaware, Florida, Illinois, Iowa, Massachusetts, New York, Oklahoma, Oregon, South Carolina, Vermont, and Wisconsin. The Phase 2 roll-out ran from November 2002 through September 2003, during which time SSA distributed Tickets in the following 20 “Phase 2” states and the District of Columbia: Alaska, Arkansas, Connecticut, Georgia, Indiana, Kansas, Kentucky, Louisiana, Michigan, Mississippi, Missouri, Montana, Nevada, New Hampshire, New Jersey, New Mexico, North Dakota, South Dakota, Tennessee, Virginia, and the District of Columbia. The Phase 3 roll-out ran from November 2003 through September 2004, during which time SSA distributed Tickets in 17 “Phase 3” states: Alabama, California, Hawaii, Idaho, Maine, Maryland, Minnesota, Nebraska, North Carolina, Ohio, Pennsylvania, Rhode Island, Texas, Utah, Washington, West Virginia, and Wyoming, as well as in American Samoa, Guam, the Northern Mariana Islands, Puerto Rico, and the Virgin Islands.

TABLE I.1
NATIONAL BENEFICIARY AND TTW PARTICIPANT SAMPLE SIZES

Sample ^a		Year 1	Year 2	Year 3	Year 4	All Years ^c
National Beneficiary Samples		7,200	4,800	2,400	1,500	15,900
Longitudinal TTW Participant Samples	Phase 1 Cohorts (1) ^b	1,000	922	850	784	3,556
	(2)		1,000			1,000
	Phase 2 Cohorts (1)		1,000	922	850	2,772
	(2)			1,000		1,000
	Phase 3 Cohorts (1)			1,000	922	1,922
	(2)				1,000	1,000
	Total		1,000	2,922	3,772	3,556
Total Sample Size		8,200	7,722	6,172	5,056	27,150

Source: NBS Sample Design Report (Bethel and Stapleton 2002)

^a Sample sizes refer to number of completed interviews

^b(1)=TTW participant longitudinal sample and (2)=TTW participant cross-sectional supplement

^c The All Years column is a tabulation of the number of interviews, not the number of sample members. Longitudinal cases may be included up to three times in these counts, depending upon the number of completed interviews for the sample member in question.

The NBS used a multi-stage sampling design (which was used for all survey rounds) with a supplemental single-stage sample for some Ticket participant populations. For the multi-stage design, data from SSA on the counts of eligible beneficiaries in each county were used to form the primary sampling units (PSUs) consisting of one or more counties. The sample of all SSA beneficiaries (the Representative Beneficiary Sample) was selected from among beneficiaries residing in these PSUs (or, in two counties with a large number of beneficiaries, secondary sampling units) using age-defined sampling strata. Separate samples of Ticket participants within each phase in the original sample design were selected from all Ticket participants in these PSUs. The Ticket Participant Sample was divided into three strata (within each phase) according to the type of payment system under which SSA paid a service provider: the traditional

vocational rehabilitation payment system, the milestone-outcome payment system, and the outcome-only payment system.³ The supplemental single stage sample for some Ticket participant populations was drawn from all Ticket participants, not just those in the PSUs, with stratification based upon payment type and whether the participant was in a PSU or not. The round 3 User's Guide (Wright, et al. 2008) contains more information on the round 3 sampling design.

In round 1 (2004), two surveys were fielded: the first national survey of all beneficiaries (the Representative Beneficiary Sample) and the first cross-sectional survey of Ticket participants in the Phase 1 states (the Ticket Participant Sample). Three cross-sectional surveys were fielded in round 2 (2005):

1. The second national survey of all beneficiaries (The Representative Beneficiary Sample),
2. The second cross-sectional survey of Ticket participants who resided in a Phase 1 state at the time of Ticket assignment (The Phase 1 Cross-Sectional Ticket Participant Sample), and
3. The first cross-sectional survey of Ticket participants who resided in a Phase 2 state at the time of Ticket assignment (The Phase 2 Cross-Sectional Ticket Participant Sample).

At round 2, we also attempted to re-interview Phase 1 Ticket Participants who were selected into the sample at round 1, whether or not they had been interviewed in round 1 (the Phase 1 Longitudinal Sample). The original sample design called for re-interviewing only those longitudinal cases that had completed the previous round. However, based on MPR's recommendation, interviews were attempted with all longitudinal cases.

³ ENs may choose to be paid under the traditional payment system or under one of two other payment systems developed specifically for the Ticket program: (a) an outcome-only payment system or (b) a milestone-outcome payment system. Under both new payment systems, SSA will make up to 60 monthly payments to the EN for each assigned beneficiary who is not receiving SSDI or SSI payments because of work or earnings. Under the milestone-outcome payment system, SSA pays smaller monthly payments in the event that the beneficiary leaves cash benefits but will also pay the EN for up to four milestones achieved by a beneficiary.

Three cross-sectional surveys were also fielded at round 3:

1. The third national survey of all beneficiaries,
2. The second cross-sectional survey of Ticket participants who resided in a Phase 2 state at the time of Ticket assignment (The Phase 2 Cross-Sectional Ticket Participant Sample), and
3. The first cross-sectional survey of Ticket participants who resided in a Phase 3 state at the time of Ticket assignment (The Phase 3 Cross-Sectional Ticket Participant Sample).

Additionally, we attempted to re-interview Phase 1 Ticket Participants who were selected into the sample at round 1, whether or not they had been interviewed in rounds 1 or 2 (the Phase 1 Longitudinal Sample) and Phase 2 cases who were selected into the sample at round 2, whether or not they had been interviewed at round 2 (the Phase 2 Longitudinal Sample).

In each first follow-up year (round 2 for Phase 1 participants, and round 3 for Phase 2 participants), a supplemental sample of those who had entered the Ticket program since the first year of rollout for each phase, or otherwise had not been sampled before, was selected to produce an expanded second-year cross-sectional sample survey. For Phase 1 and Phase 2 participants, this resulted in cross-sectional samples for two consecutive years. The cross-sectional surveys consisted of the supplemental cases, plus the longitudinal cases who were still Ticket participants at the time of sampling. In round 3, there were cross-sectional samples for Phase 2 and Phase 3 participants; however, no supplemental cross-sectional sample was selected for the Phase 1 Ticket participants at round 3. (Hence, no Phase 1 cross-sectional weights were calculated in round 3.)

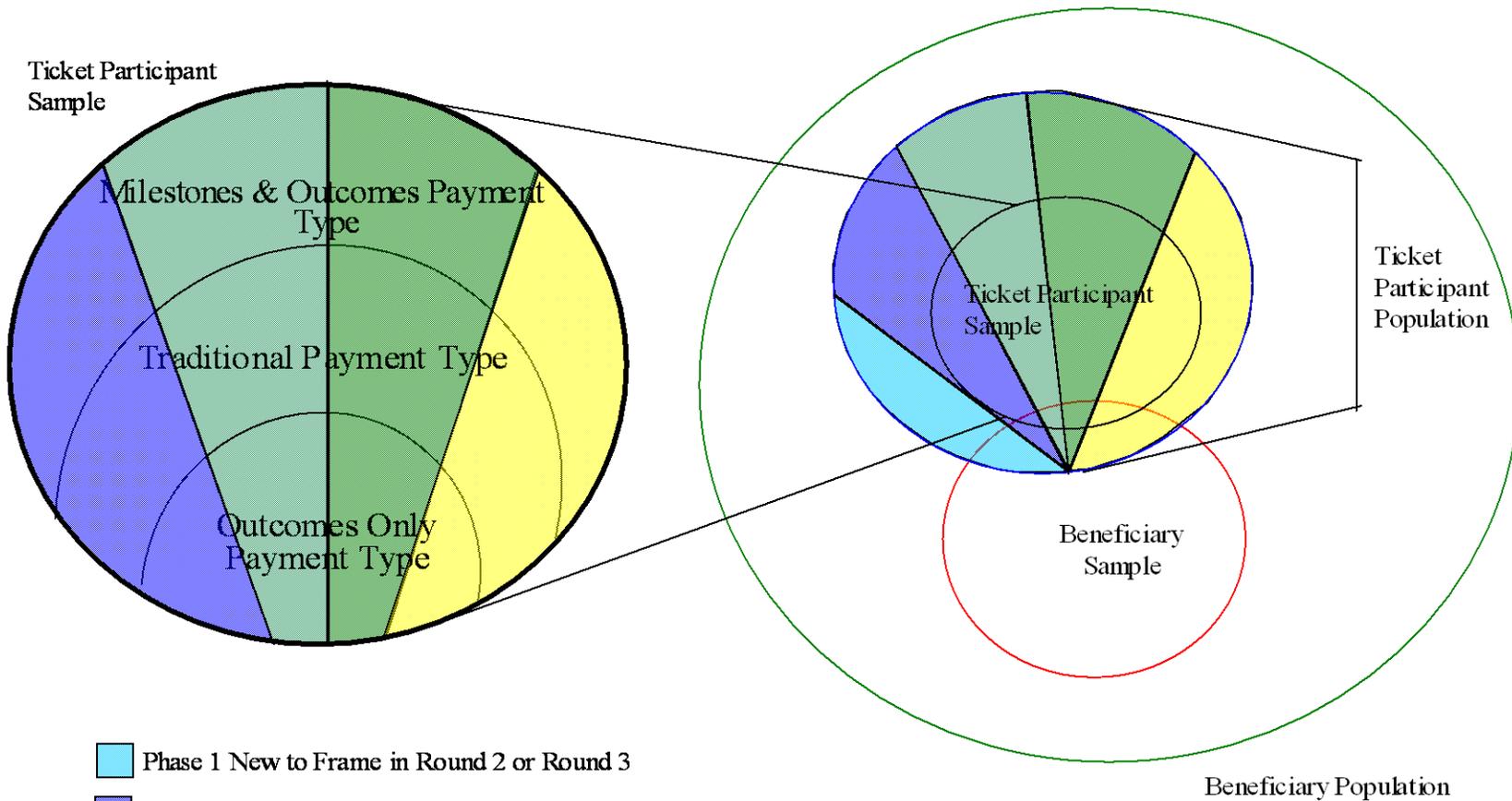
Figure I.1 is a graphic representation (not to scale) of how the various samples and populations relate to one another.⁴ The population of Ticket participants, represented by the

⁴The composition of the populations and samples represented by these circles changes from round to round. For example, a round 1 snapshot would show only Phase 1 cases in the Ticket Participant Sample; a round 2 snapshot would not include a Phase 3 subsample. Although not apparent from the diagram, the Ticket participant population represented less than 1 percent of the beneficiary population.

circle with a yellow, green, purple, and blue interior, is a subset of the population of all Supplemental Security Income (SSI) and Social Security Disability Insurance (SSDI) beneficiaries, as represented by the large green circle. The population of Ticket participants is broken into five parts: two parts for Phase 1 states, two parts for Phase 2 states, and one for Phase 3 states, four of which represented in the Ticket Participant Cross-Sectional Sample. The fifth part represents the Phase 1 cases that were new to the program after round 1 (new Phase 1 cases) and are not represented in the Ticket Participant Cross-Sectional Sample, as referenced earlier. The Ticket Participant Cross-Sectional Sample is blown up to show the various payment types, with yellow, green, and purple interior, but no blue, since blue represents Phase 1 states new to the program after round 1. The Representative Beneficiary Sample (represented by the red circle) could include some individuals who are also Ticket participants (there were 55 such cases in round 3, where the red and blue circles intersect). Moreover, it is possible for a Ticket participant to have been selected for both the Representative Beneficiary Sample and the Ticket Participant Cross-Sectional Sample (there were 17 such cases in round 3, where the red and black circles intersect). The samples taken from these populations represent a snapshot of the populations at round 3, so that the Ticket Participant Cross-Sectional Sample, as shown in Figure I.1, does not include Phase 1 or Phase 2 longitudinal sample cases who were not in the program in round 3, since they were no longer part of the Ticket participant population. Finally, the figure depicts the Ticket participant subpopulation as a much larger proportion of the beneficiary population than is actually the case. In fact, in round 3, the Ticket participant subpopulation was less than 0.6 percent of the entire beneficiary population.

FIGURE I.1

REPRESENTATIVE BENEFICIARY AND TICKET PARTICIPANT SAMPLES AND POPULATIONS AT ROUND 3



- Phase 1 New to Frame in Round 2 or Round 3
- Phase 1 Longitudinal Cases From Round 1 Still in Ticket Program in Round 3
- Phase 2 New to Frame in Round 3
- Phase 2 Longitudinal Cases From Round 2 Still in Ticket Program in Round 3
- Phase 3

Note: Samples and populations are not to scale

B. NBS OBJECTIVES

The NBS is one of several components of an evaluation of the impact of TTW relative to the current system, the SSA Vocational Rehabilitation Reimbursement Program, which has been in place since 1981. The evaluation includes a process analysis, as well as an impact and a participation analysis. Along with the NBS, the data sources include SSA administrative records and interviews with program stakeholders. The NBS collects data needed for the TTW evaluation that are not available from SSA administrative data or other sources.

The NBS has five objectives:

1. To provide critical data on the work-related activities of SSI and SSDI beneficiaries, particularly as these activities relate to TTW implementation
2. To collect data on the characteristics and program experiences of beneficiaries who use their Ticket
3. To gather information about beneficiaries who do not use their Ticket, and the reasons for this choice
4. To collect data that will allow us to evaluate the employment outcomes of Ticket users and other SSI and SSDI beneficiaries
5. To collect data on service use, barriers to work, and beneficiary perceptions about TTW and other SSA programs designed to help SSA beneficiaries with disabilities find and keep jobs

Round 3 NBS data will be combined with SSA administrative data to provide critical information on access to jobs, and on employment outcomes for beneficiaries, including those who participate in the TTW program and those who do not. Though some sections of the NBS target beneficiary activity directly related to TTW, most of the survey captures more general information on SSA beneficiaries, including their disability, interest in work, use of services, and employment. As a result, SSA and external researchers interested in disability and employment issues can use the survey data for other policymaking and program-planning efforts.

C. ROUND 3 SURVEY OVERVIEW

As in previous rounds, round 3 sample members in both the Representative Beneficiary Sample and the Ticket Participant Sample received the same survey instrument. The NBS collects data on a wide range of topics including employment, limiting conditions, experience with SSA programs, employment services, health and functional status, health insurance, income, and socio-demographic information. The survey items were developed and initially pre-tested as part of a separate contract held by Westat. Revisions were made by MPR to prepare the instrument for CATI/CAPI programming, and additional minor wording changes were made after pre-testing. More information about the questionnaire design can be found in the Round 3 User's Guide (Wright, et al. 2008). The survey instrument is available from SSA or MPR upon request.

Round 3 CATI data collection for both samples began in February 2006. Beginning in May 2006, MPR conducted in-person CAPI interviews with beneficiaries who did not respond to the CATI interview, as well as those who could not be located (and whose names and other information were sent to field interviewers for additional locating), or who requested an in-person interview to facilitate their participation in the survey. The survey instrument was identical in each mode. When possible, the interview was attempted with the sample person. If the sample person was unable to complete either a telephone or an in-person interview because of his or her disability, a proxy respondent was sought. Proxy interviews were attempted only when the sample member was unable to complete the survey himself or herself due to his/her disability. To promote response among Hispanic populations, the questionnaire was available in Spanish. For languages other than English or Spanish, interpreters conducted interviews. A number of additional accommodations were made available for those with hearing and/or speech

impairments including teletypewriter (TTY), Telecommunications Relay Service (TRS), amplifiers, and instant messaging technology.

As shown in Table I.2, the NBS round 3 sample comprised 3,382 cases selected for the Representative Beneficiary Sample and 5,697 cases for the Ticket Participant Sample (for a total of 9,079 cases).

TABLE I.2
ROUND 3 SAMPLE SIZES, TARGET COMPLETES, AND ACTUAL COMPLETES

Sampling Strata	Sample Size	Target Completed interviews	Actual Completed interviews
Representative Beneficiary Sample	3,382	2,400	2,508
Ticket Participant Sample	5,697	3,000	4,097
Phase 1 Longitudinal, complete in rounds 1, 2, and 3	1,466	850	953
Phase 2 Supplemental Participant Sample ^a	1,508	1,000	1,122
Phase 2 Longitudinal, complete in rounds 2 and 3	1,350	922	831
Phase 3 Ticket Participant Sample	1,373	1,000	1,053
Total Sample Size	9,079	6,172	6,605

Source: NBS, round 3.

The round 3 CATI and CAPI data collection was completed in September 2006. Interviews were completed with 2,508 individuals in the Representative Beneficiary Sample, and 4,097 people in the Ticket Participant Sample, for a total of 6,605 cases completed (including sample members from both the cross-sectional and longitudinal samples).⁵ An additional 215 in the Representative Beneficiary Sample and 113 in the Ticket Participant Sample were determined to

⁵ Because the clustered and unclustered samples of the Ticket Participant Sample were independent, it was not uncommon for individuals to be chosen for both. It was also possible for a sample member to be chosen for both the Representative Beneficiary Sample and the Ticket Participant Sample. Interviews for these duplicate cases were conducted only once, but recorded twice (once for each sample). The counts given above include these duplicates as separate cases.

be ineligible for the survey.⁶ Across both samples, 5,104 cases were completed by telephone, and 1,501 were completed by CAPI. Proxy interviews were completed for 1,286 sample members. The weighted response rate for the Representative Beneficiary Sample was 81.1 percent. The weighted response rate for the Ticket Participant Cross-Sectional Sample (among Phase 2 and Phase 3 cases still in the Ticket participant population) was 84.4 percent.

D. NBS DATA DOCUMENTATION REPORTS

The following reports make up the complete documentation describing the NBS, the round 3 data collection, and the data files:

- ***Editing, Coding, Imputation, and Weighting Report*** (current report). This report summarizes the editing, coding, imputation, and weighting procedures as well as the development of standard errors for the round 3 NBS. It includes an overview of the variable naming, coding, and construction conventions used in the data files and accompanying codebooks; describes how the sampling weights were computed to the final post-stratified analysis weights for both the Representative Beneficiary Sample and the Ticket Participant Sample (and describes the procedures for combining these samples); describes the procedures used to impute missing responses; and discusses procedures that should be used to estimate sampling variances for the NBS.
- ***Cleaning and Identification of Data Problems Report*** (Wright and Barrett 2008). This report describes the data processing procedures performed for round 3 of the NBS. It outlines the data coding and cleaning procedures and describes the data problems identified, their origins, and the corrections implemented to create the final data file. The report describes the data issues by sections of the interview and concludes with a summary of types of problems encountered and general recommendations.
- ***User's Guide for Restricted and Public Use Data Files*** (Wright et al. 2008). This report is designed to provide users with information about the restricted and public use data files including construction of the files; weight specification and variance estimation; masking procedures employed in the creation of the Public Use File; and a detailed overview of the questionnaire design, sampling, and NBS data collection. The report also contains some information covered in the two reports mentioned

⁶ Ineligible sample members include those who were deceased or incarcerated; those no longer living in the continental United States; and those whose benefit status was pending. For the Ticket Participant Sample, ineligibles also included sample members who left the program after sampling was completed (although those who were in the round 1 or round 2 sample and subsequently left the program were eligible for the longitudinal sample).

above including procedures for data editing, coding of open-ended responses, and variable construction; and a description of the imputation and weighting procedures and development of standard errors for the survey.

In addition the following supplemental materials are available from MPR or SSA upon request:

- ***NBS Questionnaire***. This document contains all items on the round 3 survey and includes documentation of skip patterns, question universe specifications, text fills, interviewer directives, and consistency and range checks.
- ***NBS Restricted Access and Public Use File Codebooks***. The codebooks provide extensive documentation for each variable on the file including variable name, label, position, variable type and format, question universe, question text, number of cases eligible to receive each item, constructed variable specifications, and user notes. Frequency distributions and means are also included as appropriate.

In the discussion that follows, we document the editing, coding, imputation, and weighting procedures, as well as the development of standard errors, for the round 3 NBS. Chapter II is an overview of the variable naming, coding, and construction conventions used in the data files and accompanying codebooks. Chapter III describes how the initial sampling weights were computed to the final post-stratified analysis weight for both the Representative Beneficiary Sample and the Ticket Participant Sample (for both the cross-sectional and longitudinal samples); also described are the procedures for combining the Representative Beneficiary Sample and Ticket Participant Cross-Sectional Sample. Chapter IV describes the procedures used to impute missing responses for selected questions. Chapter V discusses the procedures that should be used to estimate sampling variances for the NBS. Appendix A lists the open-ended items that were assigned additional categories, as discussed in Chapter II. Industry and occupation codes, also discussed in Chapter II, are listed in Appendices B and C. Detailed parameter estimates and standard errors for the weight adjustment models discussed in Chapter III are presented in Appendix D for the cross-sectional models, and Appendix E for the longitudinal models.

Appendix F covers the SUDAAN and SAS parameters for the national estimates from the TTW round 3 sample.

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

II. DATA EDITING AND CODING

Prior to imputation, the NBS data were edited and coded to create an NBS data file. This chapter documents the variable naming, coding, and construction conventions used in the data files and accompanying codebooks.

A. DATA EDITING

At the start of data cleaning, a systematic review of the frequency counts of the individual questionnaire items was conducted. We reviewed frequency counts by each questionnaire path to identify possible errors in skip patterns. We also reviewed interviewer notes and comments in order to flag and correct individual cases. In consultation with SSA and research analysts, we took the general approach of editing only cases for which there appeared to be an obvious data entry or respondent error. As a result, while we devoted a great deal of time to a meticulous review of individual responses, some suspect values remain on the file. (See Barrett and Wright (2008) for more detail on the editing and cleaning procedures.)

For all items with fixed field numeric responses (such as number of weeks, number of jobs and dollar amounts), we reviewed the upper and lower values assigned by interviewers. While data entry ranges were set in the CATI instrument to prevent improbable responses from being entered, these ranges intentionally were set to accommodate a wide spectrum of values to account for the diversity expected in this population, and so that the interview could continue in most situations. For these reasons, extremely high and low values were set to missing (.D=don't know) if there appeared to be an error in data entry.

The NBS instrument included several consistency edit checks to flag potential problems during the course of the interview. To minimize respondent burden, however, all consistency edit checks were suppressible. While the interviewer was instructed to probe such responses, the

interview could continue beyond a particular item if the respondent could not resolve the problem. In the post-interview stage, we manually reviewed remaining consistency problems to determine whether the responses were plausible. After investigating these cases, we corrected them, or set them to missing when an obvious error was encountered.

During data processing, we created several constructed variables to combine data across items. For these items, both the survey team and the analysis team reviewed the specifications, several reviewers checked the SAS programming code, and we reviewed all data values for the constructed variables based on the composite variable responses and frequencies.

For open-ended items that were assigned numeric codes, we examined frequencies to ensure that valid values were assigned. For health condition coding, we also examined codes to verify that the same codes were not assigned to both main and secondary conditions. Cases coded incorrectly were recoded according to the original verbatim response.

B. CODING VERBATIM RESPONSES

The NBS questionnaire includes a number of questions designed to elicit open-ended responses. To make it easier to use the data connected with these responses in an analysis, we grouped the responses and when possible, assigned them numeric codes. The methodology used to code each variable depended upon the content of the variable.

1. Coding Open-Ended, Other/Specify, and Field-Coded Responses

Three kinds of questions (described below) on the NBS did not have designated response categories; rather, the response to these questions was recorded verbatim:

- *Open-ended questions* have no response options specified (such as E43—Why are you no longer receiving services from your employment network?). For these items, interviewers recorded the verbatim response. Using common responses, we developed categories and reviewed them with analysts. Coders then attempted to code the verbatim response into an established category. If the response did not fit into one of these categories, it was coded as “other.”

- **“Other/specify”** is a response option for questions that have a finite number of possible answers that may not necessarily capture *all* possible responses. A good example is: “Did you do anything else to look for work in the last four weeks that I didn’t mention?” For questions of this type, respondents were asked to specify an answer to the question “anything else?” or “anyone else?”
- **Field-coded responses** are answers coded by interviewers into a pre-defined response category without reading the categories aloud to the respondent. If none of the response options seemed to apply, interviewers selected an “other specify” category and typed in the response.

As part of data processing at round 1, we examined a portion of all verbatim responses in an attempt to uncover dominant themes for each question. Based on this initial review, we developed a list of categories and decision rules for coding verbatim responses to open-ended items. In addition, supplemental response categories were added to some field-coded or other-specify items to facilitate coding if there were enough such responses and they could not be back-coded into pre-existing categories. (A list of all open-ended items assigned additional categories during the coding process appears in Appendix A.) Thus we categorized verbatim responses for quantitative analyses by coding responses that clustered together (for open-ended and “other/specify” responses) or by back-coding responses into existing response options if appropriate (for “field-coded” and “other/specify” items). Categories developed during round 1 and round 2 coding were applied at round 3. Additional categories were added at round 3 for a small number of items if there were a significant number of common responses that did not fit into previously developed categories. If during coding, it became apparent that changes to the coding scheme were necessary (for example adding additional categories or clarifying coding decisions), new decision rules were discussed and documented. Verbatim responses were sorted alphabetically by item for coders and could be filtered by coding status so that new decision rules could be easily applied to cases that had been previously coded. When it was impossible to code a response, when responses were invalid, or when they could not be coded into a given category,

we assigned a two-digit supplemental code to the response (see Table II.1). The verbatim responses themselves are excluded from the data files. (See Barrett and Wright (2008) for full details regarding the back-coding procedures.)

TABLE II.1
SUPPLEMENTAL CODES FOR OTHER SPECIFY CODING

Code	Label	Description
94	Invalid Response	Indicates this response should not be counted as an “other” response, but should be deleted
95	Refused	Used only if verbatim response indicates respondent refused to answer the question
96	Duplicate Response	Indicates the verbatim response already has been selected in a ‘code all that apply’ item
98	Don’t Know	Used only if the verbatim indicates that the respondent does not know the answer
99	Not Codeable	Indicates that a code cannot be assigned based on the verbatim response

Source: NBS, round 3.

2. Health Condition Coding

Responses to questions on health conditions required a specific type of open-ended coding. In Section B of the questionnaire, each respondent was asked to cite the main and secondary physical or mental conditions that limit the kind or amount of work or daily activities he or she can do. Main conditions could be reported as one of four items: B2 (main reason limited), B6 (main reason eligible for benefits), B12 (main reason was eligible for benefits if not currently eligible), and B15 (main reason limited when first started getting disability benefits). The main purpose of items B6, B12, and B15 was to collect information on a health condition from people who reported no limiting conditions in B2. For example, if respondents said that they had no limiting conditions, they were asked if they were currently receiving benefits from Social Security. If they answered “yes,” they were asked for the main reason that made them eligible for

benefits (B6). If respondents said that they were not currently receiving benefits, they were asked whether they had received disability benefits in the last five years. If they answered “yes,” they were asked for the condition that made them eligible for Social Security benefits (B12), or for the reason that first made them eligible if they no longer had that condition (B15). If respondents said that they had not received disability benefits in the last five years, they were screened out of the survey and coded as ineligible. Each response to B2, B6, B12, and B15 was assigned a value for the three constructs. Although respondents were asked to cite one “main” condition in B2, B6, B12, or B15, many listed more than one. These additional responses were maintained under the main condition variable and coded in the order in which they were recorded. Longitudinal cases that completed rounds 1 or 2 skipped items B6, B12, and B15 at round 3.

For each item on a main condition, respondents were also asked to list any other, or secondary, conditions. For example, respondents reporting a main condition at B2 were asked at B4 to list other conditions that limited the kind or amount of work or daily activities they could do. Respondents reporting the main reason they were eligible for disability benefits (at B6) were asked at B8 to list other conditions that made them eligible. Finally, respondents who reported that they were not currently receiving benefits and who reported a main condition at B12 (the condition that made them eligible to receive disability benefits in the last five years) were asked at B14 for other reasons that made them eligible for benefits. Those who reported that their current main condition was not the condition that made them eligible for benefits, and who were asked for the main reason they were first limited, were also asked if there were any other conditions that had limited them when they first started receiving benefits (B17).

As in previous rounds, the respondents’ verbatim responses were coded according to the International Classification of Diseases, 9th revision, Clinical Modification (ICD-9-CM) five-digit coding scheme. The ICD-9 is a classification of morbidity and mortality information that

was developed in 1950 to index hospital records by disease for data storage and retrieval. The ICD-9 was available in hard copy for each of the coders. Coders, many of whom had previous medical coding experience, attended an eight-hour training session before coding, and were instructed to code to the highest possible level of specificity. Responses not specific enough for a five-digit code were coded to four (subcategory) or three digits (category codes). Responses not specific enough for even three- or four-digit ICD-9 codes were coded either as a physical problem (not specified) or to broader categories representing disease groups. (See Table II.2 for a list of the broad categorical and supplementary codes.) For cases in which multiple, distinct conditions were provided by the respondent, all conditions were coded (for instance, three distinct conditions would be recorded and coded as B2_1, B2_2, and B2_3).

TABLE II.2
ICD-9 CATEGORY AND SUPPLEMENTAL CODES

Code	Label	Description of ICD-9 codes	Corresponding ICD-9 codes
00	Other	Other and unspecified infectious and parasitic disease; alcohol dependence syndrome and drug dependence; learning disorders and developmental speech or language disorders; complications of medical care, not elsewhere classified	136.0-136.9, 303.00-304.90, 315.00-315.39, 999.0-999.9
01	Infectious and parasitic diseases	Borne by a bacterium or parasite and viruses that can be passed from one human to another or from an animal/insect to a human including tuberculosis, HIV, other viral diseases, and venereal diseases (excluding other and unspecified infectious and parasitic diseases)	001.0-135, 137.0-139.8
02	Neoplasms	New abnormal growth of tissue, i.e., tumors and cancer, including malignant neoplasms, carcinoma in situ, and neoplasm of uncertain behavior	140.0-239.9
03	Endocrine/nutritional disorders	Thyroid disorders, diabetes, abnormal growth disorders, nutritional disorders, and other metabolic and immunity disorders	240.0-279.9
04	Blood/blood-forming	Diseases of blood cells and spleen	280.0-289.9

TABLE II.2 (continued)

Code	Label	Description of ICD-9 codes	Corresponding ICD-9 codes
05	Mental disorders	Psychoses, neurotic and personality disorders, and other non-psychotic mental disorders including mental retardation (excluding alcohol and drug dependence and learning, developmental, speech, or language disorders)	290.0–302.9, 305.00-314.9, 315.4-319
06	Diseases of nervous system	Disorders of brain, spinal cord, central nervous system, peripheral nervous system, and senses including paralytic syndromes, and disorders of eye and ear	320.0-389.9
07	Diseases of circulatory system	Heart disease, disorders of circulation, and diseases of arteries, veins, and capillaries	390-459.9
08	Diseases of respiratory system	Disorders of the nasal, sinus, upper respiratory tract, and lungs including chronic obstructive pulmonary disease	460-519.9
09	Diseases of digestive system	Diseases of the oral cavity, stomach, esophagus, and duodenum	520.0-579.9
10	Diseases of genitourinary system	Diseases of the kidneys, urinary system, genital organs, and breasts	580.0-629.9
11	Complications of pregnancy, child birth, and the puerperium	Complications related to pregnancy or delivery, and complications of the puerperium	630-677
12	Diseases of skin/subcutaneous tissue	Infections of the skin, inflammatory conditions, and other skin diseases	680.0-709.9
13	Diseases of musculoskeletal system	Muscle, bone, and joint problems including arthropathies, dorsopathies, rheumatism, osteopathies, and acquired musculoskeletal deformities	710.0-739.9
14	Congenital anomalies	Problems arising from abnormal fetal development, including birth defects and genetic abnormalities	740.0-759.9
15	Conditions in the perinatal period	Conditions that have origin in birth period even if disorder emerges later	760.0-779.9
16	Symptoms, signs, and ill-defined conditions	Ill-defined conditions and symptoms; used when no more specific diagnosis can be made	780.01-799.9
17	Injury and poisoning	Problems that result from accidents and injuries including fractures, brain injury, and burns (excluding complications of medical care not elsewhere classified)	800.00–998.9
18	Physical problem, NEC	The condition is physical, but no more specific code can be assigned.	No ICD-9 codes

TABLE II.2 (continued)

Code	Label	Description of ICD-9 codes	Corresponding ICD-9 codes
95	Refused	Verbatim indicates respondent refused to answer the question.	No ICD-9 codes
96	Duplicate condition reported	The condition has already been coded for the respondent.	No ICD-9 codes
97	No condition reported	The verbatim does not contain or symptom to condition to code.	No ICD-9 codes
98	Don't know	The respondent reports that he/she does not know the condition.	No ICD-9 codes
99	Uncodeable	A code cannot be assigned based on the verbatim response.	No ICD-9 codes

Source: NBS, round 3.

We employed several means to ensure that responses were coded according to the proper protocols. First, we did an initial quality assurance check, per coder, for the first several cases that were coded. In total, approximately 20 percent of all coded responses were reviewed by a supervisor, including cases flagged by coders for review because they were unable or did not know how to code them. Approximately 8 percent of all cases were recoded. In the course of this work, additional decision rules were developed to clarify and document coding protocol. These decisions were discussed with coders and posted to ensure that responses were coded consistently and accurately throughout the coding process. As for other open-ended items, when new decision rules were added, previously coded responses were reviewed and re-coded if necessary. After the ICD-9 coding was complete, we processed the health condition variables into a series of constructed variables that grouped health conditions into broad disease groups.

3. Industry and Occupation

Information about both a sample member's current employment and employment in 2005 was recorded in Section C and Section D of the questionnaire. For each job, respondents were asked to record their occupation (C2 and D4) and the type of business or industry (C3 and D5)

where they were employed. Verbatim responses were coded using the Bureau of Labor Statistics' 2000 Standard Occupational Classification (SOC).⁷ The SOC is a system for classifying all occupations in the economy, including private, public, and military occupations, in which work is performed for pay or profit. Occupations are classified on the basis of work performed, skills, education, training, and credentials. The sample member's occupation was assigned one occupation code. The first two digits of the SOC codes classify the occupation to a major group and the third digit to a minor group. For the NBS we assigned three-digit SOC codes to describe the major group the occupation belonged to and the minor groups within that classification (using the 23 major groups and 96 minor). Appendix B lists the three-digit minor groups classified within major groups.

As for previous rounds, verbatim responses to the industry items were coded according to the 2002 North American Industry Classification System (NAICS).⁸ The NAICS is an industry classification system that groups establishments into categories on the basis of activities in which those establishments are primarily engaged. The NAICS uses a hierarchical coding system through which all economic activity is classified into 20 industry sectors. For the NBS, we coded NAICS industries to three digits: the first two numbers specify the industry sector, and the third number specifies the sub-sector. (Appendix C lists the broad industry sectors.) Both the SOC and the NAICS coding schemes are used in most federal surveys, thus providing uniformity and comparability across data sources.

MPR developed supplemental codes for responses to questions about occupation and industry that could not be coded to a three-digit SOC or NAICS code (See Table II.3). As we did

⁷ See *Standard Occupational Classification Manual, 2000*, or <http://www.bls.gov/soc/>, for more information.

⁸ See North American Industry Classification System, 2002 or <http://www.naics.com/info.htm> for more information.

in the health condition coding, we performed an initial quality assurance check, per coder, for the first several cases coded. In total, approximately 15 percent of all coded responses were reviewed by a supervisor, including cases flagged by coders for review because they were unable or did not know how to code them. Approximately 4 percent of all cases were recoded.

TABLE II.3
SUPPLEMENTAL CODES FOR OCCUPATION AND INDUSTRY CODING

Code	Label	Description
94	Sheltered Workshop	Code used if occupation is in sheltered workshop and the occupation cannot be coded from verbatim.
95	Refused	The respondent refuses to give his/her occupation or type of business.
97	No occupation or industry reported	No valid occupation or industry is reported in the verbatim.
98	Don't know	The respondent reports that he/she does not know the occupation or industry.
99	Uncodeable	A code cannot be assigned based on the verbatim response.

Source: NBS, round 3.

III. SAMPLING WEIGHTS

The final analysis weights for the Representative Beneficiary Sample and the Ticket Participant Sample⁹ were determined via a four-step process: (1) calculate the initial weights, (2) adjust the weights for two phases of nonresponse (location and completion), (3) trim the weights to reduce the variance, and (4) post-stratification. This chapter describes these computations for both the Representative Beneficiary Sample and the Ticket Participant Sample. Section A summarizes the procedures used to compute and adjust the sampling weights, and the procedure for creating composite weights. (Composite weights were used in all rounds to combine the Representative Beneficiary Sample and the Ticket Participant Cross-Sectional Sample, and to combine two samples in the Ticket Participant Sample.) Procedures for computing the weights for the Representative Beneficiary Sample are described in detail in Section B. Sections C and D cover the same information for the Ticket Participant Cross-Sectional Sample and the Ticket Participant Longitudinal Sample, respectively.

A. COMPUTING AND ADJUSTING THE SAMPLING WEIGHTS: A SUMMARY

1. Representative Beneficiary Sample

The sampling weights for any survey are computed from the inverse selection probability that incorporates the stages of sampling in the survey. The Representative Beneficiary Sample was selected in two stages: primary sampling units (PSUs) were selected as part of the round 1 sampling activities, and the individuals within the PSUs were selected from a current database of

⁹ The “Ticket Participant Sample” in this chapter refers to both the Ticket Participant Cross-Sectional and Ticket Participant Longitudinal Samples.

beneficiaries.¹⁰ We used four age-based strata in each PSU. In particular, beneficiaries were stratified into the following age groups: 18- to 29-year-olds, 30- to 39-year-olds, 40- to 49-year-olds, and 50- to 64-year-olds. Because we used a composite size measure to select the PSUs, we could achieve equal probability samples in the age strata and nearly equal workload in each PSU for the Representative Beneficiary Sample.¹¹

For the initial beneficiary sample, we selected more individuals than we expected to need, to account for differential response and eligibility rates in both the PSUs and the sampling strata. This “augmented” sample was randomly partitioned into subsamples (called waves), where only some of the waves were used to form the actual final sample. We released an initial set of waves and then monitored data collection to identify which PSUs and strata required additional sample members. After the sample members in the initial waves were released for the final sample, we were able to limit the number of additional sample members (in subsequent released waves) only to those PSUs and strata requiring them, and so were able to achieve sample sizes that were close to our targets. Controlling the release of the sample also allowed us to control the balance between data collection costs and response rates. The initial sampling weights were computed on the basis of the inverse of the selection probability for the augmented sample. Naturally, only a subset of the augmented sample was actually released, so these initial weights were adjusted for the actual sample size. The release-adjusted weights were post-stratified to population totals obtained from SSA.¹²

¹⁰ An intermediate stage of sampling of secondary sampling units (SSUs) was used in two PSUs, but for the sake of simplicity, these generally are treated as equivalent to PSUs in this description. All PSUs and SSUs were selected during the round 1 sampling activities.

¹¹ The composite size measure was computed from the sum of the products of the sampling fraction for a stratum and the estimated count of beneficiaries in that stratum and PSU (Folsom et al. 1987).

¹² These totals were obtained from a frame file provided by SSA that contains basic demographics for all SSI and SSDI beneficiaries.

The initial sampling weights then needed to be adjusted for nonresponse. A commonly used method to compute weight adjustments is to form classes of sample members with similar characteristics, and use the inverse of the class response rate as the adjustment factor in that class. The adjusted weight is the product of the sampling weight and the adjustment factor. The “weighting classes” are formed to ensure that there are sufficient counts in each class to make the adjustment more stable (that is, to have a smaller variance). The natural extension to the weighting class procedure is to use logistic regression with the weighting class definitions used as covariates, provided each level of the model covariates has a sufficient number of sample members to ensure a stable adjustment. The logistic regression approach also has the ability to include both continuous and categorical variables, and standard statistical tests are available to evaluate the selection of variables for the model. For the location and the cooperation weight adjustments, we used logistic models to estimate the propensity for a sample member to be located and to cooperate. The inverse of the propensity score was used as the adjustment factor. The adjusted weight for each sample case is the product of the initial sampling weight and the adjustment factor.

We calculated this adjustment factor in two stages: (1) estimating a propensity score for locating a sample member, and (2) estimating a propensity score for response among located sample members. In our experience with this survey, factors associated with the inability to locate a person tend to be different from factors associated with cooperation. The unlocated person cannot deliberately avoid or otherwise refuse to cooperate. For instance, that person may have chosen not to list his or her number, or may frequently move from one address to another, but he or she has not shown a specific unwillingness to cooperate with the survey itself. Located nonrespondents may deliberately avoid the interviewer, or may be expressing displeasure or hostility toward surveys in general, or SSA in particular.

To develop the logistic propensity models for round 3, we used information from the SSA data files and geographic information (such as urban/rural or region) as covariates. Using a liberal level of statistical significance (0.3) in forward and backward stepwise regression models, we made an initial attempt to reduce the pool of covariates and interactions. We used a higher significance level because the purpose of the model was to improve the estimation of the propensity score, not to identify statistically significant factors related to response. In addition, the information sometimes reflected proxy variables for some underlying variable that was both unknown and unmeasured. Any covariate or interaction that clearly was unrelated to locating the respondent, or to response propensity, was excluded from the pool.

The next step was to carefully evaluate a series of models by comparing the following measures of predictive ability and goodness of fit: the R-squared statistic,¹³ Akaike's Information Criterion (AIC),¹⁴ percentage of concordant and discordant pairs,¹⁵ and the Hosmer-Lemeshow goodness-of-fit test.¹⁶ Model-fitting also involved reviewing the statistical significance of the coefficients of the covariates in the model and avoiding any unusually large

¹³ The Generalized Coefficient of Determination (Cox and Snell 1989) is a measure of the adequacy of the model, where higher numbers indicate a greater difference between the likelihood of the model in question and the null model likelihood. The "Max rescaled R-Square" scales this value to have a maximum of 1.

¹⁴ Akaike's Information Criterion is defined as $AIC = -2\text{LogL} + 2(k+s)$, where LogL is the loglikelihood of the binomial distribution using the parameters from the given model, k is the total number of response levels minus one, and s is the number of explanatory effects (Akaike 1974). AIC is a relative number, and has no meaning on its own. For a given model, smaller values of AIC are better than larger values.

¹⁵ A pair of observations is concordant if a responding subject has a higher predicted value than the nonresponding subject, discordant if not, and tied if both members of the pair are either respondents, nonrespondents, or have the same predicted values. It is desirable to have as many concordant and as few discordant pairs as is possible (Agresti 1990).

¹⁶ The Hosmer-Lemeshow Goodness-of-Fit Test is a test for goodness of fit of logistic regression models. Unlike the Pearson and deviance goodness-of-fit tests, it can be used to test goodness of fit even when some of the covariates are continuous (Hosmer and Lemeshow 1989).

adjustment factors. In addition, we also avoided data warnings in SUDAAN.¹⁷ We then used the specific covariate values for each located person (cooperating person) to estimate a propensity to be located (to cooperate), from which we calculated the adjusted weights. The location-adjusted weight is the product of the released adjusted weight and the inverse of the location propensity score; the nonresponse-adjusted weight is the product of the location-adjusted weight and the inverse of the cooperation propensity score.

Once the adjustments were made, we trimmed the survey weights (if necessary) to avoid unusually large weights, which would make the survey estimates less precise. We used the design effect attributed to the variation in the sampling weights as a statistical measure to determine both the necessity and the amount of trimming. The design effect attributed to weighting is a measure of the potential loss in precision caused by the variation in the sampling weights relative to a sample of the same size with equal weights. We also wanted to minimize the extent of trimming to avoid the potential for bias in the survey estimates. For the Representative Beneficiary Sample, the design effect due to unequal weighting was checked within the age-related sampling strata, and trimming to reduce the design effect was employed in only one age stratum (the 30- to 39-year-old age group). The design effect was reduced from 1.09 to 1.06, which was the maximum design effect among all the age strata in the Representative Beneficiary Sample.

The final step is a series of post-stratification adjustments through which the weights sum to known totals obtained from SSA on various dimensions (specifically, gender, age grouping, and

¹⁷ SUDAAN data warnings usually included one or more of the following: (1) an indication of a response cell with zero count; (2) one or more parameters approaching infinity (which may not be readily observable with the parameter estimates themselves); and (3) degrees of freedom for overall contrast less than the maximum number of estimable parameters. We tried to avoid all of these warnings, although avoiding the first two was of the highest priority. These warnings almost always were caused by a response cell with a count that was too small, which required dropping covariates or collapsing categories in covariates.

for beneficiaries only, recipient status¹⁸). After post-stratification, we checked the survey weights again to determine whether more trimming was necessary. In round 3, trimming was not necessary after post-stratification in the Representative Beneficiary Sample.

2. Ticket Participant Cross-Sectional Sample

The initial sampling cross-sectional weights for the Ticket Participant Cross-Sectional Sample were computed on the basis of the inverse of the selection probability for the participant. As with the Representative Beneficiary Sample, we used the PSUs as the primary source of the sample members and, when possible, selected an initial larger (augmented) sample. For participants in Phase 2 states using either the milestone-outcome or the outcome-only payment system, and for participants in Phase 3 states using the outcome-only payment system, the PSUs in the initial sampling design did not have enough participants to support analysis tasks—even with all participants in the PSUs from these two payment types selected for the sample. As a result, it was necessary to supplement the sample from the PSUs with a second independent sample of Ticket participants from two geographic strata defined by the PSUs (participants residing in a PSU, or not residing in any of the PSUs). The sample members within the initial sample design are referred to as the clustered sample; members of the second independent sample are referred to as the unclustered sample. Sample members in the unclustered sample were randomly selected from the entire population of milestone-outcome and outcome-only participants in Phase 2 states, and from the entire population of outcome-only participants in Phase 3 states, in the two aforementioned geographic strata.¹⁹ The combination of data from the

¹⁸ Disability payments were made in the form of Supplemental Security Income (SSI), Social Security Disability Insurance (SSDI), or both.

¹⁹ Because of the small populations for the payment types where the paired sample design was required, Ticket participants who resided in the selected PSUs for these payment types often were selected for both the clustered and the in-PSU strata of the unclustered samples. Hence, these duplicate cases had to be accounted for in the weighting process, as discussed later.

clustered and unclustered samples to calculate estimates is referred to as a “paired sample design,” and is discussed later in this document.

As with the Representative Beneficiary Sample, we computed the weights for the augmented sample and then adjusted them for the number of sample members that were in the final sample.²⁰ We adjusted for nonresponse separately for located sample members, and then for response among these sample members. Because Ticket participants were generally easier to locate due to their participation in the Ticket program, the number of Ticket participants who could not be located was very small. Hence, for all Ticket participants except those in Phase 2 states using the traditional payment system, we calculated the location adjustment using the weighting class method. However, the location adjustment for Phase 2 Ticket participants using the traditional payment system, and the response adjustments for located Ticket participants of both phases and all three payment types, were calculated using logistic propensity models. The modeling procedures were similar to those used with the Representative Beneficiary Sample.

The size of the sample for the three payment types was similar, but the size of the population for each was very different. (More than 80 percent of the population of Ticket participants used the traditional payment system. Specific percentages for each phase and payment type are given in Section C.) Hence, the sampling weights differed substantially in magnitude from one payment system to the next. As a result, we conducted the weight adjustments separately for each payment type. For the subsamples associated with each phase and payment type within the Ticket Participant Cross-Sectional Sample, we trimmed the weights to ensure that the design effect due to unequal weighting was not substantially greater than 3.0 (less than 3.0, if possible). (More details about the trimming of the participants’ weights, and the design effects due to

²⁰ For the clustered sample of participants using the Outcomes-Only payment system, all participants in the PSUs were selected and released for data collection.

unequal weighting before and after trimming, are given in Section C.) The final adjustment for the participants' weights was a post-stratification adjustment to the counts of participants within subgroups defined by age and gender in the sampling frame. After post-stratification, we checked the survey again to determine whether more trimming was necessary. In round 3, although trimming was required before post-stratification in the Ticket Participant Cross-Sectional Sample, no trimming was required after post-stratification.

3. Composite Cross-Sectional Weights

Although the Ticket participant population constitutes a small subset of the beneficiary population, some analyses require a sample with a substantial number of individuals both within and outside the Ticket participant population. This can be accomplished by combining the Ticket Participant Cross-Sectional Sample and Representative Beneficiary Sample and using composite weights to account for the fact that the samples have been combined. When conducting analyses representing the beneficiary population, these weights can be used to make estimates about participants within the beneficiary population. (Analyses limited to the participants subpopulation use weights from the Ticket Participant Cross-Sectional Sample only.)

In round 1, we used a sophisticated procedure to create these weights, such that the variance of survey estimates was minimized. This procedure allowed for weights to be applied to observations that were duplicated across the two samples.²¹ However, because the Ticket participants were such a small fraction of the beneficiary sample frame, we used a simpler alternative method in rounds 2 and 3.

²¹ A complex procedure also was used to combine the clustered and unclustered samples of the Ticket Participant Sample in all rounds. This procedure is described in Section C of this chapter.

In round 3, composite weights were developed only for the Phase 2 and Phase 3 Ticket participants in the Representative Beneficiary Sample. As indicated earlier, the round 3 Ticket Participant Sample included only Phase 1 Ticket participants who were selected in round 1. This meant that no cross-sectional sample of all Phase 1 participants at round 3 was available. Of the 55 Ticket participants in the Representative Beneficiary Sample, only 33 were Phase 2 or Phase 3 cases. Of these 33 cases, 31 were respondents (that is, they had completed interviews, or were ineligible after sample selection in round 3). These respondents included 14 from Phase 2 states and 17 from Phase 3 states. We replaced the original Representative Beneficiary Sample weights with a value of zero among these 31 cases. To ensure that the Ticket participant population would be represented, we replaced these members of the Representative Beneficiary Sample with the 3,161 members of the Ticket Participant Cross-Sectional Sample (2,091 from Phase 2 states and 1,070 from Phase 3 states) with completed interviews (or ineligible dispositions after sample selection).²² From the Ticket participants sampling frame, there were 33,500 participants from Phase 2 states and 31,023 participants from Phase 3 states. Because the sum of the weights of the 31 Ticket participants from Phase 2 or Phase 3 states in the Representative Beneficiary Sample did not equal the sampling frame totals for each phase, we computed post-stratification adjustments for the remaining beneficiary weights. The sum of the weights for the 31 participants in the Representative Beneficiary Sample is an unbiased estimate of the number of participants in the sampling frame. Because of the relatively small sample size, this estimate did not equal the known total in the sampling frame, as was expected. The post-stratification adjustment realigned the population totals.

²² This does not include sample members who were selected for the rounds 1 or 2 Ticket Participant Samples, were no longer Ticket participants in round 3, but were sampled in round 3 anyway for longitudinal purposes.

4. Ticket Participant Longitudinal Sample

For longitudinal analyses, the inferential population is defined by the population at the time the Ticket program was rolled out for the group of states in question, and not by the Ticket participant population as it was constituted in round 3. For Phase 1 longitudinal cases, the inferential population is the set of Phase 1 Ticket participants at round 1. For Phase 2 longitudinal cases, the inferential population is the set of Phase 2 Ticket participants at round 2. We conducted a nonresponse bias analysis to evaluate the differences between Phase 1 cases who responded in various combinations of rounds (round 1, round 2, and/or round 3). We concluded that no systematic differences in selected key variables were apparent between groups of Phase 1 cases differentiated by their response patterns. Based on this result, and on consultations with SSA and TTW Project Part A contract staff,²³ we determined that three sets of Ticket participant longitudinal weights were sufficient for anticipated longitudinal analyses. These longitudinal weights included: a set for Phase 1 participants who responded²⁴ in rounds 1, 2, and 3; a set for Phase 1 participants who responded in rounds 1 and 2, and a set for Phase 2 participants who responded in rounds 2 and 3. The initial sampling longitudinal weights for the Ticket Participant Longitudinal Sample were computed on the basis of the inverse of the selection probability for the participants in the first round for Phase 1 longitudinal weights and the second round for Phase 2 longitudinal weights. For the calculation of longitudinal weights, Ticket participants from Phase 2 states using the outcome-only and milestone-and-outcome payment systems needed a paired sampling design. Similarly, a paired sampling design was also

²³ The TTW contract was split into two parts, Part A and Part B. Sampling, weighting, and imputation procedures were conducted under Part B of the contract and analyses under Part A.

²⁴ “Respondents” include individuals who had completed interviews or were ineligible after sample selection (i.e., when surveyed).

required for the calculation of longitudinal weights for Ticket participants from Phase 1 states using the outcome-only payment system.

As with the cross-sectional weights, we calculated adjustments for nonresponse in two stages: (1) a location adjustment for locating a sample member, and (2) a cooperation adjustment for response among located sample members. However, unlike the cross-sectional weight nonresponse adjustments, we used logistic models for the location and cooperation adjustments for all payment types and phases. The inverse of the propensity score was used as the adjustment factor. The adjusted weight for each sample case is the product of the initial sampling weight and the adjustment factor.

We trimmed the weights so that the design effect due to unequal weighting was not substantially greater than 3.0 (less than 3.0, if possible), and post-stratified them to add up to the round 1 frame totals for Phase 1 cases, and round 2 frame totals for Phase 2 cases. (Details about the trimming used with longitudinal weights, and the design effects before and after trimming, are given in Section C.)

5. Quality Assurance

To ensure that the methods used to compute the weights at each step were sound, a senior statistician conducted a final quality assurance check of the weights from the Representative Beneficiary and Ticket Participant cross-sectional and longitudinal samples, as well as the composite weights. For the sake of objectivity, we chose a statistician who was not directly involved in the project.

B. REPRESENTATIVE BENEFICIARY SAMPLE

1. Initial Weights

The initial weights were computed using the inverse of the probability of selection. For the Representative Beneficiary Sample, samples were selected independently in each of four age strata in each geographic unit or PSU.²⁵ The number of sample members selected in each stratum and PSU for the augmented sample was determined by allocating three times the target sample size across the 84 geographic units (PSUs and secondary sampling units) independently for each stratum.²⁶ This ensured that plenty of reserve sample units were available in case response or eligibility rates were lower than expected. The augmented sample size for the three younger age strata (18 to 29 years, 30 to 39 years, and 40 to 49 years) was 4,000 sample members (roughly three times the target sample size of 1,333); for beneficiaries 50 to 64 years, the augmented sample size was 3,000 (again, three times the target sample size of 1,000). By using the composite size measure described previously, the initial weights for the full augmented sample of 15,000 sample members were calculated by taking the inverse of the global sampling rate (F_i) for each stratum. The global sampling rates and initial weights are given in Table III.1.

As described previously, the full sample was randomly partitioned into subsamples called waves that mirrored the characteristics of the full sample. The waves were formed in each of the four sampling strata in the 84 geographic units (a total of 336 combinations of PSU and sampling strata). At the start of data collection, a preliminary sample was assigned to the data collection

²⁵ The sample of PSUs contained 80 selections. Because of the size of its beneficiary population, the PSU representing Los Angeles County (LA) received two selections. Within the LA PSU, secondary sampling units (SSUs) were formed, and four SSUs were selected. In the PSU representing Cook County, IL (Chicago), SSUs also were formed to decrease travel costs, and two SSUs were selected. These six SSUs and the other 78 PSUs (84 units) were treated as PSUs for the beneficiary sample.

²⁶ An augmented sample that was three times as large as needed was selected to allow for an adequate supplemental sample in all PSUs and sampling strata within the PSUs, and to account for expected variation in the response and eligibility rates across PSUs and sampling strata.

effort, and additional waves were assigned as needed, based on experience with eligibility and response rates. Within the 336 combinations of PSU and sampling strata, the initial weights were adjusted to account for the number of waves assigned to data collection. The final sample size for the Representative Beneficiary Sample was 3,382 beneficiaries, as shown under “Released Sample” in Table III.1.

2. Nonresponse Adjustment

As in virtually all surveys, the sampling weights must be adjusted to compensate for sample members that cannot be located or who, once located, refuse to respond. First, weighted logistic regression models were fitted where the binary response was whether the sample member could be located. Using variables obtained from SSA databases, a pool of covariates from which to choose a final location model was selected through stepwise regression. This pool included both main effects and interactions. From this pool of covariates, candidate models were compared using various measures of goodness of fit and predictive ability, while avoiding large adjustments. This process was repeated for interview respondents among the located sample members, where another weighted logistic regression model was fitted. The two levels in the binary response for this model were “respondent” or “nonrespondent.” For the Representative Beneficiary Sample, a sample member was classified as a respondent if the sample member or the person responding for the sample member completed the interview (that is, an eligible respondent), or if the sample member was determined to be ineligible after sample selection (an ineligible respondent). Ineligible sample members included persons who were never SSA beneficiaries, were in the military service at the time of the survey, were incarcerated, had moved outside of the United States, or were deceased at the time of the survey.

TABLE III.1

SURVEY POPULATION AS OF JUNE 30, 2005, INITIAL AUGMENTED SAMPLE SIZES AND INITIAL WEIGHTS BY SAMPLING STRATA IN THE NATIONAL BENEFICIARY SURVEY

Sampling Strata (ages as of June 30, 2005)	Survey Population ^a	Augmented Sample Size	Global Sampling Rate (Fj)	Initial Sample Weights	Released Sample
Beneficiaries between 18 and 29 years old	1,064,845	4,000	0.003756	266.2	943
Beneficiaries between 30 and 39 years old	1,271,121	4,000	0.003147	317.8	941
Beneficiaries between 40 and 49 years old	2,514,758	4,000	0.001591	628.7	935
Beneficiaries between 50 and 64 years old	5,534,098	3,000	0.000542	1844.7	563
Total	10,384,822	15,000			3,382

Source: Sample allocation and counts computed by MPR.

^aThe survey population represents all SSI and SSDI beneficiaries.

Using the procedures outlined above, the main factors or attributes affecting our ability to locate and interview the sample member included the personal characteristics of the sample member (race, ethnicity, gender, and age), the type of beneficiary (recipient of SSI, SSDI, or both), identity of the payee with respect to the beneficiary, whether the beneficiary and the applicant for benefits lived in the same location, primary disability classification, type of disability claim (a person with a disability, a survivor, or other), living situation of beneficiary, and geographic characteristics.

a. Coding of Survey Dispositions

The status of each sample member was maintained in the MPR Survey Management System during the survey, and a final status code was assigned after the completion of all locating and interviewing efforts on a given sample member, or at the end of data collection. For the nonresponse adjustments, we classified the final status codes into four categories:

1. Eligible respondents
2. Ineligible respondents (sample members who were ineligible after sample selection, including deceased, sample members in the military or incarcerated, sample members living outside of the United States, and other ineligibles)
3. Located nonrespondents (including active or passive refusals and language barrier situations)
4. Unlocated sample members (sample members who could not be located using either central office tracing procedures or in-field searches)

This classification of the final status code allowed us to measure the overall response rate, the completion rate among located sample members, and the location rate among all sample members.²⁷

b. Response Rates

The 81.1 percent response rate for the Representative Beneficiary Sample quoted in the introduction to this document is the **weighted overall completion rate**, given in the first line of Table III.2. This response rate is the weighted count of sample members for whom a completed interview was obtained or who were determined to be ineligible, divided by the weighted sample count of all sample members.²⁸ It can be determined by taking the product of the weighted

²⁷ Disposition codes 420 (institutionalized) and 430 (unavailable during field period) were classified as nonrespondent codes in rounds 2 and 3, even though they were considered ineligible codes in round 1. This affected 8 cases in the round 2 beneficiary sample and 6 cases in the round 3 beneficiary sample. As a result, the nonresponse adjusted weight for these cases was 0 in rounds 2 and 3, even though a similar response in round 1 would have resulted in a positive weight. Because of the small numbers, the effect on response rates was very small.

²⁸ This response rate is the weighted count of sample members for whom a completed interview was obtained or who were determined to be ineligible divided by the weighted sample count of all sample members (# of completed interviews + # partially completed + # of ineligibles) / # of cases in the sample). It can be determined by taking the product of the weighted location rate and the weighted cooperation rate, also known as the weighted completion rate among located sample members. This response rate is basically equivalent to the AAPOR standard response rate calculation: $RR_{AAPOR} = \# \text{ of completed interviews} / (\# \text{ of cases in the sample} - \text{estimated } \# \text{ of ineligible cases})$. Ineligible cases are included in the numerator for two reasons: (1) the cases classified as ineligible are part of the original sampling frame (and hence the study population). We obtained complete information to fully classify these cases (i.e., their responses to the eligibility questions in the questionnaire are complete) and therefore classify them as respondents; (2) incorporating the ineligibles in the numerator and denominator of the response rate is essentially equivalent to the definition of a response rate with these cases excluded if the persons with an additional estimation of the number of eligible cases among those with eligibility unknown. By including the ineligible cases in the numerator and denominator, we avoid using this estimation stage and the response rate computation is more clearly explicated.

location rate and the weighted cooperation rate, also known as the weighted completion rate, among located sample members.

TABLE III.2
WEIGHTED LOCATION AND RESPONSE RATES FOR REPRESENTATIVE BENEFICIARY SAMPLE,
BY SELECTED CHARACTERISTICS

	Sample	Located Sample	Response among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate
All	3,382	3,186	95.2	2,723	81.1
SSI Only, SSDI Only, or Both SSI and SSDI					
SSI only	1,507	1,393	92.8	1,195	78.2
SSDI only	1,181	1,137	97.2	947	81.7
Both SSI and SSDI	694	656	94.1	581	85.4
SSI or SSDI					
SSI only, or in both SSI & SSDI programs	2,201	2,049	93.2	1,776	80.5
SSDI only, or in both SSI & SSDI programs	1,875	1,793	96.4	1,528	82.6
Constructed Disability Status					
Deaf	39	37	97.8	28	73.7
Mental	1,822	1,706	94.3	1,442	78.3
Physical	1,360	1,296	96.5	1,124	83.6
Unknown	161	147	87.1	129	78.6
Beneficiary's Age (Four Categories)					
18-29 years	943	874	92.7	775	82.2
30-39 years	941	878	93.3	728	77.4
40-49 years	935	894	95.6	760	81.3
50-64 years	563	540	95.9	460	81.7
Sex					
Male	1,762	1,650	94.2	1,405	80.4
Female	1,620	1,536	96.2	1,318	81.9
Hispanicity					
Non-Hispanic	3,215	3,039	95.5	2,588	81.0
Hispanic	167	147	88.1	135	83.9
Race					
White	1,989	1,895	96.3	1,630	82.4
Black	781	725	93.8	621	81.2
Unknown	547	505	92.8	431	78.8
Asian American, Pacific Islander	44	42	97.8	24	39.2
North American Indian or Alaskan Native	21	19	92.4	17	87.1

TABLE III.2 (continued)

	Sample	Located Sample		Response among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Living Situation						
Living alone	1,941	1,812	93.5	1,571	86.3	80.6
Living with others	1,40	129	94.1	105	79.4	75.1
Living with parents	34	31	91.3	28	90.5	82.6
In institution or unknown	1,267	1,214	96.8	1,019	84.5	81.8
Did the Applicant for Benefits Live in Same Zip Code as Beneficiary?						
No	303	265	86.8	224	85.3	73.8
Yes	1,596	1,507	95.0	1,319	88.2	83.8
No information	1,483	1,414	96.4	1,180	83.0	79.9
Identity of the Payee with Respect to the Beneficiary						
Beneficiary received beneficiary payments himself or herself	1,964	1,846	95.2	1,567	84.7	80.6
Payee is a family member	1,106	1,043	95.6	902	87.1	83.3
Payee is an institution	220	210	94.3	183	86.8	82.1
Other	92	87	93.2	71	84.3	78.2
Changes in Telephone Number						
No changes in last 5 years	58	55	95.6	47	88.2	84.0
One change in last 5 years	5	5	100.0	4	70.8	70.8
Two or more changes in last 5 years	2	2	100.0	2	100.0	100.0
No information on phone number	3,317	3,124	95.2	2,670	85.2	81.1
Number of Moves in Last 5 Years						
No moves in last 5 years	39	37	96.6	34	92.6	89.3
One or more moves in last 5 years	2	2	100.0	2	100.0	100.0
No information on number of moves	3,341	3,147	95.2	2,687	85.2	81.1
Type of Claim						
Survivor	310	299	94.4	261	88.5	83.4
Disabled	1,610	1,533	96.6	1,300	85.3	82.4
Unknown	1,462	1,354	92.9	1,162	84.4	78.3
Census Region						
Midwest	801	763	95.6	674	88.6	84.7
Northeast	550	517	94.6	417	81.7	77.2
South	1,370	1,289	94.6	1,117	86.1	81.4
West	661	617	96.4	515	82.4	79.5
Census Division						
East North Central	594	565	94.9	502	89.1	84.5
East South Central	306	294	97.3	249	86.3	83.8
Middle Atlantic	382	358	94.2	296	84.7	79.7
Mountain	175	163	96.5	148	89.1	86.0
New England	168	159	95.7	121	74.9	71.5
Pacific	486	454	96.4	367	80.0	77.1
South Atlantic	729	680	93.4	590	84.8	79.2
West North Central	207	198	97.6	172	87.1	85.3
West South Central	335	315	94.8	278	89.1	84.4

TABLE III.2 (continued)

	Sample	Located Sample	Response among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate
Metropolitan					
Metropolitan areas of 1 million population or more	1,474	1,378	94.5	1,148	84.3
Metropolitan areas of 250,000 to 999,999 population	864	809	95.1	695	83.6
Metropolitan areas of less than 250,000 population	381	362	93.3	316	88.1
Nonmetropolitan areas adjacent to large metropolitan areas	242	233	98.4	209	90.7
Nonmetropolitan areas adjacent to medium or small metropolitan areas	252	244	98.3	214	88.6
Nonmetropolitan areas not adjacent to metropolitan areas	169	160	97.3	141	82.8

Source: NBS, round 3.

The **weighted location rate** is the ratio of the weighted sample count for located sample members to the weighted count of all sample members, given in Table III.2 as 95.2 percent. The **weighted cooperation rate** (the weighted completion rate among located sample members), 85.2 percent in Table III.2, is the weighted count of sample members for whom a completed interview was obtained, or who were determined to be ineligible, divided by the weighted sample count of all located sample members. Weighted cooperation rates reflect the common survey situation that once a person is located, repeated contact efforts often will result in a completed interview.

The weighted rates are used because (1) the sampling rates (therefore the sampling weights) vary substantially across the sampling strata, as seen in Table III.1; and (2) the weighted rates better reflect the potential for nonresponse bias. The weighted rates represent the percentage of the full survey population for which we were able to obtain information sufficient to use either in the data analysis or to determine as ineligible for the analysis.

c. Factors Related to Location and Response

In addition to overall response rate information, Table III.2 also provides information for selected factors associated with locating a sample member, and factors associated with response among located sample members. The table includes the unweighted counts of all sample members, counts of located sample members, and counts of sample members for whom a completed interview was obtained, or who were determined to be ineligible. The table also includes the weighted location rate, the weighted completion rate among the located sample members, and the weighted overall completion rate for these factors, which helped to inform the decision about the final set of variables used in the nonresponse adjustment models.

d. Propensity Models for Weight Adjustments

The response propensity models used to determine the nonresponse adjustments were developed using the main effects described previously, plus selected interactions. To identify candidate interactions among these variables for the modeling, we first ran a chi-squared automatic interaction detector (CHAID) analysis in SPSS to find possible significant interactions. CHAID normally is attributed to Kass (1980) and Biggs et al. (1991), and its application in SPSS is described in Magidson (1993). The CHAID procedure iteratively segments a data set into mutually exclusive subgroups that share similar characteristics based on their effect on nominal or ordinal dependent variables. It automatically checks all variables in the data set and creates a hierarchy that shows all statistically significant subgroups. The algorithm finds splits in the population, which are as different as possible based on a chi-square statistic. It is a forward stepwise procedure; it finds the most diverse subgrouping, and then each of these subgroups is split further into more diverse sub-subgroups. Sample size limitations are set to avoid generating cells with small counts. It stops when splits no longer are significant; that is,

that group is homogeneous with respect to variables not yet used, or when the cells contain too few cases. The CHAID procedure results in a tree that identifies the set of variables and interactions among the variables that have an association with the ability to locate a sample member (and the propensity of a located sample member to either respond or be ineligible). CHAID first was run with all covariates, then rerun a few times with the top variable in the tree removed, to ensure all potentially important interactions were retained for further consideration. The resulting pool of covariates was reduced further by evaluating tabulations of all the main effects and the interactions identified by CHAID. At a particular level of a given covariate or interaction, if all respondents either were located or unlocated (for the location models), complete or not complete (for the cooperation models), or the total number of sample members at that level was fewer than 20, then levels were collapsed if collapsing was possible. If collapsing was not possible, then the covariate or interaction was excluded from the pool.²⁹

All of the resulting candidate main effects, and the interactions identified using CHAID, were then processed using forward and backward stepwise regression (using SAS Logistic procedure with weights normalized to the sample size) to further refine the candidate variables and interaction terms.³⁰ After identifying a smaller pool of main effects and interactions for potential inclusion in the final model, a set of models was evaluated carefully to determine the final model. Because the SAS logistic procedure does not incorporate the sampling design, the final selection of the covariates was accomplished using the logistic regression procedure in SUDAAN.

²⁹ Deafness historically has been shown to be an important indicator of both locating a sample member, and of whether the sample member completed the interview. For that reason, deafness was allowed to remain in the covariate pool even though the number of deaf cases was sometimes as low as 18.

³⁰ Because no automated stepwise procedures are available in SUDAAN, the stepwise procedures described here were performed using SAS.

For selecting variables or interactions in the stepwise procedures, we included variables or interactions that had a statistical significance level (alpha level) of 0.30 or lower (instead of the commonly used 0.05).³¹ Once the candidate list of main effects and interactions was determined, a thorough model-fitting process was used to determine a parsimonious model with few very small propensities. Model selection criteria were described in the overview of this chapter (Section A). The variables used in the model as main effects and interactions are summarized in Table III.3 for locating a sample member and in Table III.4 for cooperation among located sample members. The R-squared is 0.054 (0.168 when rescaled to have a maximum of 1) for the location model and 0.078 (0.138 when rescaled) for the cooperation model.³² These values are similar to those observed for other response propensity modeling efforts using logistic regression with design-based sampling weights. For the location model, the percentage of concordant pairs is 69.1 percent, 29.7 percent of the pairs are discordant³³, and the p-value for the chi-square statistic from the Hosmer-Lemeshow (H-L) goodness-of-fit test is 0.184³⁴; these values indicate a reasonably good fit of the model to the data. For the cooperation model, the percentage of concordant pairs is 65.7 percent, and 33.7 percent of pairs are discordant. The p-value for the chi-square statistic for the (H-L) goodness-of-fit test is 0.420 for this model. Since the Akaike's

³¹ As stated earlier, we used a higher significance level because the purpose of the model was to improve the estimation of the propensity score, and not to identify statistically significant factors related to response. In addition, the information sometimes reflected proxy variables for some underlying variable that was both unknown and unmeasured.

³² The Generalized Coefficient of Determination (Cox and Snell 1989) is a measure of the adequacy of the model, where higher numbers indicate a greater difference between the likelihood of the model in question and the null model likelihood. The "Max rescaled R-Square" scales this value to have a maximum of 1.

³³ A pair of observations is concordant if a responding subject has a higher predicted value than the nonresponding subject, discordant if not, and tied if both members of the pair are either respondents, nonrespondents, or have the same predicted values. It is desirable to have as many concordant pairs and as few discordant pairs as is possible (Agresti 1996).

³⁴ The Hosmer-Lemeshow Goodness-of-Fit Test is a test for goodness of fit of logistic regression models. Unlike the Pearson and deviance goodness-of-fit tests, it can be used to test goodness of fit even when some of the covariates are continuous (Hosmer and Lemeshow 1989).

Information Criterion (AIC) is a relative number, and has no meaning on its own, values for the AIC are not provided here.³⁵

The primary factors are identified by the base variable, often followed by the suffix “_1.” If the levels associated with the variable as it is used in the location model correspond directly to those in Table III.2, no suffix is given. However, if levels of the variable used in the location model are collapsed from those shown in Table III.2, the base variable name is followed by the suffix “_1.”

TABLE III.3

LOCATION LOGISTIC PROPENSITY MODEL: REPRESENTATIVE BENEFICIARY SAMPLE

Factors in the Location Model

Main Effects
DIG_1
REPREPAYEE_1
GENDER (SEX)
METRO_1
DIVISION_1
SSI_SSDI
RACE_1
TOC_1
AGECAT
PDZIPSAME_1

Two-Factor Interactions
DIVISION_1*RACE_1
DIG_1*PDZIPSAME_1
DIG_1*TOC_1
DIVISION_1*DIG_1
SSI_SSDI*DIVISION_1
SSI_SSDI*REPREPAYEE_1
DIVISION_1*SEX
AGECAT*TOC_1
AGECAT*DIG_1
AGECAT*DIVISION_1

³⁵ Akaike’s Information Criterion is defined as $AIC = -2\text{LogL} + 2(k+s)$, where LogL is the loglikelihood of the binomial distribution using the parameters from the given model, k is the total number of response levels minus one, and s is the number of explanatory effects (Akaike, 1974). AIC is a relative number, and has no meaning on its own. For a given model, smaller values of AIC are better than larger values.

TABLE III.4

COOPERATION LOGISTIC PROPENSITY MODEL: REPRESENTATIVE BENEFICIARY SAMPLE

Factors in the Cooperation Model

Main Effects
 AGECAT_2
 RACE_2
 HISPANICITY
 METRO_2
 DIVISION_2
 SSI_SSDI_
 GENDER (SEX)
 REPREPAYEE_2
 PDZIPSAME_2
 DIG_2
 TOC_2
 LIVING_2

Two-Factor Interactions
 DIG_2*RACE_2
 DIVISION_2* RACE_2
 SSI_SSDI * RACE_2
 DIVISION_2* PDZIPSAME_2
 DIG_2* PDZIPSAME_2
 METRO_2* TOC_2
 DIVISION_2*METRO_2
 DIG_2*DIVISION_2
 TOC_2* DIVISION_2
 SSI_SSDI* DIVISION_2
 SSI_SSDI* METRO_2
 METRO_2*SEX
 DIVISION_2*SEX
 PDZIPSAME_2*METRO_2
 REPREPAYEE_2*METRO_2
 DIG_2*SSI_SSDI
 SSI_SSDI*SEX

The factors with levels used in the location model include:

1. DIG_1: disability diagnostic classification; three levels: (1) mental disability, (2) physical disability (including deaf cases), and (3) unknown.
2. REPREPAYEE_1: the identity of the payee with respect to the beneficiary; two levels: (1) the beneficiary received benefit payments from a family member, and (2) an institution received payments on behalf of the beneficiary, or the beneficiary received benefit payments himself or herself, or identity of payee not known.
3. GENDER (SEX): two levels: (1) Male, and (2) Female.

4. ***METRO_1***. Urbanicity of beneficiary's place of residence; four levels: (1) beneficiary lived in metropolitan area, (2) beneficiary lived in nonmetropolitan area adjacent to a metropolitan area of 1 million or more, and (3) beneficiary lived in nonmetropolitan area adjacent to a metropolitan area of less than 1 million, (4) beneficiary lived in nonmetropolitan area not adjacent to metropolitan area.
5. ***DIVISION_1***. Geographic region (based on U.S. Census divisions) of beneficiary's place of residence; four levels: (1) South Atlantic, (2) West (Mountain and Pacific), (3) Midwest (East North Central and West South Central), (4) all other divisions.
6. ***SSI_SSDI***. Beneficiary status; three levels: (1) SSI only, (2) SSDI only, (3) Both SSI and SSDI.
7. ***RACE_1***. Race; two levels: (1) White, (2) Not white or not known to be white.
8. ***TOC_1***. Type of claim; 2 levels: (1) Disability claim, (2) Survivor claim or unknown.
9. ***AGECAT***. Beneficiary's age category; four levels: (1) age in range 18 to 29 years, (2) age in range 30 to 39 years, (3) age in range 40 to 49 years, and (4) age in range 50 to 64 years.
10. ***PDZIPSAME_1***. Whether the beneficiary and the applicant for benefits lived in the same zip code; two levels: (1) beneficiary and applicant lived in the same zip code, (2) beneficiary and applicant lived in different zip codes, or information unknown.

Various interactions among these variables were also included in the model for locating the sample member. The main effects using the variable names listed above, as well as interactions, are provided in Table III.3. An expanded form of Table III.3, showing the specific levels of the interactions shown in Table III.3, along with parameter estimates and their standard errors, is provided in Appendix D.

For the cooperation models, the primary factors are identified by the base variable, often followed by the suffix “_2.” As with the location model, if the levels associated with the variable used in the cooperation model are collapsed from those given in Table III.2, the base variable

name is accompanied by the suffix, but if no collapsing was necessary, no suffix is given. The factors include³⁶:

1. **AGECAT_2.** Beneficiary's age category; 3 levels: (1) age in range 18 to 29 years, (2) age in range 30 to 39 years, and (3) age in range 40 to 64 years
2. **RACE_2.** Race of the beneficiary; four levels: (1) white, (2) black, (3) Asian or Pacific Islander, and (4) race known to be neither white nor black nor Asian/Pacific Islander, or unknown.
3. **HISPANICITY.** Whether the beneficiary was Hispanic or not; two levels: (1) Hispanic, and (2) not Hispanic, or unknown.
4. **METRO_2.** Urbanicity of beneficiary's place of residence; four levels: (1) beneficiary lived in metropolitan area with population of 1 million or more, (2) beneficiary lived in metropolitan area with population between 250,000 and 1 million, (3) beneficiary lived in metropolitan area with population less than 250,000, (4) beneficiary lived in nonmetropolitan area.
5. **DIVISION_2.** Geographic region (based on U.S. Census divisions) of beneficiary's place of residence; seven levels: (1) New England, (2) Middle Atlantic, (3) South Atlantic, (4) West South Central, (5) Mountain, (6) Pacific and (7) Midwest (East North Central and West North Central), and (8) East South Central.
6. **SSI_SSDI.** Beneficiary status; three levels: (1) SSI only, (2) SSDI only, (3) Both SSI and SSDI.
7. **GENDER (SEX).** Two levels: (1) Male, and (2) Female.
4. **REPREPAYEE_2.** The identity of the payee with respect to the beneficiary; two levels: (1) the beneficiary received benefit payments from a family member, and (2) an institution received payments on behalf of the beneficiary, or the beneficiary received benefit payments himself or herself, or identity of payee not known.
5. **PDZIPSAME_2.** Whether the beneficiary and the applicant for benefits lived in the same zip code; two levels: (1) beneficiary and applicant lived in the same zip code, (2) beneficiary and applicant lived in different zip codes/information unknown.
8. **DIG_2.** Disability diagnostic classification; three levels: (1) mental disability, (2) physical disability (excluding deaf cases), and (3) deafness/unknown.
9. **TOC_2.** Beneficiary's type of claim; two levels: (1) disability claim, and (2) survivor claim, or unknown.

³⁶ Primary factors based on the same base variable as those given in the location model, but with different collapsing of categories, are given the same name except that they are followed by an “_2”.

10. **LIVING_2**. Beneficiary's living situation: three levels: (1) beneficiary lives alone, (2) beneficiary lives with others and (3) others/unknown.

Once again, various interactions among these variables were also included in the model for the cooperation of the sample members. The main effects using these variable names, as well as interactions, are provided in Table III.4. An expanded form of Table III.4, with the specific levels of the interactions shown in Table III.4, along with parameter estimates and their standard errors, is provided in Appendix D.

After adjustments were applied to the sampling weights, the distribution of weights was reviewed to determine if trimming of the sampling weights was necessary. Prior to trimming, the maximum design effect due to unequal weighting was 1.09, observed with the second youngest age group stratum. Trimming reduced this design effect to 1.06, which was still the maximum design effect, due to unequal weighting among all the strata.

3. Post-Stratification

Post-stratification is the procedure in which the weighted sums of the response-adjusted weights are aligned to known totals external to the survey. This process offers face-validity for reporting population counts and has some statistical benefits. For the Representative Beneficiary Sample, we post-stratified to the 24 population totals obtained from SSA.³⁷ In particular, the totals were the total number of SSI/SSDI beneficiaries by age (four categories), gender, and recipient status (SSI only, SSDI only, and both). No trimming was conducted after post-stratification.

³⁷ These totals were obtained from a frame file provided by the SSA, giving information on basic demographics for all SSI and SSDI beneficiaries.

C. TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE

As noted earlier, the Ticket Participant Cross-Sectional Sample was selected from the round 3 population of Ticket-to-Work participants in Phase 2 and Phase 3 states, a subset of all SSI/SSDI beneficiaries, which was partitioned based on different payment types in the Ticket-To-Work payment system (traditional vocational rehabilitation, milestone-outcomes, and outcome-only). Ticket participants using the traditional payment system accounted for 84 percent (28,170 of 33,500) of Phase 2 participants and 84 percent (25,913 of 31,023) of Phase 3 participants at the time the sampling frame was developed. Participants using the milestone-outcomes payment system totaled 4,138 Phase 2 participants (12 percent of all Phase 2 participants) and 4,410 Phase 3 participants (14 percent of all Phase 2 participants). Phase 2 participants using the outcome-only payment system totaled only 1,192 Phase 2 participants (4 percent of all Phase 2 participants) and 700 Phase 3 participants (2 percent of all Phase 3 participants). As was also noted earlier, the PSUs in the initial sampling design did not contain a sufficient number of participants in the milestone-outcome payment type in Phase 2 states, or the outcome-only payment type in Phase 2 or Phase 3 states, to support analysis tasks. As a result, the clustered sample, consisting of respondents selected within the initial sample design, was supplemented by a sample randomly selected from the entire population of milestone-outcome and outcome-only participants in Phase 2 states, and from the entire population of outcome-only participants in Phase 3 states (the unclustered sample).

The clustered sample was part of the original sample design, so all of the respondents in the clustered sample were selected from within PSUs, whereas the unclustered sample included units that may or may not have been in the selected PSUs. The unclustered sample was therefore organized into two strata: in the PSU or not in the PSU. In most cases, the respondents who were selected for the in-PSU stratum of the unclustered sample were also in the clustered sample. The

weights for these duplicate cases had to be appropriately adjusted to account for a single respondent's appearance in two independent samples. The compositing scheme used to do this is discussed in the next subsection. In addition, respondents who could not be located by the central office³⁸ based on sample frame information were treated differently in the clustered and unclustered samples. In the clustered sample, potential respondents who could not be located were sent to the field for further follow-up so that personal interviews could be attempted. In the unclustered sample, no further attempt was made to locate potential respondents who could not be located by the central office. If a sample member was selected as part of both the clustered and unclustered samples, and was sent to the field for further follow-up and located in the field, the response had to be treated differently between the two samples. For the sample respondent, the value in the clustered sample was recorded according to its final status in the field, whereas the value in the unclustered sample was recorded as "ineligible for field follow-up." Sample members with no field follow-up (in the unclustered sample) were not "selected" for field follow-up. This process is analogous to the accepted practice of subsampling of nonrespondents for more intensive effort—in this case, we subsampled cases in the clustered sample for field follow-up. Ineligible-for-field-follow-up cases in the unclustered sample were treated differently than other ineligible cases, regardless of whether the observation was duplicated with a clustered observation. The procedure used to create composite weights (described in the next subsection) was not applied to these cases. Rather, such a case in the unclustered sample would have its weight zeroed out. If such a case was duplicated with one in the clustered sample, the clustered sample case kept its original weight, appropriately adjusted so that the sum of weights was kept the same. The final sample sizes for the participants cross-sectional sample are in Table III.5.

³⁸ The "central office" is the MPR Survey Operations Center (SOC).

Although a portion of the Ticket participant population in round 3 included Phase 1 cases, no cross-sectional sample of these Phase 1 cases was selected in round 3, as indicated by the N/A entries in Table III.5.

TABLE III.5

SURVEY POPULATION AND INITIAL AUGMENTED AND FINAL CROSS-SECTIONAL SAMPLE SIZES, BY SAMPLING STRATA IN THE PARTICIPANT SURVEY

Sampling Strata (Payment System)	Survey Population ^a	Initial Augmented Sample Size ^b	Released Sample
Total Phase 1	44,265	N/A	N/A
1. Traditional payment type	39,357	N/A	N/A
2. Milestone-outcome payment type	3,613	N/A	N/A
3. Outcome-only payment type	1,295	N/A	N/A
Total Phase 2	33,500	3,388	2,797
1. Traditional payment type	28,170	1,000	867
2. Milestone-outcome payment type		1,500	937
Clustered sample	4,138	389	389
Unclustered sample	4,138	1,111	548
In PSUs	403	109	59
Not in PSUs	3,735	1,002	489
3. Outcome-only payment type		888	993
Clustered sample	1,192	123	123
Unclustered sample	1,192	765	870
In PSUs	123	79	88
Not in PSUs	1,069	686	782
Total Phase 3	31,023	2,858	1,373
1. Traditional payment type	25,913	1,000	444
2. Milestone-outcome payment type	4,410	1,000	444
3. Outcome-only payment type		858	485
Clustered sample	700	237	237
Unclustered sample	700	621	248
In PSUs	237	210	84
Not in PSUs	463	411	164

Source: Sample allocation and counts computed by MPR.

^aThis column reflects weighted totals before compositing.

^bThe initial (augmented) and final (released) sample sizes include participants for whom the number obtained from the original sample design was insufficient for analysis. For Phase 2 participants using the milestone-outcome or outcome-only payment types, and for Phase 3 participants using the outcome-only payment type, a paired sample design was employed, whereby the participants who were in the PSUs potentially could be selected for both samples.

For the clustered samples for TTW participants, the sample was allocated across the 79 PSUs, with the Los Angeles PSU receiving a double allocation because it had two selections. Because of the smaller population sizes, we used only the full PSUs; we did not use the SSUs in the Los Angeles PSU (four SSUs) or the Cook County (Chicago) PSU (two SSUs), which were used for the Representative Beneficiary Sample.

1. Initial Weights

The initial weights were computed based on the probability of selection within the PSU of the augmented sample and the probability of selection for the PSU. For the unclustered sample for the milestone-outcome and outcome-only participants in Phase 2 states, and for the outcome-only participants in Phase 3 states, we computed the initial weights based on the selection probability within the two sampling strata (in one of the PSUs, or not in any PSU). Since only a portion of the augmented sample was actually released for use, the initial weights then were adjusted for the sample actually used in the survey.

2. Dual Frame Estimation

To obtain estimates for the outcome-only Ticket Participant Samples in Phase 2 and Phase 3 states, and to obtain estimates for the milestone-outcome Ticket Participant Sample in Phase 2 states, it was necessary to combine the clustered and unclustered samples using a “paired sample design.” As noted earlier, if a potential respondent in the unclustered sample could not be located by the central office, he or she was considered “ineligible for field follow-up” and no further attempts were made on that case. However, if a potential respondent was in the clustered sample and could not be located by the central office, the case was sent to the field for additional locating efforts (field follow-up). The paired sample design is the methodology used to combine

the samples while accounting for these different rules of field follow-up. This requires the creation of composite weights that can be applied to the combined samples.

a. Conceptual Framework for Composite Weights

To compute a survey estimate, $Est(Y)$, using information from both samples (such as the proportion who are currently working), one cannot simply combine the two samples without adjusting the weights, since the clustered and unclustered samples in the Ticket Participant Sample represent the same target population among the Ticket Participants. Separate estimates can be computed from each sample, within each payment type, and combined, using the equation

$$(1) \ Est(Y) = \lambda Y(\text{clustered}) + (1 - \lambda) Y(\text{unclustered})$$

where $Y(\text{clustered})$ is the survey estimate from the clustered sample for the given payment type, $Y(\text{unclustered})$ is the survey estimate from the unclustered sample for the given payment type, and λ is an arbitrary constant between 0 and 1. For example, for the Phase 2 milestone-outcomes payment type in the round 3 data, there were 389 in the clustered sample and 548 in the unclustered sample. The estimates to be combined are the proportion of the 389 in the clustered sample who are currently working and the proportion of the 548 in the unclustered sample who are currently working. In practice, of course, it is more complicated than this, because we have to account for the different rules used in the two samples for following up with nonrespondents or unlocated sample members, as will be discussed later. For the sampling variance, $V(Y)$, the estimate is computed using the equation

$$(2) \ V(Y) = \lambda^2 V(Y(\text{clustered})) + (1 - \lambda)^2 V(Y(\text{unclustered}))$$

where $V(Y(\textit{clustered}))$ is the sampling variance for the estimate from the clustered sample, and $V(Y(\textit{unclustered}))$ is the sampling variance for the estimate from the unclustered sample. Any value of λ will result in an unbiased estimate of the survey estimate, but not necessarily an estimate with the minimum sampling variance. A lambda value producing a sampling variance at its minimum value results in the shortest confidence interval and, by implication, the most precise point estimate.

A value of lambda that minimizes the variance can be calculated as:

$$(3) \lambda = 1/V(Y(\textit{clustered})) / [1/V(Y(\textit{clustered})) + 1/V(Y(\textit{unclustered}))]$$

$$= V(Y(\textit{unclustered})) / [V(Y(\textit{clustered})) + V(Y(\textit{unclustered}))]$$

In this case, the minimum variance is:

$$(4) V(Y) = [V(Y(\textit{clustered})) * V(Y(\textit{unclustered}))] / [V(Y(\textit{clustered})) + V(Y(\textit{unclustered}))]$$

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

For this survey round, we used an alternative approach, which was to identify a single lambda that was calculated using sample sizes and design effects due to unequal weighting for the two samples. In particular, λ acts as a weighting factor, with more weight given to the larger sample, with the sample sizes adjusted by the design effect due to unequal weighting. The formula for λ is given by:

$$(5)\lambda = \frac{n(\text{clustered}) / \text{deff}(\text{clustered})}{n(\text{clustered}) / \text{deff}(\text{clustered}) + n(\text{unclustered}) / \text{deff}(\text{unclustered})}$$

where $n(\text{clustered})$ and $n(\text{unclustered})$ are the sample sizes of the clustered and unclustered central office-located samples respectively, and $\text{deff}(\text{clustered})$ and $\text{deff}(\text{unclustered})$ are the design effects due to unequal weighting for the clustered and unclustered central office-located samples, respectively.

b. Application of Composite Weights to Ticket Participant Sample

The population of participants in the relevant payment type can be separated into two parts: the portion that requires field follow-up and the portion that does not. For the portion of the target population that does not require field follow-up (that is, those who can be located by central office locating efforts), both the clustered and unclustered samples are independent samples that can provide unbiased estimates for this subpopulation. However, for the other portion of the target population that does require field follow-up (that is, those who cannot be located by central office locating efforts), only the clustered sample can provide unbiased estimates for this subpopulation, since unclustered sample cases were not eligible for field follow-up.

For the subpopulation that can be located by central office locating efforts, the clustered and unclustered samples can be combined using the compositing method (called a “dual frame” estimation procedure). To compute the composite weight for each sample member in the clustered central office-located sample:

$$(6)WT = \lambda \text{ WT}(\text{unclustered central office-located sample weight})$$

For units in the unclustered central office-located sample:

$$(7)WT = (1 - \lambda) WT(\text{clustered central office-located sample weight})$$

Conversely, for the subpopulation of persons who could not be found by central office locating efforts, only the clustered sample can be used. In this case, no combining is required, and the clustered weight is used directly:

$$(8)WT = 1 * WT(\text{clustered field-located sample weight})$$

The sum of weights among cases that were field-located in the clustered sample was adjusted so that the total sum matched the original total sum. Because the weights for each subpopulation sum to the total number of individuals in each subpopulation, the two subpopulations simply can be combined to form the entire target population.

Because of the paucity of sample members in the PSUs in some cases, it was not uncommon for the unclustered sample to be much larger than the clustered sample. When combining samples and creating composite weights, this sometimes resulted in weights with unacceptably high levels of variation. This made trimming necessary to reduce this variation, which is described in a later section.

3. Nonresponse Adjustment

As with the Representative Beneficiary Survey, the sampling weights were adjusted in two stages, one stage for the sample members who could not be located and another stage for those who, once located, refused to respond. Due to the small number of unlocated Ticket participants, the location adjustment was calculated using the weighting class method for all except traditional Phase 2 participants. However, the location adjustment for traditional Phase 2 Ticket participants and the response adjustment for located Ticket participants of both phases and all three payment types was calculated using logistic propensity models. For the milestone-outcome and outcome-

only payment types in Phase 2 states, and for the outcome-only payment type in Phase 3 states, the nonresponse adjustments were applied to the composite weights for the clustered and unclustered samples. Roughly equal sample sizes with vastly different population sizes for the three payment types resulted in substantial differences in the magnitude of the weights. Thus, it was necessary to calculate separate adjustments for each payment type and phase, first for the location adjustment and subsequently for the cooperation adjustment. This resulted in a total of 12 weight adjustments, including 5 location adjustments using the weighting class method, and 7 adjustments using logistic propensity models. The models were fitted in the same way as the adjustment models for the Representative Beneficiary Sample, as described in Section B.2 of this chapter. The main factors or attributes affecting our ability to locate and interview Ticket Participant sample members were the same as those used to locate and interview Representative Beneficiaries, where the specific covariates for each of the 12 weight adjustments varied as described in subsequent sections.

a. Coding of Survey Dispositions

The scheme used to code respondents included the four general categories described in Section B.2: eligible respondents, ineligible respondents, located nonrespondents, and unlocated sample members.³⁹

³⁹ Disposition codes 420 (institutionalized) and 430 (unavailable during field period) were classified as nonrespondent codes in round 3, even though they were considered ineligible codes in round 1. This affected 4 cases in the round 3 participant sample. As a result, the nonresponse adjusted weight for these 4 cases was 0 in round 3, even though a similar response in round 1 would have resulted in a positive weight. Because of the small numbers, the effect on response rates was very small.

b. Response Rates

The response rate for the Ticket Participant Cross-Sectional Sample is 84.4 percent, which is the weighted overall completion rate for Phase 2 and Phase 3 cases. This rate is a combination of the Phase 2 weighted overall completion rate (84.5 percent) and the Phase 3 weighted overall completion rate (84.2 percent). It is also the product of the weighted location rate and the weighted completion rate among located sample members. The weighted location rate is 97.0 percent, the combination of the Phase 2 and Phase 3 location rates (97.1 percent and 96.8 percent, respectively). The weighted cooperation rate (the weighted completion rate among located sample members), is 87.0 percent, the combination of the Phase 2 and Phase 3 weighted completion rates (87.1 percent and 86.9 percent, respectively).

Analogous to the beneficiary sample, the weighted rates are used because the sampling weights vary substantially across the sampling strata, and the weighted rates better reflect the potential for nonresponse bias.

c. Factors Related to Location and Response

Tables III.6 – III.11 provide information for selected factors associated with locating a sample member within each phase-payment type combination, and factors associated with response among located sample members. The tables include unweighted counts of all sample members, counts of located sample members, and counts of the sample members for whom a completed interview was obtained, or who were determined to be ineligible. The tables also include the weighted location rate, the weighted completion rate among located sample members, and the weighted overall completion rate for these factors, which helped inform the decision about the final set of variables used to define the weighting classes and in the nonresponse adjustment models.

TABLE III.6

WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 2 STATES, MILESTONE-OUTCOME PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS

	Sample		Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate	
All	841	824	98.2	674	81.2	79.8	
SSI Only, SSDI Only, or Both SSI and SSDI							
SSI only	245	234	95.2	191	78.9	75.2	
SSDI only	375	369	98.7	306	81.9	80.9	
Both SSI and SSDI	221	221	100.0	177	82.1	82.0	
SSI or SSDI							
SSI only, or in both SSI & SSDI programs	466	455	97.7	368	80.6	78.8	
SSDI only, or in both SSI & SSDI programs	596	590	99.2	483	82.0	81.3	
Constructed Disability Status							
Deaf	32	31	99.1	17	58.7	58.6	
Mental	498	482	96.8	390	80.4	77.9	
Physical	297	297	100.0	254	84.4	84.4	
Unknown	14	14	100.0	13	90.8	90.6	
Beneficiary's Age (Four Categories)							
18-29 years	228	224	98.4	173	76.8	75.5	
30-39 years	189	180	95.5	148	83.1	79.4	
40-49 years	227	223	98.7	188	83.7	82.7	
50-64 years	197	197	100.0	165	82.1	82.0	
Sex							
Male	422	414	98.7	333	81.7	80.7	
Female	419	410	97.7	341	80.8	78.9	
Hispanicity							
Hispanic	11	10	88.0	9	86.3	76.3	
Non-Hispanic/unknown	830	814	98.3	665	81.2	79.8	
Race							
White	431	426	98.7	348	81.5	80.5	
Black	323	312	97.3	261	82.5	80.1	
Other/unknown	87	86	98.5	65	75.6	74.5	
Living Situation							
Living alone	457	446	97.7	366	83.1	81.2	
Living with others/unknown	384	378	98.7	308	79.2	78.2	
Did the Applicant for Benefits Live in Same Zip Code as Beneficiary?							
No	98	95	97.2	81	82.2	80.1	
Yes	472	460	98.0	371	81.9	80.3	
No information	271	269	98.8	222	79.8	78.9	

TABLE III.6 (continued)

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Identity of the Payee with Respect to the Beneficiary						
Beneficiary received beneficiary payments himself or herself	508	496	98.2	424	85.6	84.0
Payee is a family member	261	258	98.5	197	75.8	74.8
Payee is an institution	54	53	98.2	41	73.7	72.4
Other	18	17	93.6	12	60.0	56.9
Changes in Telephone Number						
No changes in last 5 years	249	245	97.9	205	87.4	85.4
One or more changes in last 5 years	13	12	90.8	10	80.2	72.8
No information/other	579	567	98.6	459	78.0	77.0
Number of Moves in Last 5 Years						
No moves in last 5 years	123	120	97.3	106	90.0	87.4
One or more moves in last 5 years	10	10	100.0	7	60.0	59.6
No information on number of moves	708	694	98.4	561	79.8	78.6
Type of Claim						
Survivor	90	90	100.0	71	78.6	78.7
Disabled	483	476	98.8	393	82.8	81.8
Unknown	268	258	96.3	210	78.9	76.0
Census Region						
Midwest	381	372	98.4	307	81.7	80.3
Northeast	33	33	100.0	24	71.1	71.2
South	413	405	97.7	333	82.1	80.3
West	14	14	100.0	10	71.4	71.4
Census Division						
East North Central	338	331	99.2	276	82.1	81.4
East South Central	76	75	99.0	65	85.5	84.6
Middle Atlantic	16	16	100.0	13	82.4	82.2
Mountain	12	12	100.0	8	66.6	66.7
New England	17	17	100.0	11	60.4	60.5
Pacific	2	2	100.0	2	100.0	100.0
South Atlantic	47	47	100.0	40	84.9	84.9
West North Central	43	41	94.0	31	79.4	74.1
West South Central	290	283	96.5	228	79.4	76.8

TABLE III.6 (continued)

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Metropolitan						
Metropolitan areas of 1 million population or more	523	506	96.0	415	80.3	77.3
Metropolitan areas of 250,000 to 999,999 population	89	89	100.0	73	83.2	83.3
Metropolitan areas of less than 250,000 population	101	101	100.0	84	80.8	80.9
Nonmetropolitan areas adjacent to large metropolitan areas	56	56	100.0	49	88.0	87.8
Nonmetropolitan areas adjacent to medium or small metropolitan areas	42	42	100.0	29	68.8	69.0
Nonmetropolitan areas not adjacent to metropolitan areas	30	30	100.0	24	80.4	80.3
Longitudinal Sample Case						
Yes	372	362	97.2	293	82.2	79.8
No	469	462	99.1	381	80.4	79.7

Source: NBS, round 3.

TABLE III.7

WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE,
PHASE 2 STATES, OUTCOME-ONLY PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS

	Sample		Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate	
All	870	857	95.8	675	77.2		74.1
SSI Only, SSDI Only, or Both SSI and SSDI							
SSI only	111	108	97.3	81	75.9		73.4
SSDI only	665	658	98.2	516	76.0		75.0
Both SSI and SSDI	94	91	80.4	78	86.6		69.6
SSI or SSDI							
SSI only, or in both SSI & SSDI programs	205	199	88.4	159	81.0		71.4
SSDI only, or in both SSI & SSDI programs	759	749	95.6	594	77.3		74.2
Constructed Disability Status							
Deaf	21	21	100.0	16	79.8		78.0
Mental	457	449	98.0	360	75.7		74.8
Physical	385	380	92.6	294	79.1		73.1
Unknown	7	7	100.0	5	72.2		71.3
Beneficiary's Age (Four Categories)							
18-29 years	89	87	98.1	73	83.9		82.3
30-39 years	195	192	98.5	140	71.1		70.5
40-49 years	328	323	92.3	249	80.3		73.8
50-64 years	258	255	97.6	213	75.4		74.8
Sex							
Male	436	433	99.3	326	75.6		75.1
Female	434	424	92.5	349	78.5		73.2
Hispanicity							
Hispanic	15	14	92.8	13	92.3		85.5
Non-Hispanic/unknown	855	843	95.8	662	76.9		73.9
Race							
White	594	588	99.2	460	77.2		76.8
Black	173	170	94.4	131	72.9		69.1
Unknown	103	99	78.8	84	84.2		66.6
Living Situation							
Living alone	229	223	89.0	181	81.2		72.3
Living with others/unknown	641	634	98.1	494	75.9		74.8
Did the Applicant for Benefits Live in Same Zip Code as Beneficiary?							
No	65	65	100.0	58	90.4		90.5
Yes	330	323	90.4	255	78.6		71.0
No information	475	469	99.0	362	74.7		74.3

TABLE III.7 (continued)

	Sample	Located Sample		Response Among Located Sample	Overall Respondents	
	Count	Count	Location Rate	Count	Response Rate	
Identity of the Payee with Respect to the Beneficiary						
Beneficiary received beneficiary payments himself or herself	716	704	95.0	555	77.2	73.6
Payee is a family member	105	104	99.1	79	74.3	73.9
Payee is an institution	29	29	100.0	28	96.5	96.4
Other	20	20	100.0	13	65.2	65.2
Changes in Telephone Number						
No changes in last 5 years	339	334	93.6	245	70.5	66.4
One or more changes in last 5 years	5	5	100.0	4	80.6	80.0
No information/other	526	518	97.3	426	82.2	79.9
Number of Moves in Last 5 Years						
No moves in last 5 years	91	90	98.9	70	78.1	77.2
One or more moves in last 5 years	5	5	100.0	4	80.2	80.0
No information on number of moves	774	762	95.4	601	77.0	73.8
Type of Claim						
Survivor	48	46	96.5	39	83.2	80.5
Disabled	722	714	95.6	559	76.4	73.3
Unknown	100	97	97.0	77	80.5	77.6
Census Region						
Midwest	160	156	88.0	127	72.7	64.8
Northeast	499	493	98.9	385	78.0	77.2
South	183	180	96.4	141	79.6	76.7
West	28	28	100.0	22	78.9	78.5
Census Division						
East North Central	71	68	94.1	56	56.9	55.7
East South Central	77	76	98.8	57	78.5	76.9
Middle Atlantic	92	90	97.7	69	76.3	74.5
Mountain	24	24	100.0	19	79.2	79.1
New England	407	403	99.1	316	78.4	77.8
Pacific	4	4	100.0	3	77.1	75.0
South Atlantic	61	60	98.4	48	78.7	77.9
West North Central	89	88	83.7	71	85.5	71.1
West South Central	45	44	89.3	36	83.2	74.5

TABLE III.7 (continued)

	Sample	Located Sample		Response Among Located Sample	Overall Respondents	
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Metropolitan						
Metropolitan areas of 1 million population or more	320	315	96.3	252	80.2	77.2
Metropolitan areas of 250,000 to 999,999 population	355	353	99.6	274	77.6	77.4
Metropolitan areas of less than 250,000 population	48	48	100.0	37	81.0	80.9
Nonmetropolitan areas adjacent to large metropolitan areas	19	17	64.8	13	91.5	58.6
Nonmetropolitan areas adjacent to medium or small metropolitan areas	82	79	96.4	62	78.8	75.6
Nonmetropolitan areas not adjacent to metropolitan areas	46	45	98.6	37	49.8	52.1
Longitudinal Sample Case						
Yes	412	405	94.1	300	71.6	67.7
No	458	452	97.4	375	82.8	80.6

Source: NBS, round 3.

TABLE III.8

WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE,
PHASE 2 STATES, TRADITIONAL PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS

	Sample	Located Sample		Response Among Located Sample	Overall Respondents
	Count	Count	Location Rate	Count	Response Rate
All	867	840	97.0	742	85.7
SSI Only, SSDI Only, or Both SSI and SSDI					
SSI only	234	226	96.5	199	84.3
SSDI only	416	407	98.0	359	86.6
Both SSI and SSDI	217	207	95.7	184	85.2
SSI or SSDI					
SSI only, or in both SSI & SSDI programs	451	433	96.1	383	84.7
SSDI only, or in both SSI & SSDI programs	633	614	97.2	543	86.2
Constructed Disability Status					
Deaf	35	31	88.9	24	68.7
Mental	486	471	96.8	418	85.9
Physical	332	324	97.9	290	87.9
Unknown	14	14	100.0	10	69.0
Beneficiary's Age (Four Categories)					
18-29 years	246	240	97.5	214	87.0
30-39 years	182	175	96.7	151	82.5
40-49 years	249	240	96.2	208	83.6
50-64 years	190	185	97.6	169	89.6
Sex					
Male	454	442	97.4	387	85.5
Female	413	398	96.5	355	85.9
Hispanicity					
Hispanic	10	10	100.0	8	80.6
Non-Hispanic/unknown	857	830	97.0	734	85.7
Race					
White	552	536	97.1	472	85.7
Black	201	196	97.7	175	86.4
Other/unknown	114	108	95.4	95	84.2
Living Situation					
Living alone	444	426	96.0	381	85.7
Living with others/unknown	423	414	98.0	361	85.6
Did the Applicant for Benefits Live in Same Zip Code as Beneficiary?					
No	78	71	91.2	61	77.5
Yes	473	462	97.6	412	86.9
No information	316	307	97.4	269	85.8

TABLE III.8 (continued)

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Identity of the Payee with Respect to the Beneficiary						
Beneficiary received beneficiary payments himself or herself	534	518	97.0	457	88.3	85.7
Payee is a family member	261	252	96.7	227	90.0	86.9
Payee is an institution	51	51	100.0	47	92.9	92.8
Other	21	19	93.1	11	60.1	55.7
Changes in Telephone Number						
No changes in last 5 years	311	301	97.0	262	87.2	84.6
One or more changes in last 5 years	8	8	100.0	8	100.0	100.0
No information/other	548	531	96.9	472	88.8	86.1
Number of Moves in Last 5 Years						
No moves in last 5 years	122	118	97.0	103	87.9	85.2
One or more moves in last 5 years	7	7	100.0	7	100.0	100.0
No information/Other	738	715	97.0	632	88.3	85.6
Type of Claim						
Survivor	90	89	99.2	81	91.2	90.4
Disabled	553	534	96.8	471	88.5	85.7
Unknown	224	217	96.5	190	86.8	83.8
Census Region						
Midwest	449	435	96.8	383	88.5	85.6
Northeast	46	46	100.0	38	82.9	83.2
South	324	311	96.4	276	88.0	84.7
West	48	48	100.0	45	93.8	93.9
Census Division						
East North Central	368	358	97.3	314	87.9	85.5
East South Central	82	76	93.6	68	87.0	81.4
Middle Atlantic	12	12	100.0	10	84.2	84.2
Mountain	48	48	100.0	45	93.8	93.9
New England	34	34	100.0	28	82.6	82.9
Pacific	0	0	N/A	0	0	N/A
South Atlantic	158	156	98.7	134	86.0	84.8
West North Central	81	77	95.0	69	90.6	86.1
West South Central	84	79	94.1	74	93.7	88.2

TABLE III.8 (continued)

	Sample	Located Sample		Response Among Located Sample	Overall Respondents	
	Count	Count	Location Rate	Count	Response Rate	
Metropolitan						
Metropolitan areas of 1 million population or more	386	377	97.8	323	84.1	
Metropolitan areas of 250,000 to 999,999 population	123	120	97.5	107	86.7	
Metropolitan areas of less than 250,000 population	118	112	95.8	106	90.5	
Nonmetropolitan areas adjacent to large metropolitan areas	119	113	94.8	100	83.7	
Nonmetropolitan areas adjacent to medium or small metropolitan areas	26	25	95.3	22	82.0	
Nonmetropolitan areas not adjacent to metropolitan areas	95	93	97.9	84	88.4	
Longitudinal Sample Case						
Yes	432	420	97.5	367	85.0	
No	435	420	96.5	375	86.3	

Source: NBS, round 3.

TABLE III.9

WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE,
PHASE 3 STATES, MILESTONE-OUTCOME PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS

	Sample	Located Sample		Response Among Located Sample	Overall Respondents
	Count	Count	Location Rate	Count	Response Rate
All	444	426	95.9	369	84.3
SSI Only, SSDI Only, or Both SSI and SSDI					
SSI only	157	151	96.5	132	85.6
SSDI only	169	162	95.1	141	83.8
Both SSI and SSDI	118	113	96.2	96	83.2
SSI or SSDI					
SSI only, or in both SSI & SSDI programs	275	264	96.4	228	84.5
SSDI only, or in both SSI & SSDI programs	287	275	95.6	237	83.6
Constructed Disability Status					
Deaf	8	8	100.0	4	63.2
Mental	246	232	95.0	196	82.1
Physical	182	178	96.8	162	88.4
Unknown	8	8	100.0	7	88.1
Beneficiary's Age (Four Categories)					
18-29 years	85	81	95.9	72	86.9
30-39 years	104	100	96.5	80	79.0
40-49 years	145	138	94.3	121	84.1
50-64 years	110	107	97.5	96	87.5
Sex					
Male	223	213	95.0	182	81.9
Female	221	213	96.8	187	86.5
Hispanicity					
Hispanic	14	14	100.0	13	92.9
Non-Hispanic/unknown	430	412	95.8	356	84.0
Race					
White	179	171	95.5	150	85.9
Black	186	178	95.7	152	81.9
Other/unknown	79	77	97.4	67	85.0
Living Situation					
Living alone	264	254	96.5	221	85.1
Living with others/unknown	180	172	95.0	148	83.0
Did the Applicant for Benefits Live in Same Zip Code as Beneficiary?					
No	42	37	88.5	30	72.0
Yes	225	220	98.0	194	87.9
No information	177	169	94.7	145	82.2

TABLE III.9 (continued)

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Identity of the Payee with Respect to the Beneficiary						
Beneficiary received beneficiary payments himself or herself	318	306	96.0	267	88.5	85.0
Payee is a family member	95	93	98.2	75	82.4	80.9
Payee is an institution	27	24	88.5	24	100.0	88.5
Other	4	3	75.6	3	100.0	75.6
Type of Claim						
Survivor	23	22	96.1	19	87.1	84.0
Disabled	269	258	95.6	224	87.8	84.0
Unknown	152	146	96.5	126	87.9	84.8
Census Region						
Midwest	50	47	94.6	42	88.6	83.8
Northeast	55	55	100.0	47	89.2	89.3
South	193	184	94.9	170	92.9	88.2
West	146	140	95.9	110	79.0	75.8
Census Division						
East North Central	33	30	90.9	28	93.4	84.9
East South Central	40	38	95.0	37	97.3	92.5
Middle Atlantic	55	55	100.0	47	89.2	89.3
Mountain	2	2	100.0	1	50.6	50.0
New England	0	0	N/A	0	0	N/A
Pacific	144	138	95.8	109	79.3	76.0
South Atlantic	119	114	94.8	105	92.8	88.0
West North Central	17	17	100.0	14	82.4	82.4
West South Central	34	32	94.9	28	88.9	84.3
Metropolitan						
Metropolitan areas of 1 million population or more	242	233	96.3	201	86.2	83.1
Metropolitan areas of 250,000 to 999,999 population	188	180	96.0	155	86.5	83.0
Metropolitan areas of less than 250,000 population	5	5	100.0	5	100.0	100.0
Nonmetropolitan areas adjacent to large metropolitan areas	2	1	56.1	1	100.0	56.1
Nonmetropolitan areas adjacent to medium or small metropolitan areas	7	7	100.0	7	100.0	100.0
Nonmetropolitan areas not adjacent to metropolitan areas	0	0	N/A	0	0	N/A

Source: NBS, round 3.

TABLE III.10

WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 3 STATES, OUTCOME-ONLY PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
All	427	409	94.9	326	78.2	74.0
SSI Only, SSDI Only, or Both SSI and SSDI						
SSI only	111	104	93.0	87	82.6	76.7
SSDI only	221	217	97.4	171	76.7	74.4
Both SSI and SSDI	95	88	90.0	68	77.7	70.1
SSI or SSDI						
SSI only, or in both SSI & SSDI programs	206	192	91.5	155	80.2	73.5
SSDI only, or in both SSI & SSDI programs	316	305	95.4	239	77.0	73.2
Constructed Disability Status						
Deaf	7	7	100.0	2	30.7	32.1
Mental	235	221	93.5	172	75.6	71.0
Physical	174	171	98.3	142	82.3	80.0
Unknown	11	10	76.6	10	100.0	76.6
Beneficiary's Age (Four Categories)						
18-29 years	57	52	92.5	43	83.6	77.1
30-39 years	85	79	90.0	58	77.8	70.0
40-49 years	148	145	98.0	123	83.6	81.8
50-64 years	137	133	95.8	102	70.2	67.0
Sex						
Male	222	209	92.8	174	79.4	73.6
Female	205	200	97.3	152	77.0	74.5
Hispanicity						
Hispanic	17	15	87.0	14	88.2	76.9
Non-Hispanic/unknown	410	394	95.1	312	78.0	73.9
Race						
White	200	197	98.3	157	76.2	74.9
Black	131	120	87.0	100	85.7	74.7
Unknown	96	92	96.0	69	74.3	71.3
Living Situation						
Living alone	206	192	91.8	152	78.5	72.0
Living with others/unknown	221	217	97.3	174	78.0	75.7
Did the Applicant for Benefits Live in Same Zip Code as Beneficiary?						
No	33	27	79.0	19	73.8	57.1
Yes	216	206	95.1	174	80.9	77.2
No information	178	176	98.3	133	75.6	74.0

TABLE III.10 (continued)

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Identity of the Payee with Respect to the Beneficiary						
Beneficiary received beneficiary payments himself or herself	348	334	94.9	266	77.6	73.4
Payee is a family member	62	59	95.0	45	77.1	73.1
Payee is an institution	9	8	89.4	8	100.0	89.4
Other	8	8	100.0	7	96.3	96.7
Type of Claim						
Survivor	16	16	100.0	13	76.1	75.8
Disabled	297	285	94.8	226	78.0	73.6
Unknown	114	108	94.3	87	79.4	74.9
Census Region						
Midwest	31	30	95.0	24	78.9	74.8
Northeast	113	107	94.4	94	79.3	75.1
South	94	94	100.0	72	82.3	81.3
West	189	178	91.2	136	73.9	67.6
Census Division						
East North Central	25	24	93.7	20	82.4	77.1
East South Central	7	7	100.0	7	100.0	100.0
Middle Atlantic	112	106	94.3	94	80.4	76.1
Mountain	7	7	100.0	5	70.8	69.9
New England	1	1	100.0	0	0.0	0.0
Pacific	182	171	90.7	131	74.1	67.5
South Atlantic	38	38	100.0	29	81.5	80.3
West North Central	6	6	100.0	4	66.9	66.7
West South Central	49	49	100.0	36	79.6	78.9
Metropolitan						
Metropolitan areas of 1 million population or more	333	315	92.2	262	80.7	74.7
Metropolitan areas of 250,000 to 999,999 population	54	54	100.0	38	76.3	75.3
Metropolitan areas of less than 250,000 population	18	18	100.0	13	79.4	78.3
Nonmetropolitan areas adjacent to large metropolitan areas	5	5	100.0	4	79.6	80.0
Nonmetropolitan areas adjacent to medium or small metropolitan areas	12	12	100.0	6	50.3	50.0
Nonmetropolitan areas not adjacent to metropolitan areas	5	5	100.0	3	61.0	60.0

Source: NBS, round 3.

TABLE III.11

WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE,
PHASE 3 STATES, TRADITIONAL PAYMENT SYSTEM, BY SELECTED CHARACTERISTICS

	Sample	Located Sample		Response Among Located Sample	Overall Respondents	
	Count	Count	Location Rate	Count	Response Rate	Response Rate
All	444	431	97.1	375	86.9	84.4
SSI Only, SSDI Only, or Both SSI and SSDI						
SSI only	149	143	96.0	124	86.7	83.3
SSDI only	183	177	96.7	155	87.5	84.7
Both SSI and SSDI	112	111	99.1	96	86.5	85.7
SSI or SSDI						
SSI only, or in both SSI & SSDI programs	261	254	97.3	220	86.6	84.3
SSDI only, or in both SSI & SSDI programs	295	288	97.6	251	87.1	85.0
Constructed Disability Status						
Deaf	16	16	100.0	13	81.3	81.4
Mental	250	242	96.8	212	87.5	84.8
Physical	173	168	97.1	146	86.9	84.4
Unknown	5	5	100.0	4	79.5	79.7
Beneficiary's Age (Four Categories)						
18-29 years	122	119	97.5	104	87.3	85.2
30-39 years	88	84	95.4	74	88.0	84.1
40-49 years	136	132	97.0	113	85.5	83.1
50-64 years	98	96	97.9	84	87.4	85.7
Sex						
Male	241	234	97.1	206	87.9	85.5
Female	203	197	97.0	169	85.8	83.2
Hispanicity						
Hispanic	30	30	100.0	29	96.6	96.7
Non-Hispanic/unknown	414	401	96.8	346	86.2	83.5
Race						
White	268	263	98.1	236	89.8	88.1
Black	77	73	94.8	64	87.6	83.1
Other/unknown	99	95	95.9	75	78.9	75.8
Living Situation						
Living alone	255	248	97.2	216	87.1	84.7
Living with others/unknown	189	183	96.8	159	86.7	84.1
Did the Applicant for Benefits Live in Same Zip Code as Beneficiary?						
No	34	31	91.1	31	100.0	91.1
Yes	249	246	98.8	215	87.4	86.4
No information	161	154	95.7	129	83.6	80.1

TABLE III.11 (continued)

	Sample		Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate	
Identity of the Payee with Respect to the Beneficiary							
Beneficiary received beneficiary payments himself or herself	281	272	96.8	238	87.4		84.7
Payee is a family member	133	129	97.0	113	87.6		85.0
Payee is an institution	25	25	100.0	19	75.6		75.7
Other	5	5	100.0	5	100.0		100.0
Type of Claim							
Survivor	31	29	93.7	24	82.6		77.5
Disabled	280	275	98.2	243	88.3		86.7
Unknown	133	127	95.5	108	85.0		81.3
Census Region							
Midwest	72	69	95.8	64	92.9		89.1
Northeast	28	28	100.0	26	92.7		92.8
South	127	124	97.6	113	91.1		89.0
West	217	210	96.8	172	81.8		79.3
Census Division							
East North Central	48	47	97.9	45	95.8		93.9
East South Central	43	41	95.4	37	90.3		86.1
Middle Atlantic	28	28	100.0	26	92.7		92.8
Mountain	27	26	96.3	19	72.8		70.4
New England	0	0	N/A	0	0		N/A
Pacific	190	184	96.8	153	83.1		80.5
South Atlantic	49	48	98.0	44	91.6		89.7
West North Central	24	22	91.7	19	86.4		79.2
West South Central	35	35	100.0	32	91.6		91.6
Metropolitan							
Metropolitan areas of 1 million population or more	166	159	95.8	132	82.9		79.5
Metropolitan areas of 250,000 to 999,999 population	183	177	96.7	154	87.0		84.2
Metropolitan areas of less than 250,000 population	39	39	100.0	37	95.0		94.9
Nonmetropolitan areas adjacent to large metropolitan areas	9	9	100.0	8	88.4		88.2
Nonmetropolitan areas adjacent to medium or small metropolitan areas	44	44	100.0	42	95.5		95.5
Nonmetropolitan areas not adjacent to metropolitan areas	3	3	100.0	2	65.6		66.7

Source: NBS, round 3.

d. Weighting Classes and Propensity Models for Weight Adjustments

The weight adjustments used in the Ticket Participant Cross-Sectional Sample were based on weighting classes and predicted propensities from a logistic regression model. For the location weighting classes, each class is defined by the levels of the variables associated with locating the sample member. The response probability for a given class c is estimated as follows:

$$\phi(c) = \frac{\textit{sum of weights for respondents in class}}{\textit{sum of weights for all sample members in class}}$$

For a nonresponding unit within class c , the location adjustment is simply the inverse of this quantity. We calculated the location propensity for traditional Phase 2 participants and all of the cooperation propensities using logistic models. The adjustments for these were simply the inverse of the predicted propensities. The adjusted weight for each sample case is the product of the initial sampling weight and the adjustment factor.

The weighting classes were developed by reviewing Tables III.6 – III.11 to determine the variables most closely associated with locating a sample member. The models were developed using the main effects described previously, in addition to two main effects not used in the beneficiary models (the number of times the beneficiary moved in the past five years and the number of changes in the beneficiary’s phone number in the past five years), plus selected interactions. Interactions to be considered for inclusion in model development were identified using CHAID, as described in the model-fitting section for the Representative Beneficiary Sample.

The primary factors used to calculate the location adjustments are given below, with the potential levels used in the models. Details about how these levels were collapsed for each model are given in Appendix D.

1. ***PDZIPSAME***. Whether the beneficiary and the applicant for benefits lived in the same zip code; possible levels: (1) beneficiary and applicant lived in the same zip code, (2) beneficiary and applicant lived in different zip codes, and (3) information unknown.
2. ***METRO***. Urbanicity of beneficiary's place of residence; possible levels: (1) beneficiary lived in metropolitan area of 1 million or more residents, (2) beneficiary lived in metropolitan area of 250,000 to 1 million residents, (3) beneficiary lived in metropolitan area of less than 250,000 residents, (4) beneficiary lived in nonmetropolitan area adjacent to a metropolitan area of 1 million or more, (5) beneficiary lived in nonmetropolitan area adjacent to a metropolitan area of less than 1 million, and (6) beneficiary lived in nonmetropolitan area not adjacent to any metropolitan area.
3. ***GENDER (SEX)***. Possible levels: (1) male, and (2) female.
4. ***LIVING***. Beneficiary's living situation; possible levels: (1) beneficiary lives alone, (2) beneficiary lives with his or her parents, and (3) beneficiary does not live alone or with his or her parents, or information unknown.
5. ***AGECAT***. Beneficiary's age category; possible levels: (1) age in range 18 to 29 years, (2) age in range 30 to 39 years, (3) age in range 40 to 49 years, and (4) age in range 50 to 64 years.
6. ***SSI_SSDI***. Beneficiary status; possible levels: (1) SSI only, (2) SSDI only, or (3) both SSI and SSDI.
7. ***TOC***. Type of claim; possible levels: (1) survivor claim, (2) disability claim, and (3) type of claim unknown.
8. ***RACE***. Possible levels: (1) white, (2) black, (3) Asian or Pacific Islander, and (4) not white, black, or Asian/Pacific Islander, or unknown.

A list of the main effects using the variable names listed above for the weight adjustment procedure associated with each payment type-phase combination is provided in Tables III.12 and III.13. (An expanded form of Table III.12, with the specific levels of the main effects for each weighting class or model shown in Tables III.12 and III.13, along with parameter estimates and their standard errors, is provided in Appendix D.)

TABLE III.12

VARIABLES USED IN THE LOCATION ADJUSTMENTS: PHASE 2 TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE

Variables Defining the Location Weighting Class Adjustment for Participants Using the Milestone-Outcome Payment System

Main Effects
SEX
AGECAT
SSI_SSDI

Variables Defining the Location Weighting Class Adjustment for Participants Using the Outcome-Only Payment System

Main Effects
SEX
AGECAT

Variables in the Location Model for Participants Using the Traditional Payment System

Main Effects
LIVING
SSI_SSDI
METRO
TOC
PDZIPSAME

TABLE III.13

VARIABLES USED IN THE LOCATION ADJUSTMENTS: PHASE 3 TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE

Variables Defining the Location Weighting Class Adjustment for Participants Using the Milestone-Outcome Payment System
Main Effects PDZIPSAME SEX
Variables Defining the Location Weighting Class Adjustment for Participants Using the Outcome-Only Payment System
Main Effects SSL_SSDI RACE
Variables Defining the Location Weighting Class Adjustment for Participants Using the Traditional Payment System
Main Effects TOC RACE

The primary factors in the cooperation models are given below. Only the base variables with all possible levels are given. Some of the base variables below were also listed in the discussion of location adjustments. For these base variables, details about the levels are not given below, since they were provided earlier. (The description of how these levels were collapsed for each model is given in Appendix D.)

1. **MOVE.** The number of address changes in the past five years; possible levels: (1) no moves, (2) one move, (3) two or more moves, and (4) information older than five years, or no information.
2. **DIG.** Disability diagnostic classification; possible levels: (1) mental disability, (2) physical disability (excluding deaf cases), (3) deaf, and (4) unknown.
3. **REPREPAYEE.** The identity of the payee with respect to the beneficiary; possible levels: (1) the beneficiary received payments himself or herself, (2) a family member received benefits on behalf of the beneficiary, and (3) an institution received payments on behalf of the beneficiary, or identity of payee not known.
4. **PDZIPSAME.** Whether the beneficiary and the applicant for benefits lived in the same zip code.

5. **METRO.** Urbanicity of beneficiary's place of residence.
6. **GENDER (SEX).**
7. **REGION or DIVISION.** Geographic region of beneficiary's place of residence: DIVISION is based on U.S. Census divisions, with nine levels: (1) Pacific, (2) Mountain, (3) East North Central, (4) West North Central, (5) East South Central, (6) West South Central, (7) South Atlantic, (8) Middle Atlantic, and (9) New England. REGION is based on U.S. Census regions with four levels, which can be collapsed from the nine levels of DIVISION: (1) West is Pacific + Mountain, (2) Midwest is East North Central + West North Central, (3) South is East South Central + West South Central + South Atlantic, and (4) Northeast is Middle Atlantic + New England.⁴⁰
8. **LIVING.** Beneficiary's living situation.
9. **PHONE.** Number of phone numbers on SSA file over past five years; possible levels: (1) only one phone number on file, (2) one change in phone number on SSA file, (3) two or more changes in phone number on SSA file, and (4) information unknown.
10. **AGECAT.** Beneficiary's age category.
11. **SSI_SSDI.** Beneficiary status.
12. **TOC.** Type of claim.
13. **RACE.**
14. **R2LONG.** Phase 2 longitudinal sample case; possible levels: (1) Phase 2 longitudinal case, and (2) not a Phase 2 longitudinal case.

Various interactions among these variables were also included in the model for the cooperation of the sample members. A list of the main effects using the variable names listed above, as well as the interactions, is provided in Tables III.14 and III.15. (An expanded form of Tables III.14 and III.15, with levels appropriately collapsed for each model and the specific levels of the interactions shown in Tables III.14 and III.15, along with parameter estimates and their standard errors, is provided in Appendix D.)

⁴⁰ Many of the cooperation models used REGION instead of DIVISION. If a U.S. Census Division was used in a model, then the U.S. Census Region corresponding to that Division could not be in the model.

TABLE III.14

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODELS:
PHASE 2 TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE

Variables in the Milestone-Outcome Cooperation Model

Main Effects

REPREPAYEE
MOVE
LIVING
SSI_SSDI
DIVISION
TOC
DIG
RACE
AGECAT

Two-Factor Interactions

DIG*REPREPAYEE
TOC*AGECAT
TOC*RACE
DIG*DIVISION

Variables in the Outcome-Only Cooperation Model

Main Effects

RACE
LIVING
DIVISION
SEX
SSI_SSDI
PHONE
PDZIPSAME
AGECAT
R2LONG
METRO

Two-Factor Interactions

REGION*PHONE
R2LONG*AGECAT
R2LONG*REGION
R2LONG*SEX
REGION*AGECAT
AGECAT*R2LONG
AGECAT*SEX
METRO*AGECAT
REGION*SEX

Three-Factor Interactions

R2LONG*SEX*REGION

TABLE III.14 (continued)

Variables in the Traditional Cooperation Model
Main Effects
SSI_SSDI
DIVISION
LIVING
DIG
TOC
AGECAT
REPREPAYEE
METRO
Two-Factor Interactions
DIG*REPREPAYEE
TOC*AGECAT

TABLE III.15

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODELS: PHASE 3 TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE

Variables in the Milestone-Outcome Cooperation Model
Main Effects
REGION
RACE
SEX
DIG
REPREPAYEE
Two-Factor Interactions
REGION*SEX
Variables in the Outcome-Only Cooperation Model
Main Effects
AGECAT
SEX
REGION
METRO
Two-Factor Interactions
REGION*SEX
Variables in the Traditional Cooperation Model
Main Effects
RACE
REGION
REPREPAYEE
LIVING
Two-Factor Interactions
REPREPAYEE * LIVING

The process for determining which variables would be included in the weighting class adjustments was relatively straightforward. We reviewed Tables III.6-III.11 to determine two or three of the variables most closely associated with locating a sample member for each phase-payment type combination, and those variables were used to define the weighting classes. The model-fitting process, however, was considerably more involved. After identifying a smaller pool of main effects and interactions for potential inclusion in the final model, using backward and forward stepwise regressions, a set of models was statistically evaluated to determine the final model. Because the SAS logistic procedure does not incorporate the sampling design, the final selection of the covariates was accomplished using the logistic regression procedure in SUDAAN.

For selecting variables or interactions in the stepwise procedures, we again included variables or interactions that had a statistical significance level (alpha level) of 0.30 or lower (instead of the commonly used 0.05). Once the candidate list of main effects and interactions was determined, a thorough model-fitting process was used to determine a parsimonious model with few very small propensities.

The main effects used to calculate the location adjustments (for both weighting classes and models) are summarized in Tables III.12 and III.13, and the main effects and interactions in the models for cooperation among located sample members are summarized in Tables III.14 and III.15. The R-squared values for the 7 logistic models are given in Table III.16. The unadjusted R-squared value for the Phase 2 traditional location model was 0.013 (0.054 when rescaled to have a maximum of 1). The unadjusted R-squared value for the nonresponse models ranged from a low of 0.045 (0.084 when rescaled as above) up to 0.127 (0.193 when rescaled). These values are similar to those observed for other response propensity modeling efforts using logistic regression with design-based sampling weights. The percentages of concordant and discordant pairs, and the p-values for the Hosmer-Lemeshow goodness-of-fit test, are given in Table III.17.

TABLE III.16

UNADJUSTED AND ADJUSTED R-SQUARED VALUES FOR LOGISTIC PROPENSITY MODELS IN TICKET PARTICIPANT CROSS-SECTIONAL SAMPLES

Model			Unadjusted R-Squared Value	Adjusted R-Squared Value
Phase	Payment Type	Location or Cooperation		
2	Milestone-Outcome	Location	N/A	N/A
2	Milestone-Outcome	Cooperation	0.076	0.123
2	Outcome-Only	Location	N/A	N/A
2	Outcome-Only	Cooperation	0.127	0.193
2	Traditional	Location	0.013	0.054
2	Traditional	Cooperation	0.052	0.101
3	Milestone-Outcome	Location	N/A	N/A
3	Milestone-Outcome	Cooperation	0.062	0.118
3	Outcome-Only	Location	N/A	N/A
3	Outcome-Only	Cooperation	0.102	0.157
3	Traditional	Location	N/A	N/A
3	Traditional	Cooperation	0.045	0.084

TABLE III.17

PERCENTAGES OF CONCORDANT AND DISCORDANT PAIRS AND HOSMER-LEMESHOW P-VALUES FOR LOGISTIC PROPENSITY MODELS IN TICKET PARTICIPANT CROSS-SECTIONAL SAMPLES

Model			Percentage Concordant	Percentage Discordant	Hosmer-Lemeshow p-Value
Phase	Payment Type	Location or Cooperation			
2	Milestone-Outcome	Location	N/A	N/A	N/A
2	Milestone-Outcome	Cooperation	66.5	32.2	0.751
2	Outcome-Only	Location	N/A	N/A	N/A
2	Outcome-Only	Cooperation	63.1	36.1	0.179
2	Traditional	Location	63.2	29.2	0.740
2	Traditional	Cooperation	66.8	31.8	0.464
3	Milestone-Outcome	Location	N/A	N/A	N/A
3	Milestone-Outcome	Cooperation	68.5	27.3	0.969
3	Outcome-Only	Location	N/A	N/A	N/A
3	Outcome-Only	Cooperation	65.3	29.2	0.319
3	Traditional	Location	N/A	N/A	N/A
3	Traditional	Cooperation	63.1	28.9	0.389

Although the minimum difference between the percentages of concordant pairs and discordant pairs is 27 percent (Phase 2 outcome-only cooperation model), for the remainder of the models the difference between these percentages is at least 34 percent. The minimum p-value

associated with the Hosmer-Lemeshow goodness-of-fit test is 0.179, indicating no evidence of lack of fit for any of the models.

4. Trimming

After adjustments were applied to the sampling weights, the distribution of weights was reviewed to determine if trimming of the sampling weights was necessary. Because of the wide variation in the magnitude of the weights, due to the use of the composite weights in the outcome-only payment type in Phase 2 and Phase 3 states, and in the milestone-outcome payment type in Phase 2 states, trimming was sometimes necessary to increase the precision of survey estimates. However, we minimized the extent of trimming to reduce the potential for bias in the survey estimates. The design effects due to unequal weighting associated with each of the six phase-payment type combinations before and after trimming, before post-stratification, are presented in Table III.18. Design effects were calculated separately within trimming strata, which in turn were defined within phase-payment type combinations. In general, the trimming strata were defined according to whether the observation was in the clustered or unclustered sample, and whether the sample was part of the longitudinal or supplemental sample. For unclustered cases, the trimming strata were further subdivided according to whether the sample case was in a PSU or not, and whether the frame used to select the sample value was the round 2 or round 3 frames. The strata within which trimming was employed are given in Table III.18. If no trimming was employed for a phase-payment type combination, the maximum design effect across all trimming strata is presented. In that instance, the stratum associated with that maximum design effect is not presented since, in most cases, when no trimming is required, the design effects do not differ significantly across trimming strata.

TABLE III.18

DESIGN EFFECTS DUE TO UNEQUAL WEIGHTS BEFORE AND AFTER TRIMMING, WITHIN TRIMMING STRATA, FOR PHASES AND PAYMENT TYPES IN THE ROUND 3 TICKET PARTICIPANT CROSS-SECTIONAL SAMPLES

Phase and Payment Type	Trimming Stratum in which Trimming Occurred	Design Effect due to Unequal Weights	
		Before Trimming	After Trimming
Phase 2 Participants		Before Trimming	After Trimming
Milestone-Outcome	Clustered R2 frame supplemental	2.97	2.34
Milestone-Outcome	Clustered R2 frame longitudinal	3.26	2.94
Milestone-Outcome	Clustered R3 frame supplemental	3.47	2.81
Outcome-Only	Clustered R2 frame longitudinal	4.94	2.68
Outcome-Only	Clustered R3 frame supplemental	3.79	3.05
Traditional	No Trimming (three trimming strata)	1.05 (maximum)	1.05(maximum)
Phase 3 Participants		Before Trimming	After Trimming
Milestone-Outcome	No Trimming (one trimming stratum)	1.22	1.22
Outcome-Only	Clustered R3 frame longitudinal	2.73	2.03
Traditional	No Trimming (one trimming stratum)	1.01	1.01

Design Effect due to Unequal Weights = $n\Sigma w^2/(\Sigma w)^2$

5. Post-Stratification

After the nonresponse adjustment and trimming, the weights were post-stratified to the population age and gender totals for each payment type obtained from the SSA sampling frame. This sampling frame included all SSI or SSDI beneficiaries for each payment type within the Ticket Participant population. The distributions of weights within each phase and payment type combination were rechecked to determine if more weight trimming was necessary. No extreme weights were found after post-stratification.

D. TICKET PARTICIPANT LONGITUDINAL SAMPLE

For Phase 1 participants who responded in rounds 1, 2, and 3, or only in rounds 1 and 2, the inferential population is defined by the population of Ticket participants at the time of round 1 sampling. Of the 21,477 Phase 1 participants at round 1, 18,100 were participants who used the traditional payment system (84 percent), 2,809 used the milestone-outcome payment system (13

percent) and 568 used the outcome-only payment system (3 percent). In this population, only the participants who used the outcome-only payment system required a supplemental unclustered sample. For Phase 2 participants who responded in rounds 2 and 3, the inferential population is defined by the population of Ticket participants at the time of round 2 sampling. Of the 21,196 Phase 2 participants at round 2, 17,081 were participants who used the traditional payment system (81 percent), 3,208 used the milestone-outcome payment system (15 percent) and 907 used the outcome-only payment system (4 percent). As with the round 3 cross-sectional weights, Phase 2 participants using either the milestone-outcome or outcome-only payment systems required a supplemental unclustered sample. For both the Phase 1 and Phase 2 longitudinal samples, the clustered and unclustered samples were combined using composite weights, the creation of which was described in Section C.

The final sample sizes for the participant longitudinal samples are given in Table III.19. As stated earlier, the longitudinal samples are determined by the sample selected in round 1 for Phase 1 cases, and by the sample selected in round 2 for Phase 2 cases.

1. Initial Weights

The initial weights were computed based on the probability of selection within the PSU of the released participant sample of Phase 1 cases in round 1 and Phase 2 cases in round 2, and the probability of selection for the PSU. For the unclustered sample for each phase, we computed the initial weights based on the selection probability within the two sampling strata (in one of the PSUs, or not in any PSU).

TABLE III.19

SURVEY POPULATION AND FINAL LONGITUDINAL SAMPLE SIZES, BY SAMPLING STRATA IN THE PARTICIPANT SURVEY

Sampling Strata (Payment System)	Survey Population	Released Sample ^a
Total Phase 1	21,477	1,466
1. Traditional Payment Type	18,100	441
2. Milestone-Outcome Payment Type	2,809	455
3. Outcome-Only Payment Type		570
Clustered Sample	568	123
Unclustered Sample	568	447
In PSUs	123	123
Not in PSUs	445	324
Total Phase 2	21,196	1,350
1. Traditional Payment Type	17,081	437
2. Milestone-Outcome Payment Type		436
Clustered Sample	3,208	216
Unclustered Sample	3,208	220
In PSUs	273	19
Not in PSUs	2,935	201
3. Outcome-Only Payment Type		477
Clustered Sample	907	86
Unclustered Sample	907	391
In PSUs	86	44
Not in PSUs	821	347

Source: Sample allocation and counts computed by MPR.

^a The final (released) sample size includes participants for whom the number obtained from the original sample design was insufficient for analysis. For Phase 1 participants using the outcome-only payment type, and for Phase 2 participants using the milestone-outcome or outcome-only payment types, a paired sample design was employed, whereby the participants who were in the PSUs potentially could be selected for both samples.

2. Dual Frame Estimation

To obtain estimates for the paired Ticket Participant Cross-Sectional Samples, it was necessary to combine the clustered and unclustered samples using a paired sample design. As noted earlier, if a potential respondent in the unclustered sample could not be located by the central office, he or she was considered ineligible for field follow-up, and no further attempts were made on that case. However, if a potential respondent was in the clustered sample and could not be located by the central office, the case was sent to the field for additional locating efforts (field follow-up). The paired sample design is the methodology used to combine the

samples while accounting for these different rules of field follow-up. This requires the creation of composite weights that can be applied to the combined samples. This same logic is applied to the paired Ticket Participant Longitudinal Samples. The difference, however, is that a sample member may be ineligible for field follow-up in one round, but he or she might be a completed respondent at the central office in the next round. Hence, to create the longitudinal composite weights, different rules were required for defining who was ineligible for field follow-up across rounds. This, of course, depended upon the longitudinal analysis in question. In general, the following rule was applied: if the sample member was ineligible for field follow-up in any of the rounds associated with the longitudinal analysis in question, then the sample member was ineligible for field follow-up in all of the rounds in question. For example, for a longitudinal analysis involving Phase 1 cases where sample members were respondents⁴¹ in rounds 1 or 2, suppose a sample member had a completed interview in round 1, but was ineligible for field follow-up in round 2. For this longitudinal analysis, the sample member would be considered ineligible for field follow-up across the two rounds.

The conceptual framework for composite weights is identical to that given in Section C.2.

3. Nonresponse Adjustment

For the Ticket Participant Longitudinal Sample, we calculated three separate nonresponse adjustments to create three sets of longitudinal weights corresponding to the longitudinal analyses envisioned. In particular, we created two sets of longitudinal weights using the initial Phase 1 longitudinal weights: one for Phase 1 participants who responded in rounds 1 and 2, and the other for Phase 1 participants who responded in all three rounds. In addition, a third set of

⁴¹ “Respondents” include sample members who had completed interviews, or were ineligible after sample selection (i.e., when surveyed).

longitudinal weights was created using the initial Phase 2 longitudinal weights for Phase 2 participants who responded in rounds 2 and 3. As with the Representative Beneficiary and Ticket Participant Cross-Sectional Samples, the nonresponse adjustments were accomplished in two stages for each set of weights: one stage for the sample members who could not be located, and another stage for those who, once located, refused to respond. All adjustments were calculated using logistic regression propensity models. We applied separate adjustments for each payment type, applying the nonresponse adjustments to the composite weights for the unclustered and clustered samples, where appropriate. This resulted in a total of 18 logistic regression models. These models were fitted in the same way as the adjustment models for the Representative Beneficiary Sample, described in Section B.2, and the adjustment models for the Ticket Participant Cross-Sectional Sample, described in Section C.3. The specific covariates for each of the 18 logistic models are described in subsequent sections.

a. Coding of Survey Dispositions

For cross-sectional estimates, identifying completed cases and calculating response rates was straightforward once sample members were categorized into one of the following four groups:

1. Eligible respondents
2. Ineligible respondents (sample members who were ineligible after sample selection, including deceased, sample members in the military or incarcerated, sample members living outside of the United States, and other ineligibles)
3. Located nonrespondents (including active or passive refusals, and language barrier situations)
4. Unlocated sample members (sample members who could not be located using either central office tracing procedures or in-field searches)

However, for longitudinal estimates, rules were necessary to assign sample members to one of the four categories for all rounds covered by the longitudinal weight, based upon the categories into which they were classified for each round. For example, for longitudinal analyses involving Phase 1 cases for all three rounds, a complete case was defined as someone who responded in all three rounds. In the list of rules for classifying respondents into the four categories for longitudinal analyses, a “round-specific” respondent or nonrespondent indicates that the sample member did or did not respond in a particular round. The definition of “across rounds” depends upon the longitudinal analysis in question. In particular, it means either (1) “across rounds 1 and 2” for longitudinal analyses involving Phase 1 cases across rounds 1 and 2; or (2) “across rounds 1, 2, and 3” for longitudinal analyses involving Phase 1 cases across rounds 1, 2, and 3; or (3) “across rounds 2 and 3” for longitudinal analyses involving Phase 2 cases across rounds 2 and 3. The list of rules is given below:

1. Longitudinal eligible respondents: eligible respondents across rounds
2. Longitudinal ineligible respondents: sample members who were either ineligible or eligible respondents across rounds, and were ineligible respondents in at least one round (they were not considered longitudinal eligible respondents).
3. Longitudinal unlocated sample members: sample members who were either ineligible respondents, eligible respondents, or were not located across rounds, and were not located in at least one round (they were neither longitudinal eligible nor longitudinal ineligible respondents).
4. Longitudinal located nonrespondents: sample members who were located and did not respond in at least one round, and in all other rounds were either eligible respondents, ineligible respondents, or unlocated sample members.

b. Response Rates

We would expect the response rates for the Ticket Participant Longitudinal Sample to be lower than those obtained by the cross-sectional sample, since response is defined across rounds, and sample attrition would result in fewer responses in each round. When response is defined across all three rounds for Phase 1 cases, the response rate is 62.8 percent. This rate is a product

of the weighted location rate (92.1 percent) and the weighted completion rate among located sample members (68.2 percent). The same Phase 1 initial longitudinal weights (adjusted for the released sample) are used to calculate the response rate for longitudinal analyses across the first two rounds (70.5 percent), which is higher because response is required only across two rounds, instead of three. This is a product of the weighted location rate (93.6 percent) and the weighted completion rate among located sample members (75.3 percent). Finally, the response rate for Phase 2 cases for analyses across rounds 2 and 3 is 71.7 percent, the product of the weighted location rate (91.4 percent) and the weighted completion rate among located cases (78.4 percent).

c. Factors Related to Location and Response

Tables III.20 – III.28 provide information for selected factors associated with locating a sample member within each payment type associated with each set of longitudinal weights. The tables include unweighted counts of all sample members, counts of located sample members across rounds, and counts of the sample members for whom a completed interview was obtained, or who were determined to be ineligible when surveyed across each round. The tables also include the weighted location rate, the weighted completion rate among located sample members, and the weighted overall completion rate for these factors, which helped inform the decision about the final set of variables used in the nonresponse adjustment models.

d. Propensity Models for Weight Adjustments

The weight adjustments used in the Ticket Participant Longitudinal Sample were based on predicted propensities from a logistic regression model. For the location and cooperation weight adjustments, we used logistic models to estimate the propensity for a sample member to be located and cooperate across rounds. The inverse of the propensity score was used as the adjustment factor. The adjusted weight for each sample case is the product of the initial sampling weight and the adjustment factor.

TABLE III.20

WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL
SAMPLE, PHASE 1 STATES, ROUNDS 1 AND 2 RESPONDENTS, MILESTONE-OUTCOME PAYMENT SYSTEM,
BY SELECTED CHARACTERISTICS

	Sample	Located Sample	Response Among Located Sample		Overall Respondents	
	Count	Count	Location Rate	Count	Response Rate	
All	455	398	90.3	300	76.8	
SSI Only, SSDI Only, or Both SSI and SSDI						
SSI only	166	144	88.4	109	78.2	
SSDI only	197	178	92.7	128	72.5	
Both SSI and SSDI	92	76	88.3	63	84.2	
SSI or SSDI						
SSI only, or in both SSI & SSDI programs	258	220	88.3	172	80.6	
SSDI only, or in both SSI & SSDI programs	289	254	91.3	191	76.1	
Constructed Disability Status						
Deaf	6	5	84.6	4	76.3	
Mental	257	224	90.5	179	80.6	
Physical	179	159	91.1	110	71.9	
Unknown	13	10	76.7	7	66.1	
Beneficiary's Age (Four Categories)						
18-29 years	79	61	80.2	49	79.0	
30-39 years	114	104	93.4	81	76.9	
40-49 years	122	112	93.8	79	73.1	
50-64 years	140	121	89.6	91	79.1	
Sex						
Male	331	290	89.8	214	73.5	
Female	124	108	91.7	86	84.6	
Hispanicity						
Hispanic	19	12	64.9	11	89.4	
Non-Hispanic/unknown	436	386	91.1	289	76.5	
Race						
White	131	117	93.1	88	80.2	
Black	115	103	91.4	85	85.7	
Other/unknown	209	178	87.5	127	69.0	
Living Situation						
Living alone	264	228	89.5	179	80.6	
Living with others/unknown	191	170	91.4	121	72.2	
Did the Applicant for Benefits Live in Same Zip Code as Beneficiary?						
No	43	32	79.6	23	76.3	
Yes	205	179	90.3	145	83.1	
No information	207	187	92.6	132	70.9	

TABLE III.20 (continued)

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Identity of the Payee with Respect to the Beneficiary						
Beneficiary received beneficiary payments himself or herself	351	307	90.1	231	76.4	69.1
Payee is a family member	80	70	91.5	55	80.3	73.7
Payee is an institution	18	15	86.9	8	58.4	48.7
Other	6	6	100.0	6	100.0	100.0
Changes in Telephone Number						
No changes in last 5 years	314	277	91.3	210	77.6	70.9
One or more changes in last 5 years	12	9	76.3	9	100.0	76.3
No information/other	129	112	88.7	81	72.7	65.0
Number of Moves in Last 5 Years						
No moves last 5 years	130	109	86.5	87	79.1	68.8
One or more moves in last 5 years	11	6	59.5	5	77.9	45.6
No information on number of moves	314	283	92.7	208	76.0	70.6
Type of Claim						
Survivor	36	31	90.3	24	80.2	72.3
Disabled	262	232	91.7	174	75.8	69.7
Unknown	157	135	87.5	102	78.0	68.7
Census Region						
Midwest	75	67	92.0	55	82.1	75.2
Northeast	124	106	87.9	76	74.3	65.5
South	148	133	93.3	99	76.8	72.0
West	108	92	86.0	70	75.3	64.9
Census Division						
East North Central	73	65	91.0	54	86.2	78.5
East South Central	0	0	N/A	0	0	N/A
Middle Atlantic	77	66	89.1	49	77.7	69.6
Mountain	91	78	86.8	59	74.5	65.0
New England	47	40	85.8	27	67.9	58.0
Pacific	17	14	82.4	11	79.4	64.7
South Atlantic	147	132	92.9	98	75.6	70.7
West North Central	2	2	100.0	1	50.0	50.0
West South Central	1	1	100.0	1	100.0	100.0
Metropolitan						
Metropolitan areas of 1 million population or more	404	349	87.0	263	75.2	65.5
Metropolitan areas of 250,000 to 999,999 population	36	34	96.3	25	81.1	78.2
Metropolitan areas of less than 250,000 population	11	11	100.0	8	73.3	73.1
Nonmetropolitan areas adjacent to large metropolitan areas	0	0	N/A	0	0	N/A
Nonmetropolitan areas adjacent to medium or small metropolitan areas	4	4	100.0	4	100.0	100.0
Nonmetropolitan areas not adjacent to metropolitan areas	0	0	N/A	0	0	N/A

Source: NBS, round 3.

TABLE III.21

WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL
SAMPLE, PHASE 1 STATES, ROUNDS 1 AND 2 RESPONDENTS, OUTCOME-ONLY PAYMENT SYSTEM,
BY SELECTED CHARACTERISTICS

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
All	489	457	90.9	326	72.0	64.7
SSI Only, SSDI Only, or Both SSI and SSDI						
SSI only	94	84	82.4	57	70.8	56.8
SSDI only	319	301	92.2	219	74.6	68.0
Both SSI and SSDI	76	72	95.2	50	62.7	60.2
SSI or SSDI						
SSI only, or in both SSI & SSDI programs	170	156	88.4	107	67.0	58.4
SSDI only, or in both SSI & SSDI programs	395	373	92.8	269	72.3	66.5
Constructed Disability Status						
Deaf	4	3	77.3	1	29.0	22.7
Mental	258	241	90.4	164	68.7	61.3
Physical	214	201	91.9	155	77.6	70.9
Unknown	13	12	90.7	6	64.6	51.4
Beneficiary's Age (Four Categories)						
18-29 years	35	35	100.0	25	70.5	70.3
30-39 years	106	93	86.1	69	73.0	61.5
40-49 years	167	155	91.0	107	67.4	61.2
50-64 years	181	174	92.2	125	76.5	69.5
Sex						
Male	297	276	91.8	193	70.5	64.4
Female	192	181	89.3	133	74.5	65.2
Hispanicity						
Hispanic	13	12	90.6	5	42.3	39.9
Non-Hispanic/unknown	476	445	90.9	321	72.6	65.3
Race						
White	264	254	93.0	193	76.6	70.3
Black	65	56	87.2	39	66.5	58.6
Other/unknown	160	147	88.4	94	65.4	57.1
Living Situation						
Living alone	183	171	90.3	122	71.2	63.5
Living with others/unknown	306	286	91.2	204	72.5	65.5
Did the Applicant for Benefits Live in Same Zip Code as Beneficiary?						
No	39	36	93.9	21	57.2	54.8
Yes	191	181	91.3	140	74.0	66.7
No information	259	240	90.2	165	72.5	64.6

TABLE III.21 (continued)

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Identity of the Payee with Respect to the Beneficiary						
Beneficiary received beneficiary payments himself or herself	409	387	91.0	281	71.7	64.7
Payee is a family member	58	50	86.5	28	59.4	47.3
Payee is an institution	11	11	100.0	8	87.6	88.7
Other	11	9	89.5	9	100.0	89.5
Changes in Telephone Number						
No changes in last 5 years	416	395	92.2	281	70.4	64.5
One or more changes in last 5 years	11	11	100.0	11	100.0	100.0
No information/other	62	51	81.6	34	77.9	62.0
Number of Moves in Last 5 Years						
No moves last 5 years	118	108	89.7	80	67.2	59.8
One or more moves in last 5 years	8	8	100.0	7	85.6	85.4
No information on number of moves	363	341	91.1	239	73.2	65.9
Type of Claim						
Survivor	19	17	89.6	10	62.8	56.3
Disabled	384	361	92.6	260	71.7	66.2
Unknown	86	79	83.2	56	74.9	59.9
Census Region						
Midwest	98	89	90.4	66	74.6	67.0
Northeast	144	137	91.9	98	73.4	68.0
South	125	123	98.5	75	56.3	56.3
West	122	108	82.9	87	83.9	68.5
Census Division						
East North Central	85	76	88.4	58	77.8	68.2
East South Central	0	0	N/A	0	0	N/A
Middle Atlantic	53	50	88.5	30	66.9	59.2
Mountain	57	54	93.7	41	83.9	76.7
New England	91	87	93.8	68	77.0	72.9
Pacific	65	54	75.0	46	84.0	62.5
South Atlantic	121	119	98.4	73	56.5	56.4
West North Central	13	13	100.0	8	60.7	61.5
West South Central	4	4	100.0	2	49.2	50.0
Metropolitan						
Metropolitan areas of 1 million population or more	300	273	86.3	195	73.3	62.2
Metropolitan areas of 250,000 to 999,999 population	72	70	97.5	51	68.8	67.5
Metropolitan areas of less than 250,000 population	42	41	97.6	26	65.0	63.6
Nonmetropolitan areas adjacent to large metropolitan areas	4	4	100.0	4	100.0	100.0
Nonmetropolitan areas adjacent to medium or small metropolitan areas	33	32	97.0	23	72.0	69.7
Nonmetropolitan areas not adjacent to metropolitan areas	38	37	97.4	27	72.8	71.1

Source: NBS, round 3.

TABLE III.22

WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL
SAMPLE, PHASE 1 STATES, ROUNDS 1 AND 2 RESPONDENTS, TRADITIONAL PAYMENT SYSTEM,
BY SELECTED CHARACTERISTICS

	Sample	Located Sample	Response Among Located Sample		Overall Respondents	
	Count	Count	Location Rate	Count	Response Rate	
All	441	414	94.2	310	75.2	
SSI Only, SSDI Only, or Both SSI and SSDI						
SSI only	118	107	90.9	77	71.7	
SSDI only	214	204	95.7	149	73.8	
Both SSI and SSDI	109	103	94.6	84	81.4	
SSI or SSDI						
SSI only, or in both SSI & SSDI programs	227	210	92.7	161	76.4	
SSDI only, or in both SSI & SSDI programs	323	307	95.4	233	76.5	
Constructed Disability Status						
Deaf	21	20	95.8	13	65.1	
Mental	221	209	94.6	148	71.5	
Physical	196	183	94.0	148	80.7	
Unknown	3	2	59.0	1	62.4	
Beneficiary's Age (Four Categories)						
18-29 years	77	73	95.2	51	69.4	
30-39 years	93	84	91.4	65	77.3	
40-49 years	140	129	92.3	97	77.1	
50-64 years	131	128	97.5	97	75.3	
Sex						
Male	242	225	93.4	175	77.7	
Female	199	189	95.2	135	72.0	
Hispanicity						
Hispanic	24	22	91.8	20	92.2	
Non-Hispanic/unknown	417	392	94.3	290	74.2	
Race						
White	226	213	94.8	162	76.2	
Black	115	105	91.4	79	75.5	
Other/unknown	100	96	95.9	69	72.5	
Living Situation						
Living alone	222	207	93.5	160	76.8	
Living with others/unknown	219	207	94.9	150	73.5	
Did the Applicant for Benefits Live in Same Zip Code as Beneficiary?						
No	33	29	89.7	22	77.1	
Yes	224	211	94.1	167	79.5	
No information	184	174	94.9	121	69.6	

TABLE III.22 (continued)

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Identity of the Payee with Respect to the Beneficiary						
Beneficiary received beneficiary payments himself or herself	307	289	94.4	213	74.2	70.1
Payee is a family member	99	92	93.6	66	71.7	66.6
Payee is an institution	26	25	95.2	23	91.1	86.6
Other	9	8	89.4	8	100.0	89.4
Changes in Telephone Number						
No changes in last 5 years	316	299	94.8	228	76.8	72.7
One or more changes in last 5 years	6	6	100.0	5	83.9	84.1
No information/other	119	109	92.3	77	70.6	65.1
Number of Moves in Last 5 Years						
No moves last 5 years	137	128	93.5	102	80.3	75.2
One or more moves in last 5 years	6	6	100.0	6	100.0	100.0
No information on number of moves	298	280	94.3	202	72.3	68.1
Type of Claim						
Survivor	43	42	97.6	31	75.2	72.6
Disabled	286	272	95.5	206	76.2	72.6
Unknown	112	100	89.6	73	72.6	65.4
Census Region						
Midwest	166	158	95.2	126	79.9	75.9
Northeast	127	118	93.4	80	68.4	64.1
South	119	113	94.8	83	73.8	70.0
West	29	25	86.2	21	85.3	73.2
Census Division						
East North Central	160	152	95.0	121	79.8	75.6
East South Central	0	0	N/A	0	0	N/A
Middle Atlantic	116	108	93.5	72	67.5	63.3
Mountain	23	20	87.2	16	81.4	69.9
New England	11	10	91.7	8	81.7	75.1
Pacific	6	5	83.3	5	100.0	83.3
South Atlantic	117	111	94.7	81	73.3	69.5
West North Central	6	6	100.0	5	83.0	83.0
West South Central	2	2	100.0	2	100.0	100.0

TABLE III.22 (continued)

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Metropolitan						
Metropolitan areas of 1 million population or more	206	190	92.8	133	70.3	64.9
Metropolitan areas of 250,000 to 999,999 population	137	130	94.8	104	80.3	76.1
Metropolitan areas of less than 250,000 population	64	62	97.0	46	74.5	72.2
Nonmetropolitan areas adjacent to large metropolitan areas	0	0	N/A	0	0	N/A
Nonmetropolitan areas adjacent to medium or small metropolitan areas	19	17	89.4	13	75.3	68.6
Nonmetropolitan areas not adjacent to metropolitan areas	15	15	100.0	14	93.2	93.3

Source: NBS, round 3.

TABLE III.23

WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL
SAMPLE, PHASE 1 STATES, ROUNDS 1, 2, AND 3 RESPONDENTS, MILESTONE-OUTCOME PAYMENT SYSTEM,
BY SELECTED CHARACTERISTICS

	Sample	Located Sample	Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate
All	455	390	89.0	264	70.9
SSI Only, SSDI Only, or Both SSI and SSDI					
SSI only	166	139	85.7	95	72.0
SSDI only	197	176	92.0	111	66.0
Both SSI and SSDI	92	75	87.4	58	80.0
SSI or SSDI					
SSI only, or in both SSI & SSDI programs	258	214	86.4	153	75.2
SSDI only, or in both SSI & SSDI programs	289	251	90.5	169	70.4
Constructed Disability Status					
Deaf	6	5	84.6	4	80.7
Mental	257	219	89.0	155	73.9
Physical	179	157	90.2	98	66.2
Unknown	13	9	69.3	7	74.2
Beneficiary's Age (Four Categories)					
18-29 years	79	61	80.2	47	75.7
30-39 years	114	103	92.8	66	67.3
40-49 years	122	109	91.9	69	67.4
50-64 years	140	117	87.4	82	75.1
Sex					
Male	331	284	88.3	191	68.2
Female	124	106	90.5	73	77.6
Hispanicity					
Hispanic	19	11	59.0	10	93.7
Non-Hispanic/unknown	436	379	89.9	254	70.4
Race					
White	131	116	92.6	76	73.6
Black	115	102	90.7	76	79.1
Other/unknown	209	172	85.0	112	64.3
Living Situation					
Living alone	264	222	87.6	155	73.6
Living with others/unknown	191	168	90.7	109	67.5
Did the Applicant for Benefits Live in Same Zip Code as Beneficiary?					
No	43	29	73.8	18	68.3
Yes	205	176	89.1	128	77.3
No information	207	185	91.9	118	65.3

TABLE III.23 (continued)

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Identity of the Payee with Respect to the Beneficiary						
Beneficiary received beneficiary payments himself or herself	351	303	89.1	203	70.0	63.0
Payee is a family member	80	68	89.7	49	75.4	67.7
Payee is an institution	18	14	82.5	8	62.1	48.7
Other	6	5	83.7	4	81.8	67.4
Changes in Telephone Number						
No changes in last 5 years	314	271	89.9	183	71.5	64.8
One or more changes in last 5 years	12	9	76.3	9	100.0	76.3
No information/other	129	110	87.3	72	66.6	58.2
Number of Moves in Last 5 Years						
No moves last 5 years	130	106	84.6	76	72.5	62.2
One or more moves in last 5 years	11	6	59.5	4	67.2	38.6
No information on number of moves	314	278	91.5	184	70.4	64.7
Type of Claim						
Survivor	36	31	90.3	22	75.5	68.2
Disabled	262	229	90.9	152	69.4	63.7
Unknown	157	130	84.6	90	72.7	61.7
Census Region						
Midwest	75	66	91.1	50	77.2	70.2
Northeast	124	104	86.4	64	65.7	57.1
South	148	130	92.0	88	72.0	67.0
West	108	90	84.3	62	69.0	58.0
Census Division						
East North Central	73	64	89.9	49	80.8	72.9
East South Central	0	0	N/A	0	0	N/A
Middle Atlantic	77	65	87.8	40	67.4	59.9
Mountain	91	76	84.8	52	68.5	57.9
New England	47	39	83.7	24	62.6	51.9
Pacific	17	14	82.4	10	71.8	58.8
South Atlantic	147	129	91.6	87	70.6	65.4
West North Central	2	2	100.0	1	49.8	50.0
West South Central	1	1	100.0	1	100.0	100.0
Metropolitan						
Metropolitan areas of 1 million population or more	404	341	85.1	229	67.5	57.5
Metropolitan areas of 250,000 to 999,999 population	36	34	96.3	23	77.6	74.8
Metropolitan areas of less than 250,000 population	11	11	100.0	8	73.2	73.1
Nonmetropolitan areas adjacent to large metropolitan areas	0	0	N/A	0	0	N/A
Nonmetropolitan areas adjacent to medium or small metropolitan areas	4	4	100.0	4	100.0	100.0
Nonmetropolitan areas not adjacent to metropolitan areas	0	0	N/A	0	0	N/A

Source: NBS, round 3.

TABLE III.24

WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL
SAMPLE, PHASE 1 STATES, ROUNDS 1, 2, AND 3 RESPONDENTS, OUTCOME-ONLY PAYMENT SYSTEM,
BY SELECTED CHARACTERISTICS

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
All	453	419	88.1	275	67.1	58.3
SSI Only, SSDI Only, or Both SSI and SSDI						
SSI only	82	72	82.5	47	71.2	57.5
SSDI only	299	283	90.7	187	66.2	59.6
Both SSI and SSDI	72	64	84.0	41	65.8	54.0
SSI or SSDI						
SSI only, or in both SSI & SSDI programs	154	136	83.2	88	68.8	56.0
SSDI only, or in both SSI & SSDI programs	371	347	89.5	228	66.2	58.5
Constructed Disability Status						
Deaf	4	3	77.2	1	28.8	22.8
Mental	234	217	86.2	132	66.5	56.1
Physical	204	188	90.3	136	69.6	62.8
Unknown	11	11	100.0	6	52.8	47.2
Beneficiary's Age (Four Categories)						
18-29 years	33	33	100.0	22	66.9	65.0
30-39 years	95	84	81.1	54	63.3	49.7
40-49 years	154	140	87.5	88	66.6	58.1
50-64 years	171	162	91.1	111	69.7	62.6
Sex						
Male	276	251	89.0	165	68.5	60.2
Female	177	168	86.6	110	64.5	55.1
Hispanicity						
Hispanic	13	12	90.5	5	43.9	39.9
Non-Hispanic/unknown	440	407	88.1	270	67.6	58.7
Race						
White	247	233	89.4	162	71.3	62.9
Black	57	50	81.1	33	73.9	58.5
Other/unknown	149	136	88.9	80	57.1	50.4
Living Situation						
Living alone	168	152	85.8	101	71.6	60.1
Living with others/unknown	285	267	89.6	174	64.2	57.2
Did the Applicant for Benefits Live in Same Zip Code as Beneficiary?						
No	36	32	87.3	18	55.8	48.9
Yes	175	160	86.8	111	69.9	59.8
No information	242	227	89.2	146	66.4	58.4

TABLE III.24 (continued)

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Identity of the Payee with Respect to the Beneficiary						
Beneficiary received beneficiary payments himself or herself	383	356	86.9	237	64.2	55.7
Payee is a family member	52	47	89.8	26	59.1	49.3
Payee is an institution	9	9	100.0	5	86.3	83.0
Other	9	7	94.3	7	100.0	94.3
Changes in Telephone Number						
No changes in last 5 years	393	369	88.6	239	64.8	56.7
One or more changes in last 5 years	9	9	100.0	9	100.0	100.0
No information/other	51	41	84.3	27	74.5	62.4
Number of Moves in Last 5 Years						
No moves last 5 years	108	97	82.6	62	60.8	49.4
One or more moves in last 5 years	7	7	100.0	5	67.8	65.9
No information on number of moves	338	315	89.5	208	68.9	60.8
Type of Claim						
Survivor	18	16	89.1	9	60.4	53.7
Disabled	360	336	89.2	220	65.8	58.0
Unknown	75	67	83.2	46	74.1	60.4
Census Region						
Midwest	89	79	87.2	53	69.4	59.3
Northeast	135	127	90.4	79	68.8	61.2
South	113	108	92.1	63	52.1	49.4
West	116	105	82.8	80	77.1	63.0
Census Division						
East North Central	77	67	84.5	46	71.8	59.5
East South Central	0	0	N/A	0	0	N/A
Middle Atlantic	49	44	88.7	21	68.6	59.0
Mountain	55	52	93.8	37	78.7	73.1
New England	86	83	91.6	58	68.9	62.7
Pacific	61	53	75.0	43	75.8	55.9
South Atlantic	109	104	91.8	61	52.2	49.4
West North Central	12	12	100.0	7	59.0	58.3
West South Central	4	4	100.0	2	48.5	50.0
Metropolitan						
Metropolitan areas of 1 million population or more	280	252	83.9	170	67.1	55.5
Metropolitan areas of 250,000 to 999,999 population	66	62	89.5	40	72.2	63.8
Metropolitan areas of less than 250,000 population	36	36	100.0	20	57.3	57.4
Nonmetropolitan areas adjacent to large metropolitan areas	4	4	100.0	3	75.2	75.0
Nonmetropolitan areas adjacent to medium or small metropolitan areas	31	30	96.8	19	63.4	61.3
Nonmetropolitan areas not adjacent to metropolitan areas	36	35	97.2	23	65.6	63.9

Source: NBS, round 3.

TABLE III.25

WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL
SAMPLE, PHASE 1 STATES, ROUNDS 1, 2, AND 3 RESPONDENTS, TRADITIONAL PAYMENT SYSTEM,
BY SELECTED CHARACTERISTICS

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
All	441	406	92.7	274	67.8	62.8
SSI Only, SSDI Only, or Both SSI and SSDI						
SSI only	118	105	89.3	71	67.2	60.1
SSDI only	214	199	94.0	128	65.6	61.4
Both SSI and SSDI	109	102	94.1	75	72.5	68.6
SSI or SSDI						
SSI only, or in both SSI & SSDI programs	227	207	91.6	146	69.8	64.2
SSDI only, or in both SSI & SSDI programs	323	301	94.0	203	67.9	63.9
Constructed Disability Status						
Deaf	21	20	95.8	10	49.6	47.5
Mental	221	205	93.1	133	64.9	60.6
Physical	196	179	92.5	130	73.3	67.5
Unknown	3	2	59.0	1	59.5	38.9
Beneficiary's Age (Four Categories)						
18-29 years	77	71	93.2	48	66.8	62.1
30-39 years	93	84	91.4	57	67.4	62.0
40-49 years	140	126	90.3	84	68.6	62.0
50-64 years	131	125	96.0	85	67.6	64.8
Sex						
Male	242	221	92.0	152	68.9	63.4
Female	199	185	93.6	122	66.3	62.1
Hispanicity						
Hispanic	24	21	87.4	18	85.2	74.2
Non-Hispanic/unknown	417	385	93.0	256	66.8	62.2
Race						
White	226	211	94.3	142	67.9	64.0
Black	115	102	89.1	72	70.4	63.0
Other/unknown	100	93	93.2	60	64.3	59.9
Living Situation						
Living alone	222	204	92.3	145	70.2	65.1
Living with others/unknown	219	202	93.1	129	65.2	60.5
Did the Applicant for Benefits Live in Same Zip Code as Beneficiary?						
No	33	29	89.7	20	70.0	63.0
Yes	224	207	92.7	148	71.5	66.6
No information	184	170	93.3	106	62.7	58.2

TABLE III.25 (continued)

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Identity of the Payee with Respect to the Beneficiary						
Beneficiary received beneficiary payments himself or herself	307	283	93.0	185	66.0	61.4
Payee is a family member	99	91	92.6	61	66.2	61.4
Payee is an institution	26	24	91.8	20	81.7	74.8
Other	9	8	89.4	8	100.0	89.4
Changes in Telephone Number						
No changes in last 5 years	316	293	93.4	200	68.6	64.1
One or more changes in last 5 years	6	5	83.6	3	60.7	51.4
No information/other	119	108	91.5	71	66.0	60.2
Number of Moves in Last 5 Years						
No moves last 5 years	137	126	92.3	93	73.8	68.5
One or more moves in last 5 years	6	5	83.7	4	77.9	67.5
No information on number of moves	298	275	93.1	177	64.8	60.2
Type of Claim						
Survivor	43	42	97.6	29	70.3	68.3
Disabled	286	266	93.9	178	67.3	63.2
Unknown	112	98	87.8	67	67.9	59.9
Census Region						
Midwest	166	157	94.6	113	71.8	68.0
Northeast	127	117	92.8	72	62.4	58.0
South	119	108	91.4	71	66.4	60.6
West	29	24	83.1	18	74.6	62.6
Census Division						
East North Central	160	151	94.3	109	72.0	68.1
East South Central	0	0	N/A	0	0	N/A
Middle Atlantic	116	108	93.5	65	61.4	57.5
Mountain	23	19	83.0	14	73.2	61.2
New England	11	9	83.4	7	78.2	65.1
Pacific	6	5	83.3	4	79.6	66.7
South Atlantic	117	106	91.2	69	65.7	59.9
West North Central	6	6	100.0	4	66.8	65.9
West South Central	2	2	100.0	2	100.0	100.0
Metropolitan						
Metropolitan areas of 1 million population or more	206	185	90.7	120	65.1	59.0
Metropolitan areas of 250,000 to 999,999 population	137	127	93.3	89	70.3	65.7
Metropolitan areas of less than 250,000 population	64	62	97.0	38	61.6	59.6
Nonmetropolitan areas adjacent to large metropolitan areas	0	0	N/A	0	0	N/A
Nonmetropolitan areas adjacent to medium or small metropolitan areas	19	17	89.4	13	76.5	68.6
Nonmetropolitan areas not adjacent to metropolitan areas	15	15	100.0	14	93.2	93.3

Source: NBS, round 3.

TABLE III.26

WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL
SAMPLE, PHASE 2 STATES, ROUNDS 2 AND 3 RESPONDENTS, MILESTONE-OUTCOME PAYMENT SYSTEM,
BY SELECTED CHARACTERISTICS

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
All	389	363	92.0	275	76.8	69.8
SSI Only, SSDI Only, or Both SSI and SSDI						
SSI only	147	132	85.9	99	67.3	56.4
SSDI only	153	146	96.1	114	84.8	80.6
Both SSI and SSDI	89	85	94.4	62	76.3	71.2
SSI or SSDI						
SSI only, or in both SSI & SSDI programs	236	217	88.9	161	70.6	61.6
SSDI only, or in both SSI & SSDI programs	242	231	95.5	176	82.1	77.6
Constructed Disability Status						
Deaf	16	12	91.8	4	29.1	25.0
Mental	227	206	87.9	154	77.0	66.0
Physical	138	137	98.4	111	83.5	82.3
Unknown	8	8	100.0	6	49.0	51.0
Beneficiary's Age (Four Categories)						
18-29 years	117	108	85.4	80	74.7	61.7
30-39 years	86	78	92.8	54	73.3	67.8
40-49 years	102	95	95.4	74	76.5	72.4
50-64 years	84	82	99.2	67	84.4	83.2
Sex						
Male	204	191	91.9	144	76.4	70.0
Female	185	172	92.1	131	77.3	69.7
Hispanicity						
Hispanic	4	3	83.8	3	100.0	83.8
Non-Hispanic/unknown	385	360	92.0	272	76.7	69.7
Race						
White	211	199	91.8	149	72.9	66.2
Black	141	130	92.3	97	79.6	72.2
Other/unknown	37	34	91.6	29	90.6	82.4
Living Situation						
Living alone	226	206	87.6	156	72.7	62.3
Living with others/unknown	163	157	97.5	119	81.8	79.0
Did the Applicant for Benefits Live in Same Zip Code as Beneficiary?						
No	38	32	85.7	24	75.1	63.4
Yes	202	189	90.2	143	74.2	65.7
No information	149	142	95.5	108	80.4	76.1

TABLE III.26 (continued)

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Identity of the Payee with Respect to the Beneficiary						
Beneficiary received beneficiary payments himself or herself	228	214	93.8	176	83.4	77.1
Payee is a family member	122	114	90.2	77	70.4	62.1
Payee is an institution	30	27	89.4	17	63.3	56.3
Other	9	8	86.1	5	74.3	63.5
Changes in Telephone Number						
No changes in last 5 years	263	250	91.7	200	83.2	75.4
One or more changes in last 5 years	12	9	84.8	7	79.0	64.3
No information/other	114	104	93.4	68	60.1	55.9
Number of Moves in Last 5 Years						
No moves last 5 years	123	112	83.6	93	81.6	66.6
One or more moves in last 5 years	11	9	85.5	6	64.4	48.6
No information on number of moves	255	242	96.0	176	75.1	72.0
Type of Claim						
Survivor	51	49	98.5	33	73.5	72.6
Disabled	201	192	95.1	151	84.0	78.9
Unknown	137	122	85.0	91	66.0	54.5
Census Region						
Midwest	131	118	91.2	91	79.7	72.5
Northeast	22	22	100.0	17	79.4	77.9
South	226	213	91.6	159	73.7	66.0
West	10	10	100.0	8	79.9	80.0
Census Division						
East North Central	109	98	92.1	78	81.5	75.3
East South Central	40	40	100.0	30	80.2	78.5
Middle Atlantic	10	10	100.0	7	71.2	70.5
Mountain	9	9	100.0	7	77.7	77.8
New England	12	12	100.0	10	87.9	86.8
Pacific	1	1	100.0	1	100.0	100.0
South Atlantic	24	24	100.0	20	82.1	82.3
West North Central	22	20	86.9	13	69.5	58.6
West South Central	162	149	85.6	109	68.5	56.1
Metropolitan						
Metropolitan areas of 1 million population or more	250	231	89.6	181	76.4	67.5
Metropolitan areas of 250,000 to 999,999 population	37	36	98.0	25	69.8	66.7
Metropolitan areas of less than 250,000 population	40	39	98.0	28	81.6	80.0
Nonmetropolitan areas adjacent to large metropolitan areas	32	28	86.2	22	81.3	69.5
Nonmetropolitan areas adjacent to medium or small metropolitan areas	14	14	100.0	7	51.6	50.0
Nonmetropolitan areas not adjacent to metropolitan areas	16	15	94.1	12	83.2	76.4

Source: NBS, round 3.

TABLE III.27

WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL
SAMPLE, PHASE 2 STATES, ROUNDS 2 AND 3 RESPONDENTS, OUTCOME-ONLY PAYMENT SYSTEM,
BY SELECTED CHARACTERISTICS

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
All	414	396	93.6	265	66.9	62.4
SSI Only, SSDI Only, or Both SSI and SSDI						
SSI only	62	58	94.6	39	75.6	70.1
SSDI only	296	285	96.5	186	63.0	61.0
Both SSI and SSDI	56	53	79.2	40	75.9	59.6
SSI or SSDI						
SSI only, or in both SSI & SSDI programs	118	111	87.4	79	75.7	65.2
SSDI only, or in both SSI & SSDI programs	352	338	93.4	226	65.0	60.8
Constructed Disability Status						
Deaf	11	11	100.0	6	55.9	54.8
Mental	210	201	96.3	138	66.1	63.8
Physical	191	182	90.3	120	68.5	61.3
Unknown	2	2	100.0	1	50.4	50.0
Beneficiary's Age (Four Categories)						
18-29 years	45	43	95.5	27	64.0	60.4
30-39 years	101	96	94.9	58	61.3	57.6
40-49 years	158	151	89.7	101	71.8	63.8
50-64 years	110	106	98.0	79	65.4	65.0
Sex						
Male	218	209	95.5	130	62.6	59.8
Female	196	187	91.7	135	71.1	64.8
Hispanicity						
Hispanic	9	9	100.0	7	74.0	73.7
Non-Hispanic/unknown	405	387	93.5	258	66.7	62.2
Race						
White	278	267	96.3	176	65.6	63.4
Black	83	80	96.2	51	63.7	60.8
Other/unknown	53	49	75.8	38	79.7	59.1
Living Situation						
Living alone	131	124	88.4	87	74.6	65.0
Living with others/unknown	283	272	96.4	178	63.0	60.9
Did the Applicant for Benefits Live in Same Zip Code as Beneficiary?						
No	33	32	98.9	24	70.1	70.4
Yes	150	143	89.2	100	71.4	63.2
No information	231	221	95.6	141	63.9	60.8

TABLE III.27 (continued)

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Identity of the Payee with Respect to the Beneficiary						
Beneficiary received beneficiary payments himself or herself	334	320	93.2	215	67.5	62.7
Payee is a family member	56	54	96.3	33	58.2	55.3
Payee is an institution	13	12	96.5	11	95.8	93.2
Other	11	10	89.2	6	57.4	52.7
Changes in Telephone Number						
No changes in last 5 years	341	328	93.4	216	64.3	60.0
One or more changes in last 5 years	5	4	80.0	4	100.0	80.0
No information/other	68	64	95.4	45	76.8	72.5
Number of Moves in Last 5 Years						
No moves last 5 years	89	87	98.4	59	69.3	68.0
One or more moves in last 5 years	5	4	80.0	4	100.0	80.0
No information on number of moves	320	305	92.6	202	65.8	60.7
Type of Claim						
Survivor	32	31	96.7	22	71.5	68.8
Disabled	324	310	92.8	205	64.1	59.6
Unknown	58	55	95.7	38	77.3	72.7
Census Region						
Midwest	77	71	86.6	53	73.3	62.4
Northeast	229	220	95.2	141	63.1	60.4
South	94	92	98.6	62	67.1	66.4
West	14	13	92.9	9	69.8	64.3
Census Division						
East North Central	40	37	95.7	25	52.6	46.2
East South Central	37	37	100.0	24	64.8	64.9
Middle Atlantic	53	48	90.8	30	61.3	55.3
Mountain	11	10	90.9	7	70.4	63.6
New England	176	172	96.5	111	63.6	62.0
Pacific	3	3	100.0	2	67.6	66.7
South Atlantic	31	30	96.9	20	65.7	63.8
West North Central	37	34	80.9	28	90.3	72.3
West South Central	26	25	98.7	18	72.3	71.6
Metropolitan						
Metropolitan areas of 1 million population or more	153	144	95.2	101	70.5	66.6
Metropolitan areas of 250,000 to 999,999 population	159	156	96.8	98	62.3	60.5
Metropolitan areas of less than 250,000 population	31	29	93.5	21	72.6	67.7
Nonmetropolitan areas adjacent to large metropolitan areas	13	11	73.1	8	92.2	66.6
Nonmetropolitan areas adjacent to medium or small metropolitan areas	37	36	97.3	24	66.9	64.9
Nonmetropolitan areas not adjacent to metropolitan areas	21	20	97.2	13	40.3	40.1

Source: NBS, round 3.

TABLE III.28

WEIGHTED LOCATION AND RESPONSE RATES FOR TICKET PARTICIPANT LONGITUDINAL
SAMPLE, PHASE 2 STATES, ROUNDS 2 AND 3 RESPONDENTS, TRADITIONAL PAYMENT SYSTEM,
BY SELECTED CHARACTERISTICS

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
All	437	397	91.2	316	79.5	72.6
SSI Only, SSDI Only, or both SSI and SSDI						
SSI only	107	101	94.9	74	72.2	68.3
SSDI only	221	198	89.8	160	81.1	72.9
Both SSI and SSDI	109	98	90.6	82	83.4	76.0
SSI or SSDI						
SSI only, or in both SSI & SSDI programs	216	199	92.7	156	77.8	72.2
SSDI only, or in both SSI & SSDI programs	330	296	90.0	242	81.9	73.9
Constructed Disability Status						
Deaf	18	14	81.9	9	63.7	52.2
Mental	241	221	91.7	179	80.4	73.8
Physical	171	155	91.2	125	81.3	74.1
Unknown	7	7	100.0	3	40.6	43.4
Beneficiary's Age (Four Categories)						
18-29 years	123	114	93.4	89	78.0	72.6
30-39 years	96	82	85.8	58	70.0	60.4
40-49 years	123	114	92.7	95	83.8	77.6
50-64 years	95	87	92.1	74	85.2	78.5
Sex						
Male	219	197	90.3	154	78.6	70.8
Female	218	200	92.1	162	80.3	74.3
Hispanicity						
Hispanic	4	3	73.8	1	42.4	27.0
Non-Hispanic/unknown	433	394	91.4	315	79.8	73.0
Race						
White	286	265	93.0	211	79.6	74.1
Black	99	87	88.2	69	78.4	69.1
Other/unknown	52	45	87.0	36	80.7	70.4
Living Situation						
Living alone	218	201	92.7	160	79.2	73.6
Living with others/unknown	219	196	89.7	156	79.8	71.5
Did the Applicant for Benefits Live in Same Zip Code as Beneficiary?						
No	41	37	90.8	27	73.4	66.7
Yes	222	200	90.6	162	80.7	73.0
No information	174	160	92.0	127	79.3	73.3

TABLE III.28 (continued)

	Sample	Located Sample		Response Among Located Sample		Overall Respondents
	Count	Count	Location Rate	Count	Response Rate	Response Rate
Identity of the Payee with Respect to the Beneficiary						
Beneficiary received beneficiary payments himself or herself	284	255	90.1	203	79.9	72.0
Payee is a family member	114	107	94.4	88	81.8	77.2
Payee is an institution	25	23	92.8	20	86.7	80.7
Other	14	12	86.3	5	39.6	34.6
Changes in Telephone Number						
No changes in last 5 years	316	283	89.9	227	80.2	72.4
One or more changes in last 5 years	8	8	100.0	6	74.3	74.0
No information/other	113	106	94.2	83	77.7	72.9
Number of Moves in Last 5 Years						
No moves last 5 years	124	109	88.8	87	80.1	71.3
One or more moves in last 5 years	7	7	100.0	7	100.0	100.0
No information on number of moves	306	281	92.0	222	78.7	72.4
Type of Claim						
Survivor	48	44	91.7	36	83.2	76.0
Disabled	288	258	90.0	209	80.9	73.1
Unknown	101	95	94.6	71	73.6	69.4
Census Region						
Midwest	217	200	92.3	154	76.8	71.2
Northeast	24	22	91.0	18	83.4	76.1
South	161	143	89.5	117	81.5	73.0
West	35	32	91.4	27	85.0	77.1
Census Division						
East North Central	202	185	91.6	141	75.9	69.9
East South Central	42	33	81.2	25	74.3	59.7
Middle Atlantic	9	9	100.0	6	66.4	66.7
Mountain	35	32	91.4	27	85.0	77.1
New England	15	13	87.1	12	91.6	80.3
Pacific	0	0	N/A	0	0	N/A
South Atlantic	72	67	93.2	57	84.8	79.1
West North Central	15	15	100.0	13	86.3	85.9
West South Central	47	43	91.0	35	82.3	74.8
Metropolitan						
Metropolitan areas of 1 million population or more	194	180	92.7	134	74.2	69.1
Metropolitan areas of 250,000 to 999,999 population	55	49	89.2	38	77.3	69.0
Metropolitan areas of less than 250,000 population	70	58	84.5	51	87.9	74.7
Nonmetropolitan areas adjacent to large metropolitan areas	44	42	96.1	36	85.2	81.6
Nonmetropolitan areas adjacent to medium or small metropolitan areas	12	11	91.9	9	78.4	71.9
Nonmetropolitan areas not adjacent to metropolitan areas	62	57	91.9	48	84.3	77.4

Source: NBS, round 3.

The models were developed with the same main effects used in the development of the nonresponse models for the beneficiary weights and Ticket participant cross-sectional weights, plus selected interactions. Interactions to be considered for inclusion in model development were identified using CHAID, as described previously.

The primary factors used in both the location and cooperation models for the longitudinal weights are given below, along with potential levels used in the models. (Details about how these levels were collapsed are given in Appendix E.)

1. **MOVE.** The number of address changes in the past five years; possible levels: (1) no moves, (2) one move, (3) two or more moves, and (4) information older than five years, or no information.
2. **DIG.** Disability diagnostic classification; possible levels: (1) mental disability, (2) physical disability (excluding deaf cases), (3) deaf, and (4) unknown.
3. **REPREPAYEE.** The identity of the payee with respect to the beneficiary; possible levels: (1) the beneficiary received payments himself or herself, (2) a family member received benefits on behalf of the beneficiary, and (3) an institution received payments on behalf of the beneficiary, or identity of payee not known
4. **PDZIPSAME.** Whether the beneficiary and the applicant for benefits lived in the same zip code; possible levels: (1) beneficiary and applicant lived in the same zip code, (2) beneficiary and applicant lived in different zip codes, and (3) information unknown.
5. **METRO.** Urbanicity of beneficiary's place of residence; possible levels: (1) beneficiary lived in metropolitan area of 1 million or more residents, (2) beneficiary lived in metropolitan area of 250,000 to 999,999, (3) beneficiary lived in metropolitan area of less than 250,000 residents, (4) beneficiary lived in nonmetropolitan area adjacent to a metropolitan area of 1 million or more, (5) beneficiary lived in nonmetropolitan area adjacent to a metropolitan area of less than 1 million, and (6) beneficiary lived in nonmetropolitan area not adjacent to any metropolitan area.
6. **GENDER (SEX).** Possible levels: (1) male, and (2) female.
7. **REGION or DIVISION.** Geographic region of beneficiary's place of residence: DIVISION is based on U.S. Census divisions with nine levels: (1) Pacific, (2) Mountain, (3) East North Central, (4) West North Central, (5) East South Central, (6) West South Central, (7) South Atlantic, (8) Middle Atlantic, and (9) New England. REGION is based on U.S. Census regions, with four levels, which can be collapsed from the nine levels of DIVISION: (1) West is Pacific and Mountain, (2) Midwest is East North Central and West North Central, (3) South is

East South Central, West South Central, and South Atlantic, and (4) Northeast is Middle Atlantic and New England.⁴²

8. **LIVING.** Beneficiary's living situation; possible levels: (1) beneficiary lives alone, (2) beneficiary lives with his or her parents, and (3) beneficiary does not live alone or with his or her parents, or information unknown.
9. **PHONE.** Number of phone numbers on SSA file over past five years; possible levels: (1) only one phone number on file, (2) one change in phone number on SSA file, (3) two or more changes in phone number on SSA file, and (4) information unknown.
10. **AGECAT.** Beneficiary's age category; possible levels; (1) age in range 18 to 29 years, (2) age in range 30 to 39 years, (3) age in range 40 to 49 years, and (4) age in range 50 to 64 years.
11. **SSI_SSDI.** Beneficiary status; possible levels: (1) SSI only, (2) SSDI only, or (3) both SSI and SSDI.
12. **TOC.** Type of claim; possible levels: (1) survivor claim, (2) disability claim, and (3) type of claim unknown
13. **RACE.** Race; possible levels: (1) white, (2) black, (3) Asian or Pacific Islander, and (4) not white, black, or Asian/Pacific Islander, or unknown.

Various interactions among these variables were also included in the model for locating the sample member. A list of the main effects using variable names listed above, as well as interactions, is provided in Tables III.29 – III.34. An expanded form of Tables III.29 – III.34, with the specific levels of the main effects for each model and the interactions shown in Tables III.29 – III.34, along with parameter estimates and their standard errors, is provided in Appendix E. In Appendix E, the variables are followed by suffixes representing the collapsing of the base variable's levels unique to each model. This follows the procedure used in the cross-sectional models. These suffixes are not shown in Tables III.29 – III.34.

⁴² Many of the models used REGION instead of DIVISION. If a U.S. Census Division was used in a model, then the U.S. Census Region corresponding to that Division could not be in the model.

TABLE III.29

VARIABLES USED IN THE LOCATION LOGISTIC PROPENSITY MODELS: TICKET PARTICIPANT
LONGITUDINAL SAMPLE, PHASE 1 STATES, ROUNDS 1 AND 2 RESPONDENTS

Variables in the Location Model for Participants Using Milestone-Outcome Payment System

Main Effects

METRO
AGECAT
PDZIPSAME
TOC
REGION

Two-Factor Interactions

REGION*TOC

Variables in the Location Model for Participants Using Outcome-Only Payment System

Main Effects

AGECAT
GENDER (SEX)
SSI_SSDI
DIVISION
REPREPAYEE
LIVING
METRO

Two-Factor Interactions

DIVISION*AGECAT

Variables in the Location Model for Participants Using Traditional Payment System

Main Effects

GENDER (SEX)
RACE
TOC
DIG
REGION
SSI_SSDI
AGECAT
REPREPAYEE

Two-Factor Interactions

SEX * TOC
SSI_SSDI* REPREPAYEE
SSI_SSDI* DIG
RACE* TOC

TABLE III.30

VARIABLES USED IN THE LOCATION LOGISTIC PROPENSITY MODELS: TICKET PARTICIPANT
LONGITUDINAL SAMPLE, PHASE 1 STATES, ROUNDS 1, 2, AND 3 RESPONDENTS

Variables in the Location Model for Participants Using Milestone-Outcome Payment System

Main Effects

MOVE
PDZIPSAME
METRO
RACE
AGECAT

Variables in the Location Model for Participants Using Outcome-Only Payment System

Main Effects

MOVE
LIVING
RACE
AGECAT
DIVISION
SSI_SSDI
METRO

Two-Factor Interactions

MOVE*SSI_SSDI
DIVISION*RACE
SSI_SSDI*RACE
AGECAT*DIVISION

Variables in the Location Model for Participants Using Traditional Payment System

Main Effects

GENDER (SEX)
SSI_SSDI
REGION
RACE
AGECAT

Two-Factor Interactions

RACE*SEX

TABLE III.31

VARIABLES USED IN THE LOCATION LOGISTIC PROPENSITY MODELS: TICKET PARTICIPANT
LONGITUDINAL SAMPLE, PHASE 2 STATES, ROUNDS 2 AND 3 RESPONDENTS

Variables in the Location Model for Participants Using Milestone-Outcome Payment System

Main Effects

AGECAT
TOC
GENDER (SEX)
REGION
DIG
PHONE
MOVE
REPREPAYEE
SSI_SSDI
LIVING

Two-Factor Interactions

MOVE*AGECAT
SEX*AGECAT
MOVE*REGION
DIG*REPREPAYEE

Variables in the Location Model for Participants Using Outcome-Only Payment System

Main Effects

AGECAT
GENDER (SEX)
REGION
MOVE
RACE

Two-Factor Interactions

RACE*REGION
REGION*SEX

Variables in the Location Model for Participants Using Traditional Payment System

Main Effects

DIVISION
RACE
PHONE
METRO
AGECAT
REPREPAYEE

Two-Factor Interactions

REPREPAYEE*PHONE
AGECAT*PHONE
METRO*AGECAT

TABLE III.32

VARIABLES USED IN THE COOPERATION LOGISTIC PROPENSITY MODELS: TICKET PARTICIPANT
LONGITUDINAL SAMPLE, PHASE 1 STATES, ROUNDS 1 AND 2 RESPONDENTS

Variables in the Cooperation Model for Participants Using Milestone-Outcome Payment System
<p>Main Effects PHONE RACE PDZIPSAME DIG</p> <p>Two-Factor Interactions PDZIPSAME* DIG</p>
Variables in the Cooperation Model Participants Using Outcome-Only Payment System
<p>Main Effects GENDER (SEX) MOVE REGION DIG REPREPAYEE LIVING SSI_SSDI RACE PDZIPSAME</p> <p>Two-Factor Interactions REGION *RACE SEX*REGION MOVE*RACE DIG*REGION</p>
Variables in the Cooperation Model for Participants Using Traditional Payment Systems
<p>Main Effects PDZIPSAME METRO REPREPAYEE DIG REGION SSI_SSDI</p> <p>Two-Factor Interactions SSI_SSDI*DIG</p>

TABLE III.33

VARIABLES USED IN THE COOPERATION LOGISTIC PROPENSITY MODELS: TICKET PARTICIPANT
LONGITUDINAL SAMPLE, PHASE 1 STATES, ROUNDS 1, 2, AND 3 RESPONDENTS

Variables in the Cooperation Model for Participants Using Milestone-Outcome Payment System
<p>Main Effects MOVE PDZIPSAME SSI_SSDI METRO RACE AGECAT DIG</p> <p>Two-Factor Interactions SSI_SSDI*DIG AGECAT*PDZIPSAME</p>
Variables in the Cooperation Model for Participants Using Outcome-Only Payment System
<p>Main Effects MOVE RACE AGECAT PDZIPSAME SSI_SSDI REGION DIG LIVING</p> <p>Two-Factor Interactions MOVE*REGION DIG*AGECAT</p>
Variables in the Cooperation Model for Participants Using Traditional Payment System
<p>Main Effects DIG PDZIPSAME AGECAT METRO REGION REPREPAYEE RACE</p> <p>Two-Factor Interactions REGION* PDZIPSAME DIG*AGECAT REGION*RACE</p>

TABLE III.34

VARIABLES USED IN THE COOPERATION LOGISTIC PROPENSITY MODELS: TICKET PARTICIPANT
LONGITUDINAL SAMPLE, PHASE 2 STATES, ROUNDS 2 AND 3 RESPONDENTS

Variables in the Cooperation Model for Participants Using Milestone-Outcome Payment System

Main Effects

AGECAT
REPREPAYEE
DIG
SSI_SSDI
PHONE
RACE
MOVE
TOC

Two-Factor Interactions

AGECAT*RACE
AGECAT*PHONE
REPREPAYEE*RACE
RACE*TOC
SSI_SSDI*REPREPAYEE
SSI_SSDI*MOVE
REPREPAYEE*TOC

Variables in the Cooperation Model for Participants Using Outcome-Only Payment System

Main Effects

GENDER (SEX)
METRO
DIVISION
DIG
PDZIPSAME

Two-Factor Interactions

DIG* PDZIPSAME

Variables in the Cooperation Model for Participants Using Traditional Payment Systems

Main Effects

GENDER (SEX)
PDZIPSAME
REPREPAYEE
DIG
REGION
SSI_SSDI
AGECAT
METRO

Two-Factor Interactions

SEX*PDZIPSAME
SSI_SSDI*REPREPAYEE
SSI_SSDI*DIG
SSI_SSDI*REGION
REGION*AGECAT

The process for identifying the final model was exactly the same as for the Representative Beneficiary Sample (described in Section B.2) and the Ticket Participant Cross-Sectional Sample (described in Section C.3). The R-squared values for the 18 logistic models are given in Table III.35. Overall, the unadjusted R-squared values ranged from a low of 0.039 (0.097 when rescaled to have a maximum of 1) to a high of 0.211 (0.449 when rescaled to have a maximum of 1). However, the models tend to be better with the Phase 2 than the Phase 1 longitudinal sample. Among Phase 1 cases, the models are better when looking only at rounds 1 and 2 respondents rather than rounds 1, 2, and 3 respondents. These values are similar to those observed for other response propensity efforts using logistic regression and design-based sampling weights.

TABLE III.35

UNADJUSTED AND ADJUSTED R-SQUARED VALUES FOR LOGISTIC PROPENSITY MODELS IN THE TICKET PARTICIPANT LONGITUDINAL SAMPLE

Model				
Respondent in Round	Payment Type	Location or Cooperation	Unadjusted R-Squared Value	Adjusted R-Squared Value
1 and 2	Milestone-Outcome	Location	0.067	0.142
1 and 2	Milestone-Outcome	Cooperation	0.077	0.116
1 and 2	Outcome-Only	Location	0.141	0.309
1 and 2	Outcome-Only	Cooperation	0.192	0.277
1 and 2	Traditional	Location	0.082	0.227
1 and 2	Traditional	Cooperation	0.078	0.116
1, 2, and 3	Milestone-Outcome	Location	0.078	0.156
1, 2, and 3	Milestone-Outcome	Cooperation	0.100	0.143
1, 2, and 3	Outcome-Only	Location	0.141	0.271
1, 2, and 3	Outcome-Only	Cooperation	0.138	0.192
1, 2, and 3	Traditional	Location	0.039	0.097
1, 2, and 3	Traditional	Cooperation	0.091	0.127
2 and 3	Milestone-Outcome	Location	0.192	0.449
2 and 3	Milestone-Outcome	Cooperation	0.211	0.319
2 and 3	Outcome-Only	Location	0.146	0.386
2 and 3	Outcome-Only	Cooperation	0.089	0.124
2 and 3	Traditional	Location	0.063	0.140
2 and 3	Traditional	Cooperation	0.155	0.243

The percentages of concordant and discordant pairs, and the p-values for the Hosmer-Lemeshow goodness-of-fit test, are given in Table III.36. Although the minimum difference between the percentages of concordant pairs and discordant pairs is 20.3 percent (Phase 2 cooperation model for participants using the outcome-only payment system), for the remainder of models the difference between these percentages is at least 26.5 percent, and all but 3 have differences greater than 30 percent. The minimum p-value associated with the Hosmer-Lemeshow goodness-of-fit test is 0.209, indicating no evidence of lack of fit for any of the models.

TABLE III.36

PERCENTAGES OF CONCORDANT AND DISCORDANT PAIRS AND HOSMER-LEMESHOW P-VALUES FOR LOGISTIC PROPENSITY MODELS IN THE TICKET PARTICIPANT LONGITUDINAL SAMPLE

Respondent in Round	Model		Percent Concordant	Percent Discordant	Hosmer- Lemeshow p- value
	Payment Type	Location or Cooperation			
1 and 2	Milestone-Outcome	Location	58.6	26.6	0.948
1 and 2	Milestone-Outcome	Cooperation	60.9	33.5	0.983
1 and 2	Outcome-Only	Location	78.7	18.7	0.631
1 and 2	Outcome-Only	Cooperation	74.1	25.0	0.425
1 and 2	Traditional	Location	77.2	20.5	0.938
1 and 2	Traditional	Cooperation	65.8	29.1	0.570
1, 2, and 3	Milestone-Outcome	Location	60.9	29.7	0.850
1, 2, and 3	Milestone-Outcome	Cooperation	61.9	35.4	0.575
1, 2, and 3	Outcome-Only	Location	68.4	28.7	0.629
1, 2, and 3	Outcome-Only	Cooperation	64.8	33.5	0.247
1, 2, and 3	Traditional	Location	65.1	28.3	0.714
1, 2, and 3	Traditional	Cooperation	68.2	29.9	0.217
2 and 3	Milestone-Outcome	Location	74.9	22.6	0.209
2 and 3	Milestone-Outcome	Cooperation	70.1	28.7	0.268
2 and 3	Outcome-Only	Location	63.0	31.4	0.314
2 and 3	Outcome-Only	Cooperation	55.7	35.4	0.496
2 and 3	Traditional	Location	71.8	22.8	0.945
2 and 3	Traditional	Cooperation	76.9	22.4	0.839

4. Trimming

After adjustments were applied to the longitudinal weights, the distribution of weights was reviewed to determine if trimming of the longitudinal weights was necessary. Because of the wide variation in the magnitude of the weights due to the use of the composite weights in some situations, trimming was sometimes necessary to increase the precision of survey estimates. However, we minimized the extent of trimming to reduce the potential for bias in the survey estimates. The design effects due to unequal weights associated with each type of longitudinal weight and payment type before and after trimming, before post-stratification, are presented in Table III.37. Design effects were calculated separately within trimming strata, which in turn were defined within longitudinal weight type and payment type. In general, weights in the clustered sample were larger than those in the unclustered sample. Therefore, the trimming strata were defined according to whether the observation was in the clustered or unclustered sample. For unclustered cases, the trimming strata were further subdivided according to whether the sample case was in a PSU or not. The strata within which trimming was employed are given in Table III.37. As with the trimming of cross-sectional weights, if no trimming was employed for a longitudinal weight type and payment type, we would have presented the maximum design effect due to unequal weights across all trimming strata. However, it always turned out that only one trimming stratum was used in these instances. In light of this, if no trimming was employed, the design effect due to unequal weighting within that single trimming stratum was presented.

TABLE III.37

DESIGN EFFECTS DUE TO UNEQUAL WEIGHTS BEFORE AND AFTER TRIMMING, WITHIN TRIMMING STRATA, FOR EACH TYPE OF LONGITUDINAL WEIGHT AND PAYMENT TYPE

Longitudinal Weight Type and Payment Type		Trimming Stratum Where Trimming Occurred	Design Effect due to Unequal Weights	
Respondent in Round	Payment Type		Before Trimming	After Trimming
1 and 2	Milestone-Outcome	No Trimming	1.66	1.66
1 and 2	Outcome-Only	Clustered Sample	3.37	2.67
1 and 2	Traditional	No Trimming	1.06	1.06
1, 2, and 3	Milestone-Outcome	No Trimming	1.54	1.54
1, 2, and 3	Outcome-Only	Clustered Sample	3.80	3.12
1, 2, and 3	Traditional	No Trimming	1.10	1.10
2 and 3	Milestone-Outcome	Clustered Sample	4.16	2.88
2 and 3	Outcome-Only	Clustered Sample	3.69	2.88
2 and 3	Traditional	No Trimming	1.15	1.15

5. Post-Stratification

After the nonresponse adjustment and trimming, the weights were post-stratified to the population age and gender totals for each payment type obtained from the SSA sampling frame as it was defined when the phase was first rolled out. This sampling frame included all SSI or SSDI beneficiaries for each payment type within the Ticket Participant population. For Phase 1 cases, the weight totals were defined by the round 1 sampling frame, which included only data from Phase 1 states. For Phase 2 cases, the weight totals were defined by the round 2 sampling frame for Phase 2 states only. The distributions of weights within each phase and payment type combination were rechecked to determine if more weight trimming was necessary. No additional weight trimming was required.

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

IV. IMPUTATIONS

In the NBS, the data collection instruments were administered using computer-assisted interviewing (CAI) technology. The CAI technology allows the use of automated routing to move the respondent to the applicable questions, and also implements checks of the entered data for consistency and reasonableness. In addition, because the program will not allow a question to be left blank, the interviewer cannot proceed unless an appropriate response has been entered (“don’t know” and “refused” are included as response options and used as necessary). These processes substantially reduce the extent of item nonresponse for a complex survey, but some item nonresponse will still exist. Item nonresponse includes cases where the question was mistakenly not asked and cases where “don’t know” or “refused” were recorded as responses.

For the NBS, imputation was used to compensate for item nonresponse. Two imputation methods were primarily used: deductive (or logical) imputation and unweighted hot-deck imputation. However, for some variables, insufficient data were available to use either of these two methods, so other specialized imputation procedures were employed to use the data available. The methods were selected based on the type of variable (dichotomous, categorical, or continuous), the amount of missing data, and the availability of data for the imputations. For some variables, imputations were processed using a combination of methods.

Where appropriate, imputed values were made consistent with pre-existing nonmissing variables by excluding donors with potentially inconsistent imputed values. After each imputation was processed, the imputed values were evaluated using a variety of quality control procedures. If the initial imputed value was out of an acceptable range or inconsistent with other data for that case, the imputation was repeated until the imputed value was in range and consistent with other reported data.

Deductive, or logical, imputation is the assignment of a value that can be deduced from other data, or for which there is a high degree of certainty that the value is correct. This method was based on a review of data related to the imputed variable.

The hot-deck imputation procedure entails the classification of sample members into mutually exclusive and exhaustive imputation classes (or imputation cells) of respondents who are assumed to be similar relative to the key population variables (such as age, disability status, and SSI recipient status). For each sample member with a missing value (a recipient), a sample member with complete data (a donor) is chosen within the same imputation class to provide a value. It is desirable to have the imputation class contain sufficient sample members to avoid the selection of a single donor for multiple sample members with missing data. The hot-deck procedure is computationally efficient and, in a National Center for Education Statistics working paper (USDE 2001), a simulation study showed that a hot-deck procedure fared well in comparison to more sophisticated imputation procedures, including multiple imputation, Bayesian bootstrap imputation, and ratio imputation. However, it should be noted that no attempt was made to estimate the component of variance due to imputation, even though such a component is always positive. Users should be aware that variance estimates using imputed data will be underestimates, with the amount of bias in the variance estimate directly related to the amount of missingness in the variable of interest. For most of the variables requiring imputation, the extent of missingness was low, so this component would be very small in these cases.

The hot-deck imputation procedure used an unweighted selection process to select a donor, with selections done within imputation classes defined by key related variables for each application. In addition to the variables defining the imputation classes, a sorting variable was included where the recipient and all donors within the imputation class were sorted together by the levels of this variable. Using the sorted data within the imputation class, a case immediately

preceding or following a sample member with missing data was randomly selected as the donor with equal probability. The hot-deck procedure was therefore unweighted and sequential, with a random component. We allowed with-replacement selection of a donor for each recipient. In other words, a sample member could have been a donor for more than one recipient. Because the extent of missing values was very low, only a few donors were used more than once.

The factors used to form the cells for each imputed variable needed to be appropriate for the population, the data collected, and the purpose of the study. The imputation classes also needed to possess a sufficient count of donors for each sample member with missing data. We used a variety of methods to form the imputation classes. These methods included bivariate cross-tabulations, step-wise regressions, and multivariate procedures such as CHAID. (Chi-squared Automatic Interaction Detection software is attributed to Kass [1980] and Biggs et al. [1991], and its application in SPSS is described in Magidson [1993].) To develop these imputation classes, we used information from both the interview and SSA data files. Classing and sorting variables were closely related to the variable being imputed (the response variable). Sorting variables were either less closely related to the response variable than classing variables, or were forms of the classing variables with finer levels. As an example of the latter situation, four age categories were sometimes used as imputation classes: (1) 18 to 29, (2) 30 to 39, (3) 40 to 49, and (4) 50 to 64. The actual age could then be used as a sorting variable, so that donors and recipients were as close together in age as possible.

If any missing values existed in variables used to define imputation classes, two different strategies were employed: (1) match recipients to donors who were also missing the value for the covariate; or (2) employ separate hot decks, depending upon the availability of the variables defining the imputation classes. In the first instance, the level defined as the missing value was treated as a separate level. In other words, if a recipient was missing a value for a variable

defining an imputation class, the donor also was missing the value for that variable. This strategy was employed if there were large numbers of donors and recipients missing the covariate in question. In the second instance, for a given recipient, a variable was used to define the imputation class for that recipient only if there was no missing value for that variable. The variables used to define an imputation class for each recipient would depend upon what values were nonmissing among those variables.

The hot-deck software automatically identified situations in which the imputation class contained only recipients and no donors. In these cases, imputation classes were collapsed, and the imputation was redone using the collapsed classes. The strategy for collapsing classes required a ranking of the variables used to define the imputation class with regard to each variable's relationship to the variable requiring imputation. If a number of covariates were used to aid in the imputation of a given variable, the covariates less closely related to the variable requiring imputation were more likely to have levels collapsed than more important covariates in the imputation. In addition, variables with many levels also were more likely to have levels collapsed. In general, if more than a very small number of imputation classes required collapsing, then one or more variables were dropped from the definition of the imputation class, and the imputation procedure was rerun.

Some variables were constructed from two or more variables. For some of the "constructed" variables, it was more efficient to impute the component variables, and then impose the recoding of the constructed variable on these imputed values. These component variables are not shown in the following tables because they were not included in the final data set.

For some of the imputed variables in the data set, the number of missing responses does not match the number of imputed responses. Often, these variables correspond to questions that follow a filter question. For example, question I33 asks if the respondent has difficulty climbing

10 steps, and if the response is “yes,” the follow-up question (I34) asks if the respondent is able to climb 10 steps at all. To be asked the follow-up question, the respondent must have answered “yes” to the screener question. If the respondent answered “no,” the follow-up question was coded a legitimate missing (“.I”), which was not imputed. However, if the respondent refused to answer the screener question, the follow-up question was also coded a legitimate missing. If the screener variable was then imputed to be “yes,” the response to the follow-up question was imputed. This caused the count of the actual number of imputed responses to be greater than the number of missing or invalid responses.

A. NBS IMPUTATIONS OF SPECIFIC VARIABLES

Included below in several tables is information about how imputation was employed in the NBS. The tables include the imputed variable names and a brief description of each. The tables also include the methods of imputation, total number of missing responses, the number of respondents eligible for the question, and the percentage of responses imputed. This information was recorded in the final file with an imputation flag, identified by the suffix “iflag,” which has the following nine levels: (.) legitimate missing or no answer, (0) self-reported data, (1) logical imputation, (2) administrative data, (3) hot-deck imputed, (4) imputation using the distribution of a variable related to the variable being imputed, (5) imputation based on specialized procedures specific to Section K, (6) constructed from other variables with imputed values, and (7) longitudinal imputation (using data from a previous round). In most cases, the logical assignments were done using imputed values.⁴³ Therefore, the distinction between “logically assigned” and “constructed from other variables with imputed values” is somewhat opaque. In general, if a logical assignment was done for variables corresponding directly to questionnaire

⁴³ No distinction was made between logical assignments using imputed values and those using self-reported values.

questions, the flag was set to 1. For variables constructed from these variables (constructed variables are prefixed with a “C_”), the flag was set to 6. In this instance, one or more of the component variables in the constructed variable were imputed.

In the sections that follow, summaries are given of the imputations conducted, organized by the sections within the questionnaire to which the variables correspond. Details of some of the imputation types are given for each section.

1. Section L: Race and Ethnicity

Several questions included in the NBS instrument gathered information on respondents’ race and ethnicity. Two of these variables, located in Section L, included imputed responses and are described in Table IV.1. In particular, L1_i corresponds to the question asking whether the respondent is Hispanic or not; C_Race_i corresponds to the question asking about the respondent’s race.

In this table, respondents who did not indicate in the questionnaire whether they were Hispanic were classified as such if the SSA administrative data so indicated; the single logical imputation was conducted by looking at the name of the respondent and comparing it to a list of Hispanic names provided by the North American Association of Central Cancer Registries (NAACCR 2003). For respondents who still had missing data, the Hispanic indicator was imputed using a hot deck with imputation classes defined by the zip code of each sample member, and race as a sorting variable.

TABLE IV.1
RACE AND ETHNICITY IMPUTATIONS

Variable Name	Description	Imputation Method	Number Missing	Number Eligible	Percent Imputed
L1_i	Hispanic/Latino Ethnic Origins	3 imputations from SSA's administrative data, 2 longitudinal imputations, 57 imputations from hot deck	62	6,605	0.94
C_Race_i	Race	15 longitudinal imputations, 93 imputations from SSA's administrative data, 126 imputations from hot deck	234	6,605	3.54

Source: NBS, round 3.

Respondents could choose from five race categories: white, black/African American, Asian, Hawaiian/Pacific Islander, and Native American/American Indian. Respondents were allowed to select more than one of these categories to identify themselves (as prescribed by the Office of Management and Budget). The final race variable on which imputation was applied had six categories, with a separate category for respondents reporting multiple races. Although the SSA administrative data did not have a category for multiple races, respondents with race information in the SSA files were categorized according to four of the five categories above (Hawaiian/Pacific Islanders were included with the respondents reporting Asian). Respondents who did not answer the race question but did have race information in the SSA files were categorized into one of the four categories. This resulted in misclassification of respondents with extant SSA administrative data who did not answer the race question in the survey, but would have identified themselves in the survey as multiple race or Hawaiian/Pacific Islander. However, we assumed that the number of respondents of this kind was small, so that misclassification was not a major problem. As with the Hispanic indicator, for respondents that still had missing data, race was imputed using a hot deck with imputation classes defined by the

zip code of each sample member, and Hispanicity as a sorting variable. If the respondent was a longitudinal case, then the imputed value from round 1 was used.

2. Section B: Disability Status Variables and Work Indicator

Table IV.2 describes five imputed variables that pertain to the sample member's disability status and an indicator of whether the respondent was currently working. These imputed variables include three that collapse and recode primary diagnosis codes from the International Classification of Diseases, Ninth Revision (ICD-9) in three different ways: C_MainConBodyGroup_i, which corresponds to the collapsing done in Table II.2; C_MainConDiagGrp_i; and C_MainConColDiagGrp_i. Additional disability status variables include age when the disability was first diagnosed (C_DisAge_i); and an indicator of childhood or adult onset of the disability (C_AdultChildOnset_i). A fourth variable with collapsed primary diagnosis codes also was imputed, with levels further collapsed from C_MainConDiagGrp_i. This variable (C_MainConInput_i) is not included in Table IV.2 because it was not released to the final file, but it was used in subsequent imputations as a classing variable. As with race and ethnicity, the age when the disability was first diagnosed cannot change from one round to the next. For 14 missing values among longitudinal cases, this age variable was obtained from round 1 data. All missing values for C_AdultChildOnset_i were "logically assigned" using the imputed values from C_DisAge_i, the age-of-onset variable. In addition, Section B contains a question asking whether the respondent was currently working (B24_i). This is a gate question for all of the work status variables in Section C.

TABLE IV.2
DISABILITY STATUS IMPUTATIONS

Variable Name	Description	Imputation Method	Number Missing	Number Eligible	Percent Imputed
C_MainConDiagGrp_i	Primary diagnosis group	56 hot deck	56	5,906	0.95
C_MainConColDiagGrp_i	Main condition diagnosis group collapsed	56 hot deck	56	5,906	0.95
C_MainConBodyGroup_i	Main condition body group	56 hot deck	56	5,906	0.95
C_Disage_i	Age at onset of disability	167 hot deck; 43 from longitudinal data	210	6,605	3.18
C_Adultchild_onset_i	Adult/child onset of disability	19 logical	19	6,605	0.29
B24_i	Currently working	6 hot deck	6	6,605	0.09

Source: NBS, round 3

All of the variables in Section B used an indicator of whether the onset of the disability was in childhood or adulthood, as well as age and gender, to define imputation classes. One of the collapsed condition code variables, C_MainConInput_i was also used as a classing variable for disability age and the work indicator. Additional classing variables were used that were specific to the variable being imputed.

3. Section C: Current Jobs Variables

Several questions in the NBS asked respondents about current employment. In Section C, these questions were asked only of respondents who indicated in question B24 that they were currently working. These include salary (C_MainCurJobHrPay_i, C_MainCurJobMnthPay_i, and C_TotCurJobMnthPay_i); usual hours worked at the job or jobs (C8_1_i, C_TotCurWkHrs_i, and C_TotCurHrMnth_i); the number of places the respondent was employed (C1_i); and job description of the place of main employment (C2_1_1d_i). These variables are identified in Table IV.3.

TABLE IV.3
CURRENT JOBS IMPUTATIONS

Variable Name	Description	Imputation Method	Number Missing	Number Eligible	Percent Imputed
C1_i	Number of current jobs	2 hot deck	2	1,739	0.12
C2_1_1d_i	Main current job SOC code to one digit	8 hot deck ^a	8	1,739	0.46
C8_1_i	Hours per week usually worked at current main job	26 hot deck ^b ; 1 imputed by distributional assumptions	27	1,739	1.55
C_TotCurWkHrs_i	Total weekly hours at all current jobs	26 hot deck ^c , 6 constructed from imputed variables	32	1,739	1.84
C_TotCurHrMnth_i	Total hours per month at all current jobs	32 constructed from imputed variables	32	1,739	1.84
C_MainCurJobHrPay_i	Hourly pay at current main job	3 logical, 183 constructed from imputed variables	186	1,739	10.70
C_MainCurJobMnthPay_i	Monthly pay at current main job	10 logical, 23 imputed by distributional assumptions, 160 constructed from imputed variables	193	1,739	11.10
C_TotCurMnthPay_i	Total monthly salary all current jobs	31 logical, 160 hot deck, 11 constructed from imputed variables	202	1,739	11.62

Source: NBS, round 3.

^aImputations for current job variables include 2 cases coded as “don’t know” or “refused” in B24, which were imputed as currently working in B24_i.

^bImputations for current job variables include 2 cases coded as “don’t know” or “refused” in B24, which were imputed as currently working in B24_i.

^cIf C8_1_i was imputed by hot deck and the respondent had only one job, the flag indicated that C_TotCurWkHrs_i was imputed by hot deck, even though this variable was not processed in the hot-deck program.

Some of the variables in this table had missing values that were not directly imputed. Rather, constituent variables not included in this table had missing values that were imputed, and

then these were combined to form the variables in the table. For example, C_TotCurWkHrs_i was constructed from the number of hours per week usually worked at the current main job plus the number of hours for each of the respondent's other jobs. In most cases, the respondent worked one job, so C_TotCurWkHrs_i was set equal to C8_1_i. However, if the respondent worked multiple jobs and the number of hours in secondary jobs was imputed, then C_TotCurWkHrs_i was constructed from imputed variables.

Other variables had values imputed by using the distribution of a variable related to the variable at hand. For example, if the take-home monthly pay of the respondent's current main job was not missing but the gross monthly pay (C_MainCurJobMnthPay_i) of this job was missing, then the relationship between gross monthly and take-home monthly pay among respondents missing neither variable was used to determine the appropriate value for gross monthly pay. In particular, a random draw was selected from the observed distribution of relative taxes, where "relative tax" is defined as the proportion of imputed gross monthly pay for 22 cases with missing data for C_MainCurJobMnthPay. As Table IV.3 indicates, hot-deck imputations were applied to only four of the jobs variables: C1_i, C2_1_1d_i, C8_1_i, and C_TotCurMnthPay_i. For these variables, the collapsed condition code variable and level of education were used as classing variables. Additional classing and sorting variables specific to each variable also were used.

4. Section I: Health Status Variables

A total of 56 health status variables where imputations were applied are in Section I of the National Beneficiary Survey questionnaire. The 56 imputed variables in this section, and the methods of imputation used in each case, are identified in Tables IV.4 and IV.5. These items cover a range of topics, from the respondent's general health to more specific questions on the instrumental activities of daily living (IADLs) and activities of daily living (ADLs), and other

health and coping indicators. Also included in this section are a series of questions pertaining to the respondent's use of illicit drugs and alcohol.

TABLE IV.4
HEALTH STATUS IMPUTATIONS, QUESTIONNAIRE VARIABLES

Variable Name	Description	Imputation Method	Number Missing	Number Eligible	Percent Imputed
I1_i	Health during the past four weeks	21 hot deck	21	6,605	0.32
I9_i	Current health	34 hot deck	34	6,605	0.51
I17a_i	Wear glasses	15 hot deck	15	6,605	0.23
I17b_i	Difficulty seeing with glasses	9 logical, 28 hot deck	37	4,408	0.84
I18_i	Difficulty seeing no glasses	33 logical, 36 hot deck	69	2,234	3.09
I19_i	Uses special equipment because of difficulty seeing	49 logical, 12 hot deck	61	2,812	2.17
I21_i	Difficulty hearing	7 logical, 31 hot deck	38	6,605	0.58
I22_i	Able to hear normal conversation	26 logical, 23 hot deck	49	1,234	3.97
I23_i	Uses special equipment because of difficulty hearing	26 logical, 1 hot deck	27	1,234	2.19
I25_i	Difficulty having speech understood	4 logical, 36 hot deck	40	6,605	0.61
I26_i	Able to have speech understood at all	31 logical, 12 hot deck	43	1,691	2.54
I27_i	Uses special equipment because of difficulty speaking	31 logical, 2 hot deck	33	1,691	1.95
I29_i	Difficulty walking without assistance	14 logical, 43 hot deck	57	6,605	0.86
I30_i	Able to walk ¼ mile	24 logical, 51 hot deck	75	2,722	2.76
I31_i	Uses special equipment because of difficulty walking	24 logical, 8 hot deck	32	2,722	1.18
I33_i	Difficulty climbing 10 steps	8 logical, 69 hot deck	77	6,605	1.17

TABLE IV.4 (continued)

Variable Name	Description	Imputation Method	Number Missing	Number Eligible	Percent Imputed
I34_i	Able to climb 10 steps at all	38 logical, 55 hot deck	93	2,852	3.26
I35_i	Difficulty lifting and carrying 10 lbs.	3 logical, 38 hot deck	41	6,605	0.62
I36_i	Able to lift or carry 10 lbs. at all	29 logical, 31 hot deck	60	2,537	2.36
I37_i	Difficulty using hands or fingers	21 hot deck	21	6,605	0.32
I38_i	Able to use hands or fingers at all	14 logical, 21 hot deck	26	1,498	1.74
I39_i	Difficulty reaching over head	40 hot deck	40	6,605	0.61
I40_i	Able to reach over head at all	28 logical, 17 hot deck	45	1,524	2.95
I41_i	Difficulty standing	59 hot deck	59	6,605	0.89
I42_i	Able to stand at all	31 logical, 9 hot deck	40	3,551	1.13
I43_i	Difficulty stooping	3 logical, 35 hot deck	38	6,605	0.58
I44_i	Able to stoop at all	14 logical, 42 hot deck	56	3,556	1.57
I45_i	Difficulty getting around inside home	2 logical, 19 hot deck	21	6,605	0.32
I46_i	Need help to get around inside home	17 logical, 3 hot deck	20	921	2.17
I47_i	Difficulty getting around inside home	3 logical, 30 hot deck	33	6,605	0.50
I48_i	Need help to get around outside home	16 logical, 21 hot deck	37	2,193	1.69
I49_i	Difficulty getting into/out of bed	4 logical, 28 hot deck	32	6,605	0.48
I50_i	Need help getting into/out of bed	20 logical, 11 hot deck	31	1,539	2.01
I51_i	Difficulty bathing or dressing	1 logical, 26 hot deck	27	6,605	0.41
I52_i	Need help bathing or dressing	17 logical, 9 hot deck	26	1,236	2.10

TABLE IV.4 (continued)

Variable Name	Description	Imputation Method	Number Missing	Number Eligible	Percent Imputed
I53_i	Difficulty shopping	11 logical, 26 hot deck	37	6,605	0.56
I54_i	Need help shopping	16 logical, 10 hot deck	26	1,739	1.50
I55_i	Difficulty preparing own meals	7 logical, 24 hot deck	31	6,605	0.47
I56_i	Need help to prepare meals	17 logical, 9 hot deck	26	1,865	1.39
I57_i	Difficulty eating	20 hot deck	20	6,605	0.30
I58_i	Need help to eat	19 logical, 2 hot deck	21	685	3.07
I59_i	Trouble concentrating	42 hot deck	42	6,605	0.64
I60_i	Trouble coping with stress	80 hot deck	80	6,605	1.21
I61_i	Trouble getting along with people	70 hot deck	70	6,605	1.06
CageScore_indicator_i	CAGE Alcohol Score	30 constructed from imputed variables	30	6,605	0.45
I72_i	Use drugs in larger amounts than prescribed	51 hot deck	51	6,605	0.77

Source: NBS, round 3.

TABLE IV.5

HEALTH STATUS IMPUTATIONS, CONSTRUCTED VARIABLES

Variable Name	Description	Imputation Method	Number Missing	Number Eligible	Percent Imputed
C_EquipFuncLim_I	Use equipment/device for functional/sensory limitation	20 constructed from imputed variables	20	6,605	0.30
C_NumSenLim_i	Number of sensory limitations	107 constructed from imputed variables	107	6,605	1.62
C_NumSevSenLim_i	Number of severe sensory limitations	52 constructed from imputed variables	52	6,605	0.79
C_NumPhyLim_i	Number of physical functional limitations	193 constructed from imputed variables	193	6,605	2.92
C_NumSevPhyLim_i	Number of severe physical functional limitations	191 constructed from imputed variables	191	6,605	2.89
C_NumEmotLim_i	Number of emotional/social limitations	149 constructed from imputed variables	149	6,605	2.26
C_NumADLs_i	Number of impaired activities of daily living (ADLs)	50 constructed from imputed variables	50	6,605	0.76
C_NumADLAssist_i	Number of ADLs requiring assistance	35 constructed from imputed variables	35	6,605	0.53
C_NumIADLs_i	Number of instrumental activities of daily living (IADL) difficulties	53 constructed from imputed variables	53	6,605	0.80
C_NumIADLAssist_i	Number of IADLs Requiring Assistance	35 constructed from imputed variables	35	6,605	0.53
C_PCS8TOT_i	Physical summary score	199 constructed from imputed variables	199	6,605	3.01

TABLE IV.5 (continued)

Variable Name	Description	Imputation Method	Number Missing	Number Eligible	Percent Imputed
C_MCS8TOT_i	Mental summary score	199 constructed from imputed variables	199	6,605	3.01
C_DrugDep_i	Drug dependence	52 constructed from imputed variables	52	6,605	0.79

Source: NBS, round 3.

An example of a logical assignment in this section: if a respondent did not answer whether they had difficulty seeing newsprint letters (I17), but indicated that he or she couldn't see newsprint letters at all (I18) or required special devices to read newsprint letters (I19), then I17_i was a logically assigned "yes."

As in previous sections, "constructed from imputed variables" refers to the fact that the constituent variables of each constructed variable were imputed.

The only classing variable common to all imputations was the collapsed condition code variable. Age and gender were also used in most imputations. The other classing and sorting variables were specific to the variable being imputed.

5. Section K: Sources of Income Other than Employment

The imputed variables presented in this section are constructed variables that pertain to nonemployment-based income. These include workers' compensation, private disability claims, unemployment, and other sources of regular income. The imputed variables in this section are described in Table IV.6.

TABLE IV.6

IMPUTATIONS ON SOURCES OF INCOME OTHER THAN EMPLOYMENT

Variable Name	Description	Imputation Method	Number Missing	Number Eligible	Percent Imputed
C_AmtPrivDis_i	Amount received from private disability last month	77 logical, 13 imputed using specialized procedures	90	6,605	1.36
C_AmtWorkComp_i	Amount received from workers' compensation last month	37 logical, 8 imputed using specialized procedures	45	6,605	0.68
C_AmtVetBen_i	Amount received from veterans' benefits last month	29 logical, 10 imputed using specialized procedures	39	6,605	0.59
C_AmtPubAssis_i	Amount received from public assistance last month	38 logical, 16 imputed using specialized procedures	54	6,605	0.82
C_AmtUnemploy_i	Amount received from unemployment benefits last month	30 logical, 1 imputed using specialized procedures	31	6,605	0.47
C_AmtPrivPen_i	Amount received from private pension last month	32 logical, 19 imputed using specialized procedures	51	6,605	0.77
C_AmtOthReg_i	Amount received from other regular sources last month	34 logical, 15 imputed using specialized procedures	49	6,605	0.74

Source: NBS, round 3.

In this section, respondents first were asked if they had received money from a specific source, and then the specific amount received from that source. If a respondent could not provide a specific value, he or she was asked a series of questions on whether the amount was above or below specific values, or was given the option of providing a range of values, where the options depended upon responses to a series of questions. After being classified according to a range of

values that he or she provided, the respondent was assigned the median of the specific values provided by others who gave responses within the same range. If a respondent could not say whether the actual value was above or below a specific threshold, we imputed first the range (using a random assignment), and then assigned the median of the values provided by respondents who gave specific values within that range. If the respondent did not know if he or she received funds from a source, we used hot-deck imputation to determine this, and then proceeded as above.

The logical assignments in this section derive from imputed values in the constituent questions. For example, if the respondent was imputed to not have received private disability insurance (K6a_i), then C_AmtPrivDis_i was a logically assigned “no.” Otherwise, if any income was derived from these sources but an imputation was required at some point in the sequence (either everything was imputed, or just the individual’s income was imputed) then the imputation flag indicated imputation by “special procedures.”

For variables requiring hot-deck imputation, the classing variables were the same for all variables: an indicator of whether the respondent was a recipient of SSI, SSDI, or both; living situation; and education. None of the variables requiring hot-deck imputation are listed in Table IV.6 because they were only component variables for the delivered variables listed in the table.

6. Section L: Personal and Household Characteristics

Other than the personal characteristics of race and ethnicity discussed earlier, most of the imputed variables in section L pertain to household characteristics. These questions include education (L3_i), marital status (L8_i), cohabitation status (C_Cohab_i), number of children in the household (C_NumChildHH_i), household size (C_Hhsize_i), and body mass index for poverty-level respondents (C_BMI_cat_i), since it is constructed of variables collected in section

L. Most of these variables were imputed early in imputation processing and were used in the imputation (FedPovertyLevel_cat1).⁴⁷ Also included in this section is the constructed variable for the work status variables; however, poverty level was imputed later. Both sets of variables are discussed in this section.

The imputation of poverty level required the imputation of annual income and household size. The annual income question was another case of a specific value being requested, and if a specific value could not be provided, the respondent was asked if the annual income fell within certain ranges. For this item, some respondents provided a specific value; some only provided a range of values, and some refused to provide any information. Although annual income was a key variable used in the imputation of poverty level, it is not included in this table since it was not released in the final file. All of the missing values in C_FedPovertyLevel_cat1⁴⁸ were derived from the imputed annual incomes; hence all missing values are “constructed from imputed variables.” Table IV.7 identifies imputed variables in section L.

Logical assignments in this section are based on related variables that also are in this section. For example, the four logical assignments for L11_i (living situation of beneficiary) are due to the fact that four respondents did not answer L11, but indicated in L16 (number of adults in household) that only one adult lived in the household, and in L17 (number in household under 18 years old) indicated the number of children living with them in the household. For these four respondents, the value for L11_i was logically assigned to 1 or 2, depending upon the response to L17.

⁴⁷ An additional variable, C_NumChildren_i, also was imputed. This variable is defined as the total number of children in the household plus the number of respondent’s children living outside the household. This variable was not used in any subsequent processing and upon further review, was not deemed necessary for analysis, but is in the final file.

⁴⁸ The name of this variable reflects that fact that the final variable was a categorical (as opposed to a continuous) measure of poverty level.

TABLE IV.7

IMPUTATIONS OF PERSONAL AND HOUSEHOLD CHARACTERISTICS

Variable Name	Description	Imputation Method	Number Missing	Number Eligible	Percent Imputed
C_BMI_Cat_i	Body Mass Index categories	246 hot deck	246	6,605	3.72
L3_i	Highest year/grade completed in school	87 hot deck	87	6,605	1.32
L8_i	Marital Status	32 hot deck	32	6,605	0.48
L11_i	Living arrangements	3 logical, 32 hot deck	35	6,605	0.53
C_NumChildhh_i	Number of children living in the household	26 hot deck, 4 constructed from imputed variables	30	6,605	0.45
C_hhsize_i	Household Size	35 hot deck	35	6,605	0.53
C_cohab_i	Cohabitation Status	3 logical, 34 hot deck	37	6,605	0.56
C_FedPovertyLevel_cat1	2005 Federal Poverty Level	2141 constructed from imputed variables	2141	6,605	32.41

Source: NBS, round 3.

The only classing variable common to all imputations for the variables listed in Table IV.7 was the collapsed condition code variable. Other classing and sorting variables were specific to the variable being imputed.

V. ESTIMATING SAMPLING VARIANCE FOR NBS

The sampling variance of an estimate derived from survey data for a statistic (such as a total, a mean or proportion, or a regression coefficient) is a measure of the random variation among estimates of the same statistic computed over repeated implementation of the same sample design, with the same sample size, on the same population. The sampling variance is a function of the population characteristics, the form of the statistic, and the nature of the sampling design. The two general forms of statistics are linear combinations of the survey data (for example, a total) and nonlinear combinations. The latter include the ratio of two estimates (for example, a mean or proportion in which both the numerator and denominator are estimated) and more complex combinations, such as regression coefficients. For linear estimates with simple sample designs (such as a stratified or unstratified simple random sample) or complex designs (such as stratified multistage designs), explicit equations are available to compute the sampling variance. For the more common nonlinear estimates with simple or complex sample designs, explicit equations generally are not available, and various approximations or computational algorithms are used to provide an essentially unbiased estimate of the sampling variance.

The NBS sample design involves stratification and unequal probabilities of selection. Variance estimates calculated from NBS data must incorporate the sample design features to obtain the correct estimate. Most procedures in standard statistical packages, such as SAS, STATA, and SPSS, are not appropriate for analyzing data from complex survey designs, such as the NBS design. These procedures assume independent, identically distributed observations or simple random sampling with replacement. Although the simple random sample (SRS) variance may approximate the true sampling variance for some surveys, it is likely to substantially underestimate the sampling variance with a design as complex as that used for the NBS.

Complex sample designs have led to the development of a variety of software options that require the user to identify essential design variables such as strata, clusters, and weights.⁴⁹

The most appropriate sampling variance estimators for complex sample designs such as the NBS are the procedures based on the Taylor series linearization of the nonlinear estimator using explicit sampling variance equations, and the procedures based on forming pseudo-replications⁵⁰ of the sample. The Taylor series linearization procedure is based on a classic statistical method in which a nonlinear statistic can be approximated by a linear combination of the components within the statistic. The accuracy of the approximation is dependent on the sample size and the complexity of the statistic. For most commonly used nonlinear statistics (such as ratios, means, proportions, and regression coefficients), the linearized form has been developed and has good statistical properties. Once a linearized form of an estimate is developed, the explicit equations for linear estimates can be used to estimate the sampling variance. Because the explicit equations can be used, the sampling variance can be estimated using many features of the sampling design (for example, finite population corrections, stratification, multiple stages of selection, and unequal selection rates within strata). This is the basic variance estimation procedure used in SUDAAN, the survey procedures in SAS, STATA, and other software packages to accommodate simple and complex sampling designs. To calculate the variance, sample design information (such as stratum, analysis weight, and so on) is needed for each sample unit.

⁴⁹ A website that reviews software for variance estimation from complex surveys, created with the encouragement of the Section on Survey Research Methods of the American Statistical Association, is available online at <http://www.fas.harvard.edu/~stats/survey-soft/survey-soft.html>. The site lists software packages available for personal computers, and provides direct links to the home pages of these packages. The site also contains articles and links to articles that provide general information about variance estimation, as well as links to articles that compare features of the software packages.

⁵⁰ Pseudo-replications of a specific survey sample, as opposed to true replications of the sampling design, entail the selection of multiple independent subsamples from the original sample data using the same sampling design. These subsamples can be random (as in a bootstrap) or restricted (as in Balanced Repeated Replication).

Currently, more survey data analysis software packages use the Taylor series linearization procedure and explicit sampling variance equations. Therefore, we developed the variance estimation specifications necessary for the Taylor series linearization (PseudoStrata and PseudoPSU). Example code for this procedure using SAS and the survey data analysis software SUDAAN is given in Appendix F.⁵¹ Details about syntax for SAS are available from SAS (SAS Institute 2004). Details about SUDAAN syntax are available from RTI International (Research Triangle Institute 2004).

⁵¹ The example code provided in Appendix F is for simple descriptive statistics using the procedures DESCRIBE in SUDAAN and SURVEYMEANS in SAS. Other procedures in SAS (SURVEYREG, SURVEYFREQ, and SURVEYLOGISTIC) and in SUDAAN (CROSSTAB, REGRESS, LOGISTIC, MULTILOG, LOGLINK, and SURVIVAL) are available for more complex analyses. Since SUDAAN was created specifically for survey data, the range of analyses that can be performed with these data in SUDAAN is much wider than in SAS.

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

VI. REFERENCES

- Agresti, A. *Categorical Data Analysis*. New York: John Wiley and Sons, 1990.
- Akaike, H. "A New Look at the Statistical Model Identification." *IEEE Transaction on Automatic Control*, AC-19, 1974, pp. 716-723.
- Biggs, D., B. deVillie, and E. Suen. "A Method of Choosing Multiway Partitions for Classification and Decision Trees." *Journal of Applied Statistics*, vol. 18, 1991, pp. 49-62.
- Cox, D. R., and E. J. Snell. *The Analysis of Binary Data*, Second Edition. London: Chapman and Hall, 1989.
- Folsom, R., F. Potter, and S. Williams. "Notes on a Composite Size Measure for Self to Weighting Samples in Multiple Domains." *Proceedings of the American Statistical Association, Section on Survey Research Methods*, 1987, pp. 792-796.
- Hosmer, D.W., Jr., and S. Lemeshow, "Goodness-of-Fit Tests for the Multiple Logistic Regression Model. *Communications in Statistics, Theory and Methods*, Vol A9(10), 1980, pp. 1043-1069.
- Kass, G. V. "An Exploratory Technique for Investigating Large Quantities of Categorical Data." *Applied Statistics*, vol. 29, 1980, pp. 119-127.
- Magidson, J. *SPSS for Windows CHAID Release 6.0*. Belmont MA: Statistical Innovations, Inc., 1993.
- NAACCR Expert Panel on Hispanic Identification. "Report of the NAACCR Expert Panel on Hispanic Identification 2003." Springfield, IL: North American Association of Central Cancer Registries 2003.
- Research Triangle Institute. *SUDAAN Language Manual, Release 9.0*. Research Triangle Park, NC: Research Triangle Institute. 2004.
- SAS[®] Institute. *SAS/STAT[®] 9.1 User's Guide*. Cary, NC: SAS Institute. 2004.
- Thornton, C., G. Livermore, D. Stapleton, J. Kregel, T. Silva, B. O'Day, T. Fraker, W.G. Revell Jr., H. Schroeder, and M. Edwards. "Evaluation of the Ticket to Work Program: Initial Evaluation Report." Prepared for the Social Security Administration. Washington DC: Mathematica Policy Research, Inc., 2004.
- Wright, D., and K. Barrett. "National Beneficiary Survey Round 3 (volume 2 of 3): Cleaning and Identification of Data Problems." Report submitted to the Social Security Administration. Washington, DC: Mathematica Policy Research, Inc., December 2008.

U.S. Department of Education. National Center for Education Statistics. "A Study of Imputation Algorithms." Working Paper No. 2001-17. Ming-xiu and Sameena Salvucci. Washington, DC. 2001.

Wright, D., E. Grau, Y. Zhang, D. Basson, F. Potter, K. Barrett. "National Beneficiary Survey Round 3 (volume 3 of 3): User's Guide for Restricted and Public Use Data Files." Report submitted to the Social Security Administration. Washington, DC: Mathematica Policy Research, Inc., January 2010.

APPENDIX A

**OTHER SPECIFY AND OPEN-ENDED ITEMS WITH ADDITIONAL CATEGORIES
CREATED DURING CODING**

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

OTHER/SPECIFY AND OPEN-ENDED ITEMS WITH ADDITIONAL CATEGORIES CREATED DURING CODING

Question #	Question Text	Current Response Options	Additional Categories Created
B27	What are they (the other reasons you are not working that I didn't mention)?	a=A physical or mental condition prevents {you/him/her} from working b={You/NAME} cannot find a job that {you are/(he/she) is} qualified for c={You do/NAME does} not have reliable transportation to and from work d={You are/NAME is} caring for someone else. f={You/NAME} cannot find a job {you want/(he/she) wants}. g={You are/NAME is} waiting to finish school or a training program. h=Workplaces are not accessible to people with {your/NAME's} disability. i={You do/NAME does} not want to lose benefits such as disability, worker's compensation, or Medicaid j={Your/NAME's} previous attempts to work have been discouraging l=Others do not think {you/NAME} can work m=Employers will not give {you/NAME} a chance to show that {you/he/she} can work. n={You/NAME} does not have the special equipment or medical devices that {you/he/she} would need in order to work. o={You/NAME} cannot get the personal assistance {you need/he needs/she needs} in order to get ready for work each day	p=Cannot find a job/job market is bad q=Lack skills
B39	Who {do you/does NAME} discuss your work goals with the most?	01=PARENT/GUARDIAN 02=SPOUSE/PARTNER 03=FRIEND 04=JOB COACH 05=EMPLOYER/SUPERVISOR 06=OTHER RELATIVE 07=CASE WORKER/COUNSELOR/PROGRAM STAFF 08=MEDICAL PROVIDER 09= OTHER (SPECIFY: <OPEN>)	10=OTHER NON-RELATIVE

Question #	Question Text	Current Response Options	Additional Categories Created
B42	Who else {do you/does NAME} discuss {your/his/her} work goals with?	01=PARENT/GUARDIAN 02=SPOUSE/PARTNER 03=FRIEND 04=JOB COACH 05=EMPLOYER/SUPERVISOR 06=OTHER RELATIVE 07=CASE WORKER/COUNSELOR/PROGRAM STAFF 08=MEDICAL PROVIDER 09= OTHER (SPECIFY: <OPEN>)	10=OTHER NON-RELATIVE
B45	Who else {do you/does NAME} discuss {your/his/her} work goals with?	01=PARENT/GUARDIAN 02=SPOUSE/PARTNER 03=FRIEND 04=JOB COACH 05=EMPLOYER/SUPERVISOR 06=OTHER RELATIVE 07=CASE WORKER/COUNSELOR/PROGRAM STAFF 08=MEDICAL PROVIDER 09= OTHER (SPECIFY: <OPEN>)	10=OTHER NON-RELATIVE
C23	What kind of special equipment {do you/does NAME} use?	01=BRACE 02=CANE/CRUTCHES/WALKER 03=WHEELCHAIR 04=MODIFIED COMPUTER HARDWARE 05=MODIFIED COMPUTER SOFTWARE 06=OTHER (SPECIFY: <OPEN>)	07=HEARING AIDS 08=GLASSES 09=SPECIAL CHAIR / BACK SUPPORT 10=SPECIAL SHOES / SUPPORT STOCKINGS
C35	Are there any changes in {your/NAME's} {main/current} job or workplace related to {your/his/her} mental or physical condition that {you need/he/she needs}, but that have <u>not</u> been made? (IF YES) What are those changes?	<OPEN>	a=Need special equipment or assistive b=Need changes in {your/NAME's} work c=Need changes to the tasks {you were/NAME was} assigned or how they are performed d=Need changes to the physical work environment e=Need co-workers or others to assist {you/NAME}?

Question #	Question Text	Current Response Options	Additional Categories Created
C39b	{ Do you/Does NAME} work fewer hours or earn less money than {you/he/she} could because {you/he/she}:	a={ Are/Is} taking care of children or others? b={ Are/Is} enrolled in school or a training program? c=Want{s} to keep Medicare or Medicaid coverage? d=Want{s} to keep cash benefits {you/he/she} need such as disability or workers' compensation? e=Just {do/does} not want to work more? f=Are there any reasons I didn't mention why {you are/NAME is} working or earning less than {you/he/she} could?	g={ Are/is} in poor health or {have/has} health concerns? h={Need/Needs} a more supportive/helpful employer and/or coworkers?
D23	Why did {you/NAME} stop working at this job?	LAYOFF, FIRED, RETIRED 1=LAYOFF, PLANT CLOSED 2=FIRED 3=RETIRED/OLD AGE 4=JOB WAS TEMPORARY AND ENDED PROBLEMS WITH JOB 5=DID NOT LIKE SUPERVISOR OR CO-WORKERS 6=DID NOT LIKE JOB DUTIES 7=DID NOT LIKE JOB EARNINGS 8=DID NOT LIKE BENEFITS 9=DID NOT LIKE OPPORTUNITIES FOR ADVANCEMENT 10=DID NOT LIKE LOCATION 11=DID NOT GET ACCOMMODATIONS THAT WERE NEEDED OTHER PROBLEMS 12=TRANSPORTATION PROBLEMS 13=DECIDED TO GO TO SCHOOL 14=CHILD CARE RESPONSIBILITIES (PREGNANT) 15=OTHER FAMILY OR PERSONAL REASONS DISABILITY 16=DISABILITY GOT WORSE 17=BECAME DISABLED 18=OTHER (SPECIFY: <OPEN>)	19= Moved to another area 20= Found another job 21=Loss or potential loss of government benefits 22=Work schedule

Question #	Question Text	Current Response Options	Additional Categories Created
D25b	Did you work fewer hours or earn less money than you could have because {you/he/she} you...	a={ Were/Was } taking care of somebody else? b={ Were/Was } enrolled in school or a training program? c=Wanted to keep Medicare or Medicaid coverage d=Wanted to keep cash benefits such as disability or workers compensation? e=Just didn't want to work more? f=Are there any reasons I didn't mention why {you/NAME} might have chosen to work or earn less than {you/he/she} could have during 2004? (SPECIFY: <OPEN>)	g=Had medical problems/complications
D26	In 2004, do you think {you/NAME} could have worked or earned more if {you/he/she} had:	a=Help caring for {your/his/her} children or others in the household? b=Help with {your/his/her} own personal care such as bathing, dressing, preparing meals, and doing housework? c=Reliable transportation to and from work? d=Better job skills? e=A job with a flexible work schedule? f=Help with finding and getting a better job? g=Any special equipment or medical devices? (SPECIFY: <OPEN>) h=Is there anything else that I didn't mention that would have helped {you/NAME} to work or earn more during 2004? (SPECIFY: <OPEN>)	i=Better health/treatment j=More supportive/helpful employer and/or coworker
E32	Who talked to {you/NAME or his/her representative} about the program?	01=SOCIAL SECURITY ADMINISTRATION 02=MAXIMUS 03=STATE VOCATIONAL REHABILITATION AGENCY, OR {VRNAME} 04=CURRENT/FORMER EMPLOYER 05=FRIEND/FAMILY MEMBER 06=INDEPENDENT LIVING CENTER 07=EMPLOYMENT NETWORK 08=OTHER AGENCY/ORGANIZATION 09=HEALTH CARE PROVIDER 10=OTHER (SPECIFY: <OPEN>)	11=CASE WORKER/SOCIAL WORKER
E37a1	Why {are you/is NAME} no longer receiving services from {EN FROM ROUND 1 E39 OR E46 WHEN E41=01 OR E45=01}?	<OPEN>	1=Never received information/case dropped/didn't help 2=Found a job 3=I cannot work for health reasons

Question #	Question Text	Current Response Options	Additional Categories Created
E43	Why {are you/is NAME} no longer receiving services from {EN IN 2004 FROM E39}?	<OPEN>	01=Never received any info/case dropped/ didn't help 02=Found a job 03=I cannot work for health reasons 04=Other reason related to personal circumstance 05=Other reason related to EN
F14	Why didn't {you/NAME or his/her representative} try to use {your/NAME's} Ticket with the State VR agency in 2004?	<OPEN>	01=Agency didn't help/ couldn't find job 02=Did not know could/did not have ticket 03=Was not healthy enough to participate
F29	After receiving information about the Employment Networks in {your/NAME's} area including the State VR agency or {STATE NAME FOR VR}, why didn't {you/NAME or his/her representative} contact any of them?	01=PHYSICAL/MENTAL CONDITION 02=CHANGED MIND 03=FAMILY RESPONSIBILITIES 04=FAMILY WOULD NOT SUPPORT 05=COULD NOT GET RELIABLE TRANSPORTATION 06=ECONOMIC CONDITIONS CHANGED – NO LONGER THINK JOB OPPORTUNITIES EXIST 07=FEARED SERVICES WOULD ENDANGER BENEFITS 08=INFORMATION TOO CONFUSING – DID NOT KNOW WHERE TO START 09=EMPLOYMENT NETWORK {NAME} WANTED WAS NOT PARTICIPATING 10=ENs TOO FAR AWAY 11=COULD NOT GET IN CONTACT WITH ENs 12=NO ENs PROVIDED SERVICES {NAME} NEEDS 13=NO ENs SERVE MY KIND OF DISABILITY 14=OTHER (SPECIFY: <OPEN>)	15=GOT A JOB OR IN SCHOOL
G7	Thinking about {PROVIDER FROM G2}, was this place:	01=A state agency 02=A private business 03=Some other type of place? (SPECIFY: <OPEN>)	04=School
G18	Thinking about {NEW PROVIDER FROM G16}, was this place:	01=A clinic, 02=A hospital, 03=A doctor's office, or 04=Some other type of place? (SPECIFY: <OPEN>)	05=A school 06=A nursing home/group home 07=A government agency 08=In home care 09=A medical equipment store 10=A rehabilitation/counseling center 11=Physical therapy center

Question #	Question Text	Current Response Options	Additional Categories Created
G22	Thinking about {NEW PROVIDER FROM G20}, was this place:	01=A mental health agency, 02=A clinic, 03=A hospital, 04=A doctor's office, or 05=Some other type of place? (SPECIFY: <OPEN>)	06=Residential treatment program/facility 07=Rehab center/counseling center/day program 08=Church or religious institution
G36	In 2004, please tell me if {you/NAME} received any of the following services from {PROVIDER FROM G30_1 DE-DUPLICATED LIST IF USED IN 2004}. Did {you/he/she} receive:	a=Physical therapy? b=Occupational therapy? c=Speech therapy? e=Special equipment or devices? f=Personal counseling or therapy? g=Group therapy? d= Medical services? h=A work or job assessment? i=Help to find a job? j=Training to learn a new job or skill? k=Advice about modifying {your/his/her} job or work place? l=On-the-job training, job coaching, or support services? m=Anything else that I didn't mention? (SPECIFY: <OPEN>)	n=Scholarships/grants/loans
G45	In 2004, who paid for the services {you/NAME} received from {PROVIDER FROM G32 DE-DUPLICATED LIST IF USED IN 2004}?	01={NAME} 02=PROVIDER FROM G32 DE-DUPLICATED LIST IF USED IN 2004 03=NO ONE 04=FAMILY 05=EMPLOYMENT NETWORK 06=MEDICARE 07=MEDICAID 08=EMPLOYER 09=NON-PROFIT ORGANIZATION SERVING PEOPLE WITH DISABILITIES 10=WORKER'S COMPENSATION 11=DISABILITY INSURANCE 12=OTHER (SPECIFY: <OPEN>)	14=SCHOOL/FINANCIAL AID/GRANT 15=STATE AGENCY/COUNTY/GOVERNMENT

Question #	Question Text	Current Response Options	Additional Categories Created
G55	Who pressured {you/NAME} to use these services?	01=PARENT/GUARDIAN 02=SPOUSE/PARTNER 03=OTHER FAMILY MEMBER 04=FRIEND/CO-WORKER 05=EMPLOYER/SUPERVISOR 06=STAFF OF EMPLOYMENT NETWORK 07=VOCATIONAL REHABILITATION CASE MANAGER 08=JOB COACH 09=SSA LETTER 10=SSA STAFF 11=BENEFIT SPECIALIST/BPAO 12=OTHER (SPECIFY: <OPEN>)	13=Health care professional 14=Court/Police
G56	How did {your/NAME's} {FILL PERSON(S) FROM G55} pressure {you/him/her} to use these services?	01=SAID {NAME} WOULD LOSE DISABILITY AND/OR HEALTH INSURANCE BENEFITS 02=ENCOURAGED/WOULD NOT TAKE "NO" FOR AN ANSWER 03=THREATENED TO WITHHOLD SERVICES 04=THREATENED TO TAKE AWAY OTHER SUPPORT 05=OTHER (SPECIFY: <OPEN>)	06=THREATENED HOSPITALIZATION/JAIL
G61	Why {were you/was NAME} unable to get these services?	<OPEN>	01=Not eligible/request refused 02=Lack information on how to get services 03=Could not afford/insurance would not cover 04=Did not try 05=Too difficult/too confusing to get services 06=Problems with the service or agency
H3	Why did {you/NAME} decide to participate in the Ticket to Work program?	<OPEN>	01=Wanted to get a job or more money/benefits 02=Wanted to do something and feel more independent 03=Recommended/told to use it/thought using it was required

Question #	Question Text	Current Response Options	Additional Categories Created
H23	Why didn't {you/NAME or his/her representative} try to use {your/NAME's} Ticket with the State VR agency in 2004?	<OPEN>	01=Signed up with another agency 02=Already receiving services from VR 03=Didn't understand ticket/didn't know what it was for
H29	Why didn't {you/NAME or (his/her) representative} try to use {your/NAME's} Ticket with {any of} the other Employment Network(s) {you/NAME or (his/her) representative} contacted in 2004?	<OPEN>	01=Location 02=Not helpful/didn't like 03=Did not like the job offered
H31	Why didn't {any of} the other { Employment Network(s) {you/NAME} tried to use {your/his/her} Ticket with accept {your/NAME's} Ticket in 2004?	01=NOT TAKING TICKETS WHEN CONTACTED 02=DID NOT OFFER SERVICES {NAME} NEEDED 03=DID NOT SERVE PEOPLE WITH {NAME'S} DISABILITY/NEEDS 04={NAME} NOT WILLING/ABLE TO WORK FULL-TIME/ENOUGH HOURS 05={NAME} NOT WILLING TO GO OFF OF DISABILITY BENEFITS 06=OTHER (SPECIFY: <OPEN>)	07=TROUBLE CONTACTING EN
H33	What information did {you/NAME} need but didn't get?	<OPEN>	01=Information on how and where to use the Ticket 02=Information about services provided
H35	Why did {you/NAME or (his/her) representative} choose {{LONGEST} EMPLOYMENT NETWORK IN 2004}?	01=STAFF WERE MOST RESPONSIVE/COURTEOUS/KNOWLEDGEABLE 02=MOST WILLING TO PROVIDE THE SERVICES {NAME} WANTED 03=SERVED PEOPLE WITH {NAME'S} DISABILITY/NEEDS 04=WAIT FOR SERVICES WAS NOT TOO LONG 05=ONLY PROVIDER NEARBY/CLOSEST PROVIDER 06=ONLY PROVIDER WILLING TO ACCEPT TICKET 07=OTHER (SPECIFY: <OPEN>)	08=KNEW ABOUT THEM OR REFERRED TO THEM 09=FINANCIAL COMPENSATION
H38	What problems did {you/NAME} have during 2004 (with the services you received from EN)?	<OPEN>	01=Trouble making/keeping contact 02=Did not receive services needed 03=Problems with counselor 04=Transportation/location problems

Question #	Question Text	Current Response Options	Additional Categories Created
H48	What was the problem about?	<OPEN>	01=Trouble making/keeping contact 02=Did not receive services wanted/needed
H50	What did {you/NAME} or someone else do to try to solve the problem?	01=REFERRED TO DOCUMENTS/INFORMATION ABOUT PROVIDER/PROGRAM 02=CONTACTED EN BY PHONE 03=CONTACTED EN IN WRITING 04=CONTACTED PROGRAM MANAGEMENT COMPANY (MAXIMUS) BY PHONE 05=CONTACTED PROGRAM MANAGEMENT COMPANY (MAXIMUS) BY WRITING 06=CONTACTED SSA BY PHONE 07=CONTACTED SSA IN WRITING 08=CONTACTED OTHER STATE/LOCAL AGENCY 09=CONTACTED {LOCAL PROTECTION & ADVOCACY AGENCY} FOR HELP 10=CONTACTED CASE WORKER/JOB COACH 11=OTHER (SPECIFY: <OPEN>)	12= QUIT/LOOKED FOR JOB ON OWN
I20	What devices, equipment, or other types of assistance {do you/does NAME} use? Anything else?	01=TELESCOPIC LENSES 02=ADAPTED COMPUTER EQUIPMENT 03=BRAILLE 04=READERS 05=GUIDE DOG 06=WHITE CANE 07=OTHER SEEING ASSISTANCE (SPECIFY: <OPEN>)	08=MAGNIFYING GLASS
I32	What devices, equipment, or other types of assistance {do you/does NAME} use? Anything else?	01=BRACES, CRUTCHES, CANE, OR WALKER 02=WHEELCHAIR OR SCOOTER 03=PROSTHETIC DEVICE 04=SPECIAL CHAIR (NOT WHEELCHAIR) 05=VEHICLE HAND CONTROLS 06=LIFT (HOME OR VEHICLE) 07=OTHER MOBILITY ASSISTANCE (SPECIFY: <OPEN>)	09=SPECIAL SHOES OR SHOE INSERTS 10=DEVICES TO AIDE IN BREATHING INCLUDING OXYGEN, INHALER, ALBUTEROL, AND/OR NEBULIZER

Question #	Question Text	Current Response Options	Additional Categories Created
J11	Now, I'd like you to think back to 2004. In 2004, what kinds of health coverage did {you/NAME} have?	01=MEDICAID/{STATMED} 02=MEDICARE 03=CHAMPUS/CHAMP-VA, TRICARE, VA, OTHER MILITARY 04=INDIAN HEALTH SERVICE 05=MEDI-GAP 06=STATE PROGRAM 07=PRIVATE INSURANCE THROUGH OWN EMPLOYER 08=PRIVATE INSURANCE THROUGH SPOUSE/PARTNER/PARENT 09=PRIVATE INSURANCE PAID BY SELF/FAMILY 10=OTHER PLAN (SPECIFY: <OPEN>)	11=PRIVATE INSURANCE, NOT SPECIFIED WHO THROUGH
K14	What other assistance did {you/NAME} receive <u>last month</u> ?	<OPEN>	01=Housing Assistance 02=Energy Assistance 03=Food assistance
M2a_rlshp	How are you related to {NAME}?	01={NAME'S} SPOUSE 02={NAME'S} MOTHER 03={NAME'S} FATHER 04={NAME'S} CHILD 05=GRANDPARENT OF {NAME} 06=BROTHER/SISTER (NATURAL/STEP) OF {NAME} 07=AUNT/UNCLE OF {NAME} 08=OTHER RELATIVE OF {NAME} (SPECIFY: <OPEN>) 09=NOT RELATED (SPECIFY: <OPEN>) 10=STAFF AT RESIDENCE	11=FRIEND 12=CASEWORKER/CAREGIVER/PAYEE 13=GIRLFRIEND/BOYFRIEND/PARTNER 14=GUARDIAN/FOSTER PARENT/STEP PARENT 15=IN-LAW
M8	How is that person related to {you/NAME}, if at all?	01={NAME'S} SPOUSE 02={NAME'S} MOTHER 03={NAME'S} FATHER 04={NAME'S} CHILD 05=GRANDPARENT OF {NAME} 06=BROTHER/SISTER (NATURAL/STEP) OF {NAME} 07=AUNT/UNCLE OF {NAME} 08=OTHER RELATIVE OF {NAME} (SPECIFY: <OPEN>) 09=NOT RELATED (SPECIFY: <OPEN>) 10=STAFF AT RESIDENCE	11=FRIEND 12=CASEWORKER/CAREGIVER/PAYEE 13=GIRLFRIEND/BOYFRIEND/PARTNER 14=GUARDIAN/FOSTER PARENT/STEP PARENT 15=IN-LAW

Question #	Question Text	Current Response Options	Additional Categories Created
M10	How is that person related to {you/NAME}, if at all?	01={NAME'S} SPOUSE 02={NAME's} MOTHER 03={NAME'S} FATHER 04={NAME'S} CHILD 05=GRANDPARENT OF {NAME} 06=BROTHER/SISTER (NATURAL/STEP) OF {NAME} 07=AUNT/UNCLE OF {NAME} 08=OTHER RELATIVE OF {NAME} (SPECIFY: <OPEN>) 09=NOT RELATED (SPECIFY: <OPEN>) 10=STAFF AT RESIDENCE	11=FRIEND 12=CASEWORKER/CAREGIVER/PAYEE 13=GIRLFRIEND/BOYFRIEND/PARTNER 14=GUARDIAN/FOSTER PARENT/STEP PARENT 15=IN-LAW
M13	How is the assistant/proxy related to (NAME)?	01={NAME'S} SPOUSE 02={NAME's} MOTHER 03={NAME'S} FATHER 04={NAME'S} CHILD 05=GRANDPARENT OF {NAME} 06=BROTHER/SISTER (NATURAL/STEP) OF {NAME} 07=AUNT/UNCLE OF {NAME} 08=OTHER RELATIVE OF {NAME} (SPECIFY: <OPEN>) 09=NOT RELATED (SPECIFY: <OPEN>) 10=STAFF AT RESIDENCE	11=FRIEND 12=CASEWORKER/CAREGIVER/PAYEE 13=GIRLFRIEND/BOYFRIEND/PARTNER 14=GUARDIAN/FOSTER PARENT/STEP PARENT 15=IN-LAW

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

APPENDIX B

SOC MAJOR AND MINOR OCCUPATION CLASSIFICATIONS

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

SOC MAJOR AND MINOR OCCUPATION CLASSIFICATIONS

Code	Occupation
	Management
111	Top Executives
112	Advertising, Marketing, PR, Sales
113	Operations Specialist Managers
119	Other Management Occupations
	Business /Financial Operations
131	Business Operations Specialist
132	Financial Specialist
	Computer and Mathematical Science
151	Computer Specialist
152	Mathematical Science Occupations
	Architecture and Engineering
171	Architects, Surveyors and Cartographers
172	Engineers
173	Drafters, Engineering and Mapping Technicians
	Life, Physical and Social Science
191	Life Scientists
192	Physical Scientists
193	Social Scientists and Related Workers
194	Life, Physical and Social Science Technicians
	Community and Social Services
211	Counselors, Social Workers and Other Community and Social Service Specialists
212	Religious Workers
	Legal
231	Lawyers, Judges and Related Workers
232	Legal Support Workers
	Education, Training and Library
251	Postsecondary Teachers
252	Primary, Secondary and Special Education School Teachers
253	Other Teachers and Instructors
254	Librarians, Curators and Archivists
259	Other Education, Training and Library Occupations
	Arts, Design, Entertainment, Sports and Media
271	Art and Design Workers
272	Entertainers and Performers, Sports and Related Workers
273	Media and Communication Workers
274	Media and Communication Equipment Workers
	Healthcare Practitioner and Technical Occupations
291	Health Diagnosing and Treating Practitioners
292	Health Technologists and Technicians
299	Other Healthcare Practitioner and Technical Occupations
	Healthcare Support
311	Nursing, Psychiatric and Home Health Aides

Code	Occupation
312	Occupational and Physical Therapist Assistants and Aides
319	Other Healthcare Support Occupations
	Protective Service
331	Supervisors, Protective Service Workers
332	Firefighting and Prevention Workers
333	Law Enforcement Workers
339	Other Protective Service Workers
	Food Preparation and Serving Related
351	Supervisors, Food Preparation and Food Serving Workers
352	Cooks and Food Preparation Workers
353	Food and Beverage Serving Workers
359	Other Food Preparation and Serving Related Workers
	Building and Grounds Cleaning and Maintenance
371	Supervisors, Building and Grounds Cleaning and Maintenance Workers
372	Building Cleaning and Pest Control Workers
373	Grounds Maintenance Workers
	Personal Care and Service Occupations
391	Supervisors, Personal Care and Service Workers
392	Animal Care and Service Workers
393	Entertainment Attendants and Related Workers
394	Funeral Service Workers
395	Personal Appearance Workers
396	Transportation, Tourism, and Lodging Attendants
399	Other Personal Care and Service Workers
	Sales and Related Occupations
411	Supervisors, Sales Workers
412	Retail Sales Workers
413	Sales Representative, Services
414	Sales Representative, Wholesale and Manufacturing
419	Other Sales and Related Workers
	Office and Administrative Support
431	Supervisors, Office and Administrative Support Workers
432	Communications Equipment Operators
433	Financial Clerks
434	Information and Record Clerks
435	Material Recording, Scheduling Dispatching, and Distribution Workers
436	Secretaries and Administrative Assistants
439	Other Office and Administrative Support Workers
	Farming, Fishing and Forestry Workers
451	Supervisors, Farming, Fishing and Forestry Workers
452	Agricultural Workers
453	Fishing and Hunting Workers
454	Forest, Conservation and Logging Workers

Code	Occupation
Construction and Extraction Occupations	
471	Supervisors, Construction and Extraction Workers
472	Construction Trade Workers
473	Helpers, Construction Trades
474	Other Construction and Related Workers
475	Extraction Workers
Installation, Maintenance and Repair Occupations	
491	Supervisors, Installation, Maintenance and Repair Workers
492	Electrical and Electronic Equipment Mechanics, Installers and Repairers
493	Vehicle and Mobile Equipment Mechanics, Installers and Repairers
494	Other Installation, Maintenance and Repair Occupations
Production Occupations	
511	Supervisors, Production Workers
512	Assemblers and Fabricators
513	Food Processing Workers
514	Metal Workers and Plastic Workers
515	Printing Workers
516	Textile, Apparel, and Furnishing Workers
517	Woodworkers
518	Plant and System Operators
519	Other Production Occupations
Transportation and Material Moving Occupations	
531	Supervisors, Transportation and Material Moving Workers
532	Air Transportation Workers
533	Motor Vehicle Operators
534	Rail Transportation Workers
535	Water Transportation Workers
536	Other Transportation Workers
537	Material Moving Workers
Military Specific Occupations	
551	Military Officer and Tactical Operations Leaders/Managers
552	First-Line Enlisted Military Supervisors/Managers
553	Military Enlisted Tactical Operations and Air/Weapons Specialists and Crew Members

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

APPENDIX C
NAICS INDUSTRY CODES

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

NAICS INDUSTRY CODES

Code	Description
11	Agriculture, Forestry Fishing and Hunting
111	Crop Production
112	Animal Production
113	Forestry and Logging
114	Fishing, Hunting and Trapping
115	Support Activities for Agriculture and Forestry
21	Mining
211	Oil and Gas Extraction
212	Mining (except Oil and Gas)
213	Support Activities for Mining
22	Utilities
221	Utilities
23	Construction
236	Construction of Buildings
237	Heavy and Civil Engineering Construction
238	Specialty Trade Contractors
31-33	Manufacturing
311	Food Manufacturing
312	Beverage and Tobacco Product Manufacturing
313	Textile Mills
314	Textile Product Mills
315	Apparel Manufacturing
316	Leather and Allied Product Manufacturing
321	Wood Product Manufacturing
322	Paper Manufacturing
323	Printing and Related Support Activities
324	Petroleum and Coal Products Manufacturing
325	Chemical Manufacturing
326	Plastics and Rubber Products Manufacturing
327	Nonmetallic Mineral Product Manufacturing
331	Primary Metal Manufacturing
332	Fabricated Metal Products Manufacturing
333	Machinery Manufacturing
334	Computer and Electronic Product Manufacturing
335	Electrical Equipment, Appliance and Component Manufacturing
336	Transportation Equipment Manufacturing
337	Furniture and Related Product Manufacturing
339	Miscellaneous Manufacturing
42	Wholesale Trade
423	Merchant Wholesalers, Durable Goods

Code	Description
424	Merchant Wholesalers, Nondurable Goods
425	Wholesale Electronic Markets and Agents and Brokers
44-45	Retail Trade
442	Furniture and Home Furnishings Stores
443	Electronics and Appliance Stores
444	Building Material and Garden Equipment and Supplies Dealers
445	Food and Beverage Stores
446	Health and Personal Care Stores
447	Gasoline Stations
448	Clothing and Clothing Accessories Stores
451	Sporting Goods, Hobby, Book, and Music Stores
452	General Merchandise Stores
453	Miscellaneous Store Retailers
454	Nonstore Retailers
48-49	Transportation and Warehousing
481	Air Transportation
482	Rail Transportation
483	Water Transportation
484	Truck Transportation
485	Transit and Ground Passenger Transportation
486	Pipeline Transportation
487	Scenic and Sightseeing Transportation
488	Support Activities for Transportation
491	Postal Service
492	Couriers and Messengers
493	Warehousing and Storage
51	Information
511	Publishing Industries (except Internet)
512	Motion Picture and Sound Recording Industries
515	Broadcasting (except Internet)
516	Internet Publishing and Broadcasting
517	Telecommunications
518	Internet Service Providers, Web Search Portals, and Data Processing Services
519	Other Information Services
52	Finance and Insurance
522	Credit Intermediation and Related Activities
523	Securities, Commodity Contracts, and Other Financial Investments and Related Activities
524	Insurance Carriers and Related Activities
525	Funds, Trusts, and Other Financial Vehicles
53	Real Estate and Rental and Leasing
531	Real Estate
532	Rental and Leasing Services

Code	Description
533	Lessors of Nonfinancial Intangible Assets (except Copyrighted Works)
54	Professional, Scientific, and Technical Services
55	Management of Companies and Enterprises
551	Management of Companies and Enterprises
56	Administrative and Supportive Waste Management and Remediation Services
561	Administrative and Support Services
562	Waste Management and Remediation Services
61	Educational Services
611	Educational Services
62	Health Care and Social Assistance
621	Ambulatory Health Care Services
622	Hospitals
623	Nursing and Residential Care Facilities
624	Social Assistance
71	Arts, Entertainment, and Recreation
711	Performing Arts Companies
712	Museums, Historical Sites, and Similar Institutions
713	Amusement, Gambling, and Recreation Industries
72	Accommodation and Food Services
721	Accommodation
722	Food Services and Drinking Places
81	Other Services (except Public Administration)
811	Repair and Maintenance
812	Personal and Laundry Services
813	Religious, Grantmaking, Civic, Professional, and Similar Organizations
814	Private Households
92	Public Administration
921	Executive, Legislative, and Other General Government Support
922	Justice, Public Order, and Safety Activities
923	Administration of Human Resources Programs
924	Administration of Environmental Quality
925	Administration of Housing Programs, Urban Planning, and Community Development
926	Administration of Economic Programs
927	Space Research and Technology
928	National Security and International Affairs

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

APPENDIX D

**PARAMETER ESTIMATES AND STANDARD ERRORS FOR
NONRESPONSE MODELS AND WEIGHTING CLASS ADJUSTMENTS IN CROSS-
SECTIONAL SAMPLES**

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

TABLE D.1

VARIABLES IN THE LOCATION LOGISTIC PROPENSITY MODEL: REPRESENTATIVE BENEFICIARY SAMPLE^a

Variables in the Beneficiary Location Model		
Main Effects	Parameter Estimate ^b	Standard Error
Primary Diagnosis Classification (DIG_1):		
Beneficiary had mental disability	0.297 [†]	0.615
Beneficiary had physical disability (including deaf cases)	0.873	0.446
Beneficiary's disability unknown	Ref. cell	
Identity of payee relative to beneficiary (REPREPAYEE_1):		
Beneficiary received benefit payments himself/herself, or institution received benefits on behalf of beneficiary, or information unknown	Ref. cell	
A family member received benefits on behalf of the beneficiary	1.839 [†]	0.410
Gender (SEX):		
Male	Ref. cell	
Female	0.656 [†]	0.250
Urbanicity of place of residence of beneficiary (METRO_1):		
Beneficiary resides in metropolitan statistical area (MSA)	Ref. cell	
Beneficiary resides in nonmetropolitan area adjacent to large metropolitan area	1.191**	0.415
Beneficiary resides in nonmetropolitan area adjacent to small metropolitan area	1.166**	0.421
Beneficiary resides in nonmetropolitan area not adjacent to metropolitan area	0.806*	0.409
Geographic region (based on U.S. Census regions) of beneficiary's place of residence (DIVISION_1):		
South Atlantic	0.404 [†]	0.0.405
All other divisions	Ref. cell	
Geographic region (based on U.S. Census divisions) of beneficiary's place of residence (REGION_1):		
Midwest	-0.575 [†]	0.424
West	1.926 [†]	0.522
Northeast/South	Ref. cell	
Beneficiary's status (SSI_SSDI):		
SSI only	0.272 [†]	0.515
SSDI only	1.165 [†]	0.484
Both SSI /SSDI	Ref. cell	
Race of the beneficiary (Race_1) :		
White	0.564 [†]	0.244
Not white or unknown	Ref. cell	
Beneficiary's type of claim (TOC_1):		
Disability claim	1.971 [†]	0.756
Survivor claim, or unknown	Ref. cell	
Beneficiary's age category (AGECAT):		
Age in range 18 to 29 years	-0.003 [†]	0.386
Age in range 30 to 39 years	-0.002 [†]	0.387
Age in range 40 to 49 years	0.656 [†]	0.473
Age in range 50 to 64 years	Ref. cell	
Indicator whether beneficiary and applicant for benefits are in same zip code (PDZIPSAME_1):		
Applicant and beneficiary live in same zip code	-0.065 [†]	0.368
Applicant and beneficiary live in different zip code, or no information	Ref. cell	

TABLE D.1 (continued)

Variables in the Beneficiary Location Model		
Main Effects	Parameter Estimate ^b	Standard Error
Two-factor interactions^c		
DIVISION_1*RACE_1		
South Atlantic*White	-0.878*	0.442
SSI_SSDI*DIVISION		
SSD only*South Atlantic	-1.157*	0.560
DIG_1*PDZIPSAME_1		
Beneficiary had mental disability*Applicant and beneficiary live in same zip code	0.617	0.463
DIG_1*TOC_1		
Beneficiary had mental disability*Disability claim	-1.650**	0.568
REGION_1*DIG_1		
Midwest*Beneficiary had mental disability	2.089	1.026
SSI_SSDI *REGION_1		
SSD only*West	-1.756**	0.547
SSI_SSDI *REPREPAYEE_1		
SSI only *A family member received benefits on behalf of the beneficiary	-1.569**	0.509
REGION_1*SEX		
West*Female	-0.878	0.449
AGECAT*TOC_1		
Age in range 18 to 39 years*Disability claim	-0.723	0.580
Age in range 40 to 49 years*Disability claim	-2.305**	0.629
AGECAT*DIG_1		
Age in range 18 to 39 years*Beneficiary had mental disability	0.153	0.506
Age in range 40 to 49 years*Beneficiary had mental disability	1.374*	0.643
AGECAT*REGION_1		
Age in range 18 to 29 years*West	-1.520**	0.504
Age in range 30 to 39 years*West	-1.216*	0.480

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_1”.

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells.

TABLE D.2

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODEL: REPRESENTATIVE
BENEFICIARY SAMPLE^a

Variables in the Beneficiary Cooperation Model		
Main Effects	Parameter Estimate ^b	Standard Error
Beneficiary's age category (AGECAT_2):		
Age in range 18 to 29 years	0.223	0.182
Age in range 30 to 39 years	-0.295*	0.138
Age in range 40 to 64 years	Ref.cell	
Race of the beneficiary (RACE_2):		
White	-0.083†	0.281
Black	0.353†	0.273
Asian or Pacific islander	-1.617**	0.515
Others/Unknown	Ref. cell	
Whether the beneficiary was Hispanic or not (HISPANICITY):		
Hispanic	1.496**	0.401
Nonhispanic/Unknown	Ref. cell	
Urbanicity of place of residence of beneficiary (METRO_2):		
Beneficiary resides in metropolitan statistical area (MSA) of 1 million or more residents	Ref. cell	
Beneficiary resides in metropolitan area of 250,000 to 1 million residents	-0.086†	0.219
Beneficiary resides in metropolitan area of less than 250,000 to residents	2.286†	0.606
Beneficiary resides in nonmetropolitan area	0.640†	0.565
Geographic region (based on U.S. Census divisions) of beneficiary's place of residence (DIVISION_2):		
New England	0.352†	0.469
Middle Atlantic	0.614†	0.403
South Atlantic	-0.091†	0.346
West South Central	-0.511†	0.830
Mountain	0.546	0.421
Pacific	-0.142	0.301
East North Central/East South Central/West North Central.	Ref. cell	
Geographic region (based on U.S. Census regions) of beneficiary's residence (REGION_2):		
Midwest	0.429†	0.345
West/South/Northeast	Ref. cell	
Beneficiary's status (SSI_SSDI):		
SSI only	-0.862†	0.385
SSDI only	-1.779†	0.493
Both SSI and SSDI	Ref. cell	
Gender (SEX):		
Male	Ref. cell	
Female	-0.050†	0.209
Identity of payee relative to beneficiary (REPREPAYEE_2):		
Beneficiary received benefit payments himself/herself, or institution received benefits on behalf of beneficiary, or information unknown	Ref. cell	
A family member received benefits on behalf of the beneficiary	0.008†	0.228

TABLE D.2 (continued)

Variables in the Beneficiary Cooperation Model		
Main Effects	Parameter Estimate ^b	Standard Error
Indicator whether beneficiary and applicant for benefits are in same zip code (PDZIPSAME_2):		
Applicant and beneficiary live in same zip code	0.333†	0.208
Applicant and beneficiary live in different zip code, or no information	Ref. cell	
Primary Diagnosis Classification (DIG_2):		
Beneficiary had mental disability	-0.671†	0.351
Beneficiary had physical disability	0.332†	0.374
Deaf/Beneficiary's disability unknown	Ref. cell	
Beneficiary's type of claim (TOC_2):		
Disability claim	-0.288†	0.267
Survivor claim, or unknown	Ref. cell	
Beneficiary's living situation (LIVING_2)		
Beneficiary lives alone	-1.369**	0.443
Beneficiary lives with others	-1.821**	0.526
Other/Unknown	Ref. cell	
Two-factor interactions^c		
DIG_2*RACE_2		
Beneficiary had mental disability *White	0.604*	0.302
DIVISION_2* RACE_2		
South Atlantic* White	0.867*	0.382
SSI_SSDI * RACE_2		
SSD only*Black	0.676	0.409
REGION_2* PDZIPSAME_2		
Midwest* Applicant and beneficiary live in same zip code	-0.640	0.361
DIVISION_2* PDZIPSAME_2		
West South Central * Applicant and beneficiary live in same zip code	-1.203*	0.558
DIG_2* PDZIPSAME_2		
Beneficiary had physical disability * Applicant and beneficiary live in same zip code	0.480	0.295
METRO_2* TOC_2		
Beneficiary resides in metropolitan area of less than 250,000 to residents * Disability claim	-1.799**	0.607
Beneficiary resides in nonmetropolitan area *Disability claim	-1.019	0.543
DIVISION_2*METRO_2		
West South Central * Beneficiary resides in nonmetropolitan area	0.908	0.523
New England*Beneficiary resides in metropolitan area of less than 250,000 to residents	-1.034	0.546
DIG_2*DIVISION_2		
Beneficiary had physical disability*New England	-0.781	0.574
TOC_2* DIVISION_2		
Disability claim *West South Central	2.468**	0.761
SSI_SSDI* DIVISION_2		
SSI only* West South Central	1.496	0.929
SSDI only* West South Central	-1.534	0.861
SSDI only*South Atlantic	-0.753	0.389
SSI_SSDI* REGION_2		
SSI only* MIDWEST	0.719	0.390
SSI_SSDI* METRO_2		
SSI only* Beneficiary resides in small/nonmetropolitan area	-1.600**	0.522

TABLE D.2 (continued)

Variables in the Beneficiary Cooperation Model		
Main Effects	Parameter Estimate ^b	Standard Error
METRO_2*SEX		
Beneficiary resides in small metropolitan area*Sex	-1.067*	0.496
Beneficiary resides in nonmetropolitan area*Sex	0.333	0.392
DIVISION_2*SEX		
Middle Atlantic*SEX	-1.115*	0.448
PDZIPSAME_2*METRO_2		
Applicant and beneficiary live in same zip code* Beneficiary resides in nonmetropolitan area	0.906*	0.432
REPREPAYEE_2*METRO_2		
A family member received benefits on behalf of the beneficiary* Beneficiary resides in small metropolitan area	0.785*	0.373
A family member received benefits on behalf of the beneficiary* Beneficiary resides in non metropolitan area	0.500	0.360
DIG_2*SSI_SSDI		
Beneficiary had physical disability *SSI only	-0.849**	0.323
SSI_SSDI*SEX		
SSI only*Sex	0.554	0.295

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_2”.

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells.

TABLE D.3

VARIABLES USED IN THE LOCATION WEIGHTING CLASS ADJUSTMENT: TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 2 MILESTONE-OUTCOME PAYMENT SYSTEM^a

Variables Defining the Location Weighting Class Adjustment for Phase 2 Participants Using the Milestone-Outcome Payment System	
Main Effects	Weighted Count
Participant recipient benefit type (SSI_SSDI_3)	
Both SSI and SSDI	1,144
SSI only, or SSDI only	2,944
Participant’s age category (AGECAT):	
Age in range 18 to 29 years	1,169
Age in range 30 to 39 years	906
Age in range 40 to 49 years	1,100
Age in range 50 to 64 years	963
Participant’s gender (SEX)	
Male	2,061
Female	2,077

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_3”.

TABLE D.4

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 2 MILESTONE-OUTCOME PAYMENT SYSTEM^a

Variables in the Cooperation Model for Phase 2 Participants Using the Milestone-Outcome Payment System		
Main Effects	Parameter Estimate ^b	Standard Error
Identity of payee relative to participant (REPREPAYEE_4):		
Participant received benefit payments himself/herself	0.098†	0.402
Participant did not receive benefit payments himself/herself, or information unknown	Ref. cell	
Number of moves in past 5 years (MOVE_4):		
No moves	0.914*	0.389
One or more moves, or old information, or no information about moves	Ref. cell	
Participant's living situation (LIVING_4):		
Participant lives alone	1.173*	0.486
Participant does not live on his or her own, or unknown	Ref. cell	
Participant recipient benefit type (SSI_SSDI_4)		
SSDI only	1.602	0.635
SSI only, or both SSI and SSDI	Ref. cell	
Geographic region (based on U.S. Census divisions) of participant's place of residence (DIVISION_4):		
West South Central	0.216†	0.386
East North Central	0.789*	0.354
New England, Middle Atlantic, South Atlantic, Mountain, East South Central, West North Central, Pacific.	Ref. cell	
Geographic region (based on U.S. Census regions) of participant's residence (REGION_4):		
South	0.736	0.405
Northeast, Midwest, or West	Ref. cell	
Participant's type of claim (TOC_4):		
Disability claim	-0.841†	0.445
Survivor claim, or unknown	Ref. cell	
Disability diagnosis classification (DIG_4):		
Participant had physical disability (excluding deaf cases)	1.593†	0.572
Participant had mental disability	0.577†	0.446
Participant was deaf, or information about disability not given	Ref. cell	
Race of the participant (RACE_4):		
Black	-0.460†	0.342
Race known not to be black, or unknown	Ref. cell	
Participant's age category (AGECAT_4):		
Age in range 18 to 29 years	-0.469†	0.344
Age in range 30 to 64 years	Ref. cell	

TABLE D.4 (continued)

Variables in the Cooperation Model for Phase 2 Participants Using the Milestone-Outcome Payment System		
Main Effects	Parameter Estimate ^b	Standard Error
Two-factor interactions^c		
DIG_4* REPREPAYEE_4		
Participant had mental disability * Participant received benefit payments himself/herself	1.095*	0.529
TOC_4* AGECAT_4		
Disability claim * Age in range 18 to 29 years	0.914	0.678
TOC_4* RACE_4		
Disability claim * Black	1.060*	0.539
DIG_4* DIVISION_4		
Participant had physical disability (excluding deaf cases)* Participant sampled in Round 1	-0.917	0.546

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_3”.

^b If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_4”.

^c Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

TABLE D.5

VARIABLES USED IN THE LOCATION WEIGHTING CLASS ADJUSTMENT: TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 2 OUTCOME-ONLY PAYMENT SYSTEM

Variables Defining the Location Weighting Class Adjustment for Phase 2 Participants Using the Outcome-Only Payment System	
Main Effects	Weighted Count
Participant's age category (AGECAT):	
Age in range 18 to 29 years	105
Age in range 30 to 39 years	257
Age in range 40 to 49 years	472
Age in range 50 to 64 years	358
Participant's gender (SEX)	
Male	565
Female	627

TABLE D.6

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 2 OUTCOME-ONLY PAYMENT SYSTEM^a

Variables in the Cooperation Model for Phase 2 Participants Using the Outcome-Only Payment System		
Main Effects	Parameter Estimate ^b	Standard Error
Race of the participant (RACE_6):		
White	-0.489	0.351
Black	-0.976*	0.436
Race known to be neither white nor black, or unknown	Ref. cell	
Participant's living situation (LIVING_6):		
Participant lives alone	0.467	0.343
Participant does not live on his or her own, or unknown	Ref. cell	
Geographic region (based on U.S. Census divisions) of participant's place of residence (DIVISION_6):		
East North Central	-0.786*	0.379
New England, Middle Atlantic, South Atlantic, Mountain, East South Central, West North Central, Pacific, East North Central.	Ref. cell	
Geographic region (based on U.S. Census regions) of participant's residence (REGION_6):		
Northeast	0.503†	0.381
South, Midwest, or West	Ref. cell	
Participant's gender (SEX)		
Male	Ref. cell	
Female	1.067†	0.464
Participant recipient benefit type (SSI_SSDI_6)		
SSI only	-0.754	0.398
SSDI only, or both SSI and SSDI	Ref. cell	
Number of phone numbers on SSA file over past five years (PHONE_6):		
No phone changes on SSA file over past five years	-1.344†	0.594
One or more changes in phone number on SSA file, or unknown	Ref. cell	
Indicator whether participant and applicant for benefits are in same zip code (PDZIPSAME_6):		
No information	-0.238	0.193
Applicant and participant live in different/same zip code	Ref. cell	
Participant's age category (AGECAT_6):		
Age in range 18 to 39 years	0.164†	0.505
Age in range 40 to 49 years	-0.255†	0.432
Age in range 50 to 64 years	Ref. cell	
Whether participant was sampled in Round 2 (R2LONG)		
Participant was not sampled in Round 2	Ref. cell	
Participant was sampled in Round 2	1.637†	0.878
Urbanicity of place of residence of beneficiary (METRO_6):		
Beneficiary resides in metropolitan statistical area (MSA) of 1 million or more residents	Ref. cell	
Beneficiary resides in metropolitan area of less than 1 million residents	0.503†	0.434
Beneficiary resides in nonmetropolitan area	-0.559†	0.351

TABLE D.6 (continued)

Variables in the Cooperation Model for Phase 2 Participants Using the Outcome-Only Payment System		
Main Effects	Parameter Estimate ^b	Standard Error
Two-factor interactions^c		
REGION_6* PHONE_6		
Northeast * No phone changes on SSA file over past five years	1.462	0.763
R2LONG * AGECA6_6		
Participant was sampled in Round 2* Age in range 18 to 39 years	-0.588	0.492
Participant was sampled in Round 2* Age in range 40 to 49 years	0.707	0.500
R2LONG * REGION_6		
Participant was sampled in Round 2* Northeast	-2.161†	1.028
R2LONG * SEX		
Participant was sampled in Round 2* Female	-2.241†	0.738
REGION_6* AGECA6_6		
Northeast* Age in range 18 to 39 years	-0.876*	0.443
AGECA6_6 * R2LONG		
Age in range 18 to 39 years * Participant was sampled in Round 2	-0.588	0.492
AGECA6_6 * SEX		
Female * Age in range 18 to 39 years	1.188**	0.442
METRO_6* AGECA6_6		
Beneficiary resides in metropolitan area of less than 1 million residents * Age in range 18 to 39 years	-0.639	0.490
Beneficiary resides in metropolitan area of less than 1 million residents * Age in range 40 to 49 years	-0.990	0.593
Beneficiary resides in nonmetropolitan area * Age in range 40 to 49 years	0.576	0.580
REGION_6* SEX		
Northeast * Female	-0.794†	0.576
Three-factor interactions^d		
R2LONG * SEX * REGION_6		
Participant was sampled in Round 2* Female* Northeast	1.866*	0.888

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_6”

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells

^d All combinations for the listed interactions that are not shown are part of the reference cells

TABLE D.7

VARIABLES IN THE LOCATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 2 TRADITIONAL PAYMENT SYSTEM^a

Variables in the Location Model for Phase 2 Participants Using the Traditional Payment System		
Main Effects	Parameter Estimate ^b	Standard Error
Participant's living situation (LIVING_7):		
Participant lives alone	-0.588	0.340
Participant does not live on his or her own, or unknown	Ref. cell	
Participant recipient benefit type (SSI_SSDI_7)		
SSDI only	0.754	0.431
SSI only, or both SSI and SSDI	Ref. cell	
Urbanicity of participant's place of residence (METRO_7):		
Participant lived in metropolitan area with population of 250,000 or more	Ref. cell	
Participant did not live in metropolitan area with population of 250,000 or more	-0.621	0.474
Participant's type of claim (TOC_7):		
Disability claim	-0.584	0.547
Survivor claim, or unknown	Ref. cell	
Indicator whether participant and applicant for benefits are in same zip code (PDZIPSAME_7):		
Applicant and participant live in same zip code	0.912	0.615
Applicant and participant live in different zip code, or no information	Ref. cell	

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_7”

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

TABLE D.8

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 2 TRADITIONAL PAYMENT SYSTEM^a

Variables in the Cooperation Model for Phase 2 Participants Using the Traditional Payment System		
	Parameter Estimate ^b	Standard Error
Main Effects		
Participant recipient benefit type (SSI_SSDI_8)		
SSDI	0.999	0.548
SSI only or both SSI/SSDI	Ref. cell	
Geographic region (based on U.S. Census divisions) of beneficiary's place of residence (DIVISION_8):		
West North Central	0.594	0.326
East North Central/East South Central/ Middle Atlantic/ South Atlantic/ West South Central/ Mountain/ Pacific/New England.	Ref. cell	
Participant's living situation (LIVING_8):		
Participant lives alone	1.134	0.551
Participant does not live on his or her own, or unknown	Ref. cell	
Disability diagnosis classification (DIG_8):		
Participant had physical disability (excluding deaf cases)	0.742*	0.357
Participant had mental disability	0.240†	0.344
Participant was deaf, or information about disability not given	Ref. cell	
Participant's type of claim (TOC_8):		
Disability claim	0.382†	0.359
Survivor claim, or unknown	Ref. cell	
Participant's age category (AGECAT):		
Age in range 18 to 29 years	0.201†	0.390
Age in range 30 to 39 years	-0.676	0.382
Age in range 40 to 49 years	-0.384	0.301
Age in range 50 to 64 years	Ref. cell	
Identity of payee relative to participant (REPREPAYEE_8):		
Participant received benefit payments himself/herself, or institution received benefits on behalf of beneficiary, or information unknown	Ref. cell	
A family member received benefits on behalf of the participant	-0.576†	0.385
Urbanicity of participant's place of residence (METRO_8):		
Participant lived in metropolitan area with population of 1 million or more	Ref. cell	
Participant lived in metropolitan area with population less than 1 million	0.742*	0.278
Participant lived in nonmetropolitan area	0.226	0.158
Two-factor interactions^c		
DIG_8 * REPREPAYEE_8		
Participant had mental disability * A family member received benefits on behalf of the participant	2.066**	0.514
TOC_8 * AGECAT_8		
Disability claim * Age in range 18 to 39 years	-0.982*	0.471

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_8”

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells

TABLE D.9

VARIABLES USED IN THE LOCATION WEIGHTING CLASS ADJUSTMENT: TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 3 MILESTONE-OUTCOME PAYMENT SYSTEM^a

Variables Defining the Location Weighting Class Adjustment for Phase 3 Participants Using the Milestone-Outcome Payment System	
Main Effects	Weighted Count
Indicator whether beneficiary and applicant for benefits are in same zip code (PDZIPSAME_9):	
Applicant and beneficiary live in different zip code	380
Applicant and beneficiary live in same zip code, or no information	4,030
Participant's gender (SEX)	
Male	2,168
Female	2,242

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_9”.

TABLE D.10

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 3 MILESTONE-OUTCOME PAYMENT SYSTEM^a

Variables in the Cooperation Model for Phase 3 Participants Using the Milestone-Outcome Payment System		
Main Effects	Parameter Estimate ^b	Standard Error
Geographic region (based on U.S. Census regions) of beneficiary's residence (REGION_10):		
West	-0.661†	0.292
Midwest/South/Northeast	Ref. cell	
Race of the participant (RACE_10):		
Black	-0.565*	0.260
Race known not to be black, or unknown	Ref. cell	
Participant's gender (SEX)		
Male	Ref. cell	
Female	0.877†	0.332
Disability diagnosis classification (DIG_10):		
Participant had physical disability	0.624	0.374
Participant had mental disability or deafness, or information about disability not given	Ref. cell	
Identity of payee relative to participant (REPREPAYEE_10):		
Family member received benefits on behalf of participant	-0.510	0.401
All other payees (including those with unknown payee identity)	Ref. cell	

Variables in the Cooperation Model for Phase 3 Participants Using the Milestone-Outcome Payment System

Main Effects	Parameter Estimate ^b	Standard Error
Two-factor interactions^c		
REGION_10* SEX		
West * Female	-1.270**	0.446

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_10”

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells

TABLE D.11

VARIABLES USED IN THE LOCATION WEIGHTING CLASS ADJUSTMENT: TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 3 OUTCOME-ONLY PAYMENT SYSTEM^a

Variables Defining the Location Weighting Class Adjustment for Phase 3 Participants Using the Outcome-Only Payment System

Main Effects	Weighted Count
Participant recipient benefit type (SSI_SSDI_11)	
Both SSI and SSDI	149
SSI only or SSDI only	551
Race of the participant (RACE_11):	
Black	180
Race known not to be black, or unknown	520

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_11”.

TABLE D.12

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT CROSS-SECTIONAL SAMPLE, PHASE 3 OUTCOME-ONLY PAYMENT SYSTEM^a

Variables in the Cooperation Model for Phase 3 Participants Using the Outcome-Only Payment System		
Main Effects	Parameter Estimate ^b	Standard Error
Beneficiary's age category (AGECAT):		
Age in range 18 to 29 years	0.818	0.516
Age in range 30 to 39 years	0.372	0.465
Age in range 40 to 49 years	0.826	0.451
Age in range 50 to 64 years	Ref. cell	
Participant's gender (SEX)		
Male	Ref. cell	
Female	1.074†	0.498
Geographic region (based on U.S. Census regions) of participant's residence (REGION_12):		
West	1.038†	0.517
Northeast or Midwest or South	Ref. cell	
Urbanicity of participant's place of residence (METRO_12):		
Participant lived in metropolitan area with population from 250,000 or more	-0.619	0.415
Participant lived not in metropolitan area with population from 250,000 or more	Ref. cell	
Two-factor interactions^c		
REGION_12* SEX		
West* Female	-2.855**	0.001

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_12”

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells

TABLE D.13

VARIABLES USED IN THE LOCATION WEIGHTING CLASS ADJUSTMENT: TICKET PARTICIPANT
CROSS-SECTIONAL SAMPLE, PHASE 3 TRADITIONAL PAYMENT SYSTEM^a

Variables Defining the Location Weighting Class Adjustment for Phase 3 Participants Using the Traditional Payment System	
Main Effects	Weighted Count
Participant's type of claim (TOC_13):	
Disability claim	16,308
Survivor claim, or unknown	9,605
Race of the participant (RACE_13):	
White	15,630
Race known not to be white, or unknown	10,283

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_13”.

TABLE D.14

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT CROSS-
SECTIONAL SAMPLE, PHASE 3 TRADITIONAL PAYMENT SYSTEM^a

Variables in the Cooperation Model for Phase 3 Participants Using the Traditional Payment System		
Main Effects	Parameter Estimate ^b	Standard Error
Race of the beneficiary (RACE_14) :		
White	0.705*	0.259
Not white or unknown	Ref. cell	
Geographic region (based on U.S. Census regions) of participant's residence (REGION_14):		
West	-0.981**	0.253
Northeast, Midwest, South,	Ref. cell	
Identity of payee relative to participant (REPREPAYEE_14):		
Participant received benefit payments himself/herself	1.060†	0.683
All other payees (including those with unknown payee identity)	Ref. cell	
Participant's living situation (LIVING_14):		
Participant lives alone	1.227†	0.735
Participant does not live on his or her own, or unknown	Ref. cell	
Two-factor interactions^c		
REPREPAYEE_14 * LIVING_14		
Participant received benefit payments himself/herself * Participant lives alone	2.932*	1.393

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_14”

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

APPENDIX E

PARAMETER ESTIMATES AND STANDARD ERRORS FOR NONRESPONSE MODELS IN LONGITUDINAL SAMPLES

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

TABLE E.1

VARIABLES IN THE LOCATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 MILESTONE-OUTCOME PAYMENT TYPE, RESPONSES IN ROUNDS 1, 2, AND 3^a

Variables in the Location Model for Phase 1 Participants Using the Milestone-Outcome Payment Type Who Responded in Rounds 1, 2, and 3		
Main Effects	Parameter Estimate ^b	Standard Error
Number of moves in past 5 years (MOVE_1):		
No moves	-0.353	0.293
One, two, three, or four moves, old information, or no information about moves	Ref. cell	
Indicator whether participant and applicant for benefits are in same zip code (PDZIPSAME_1):		
Applicant and participant live in different zip code	-1.180**	0.255
Applicant and participant live in same zip code, or information unknown	Ref. cell	
Urbanicity of place of residence of participant (METRO_1):		
Participant resides in metropolitan statistical area with population over 1 million	-2.213**	0.660
Participant not resides in metropolitan statistical area with population over 1 million	Ref. cell	
Race of the participant (RACE_1):		
White	0.408	0.498
Black	0.588	0.305
Race known to be neither White nor Black, or unknown	Ref. cell	
Participant's age category (AGECAT_1):		
Age in range 18 to 29 years	-0.541	0.281
Age in range 30 to 64 years	Ref. cell	

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_1”.

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

TABLE E.2

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT
LONGITUDINAL SAMPLE, PHASE 1 MILESTONE-OUTCOME PAYMENT TYPE, RESPONSES IN ROUNDS 1, 2,
AND 3^a

Variables in the Cooperation Model for Phase 1 Participants Using the Milestone-Outcome Payment Type Who Responded in Rounds 1, 2, and 3		
Main Effects	Parameter Estimate ^b	Standard Error
Number of address changes in the past five years (MOVE_2):		
No move	-0.501	0.341
One or more moves, or information older than five years, or no information	Ref. cell	
Indicator whether participant and applicant for benefits are in same zip code (PDZIPSAME_2):		
Applicant and participant live in same zip code	0.167†	0.363
Applicant and participant live in different zip code, or no information	Ref. cell	
Participant recipient benefit type (SSI_SSDI_2)		
SSDI only	0.503†	0.527
SSI only, or both SSI and SSDI	Ref. cell	
Urbanicity of participant's place of residence (METRO_2):		
Participant lived in metropolitan area with population of 1 million or more	-0.551	0.361
Participant not lived in metropolitan area with population of 1 million or more	Ref. cell	
Race of the participant (RACE_2):		
Black	0.504†	0.373
Race known to not be black, or unknown	Ref. cell	
Participant's age category (AGECAT_2):		
Age in range 18 to 39 years	-1.043†	0.430
Age in range 40 to 49 years	-0.556	0.295
Age in range 50 to 64 years	Ref. cell	
Disability diagnosis classification (DIG_2):		
Participant had mental disability	1.145†	0.449
Participant had physical disability, or unknown	Ref. cell	
Two-factor interactions^c		
SSI_SSDI_2* DIG_2		
SSDI only * Participant had mental disability	-1.613*	0.708
AGECAT_2* PDZIPSAME_2		
Age in range 18 to 39 years * Applicant and participant live in same zip code	1.137*	0.441

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_2”.

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells.

TABLE E.3

VARIABLES IN THE LOCATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 OUTCOME-ONLY PAYMENT TYPE, RESPONSES IN ROUNDS 1, 2, AND 3^a

Variables in the Location Model for Phase 1 Participants Using the Outcome-Only Payment Type Who Responded in Rounds 1, 2, and 3		
Main Effects	Parameter Estimate ^b	Standard Error
Number of moves in past 5 years (MOVE_3):		
No moves	-1.290†	0.646
One or more moves, old information, or no information about moves	Ref. cell	
Participant's living situation (LIVING_3):		
Participant lives alone	2.972**	1.015
Participant does not live on his or her own, or unknown	Ref. cell	
Race of the participant (RACE_3):		
White	-1.228†	0.767
Black	-1.146	0.717
Race known to be neither white or black, or unknown	Ref. cell	
Participant's age category (AGECAT_3):		
Age in range 18 to 49 years	-1.17†	0.624
Age in range 50 to 64 years	Ref. cell	
Geographic region (based on U.S. Census divisions) of participant's place of residence (DIVISION_3):		
Pacific	-2.347†	0.895
New England, Middle Atlantic, South Atlantic, Mountain, East South Central, West North Central, West South Central, East North Central	Ref. cell	
Participant recipient benefit type (SSI_SSDI_3)		
SSDI only	1.533†	1.021
SSI only, or both SSI and SSDI	Ref. cell	
Geographic region (based on U.S. Census regions) of participant's residence (REGION_3):		
South	1.548†	0.810
Northeast, Midwest, or West	Ref. cell	
Urbanicity of participant's place of residence (METRO_3):		
Participant lived in metropolitan area with population of 25,000 or more	Ref. cell	
Participant did not live in metropolitan area with population of 25,000 or more	1.881*	0.807
Two-factor interactions^c		
MOVE_3 * SSI_SSDI_3		
No moves * SSDI only	2.038	1.304
REGION_3 * RACE_3		
South * White	-1.983	1.115
SSI_SSDI_3 * RACE_3		
SSDI only * White	1.899*	0.979
AGECAT_3 * DIVISION_3		
Age in range 18 to 49 years * Pacific	2.142*	1.086

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_3”.

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells.

TABLE E.4

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT
LONGITUDINAL SAMPLE, PHASE 1 OUTCOME-ONLY PAYMENT TYPE, RESPONSES IN ROUNDS 1, 2, AND 3^a

Variables in the Cooperation Model for Phase 1 Participants Using the Outcome-Only Payment Type Who Responded in Rounds 1, 2, and 3		
	Parameter Estimate ^b	Standard Error
Main Effects		
Number of moves in past 5 years (MOVE_4):		
No moves	-1.222†	0.419
One or more moves, old information, or no information about moves	Ref. cell	
Race of the participant (RACE_4):		
White	0.632*	0.315
Black	0.818	0.461
Race known to be neither white nor black, or unknown	Ref. cell	
Participant's age category (AGECAT_4):		
Age in range 18 to 49 years	0.142†	0.373
Age in range 50 to 64 years	Ref. cell	
Indicator whether participant and applicant for benefits are in the same zip code (PDZIPSAME_4):		
Applicant and participant live in the same zip code	0.565	0.332
Applicant and participant live in different zip code, or no information	Ref. cell	
Participant recipient benefit type (SSI_SSDI_4)		
SSDI only	1.567	0.799
SSI only, or both SSI and SSDI	Ref. cell	
Geographic region (based on U.S. Census regions) of participant's residence (REGION_4):		
South	-1.265†	0.352
West	0.490	0.370
Northeast, or Midwest	Ref. cell	
Disability diagnosis classification (DIG_4):		
Participant had physical disability (excluding deaf cases)	1.263†	0.452
Participant was deaf, or mental disability, or information about disability not given	Ref. cell	
Participant's living situation (LIVING_4):		
Participant lives alone	1.981*	0.773
Participant does not live on his or her own, or unknown	Ref. cell	
Two-factor interactions^c		
MOVE_4 * REGION_4		
No moves * South	1.471*	0.700
DIG_4* AGECAT_4		
Participant had physical disability * Age in range 18 to 49 years	-1.246*	0.585

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_4”.

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells.

TABLE E.5

VARIABLES IN THE LOCATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 TRADITIONAL PAYMENT TYPE, RESPONSES IN ROUNDS 1, 2, AND 3^a

Variables in the Location Model for Phase 1 Participants Using the Traditional Payment Type Who Responded in Rounds 1, 2, and 3		
Main Effects	Parameter Estimate ^b	Standard Error
Participant's gender (SEX)		
Male	Ref. cell	
Female	-0.269†	0.445
Participant recipient benefit type (SSI_SSDI_5)		
SSI only	-0.552	0.400
SSDI only, or both SSI and SSDI	Ref. cell	
Geographic region (based on U.S. Census regions) of participant's residence (REGION_5):		
West	-1.433*	0.612
Midwest, Northeast or South	Ref. cell	
Race of the participant (RACE_5):		
Black	-1.412†	0.374
Race known to not be black, or unknown	Ref. cell	
Participant's age category (AGECAT):		
Age in range 18 to 29 years	-0.546	0.624
Age in range 30 to 39 years	-0.987	0.503
Age in range 40 to 49 years	-0.987	0.631
Age in range 50 to 64 years	Ref. cell	
Two-factor interactions^c		
RACE_5 * SEX		
Black * Female	1.702	0.878

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_5”.

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells.

TABLE E.6

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 TRADITIONAL PAYMENT TYPE, RESPONSES IN ROUNDS 1, 2, AND 3^a

Variables in the Cooperation Model for Phase 1 Participants Using the Traditional Payment Type Who Responded in Rounds 1, 2, and 3		
Main Effects	Parameter Estimate ^b	Standard Error
Disability diagnosis classification (DIG_6):		
Participant had physical disability (excluding deaf cases)	1.249†	0.336
Participant was deaf, or mental disability, or information about disability not given	Ref. cell	
Indicator whether participant and applicant for benefits are in the same zip code (PDZIPSAME_6):		
Applicant and participant live in the same zip code	0.036†	0.194
Applicant and participant live in different zip code, or no information	Ref. Cell	
Participant's age category (AGECAT_6):		
Age in range 18 to 49 years	0.428†	0.342
Age in range 50 to 64 years	Ref. cell	
Urbanicity of participant's place of residence (METRO_6):		
Participant lived in metropolitan area	Ref. cell	
Participant lived in nonmetropolitan area	1.142**	0.362
Geographic region (based on U.S. Census regions) of participant's residence (REGION_6):		
South, West, Midwest	Ref. cell	
Northeast	-1.464†	0.373
Identity of payee relative to participant (REPREPAYEE):		
Participant received benefit payments himself/herself	-1.669**	0.459
A family member received benefits on behalf of the participant	-1.555**	0.487
Information unknown	Ref. cell	
Race of the participant (RACE_6):		
White	-0.487†	0.410
Race known to not be white, or unknown	Ref. cell	
Two-factor interactions^c		
REGION_6 * PDZIPSAME_6		
Northeast * Applicant and participant live in the same zip code	1.280**	0.222
REGION_6 * RACE_6		
Northeast * White	1.155	0.656
DIG_6 * AGECAT_6		
Participant had physical disability * Age in range 18 to 49 years	-0.862*	0.401

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_6”.

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells.

TABLE E.7

VARIABLES IN THE LOCATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 MILESTONE-OUTCOME PAYMENT TYPE, RESPONSES IN ROUNDS 1 AND 2^a

Variables in the Location Model for Phase 1 Participants Using the Milestone-Outcome Payment Type Who Responded in Rounds 1 and 2		
Main Effects	Parameter Estimate ^b	Standard Error
Urbanicity of participant's place of residence (METRO_7):		
Participant lived in metropolitan area with population of 1 million or more	-1.912**	0.659
Participant not lived in metropolitan area with population of 1 million or more	Ref. cell	
Participant's age category (AGECAT_7):		
Age in range 18 to 29 years	-0.818**	0.310
Age in range 30 to 64 years	Ref. cell	
Indicator whether participant and applicant for benefits are in the same zip code (PDZIPSAME_7):		
Applicant and participant live in different zip code	-0.960**	0.251
Applicant and participant live in same zip code, or no information	Ref. cell	
Participant's type of claim (TOC_7):		
Disability claim	-0.360†	0.497
Survivor claim, or unknown	Ref. cell	
Geographic region (based on U.S. Census regions) of participant's residence (REGION_7):		
South	-0.606†	0.371
Northeast, Midwest, or West	Ref. cell	
Two-factor interactions^c		
REGION_7* TOC_7		
South* Disability claim	1.437	0.805

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_7”.

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells.

TABLE E.8

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT
LONGITUDINAL SAMPLE, PHASE 1 MILESTONE-OUTCOME PAYMENT TYPE, RESPONSES
IN ROUNDS 1 AND 2^a

Variables in the Cooperation Model for Phase 1 Participants Using the Milestone-Outcome Payment Type Who Responded in Rounds 1 and 2		
	Parameter Estimate ^b	Standard Error
Main Effects		
Number of phone numbers on SSA file over past five years (PHONE_8):		
No phone changes on SSA file over past five years	Ref. cell	
One or more changes in phone number on SSA file, or unknown	-0.362	0.262
Race of the participant (RACE_8):		
White	0.543	0.344
Black	0.943*	0.420
Race known to be neither white nor black, or unknown	Ref. cell	
Indicator whether participant and applicant for benefits are in same zip code (PDZIPSAME_8):		
Applicant and participant live in same zip code	-0.251†	0.390
Applicant and participant live in different zip code, or no information	Ref. cell	
Disability diagnosis classification (DIG_8):		
Participant had mental disability	-0.110†	0.397
Participant had physical disability, or information about disability not given	Ref. cell	
Two-factor interactions^c		
PDZIPSAME_8* DIG_8		
Applicant and participant live in same zip code * Participant had mental disability	1.704*	0.603

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_8”.

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells.

TABLE E.9

VARIABLES IN THE LOCATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 OUTCOME-ONLY PAYMENT TYPE, RESPONSES IN ROUNDS 1 AND 2^a

Variables in the Location Model for Phase 1 Participants Using the Outcome-Only Payment Type Who Responded in Rounds 1 and 2		
	Parameter Estimate ^b	Standard Error
Main Effects		
Participant's age category (AGECAT_9):		
Age in range 18 to 49 years	-1.561†	0.796
Age in range 50 to 64 years	Ref. cell	
Participant's gender (SEX)		
Female	-0.815	0.529
Male	Ref. cell	
Participant recipient benefit type (SSI_SSDI)		
SSI only	-1.706*	0.786
SSDI only	1.591*	0.774
Both SSI and SSDI	Ref. cell	
Geographic region (based on U.S. Census regions) of participant's residence (REGION_9):		
South	1.790*	0.809
West, Northeast, Midwest	Ref. cell	
Geographic region (based on U.S. Census divisions) of participant's place of residence (DIVISION_9):		
Pacific	-2.557†	0.980
New England, Middle Atlantic, South Atlantic, Mountain, East South Central, West North Central, West South Central, East North Central	Ref. cell	
Identity of payee relative to participant (REPREPAYEE_9):		
Family member received benefits on behalf of participant	-0.951	0.745
All other payees (including those with unknown payee identity)	Ref. cell	
Participant's living situation (LIVING_9):		
Participant lives alone	2.832**	0.780
Participant does not live on his or her own, or unknown	Ref. cell	
Urbanicity of place of residence of participant (METRO_9):		
Participant resides in metropolitan statistical area (MSA) with population over 1 million	Ref. cell	
Participant resides in metropolitan statistical area (MSA) with population between 250,000 to 1 million	1.515*	0.697
Participant does not reside in metropolitan area with population over 250,000	2.103**	0.665
Two-factor interactions^c		
DIVISION_9* AGECAT_9		
Pacific * Age in range 18 to 49 years	1.898	1.145

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_9”.

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells.

TABLE E.10

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT
LONGITUDINAL SAMPLE, PHASE 1 OUTCOME-ONLY PAYMENT TYPE, RESPONSES IN ROUNDS 1 AND 2^a

Variables in the Cooperation Model for Phase 1 Participants Using the Outcome-Only Payment Type Who Responded in Rounds 1 and 2		
	Parameter Estimate ^b	Standard Error
Main Effects		
Participant's gender (SEX)		
Male,	Ref. cell	
Female	0.355†	0.298
Number of address changes in the past five years (MOVE_10):		
No move	0.023†	0.479
One or more moves, or information older than five years, or no information	Ref. cell	
Geographic region (based on U.S. Census regions) of participant's residence (REGION):		
South	-1.387**	0.408
West	-0.989†	0.635
Midwest	0.776†	0.501
Northeast	Ref. cell	
Disability diagnosis classification (DIG_10):		
Participant had physical disability (excluding deaf cases)	0.158†	0.349
Participant had mental disability, was deaf, or information about disability not given	Ref. cell	
Identity of payee relative to participant (REPREPAYEE_10):		
Family member received benefits on behalf of participant	-0.738	0.395
Participant received benefit payments himself/herself , or all other payees (including those with unknown payee identity)	Ref. cell	
Participant's living situation (LIVING_10):		
Participant lives alone	2.559**	0.883
Participant does not live on his or her own, or unknown	Ref. cell	
Participant recipient benefit type (SSI_SSDI_10)		
SSDI only	2.733**	0.320
SSI only, or both SSI and SSDI	Ref. cell	
Race of the participant (RACE_10):		
White	0.205†	0.317
Race known not to be white, or unknown	Ref. cell	
Indicator whether participant and applicant for benefits are in same zip code (PDZIPSAME_10):		
Applicant and participant live in different zip code , or no information	Ref. cell	
Applicant and participant live in same zip code	0.854**	0.320
Two-factor interactions^c		
REGION * RACE_10		
West* White	2.663**	0.848
SEX* REGION		
Female*Midwest	-1.423*	0.663
MOVE_10*RACE_10		
Nomove*White	-0.986	0.619

TABLE E.10 (continued)

Variables in the Cooperation Model for Phase 1 Participants Using the Outcome-Only Payment Type Who Responded in Rounds 1 and 2		
Main Effects	Parameter Estimate ^b	Standard Error
DIG_10*REGION		
Participant had physical disability * South	0.992	0.577
Participant had physical disability * West	1.754*	0.870

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_10”.

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells.

TABLE E.11

VARIABLES IN THE LOCATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 1 TRADITIONAL PAYMENT TYPE, RESPONSES IN ROUNDS 1 AND 2^a

Variables in the Location Model for Phase 1 Participants Using the Traditional Payment Type Who Responded in Rounds 1 and 2		
	Parameter Estimate ^b	Standard Error
Main Effects		
Participant's gender (SEX)		
Male	Ref. cell	
Female	2.170†	0.744
Race of the participant (RACE_11):		
Black	-2.370†	0.825
Race known not to be white, or unknown	Ref. cell	
Participant's type of claim (TOC_11):		
Disability claim	0.328†	0.607
Survivor claim, or unknown	Ref. cell	
Disability diagnosis classification (DIG_11):		
Participant had mental disability	-0.097†	0.617
Participant had physical disability, or information about disability not given	Ref. cell	
Geographic region (based on U.S. Census regions) of participant's residence (REGION_11):		
West	-1.710**	0.426
Northeast, Midwest, or South	Ref. cell	
Participant recipient benefit type (SSI_SSDI_11)		
SSI Only	-1.869†	0.800
SSDI Only, or both SSI and SSDI	Ref. cell	
Participant's age category (AGECAT_11):		
Age in range 18 to 49 years	-1.193	0.617
Age in range 50 to 64 years	Ref. cell	
Identity of payee relative to beneficiary (REPREPAYEE_11):		
A family member received benefits on behalf of the participant	-0.909†	0.588
Participant received benefit payments himself/herself, or information unknown	Ref. cell	
Two-factor interactions^c		
SEX * TOC_11		
Female* Disability claim	-2.664*	1.033
SSI_SSDI_11* REPREPAYEE_11		
SSI only * A family member received benefits on behalf of the participant	3.668*	1.565
SSI_SSDI_11* DIG_11		
SSI only * Participant had mental disability	2.094**	0.735
RACE_11* TOC_11		
Black* Disability claim	2.705*	1.110

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_11”.

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells

TABLE E.12

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT
LONGITUDINAL SAMPLE, PHASE 1 TRADITIONAL PAYMENT TYPE, RESPONSES IN ROUNDS 1 AND 2^a

Variables in the Cooperation Model for Phase 1 Participants Using the Traditional Payment Type Who Responded in Rounds 1 and 2		
Main Effects	Parameter Estimate ^b	Standard Error
Indicator whether participant and applicant for benefits are in same zip code (PDZIPSAME_12):		
Applicant and participant live in different zip code , or no information	Ref. cell	
Applicant and participant live in same zip code	0.490*	0.203
Urbanicity of beneficiary's place of residence (METRO_12):		
Participant lived in nonmetropolitan area	0.567	0.309
Participant lived in metropolitan area	Ref. cell	
Identity of payee relative to beneficiary (REPREPAYEE_12):		
A family member received benefits on behalf of the participant	-2.027**	0.655
Participant received benefit payments himself/herself	-1.894**	0.571
Information unknown	Ref. cell	
Disability diagnosis classification (DIG_12):		
Participant had physical disability	0.952†	0.296
Participant had mental disability, deafness, or information about disability not given	Ref. cell	
Geographic region (based on U.S. Census regions) of participant's residence (REGION_12):		
West, Midwest, South	Ref. cell	
Northeast	-0.503*	0.204
Participant recipient benefit type (SSI_SSDI_12)		
SSI Only or SSDI Only	Ref. cell	
Both SSI/SSDI	0.859†	0.291
Two-factor interactions^c		
SSI_SSDI_12 * DIG_12		
Both SSI/SSDI * Participant had physical disability	-1.083	0.571

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_12.”

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells

TABLE E.13

VARIABLES IN THE LOCATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 2 MILESTONE-OUTCOME PAYMENT TYPE, RESPONSES IN ROUNDS 2 AND 3^a

Variables in the Location Model for Phase 2 Participants Using the Milestone-Outcome Payment Type Who Responded in Rounds 2 and 3		
Main Effects	Parameter Estimate ^b	Standard Error
Participant's age category (AGECAT_13):		
Age in range 18 to 29 years	-0.928†	1.481
Age in range 30 to 49 years	-1.632*	1.174
Age in range 50 to 64 years	Ref. cell	
Participant's type of claim (TOC_13):		
Disability claim	2.780**	0.936
Survivor claim, or unknown	Ref. cell	
Participant's gender (SEX)		
Male	Ref. cell	
Female	-0.551†	0.844
Geographic region (based on U.S. Census regions) of participant's residence (REGION_13):		
Midwest	-1.506†	0.733
West, South, Northeast	Ref. cell	
Disability diagnosis classification (DIG_13):		
Participant had mental disability	-2.806†	0.825
Participant had physical disability, or information about disability not given	Ref. cell	
Number of phone numbers on SSA file over past five years (PHONE_13):		
No phone changes on SSA file over past five years	Ref. cell	
One or more changes in phone number on SSA file, or unknown	-2.035*	0.929
Number of address changes in the past five years (MOVE_13):		
No move	-1.072†	0.998
One or more moves, or information older than five years, or no information	Ref. cell	
Identity of payee relative to beneficiary (REPREPAYEE_13):		
A family member received benefits on behalf of the participant	-2.232†	1.163
Participant received benefit payments himself/herself, or information unknown	Ref. cell	
Participant recipient benefit type (SSI_SSDI_13)		
SSDI Only	-2.513**	0.952
SSI Only , or both SSI and SSDI	Ref. cell	
Participant's living situation (LIVING_13):		
Participant lives alone	-1.573	0.868
Participant does not live on his or her own, or unknown	Ref. cell	

TABLE E.13 (continued)

Variables in the Location Model for Phase 2 Participants Using the Milestone-Outcome Payment Type Who Responded in Rounds 2 and 3

Main Effects	Parameter Estimate ^b	Standard Error
Two-factor interactions^c		
MOVE_13* AGECAT_13		
Nomove* Age in range 18 to 29 years	-5.128**	1.319
SEX * AGECAT_13		
Female* Age in range 18 to 29 years	1.708	1.318
MOVE_13* REGION_13		
Nomove*Midwest	2.016	1.161
DIG_13* REPREPAYEE_13		
Participant had mental disability *A family member received benefits on behalf of the participant	2.918*	1.266

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_13.”

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells

TABLE E.14

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT
LONGITUDINAL SAMPLE, PHASE 2 MILESTONE-OUTCOME PAYMENT TYPE, RESPONSES
IN ROUNDS 2 AND 3^a

Variables in the Cooperation Model for Phase 2 Participants Using the Milestone-Outcome Payment Type Who Responded in Rounds 2 and 3

Main Effects	Parameter Estimate ^b	Standard Error
Participant's age category (AGECAT):		
Age in range 18 to 29 years	2.108†	0.775
Age in range 30 to 39 years	0.918†	0.581
Age in range 40 to 49 years	1.222†	0.724
Age in range 50 to 64 years	Ref. cell	
Identity of payee relative to participant (REPREPAYEE):		
Participant received benefit payments himself/herself	0.744†	0.700
A family member received benefits on behalf of the participant	0.783†	0.671
Information unknown	Ref. cell	
Disability diagnosis classification (DIG_14):		
Participant had physical or mental disability (excluding deafness)	Ref. cell	
Participant had deafness, or information about disability not given	-1.468*	0.577
Participant recipient benefit type (SSI_SSDI_14)		
SSDI only	1.237†	0.523
SSI only, or both SSI/SSDI	Ref. cell	
Number of phone numbers on SSA file over past five years (PHONE_14):		
No phone changes on SSA file over past five years	Ref. cell	
One or more changes in phone number on SSA file, or unknown	0.926†	0.502
Race of the participant (RACE_14):		
Black	-1.294	0.854
White	0.252†	1.126
Race known not to be neither black nor white, or unknown	Ref. cell	
Number of moves in past 5 years (MOVE_14):		
No moves	0.708†	0.533
One or more moves, old information, or no information about moves	Ref. cell	
Participant's type of claim (TOC_14):		
Disability claim	2.627†	0.841
Survivor claim, or unknown	Ref. cell	
Two-factor interactions^c		
AGECAT * RACE_14		
Age within range 18 to 29 years * White	-2.231*	0.989
Age within range 30 to 39 years * White	-2.412**	0.912
Age within range 40 to 49 years * White	-2.868**	0.952
AGECAT * PHONE_14		
Age within range 18 to 29 years * One or more changes in phone number on SSA file, or unknown	-2.270**	0.742

TABLE E.14 (continued)

Variables in the Cooperation Model for Phase 2 Participants Using the Milestone-Outcome Payment Type Who Responded in Rounds 2 and 3		
Main Effects	Parameter Estimate ^b	Standard Error
REPREPAYEE*RACE_14 Participant received benefit payments himself/herself * White	2.016**	0.734
RACE_14* TOC_14 White * Disability claim	-2.035**	0.729
SSI_SSDI_14* REPREPAYEE SSDI only * A family member received benefits on behalf of the participant	-2.093*	0.938
SSI_SSDI_14 * MOVE_14 SSDI only * No moves	-1.813*	0.815
REPREPAYEE *TOC_14 Participant received benefit payments himself/herself * Disability claim	-1.888*	0.813

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_14.

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells

TABLE E.15

VARIABLES IN THE LOCATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 2 OUTCOME-ONLY PAYMENT TYPE, RESPONSES IN ROUNDS 2 AND 3^a

Variables in the Location Model for Phase 2 Participants Using the Outcome-Only Payment Type Who Responded in Rounds 2 and 3		
Main Effects	Parameter Estimate ^b	Standard Error
Participant's age category (AGECAT_15):		
Age in range 18 to 39 years	-1.219	0.962
Age in range 40 to 49 years	-1.957*	0.935
Age in range 50 to 64 years	Ref. cell	
Participant's gender (SEX)		
Female	0.931†	0.709
Male	Ref. cell	
Geographic region (based on U.S. Census regions) of participant's residence (REGION_15):		
Midwest	-0.826†	1.315
West, Northeast, South	Ref. cell	
Number of address changes in the past five years (MOVE_15):		
No move	1.192	0.940
One or more moves, or information older than five years, or no information	Ref. cell	
Race of the participant (RACE_15):		
White	-0.421†	0.708
Race known not to be white, or unknown	Ref. cell	
Two-factor interactions^c		
RACE_15* REGION_15		
White * Midwest	4.465**	1.399
REGION_15*SEX		
Midwest*Female	-3.721*	1.559

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_15.”

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells

TABLE E.16

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT
LONGITUDINAL SAMPLE, PHASE 2 OUTCOME-ONLY PAYMENT TYPE, RESPONSES IN ROUNDS 2 AND 3^a

Variables in the Cooperation Model for Phase 2 Participants Using the Outcome-Only Payment Type Who Responded in Rounds 2 and 3		
	Parameter Estimate^b	Standard Error
Main Effects		
Participant's gender (SEX)		
Male,	Ref. cell	
Female	0.499	0.266
Urbanicity of beneficiary's place of residence (METRO_16):		
Participant lived in nonmetropolitan area adjacent to metropolitan area under 1 million or in nonmetropolitan area not adjacent to metropolitan area	-0.690	0.499
Participant lived in metropolitan area or nonmetropolitan area adjacent to metropolitan area over 1 million	Ref. cell	
Geographic region (based on U.S. Census regions) of participant's residence (DIVISION_16):		
West North Central	1.727**	0.560
New England, Middle Atlantic, South Atlantic, Mountain, East South Central, Pacific, West South Central, East North Central	Ref. cell	
Disability diagnosis classification (DIG_16):		
Participant had mental disability	-0.330†	0.338
Participant had physical disability, or information about disability not given	Ref. cell	
Indicator whether participant and applicant for benefits are in same zip code (PDZIPSAME_16):		
Applicant and participant live in different zip code , or no information	Ref. cell	
Applicant and participant live in same zip code	-0.261†	0.348
Two-factor interactions^c		
DIG_16 * PDZIPSAME_16:		
Participant had mental disability * Applicant and participant live in same zip code	1.142*	0.533

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_16.”

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells

TABLE E.17

VARIABLES IN THE LOCATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT LONGITUDINAL SAMPLE, PHASE 2 TRADITIONAL PAYMENT TYPE, RESPONSES IN ROUNDS 2 AND 3^a

Variables in the Location Model for Phase 2 Participants Using the Traditional Payment Type Who Responded in Rounds 2 and 3		
Main Effects	Parameter Estimate ^b	Standard Error
Geographic region (based on U.S. Census regions) of participant's residence (DIVISION_17):		
East South Central	-1.280*	0.481
New England, Middle Atlantic, South Atlantic, Mountain, West North Central, Pacific, West South Central, East North Central	Ref. cell	
Race of the participant (RACE_17):		
White	0.617	0.363
Race known not to be white, or unknown	Ref. cell	
Number of phone numbers on SSA file over past five years (PHONE_17):		
No phone changes on SSA file over past five years	0.078†	0.608
One or more changes in phone number on SSA file, or unknown	Ref. cell	
Urbanicity of place of residence of participant (METRO_17):		
Participant lived in metropolitan area with population under 250,000 or non metropolitan area	Ref. cell	
Participant lived in metropolitan area with population with 250,000 and more	1.116†	0.432
Participant's age category (AGECAT_17):		
Age in range 18 to 39 years	-0.079†	0.424
Age in range 40 to 64 years	Ref. cell	
Identity of payee relative to beneficiary (REPREPAYEE_17):		
A family member received benefits on behalf of the participant	1.493†	0.608
Participant received benefit payments himself/herself, or information unknown	Ref. cell	
Two-factor interactions^c		
REPREPAYEE_17* PHONE_17		
A family member received benefits on behalf of the participant * No phone changes on SSA file over past five years	-2.559*	1.104
AGECAT_17* PHONE_17		
Age in range 18 to 39 years * No phone changes on SSA file over past five years	2.330	1.149
METRO_17* AGECAT_17		
Participant lived in metropolitan area with population with 250,000 and more * Age in range 18 to 39 years	-1.832**	0.508

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_17.”

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells

TABLE E.18

VARIABLES IN THE COOPERATION LOGISTIC PROPENSITY MODEL: TICKET PARTICIPANT
LONGITUDINAL SAMPLE, PHASE 2 TRADITIONAL PAYMENT TYPE, RESPONSES IN ROUNDS 2 AND 3^a

Variables in the Cooperation Model for Phase 2 Participants Using the Traditional Payment Type Who Responded in Rounds 2 and 3		
	Parameter Estimate ^b	Standard Error
Main Effects		
Participant's gender (SEX)		
Male	Ref. cell	
Female	0.713†	0.204
Indicator whether participant and applicant for benefits are in same zip code (PDZIPSAME_18):		
Applicant and participant live in different zip code , or no information	Ref. cell	
Applicant and participant live in same zip code	1.231†	0.441
Identity of payee relative to beneficiary (REPREPAYEE):		
A family member received benefits on behalf of the participant	0.815	0.417
Participant received benefit payments himself/herself	-0.323†	0.406
Information unknown	Ref. cell	
Disability diagnosis classification (DIG_18):		
Deafness or information unknown	-1.334*	0.510
Participant had physical disability (excluding the deafness)	0.487†	0.343
Participant had mental disability	Ref. cell	
Geographic region (based on U.S. Census regions) of participant's residence (REGION_18):		
Midwest	2.132†	0.490
West, South, Northeast	Ref. cell	
Participant recipient benefit type (SSI_SSDI_18)		
SSI Only	0.606†	0.444
SSDI Only, or Both SSI/SSDI	Ref. cell	
Participant's age category (AGECAT_18):		
Age in range 18 to 29 years	Ref. cell	
Age in range 30 to 39 years	1.635†	0.541
Age in range 40 to 64 years	1.592†	0.418
Urbanicity of place of residence of participant (METRO_18):		
Participant lived in metropolitan area with population under 250,000 or non metropolitan area	Ref. cell	
Participant lived in metropolitan area with population with 250,000 and more	-0.868**	0.272
Two-factor interactions^c		
SEX * PDZIPSAME_18		
Female* Applicant and participant live in same zip code	-1.434**	0.490
SSI_SSDI_18* REPREPAYEE		
SSI only * Participant received benefit payments himself/herself	1.596	0.785
SSI_SSDI_18* DIG_18		
SSI Only * Participant had physical disability (excluding the deafness)	-2.224**	0.672

TABLE E.18 (continued)

Variables in the Cooperation Model for Phase 2 Participants Using the Traditional Payment Type Who Responded in Rounds 2 and 3		
Main Effects	Parameter Estimate ^b	Standard Error
SSI_SSDI_18* REGION_18		
SSI Only *Midwest	-2.073**	0.574
REGION_18* AGECAT_18		
Midwest * Age in range 30 to 39 years	-3.595**	0.691
Midwest * Age in range 40 to 64 years	-2.205**	0.554

^a If any levels in the main effects are collapsed from all possible levels, the base variable is followed by “_18.”

^b Parameter estimates with a cross (†) are essentially meaningless because higher order terms that include the variable in question are also in the model. One star (*) and two stars (**) represent significance at the 5 percent and 1 percent levels respectively.

^c All combinations for the listed interactions that are not shown are part of the reference cells

APPENDIX F

SUDAAN AND SAS PARAMETERS USED TO OBTAIN NATIONAL ESTIMATES

PAGE IS INTENTIONALLY LEFT BLANK TO ALLOW FOR DOUBLE-SIDED COPYING

SUDAAN EXAMPLE

```
proc descriptive data="SASdatasetname" filetype=sas design=wr;
  nest      A_STRATA A_PSU / missunit;
  weight    "weight variable";
  subpopn   "response variable" = "complete";
  var       "analysis variables";
  print nsum wsum mean semean deffmean / style=nchs wsumfmt=f10.0 meanfmt=f8.4
  semeanfmt=f8.4 deffmeanfmt=f8.4;
  title     "TTW National Estimates";
```

SAS EXAMPLE

```
proc surveymeans data="SASdatasetname";
  strata  A_STRATA;
  cluster A_PSU;
  weight  "weight variable";
  where   "response variable" = "complete";
  var     "analysis variables";
  title   "TTW National Estimates";
```

Weight Variables Used for Cross-Sectional Estimates

Beneficiary sample:	Wtr3_ben
Participant sample:	Wtr3_cspar
Combined samples:	Wtr3_com

Weight Variables Used for Longitudinal Estimates

Phase 1 respondent in rounds 1, 2, and 3:	R3partfnlwtg_long123
Phase 1 respondent in rounds 1 and 2:	R3partfnlwtg_long12
Phase 2 respondent in rounds 2 and 3:	R3partfnlwtg_long23

Nest Variables Used for Cross-Sectional Estimates

A_STRATA

1. Clustered samples for both beneficiaries and participants
 - a. A_STRATA = 1000 for PSUs in Phase 1 states
 - b. A_STRATA = 2000 for PSUs in Phase 2 states
 - c. A_STRATA = 3000 for PSUs in Phase 3 states
2. Unclustered samples for participants, cross-sectional estimates
 - a. A_STRATA = 2112 Outcome-only participants in PSUs in Phase 2 states, R 2 frame

- b. A_STRATA = 2122 Outcome-only participants not in PSUs in Phase 2 states, R 2 frame
- c. A_STRATA = 2113 Outcome-only participants in PSUs in Phase 2 states, R 3 frame
- d. A_STRATA = 2123 Outcome-only participants not in PSUs in Phase 2 states, R 3 frame
- e. A_STRATA = 2212 Milestone-outcome participants in PSUs in Phase 2 states, R 2 frame
- f. A_STRATA = 2222 Milestone-outcome participants not in PSUs in Phase 2 states, R 2 frame
- g. A_STRATA = 2213 Milestone-outcome participants in PSUs in Phase 2 states, R 3 frame
- h. A_STRATA = 2223 Milestone-outcome participants not in PSUs in Phase 2 states, R 3 frame
- i. A_STRATA = 3113 Outcome-only participants in PSUs in Phase 3 states, R 3 frame
- j. A_STRATA = 3123 Outcome-only participants not in PSUs in Phase 3 states, R 3 frame

A_PSU

- 1. Clustered sample cases both beneficiaries and participants cross-sectional samples
A_PSU = PSU identifier
- 2. Unclustered sample cases in Milestone-outcome and Outcome-only Phase 2 participants and Outcome-only Phase 3 participants cross-sectional samples
A_PSU = MPR_ID

Nest Variables Used for Longitudinal Estimates

A_STRATA

- 1. Clustered samples for longitudinal participants
 - a. A_STRATA = 1000 for PSUs in Phase 1 states
 - b. A_STRATA = 2000 for PSUs in Phase 2 states
- 2. Unclustered samples for participants, longitudinal estimates
 - a. A_STRATA = 1111 Outcome-only participants in PSUs in Phase 1 states
 - b. A_STRATA = 1121 Outcome-only participants not in PSUs in Phase 1 states
 - c. A_STRATA = 2112 Outcome-only participants in PSUs in Phase 2 states
 - d. A_STRATA = 2122 Outcome-only participants not in PSUs in Phase 2 states
 - e. A_STRATA = 2212 Milestones and outcomes participants in PSUs in Phase 2 states
 - f. A_STRATA = 2222 Milestones and outcomes participants not in PSUs in Phase 2 states

A_PSU

- 1. Clustered sample cases in all longitudinal participants samples
A_PSU = PSU identifier
- 2. Unclustered sample cases in Outcome-only Phase 1 participants and Milestone-outcome and outcome-only Phase 2 participants longitudinal samples
A_PSU = MPR_ID

Notes:

1. Before each SUDAAN procedure, sort by A_STRATA and A_PSU
2. Use SUDAAN's SUBPOP statement to define population for which estimates are wanted.

For example, for estimates of SSI participant population, use SUBPOP to define SSI participants.