Series 2
No. 124

Vital and Health Statistics

From the CENTERS FOR DISEASE CONTROL AND PREVENTION / National Center for Health Statistics

Sample Design, Sampling Weights, Imputation, and Variance Estimation in the 1995 National Survey of Family Growth

February 1998

DEPARTMENT OF HEALTH & HUMAN SERVICES · USA

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

CDC
CENTERS FOR DISEASE CONTROL AND PREVENTION

Suggested citation

Potter FJ, Iannacchione VG, et al. Sample design, sampling weights, imputation, and variance estimation in the 1995 National Survey of Family Growth. National Center for Health Statistics. Vital Health Stat 2(124). 1998.

Library of Congress Catalog Card Number 90-13573

Sample design, sample weights, imputation, and variance estimation in the 1995 National Survey of Family Growth.
p. cm. — (Vital and health statistics. Series 2, Data evaluation and methods research ; no. 124) (DHHS publication ; no. (PHS) 98-1398)
Includes bibliographical references.
ISBN 0-8406-0537-4
1. National Survey of Family Growth (U.S.) 2. Family size—United States—Statistical methods. 3. Fertility, Human—United States—Statistical methods. 4 .Birth control—United States—Statistical methods. 5. Health Surveys—United States. 6. Family life surveys—United States. I. Series. II. Series: DHHS publication ; no. (PHS) 98-1398.
RA409.U45 no. 124
[HQ762.U]
362.1'0723 s—dc21 97-43925
[306.85'0973] CIP

For sale by the U.S. Government Printing Office
Superintendent of Documents
Mail Stop: SSOP
Washington, DC 20402-9328

Vital and Health Statistics

Sample Design, Sampling Weights, Imputation, and Variance Estimation in the 1995 National Survey of Family Growth

Series 2:
No. 124

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

Hyattsville, Maryland
February 1998
DHHS Publication No. (PHS) 98-1398

This report is dedicated to

Steven L. Botman

December 15, 1947–June 1, 1997

This report is dedicated to the memory of our late colleague

Steven L. Botman

of the National Center for Health Statistics,
who served as the consulting mathematical statistician
on the National Survey of Family Growth
for the past decade.
He made many contributions to the tasks described in this report
and to the 1995 National Survey of Family Growth in general.
He was a pleasure to work with, and we will miss him very much.

Contents

Abstract 1

Introduction 1
Sample Design 2
Sampling Weights 2
Item Imputation 2
Variance Estimation 3
Conclusion 3

Background 3

Design Specifications 4

Sample Design 4
Summary 4
Sample Design of the National Health Interview Survey 4
National Survey of Family Growth Sample Design 5
The 1993 NHIS Sample and 1995 NSFG Sampling Frame 5
Sampling Procedure and Allocations 5
Allocation of the Sample to PSU's 6
Sample Selection Within PSU's 6

Characteristics of the Sample 6
Designated Sample Sizes and Probabilities of Selection 6
Interview Rates 7
Sample Sizes, Clustering, and Variation in the Probability of Selection 7

Sampling Weights 8
Summary 8
Weight Adjustments for Location and Response Propensity 10
National Health Interview Survey Sampling Weights and Adjustments for Cycle 5 Subsampling 10
NHIS Sampling Weight 11
Adjustments for Cycle 5 Subsampling 11
Response Probability Modeling 11
Model Development 11
Location Propensity Model 12
Response Propensity Model 12
Factors Affecting the Proportion Located 13
Factors Affecting Response Propensity 15
Evaluation of the Combined Location and Response Models 15
Poststratification Adjustments 17

Item Imputation 19
Overview 19
Imputation Procedures for Cycle 5 19
Imputation of Family Income 22
Model Development for Imputing Income 22

Variance Estimation . . . 24
Background . . . 24
Summary of Variance Estimation . . . 25
Generalized Variance Estimation . . . 26
Comparison of Generalized Standard Error (GSE) Estimates to Direct Estimates . . . 29
Hypothesis Tests . . . 29

References . . . 30

Appendix I . . . 38
Definitions of Terms . . . 38

Appendix II . . . 40
Linkage of the National Survey of Family Growth to the National Health Interview Survey . . . 40
Previous Research on Linked Samples . . . 40
Cycle 5 Tracing Activities . . . 42
NHIS Tracing Data . . . 42
Advance Tracing . . . 43
Characteristics of Unlocated Sample Women . . . 43
Effect of Unlocated Sample Women . . . 46
Discussion . . . 48

Appendix III . . . 49
Imputation Specifications . . . 49

Appendix IV . . . 54
Variance Estimation Using Taylor Series Approaches . . . 54

Appendix V . . . 55
Example SUDAAN Program Code and Output . . . 55

Appendix VI . . . 61
How the Generalized Standard Error (GSE) Estimates Were Made . . . 61

Text Tables

A. Distribution of National Health Interview Survey clusters and households by cluster composition strata . . . 6
B. Designated sample sizes, probability of selection, and average and relative variance of weights, by race/ethnicity and number of eligible women: 1995 National Survey of Family Growth . . . 6
C. Response rates among completed cases in the National Health Interview Survey, by race/ethnicity and age: 1995 National Survey of Family Growth . . . 7
D. Clustering and weight variation among completed cases in the 1995 National Survey of Family Growth, by race/ethnicity and age . . . 8
E. Weight adjustment summary by race/ethnicity and type of adjustment . . . 10
F. National Health Interview Survey variables used to predict location rates for the 1995 National Survey of Family Growth . . . 12
G. Final location propensity logistic model for the 1995 National Survey of Family Growth . . . 13
H. Final response propensity logistic model for the 1995 National Survey of Family Growth . . . 15
J. Poststratification adjustment summary, by selected characteristics . . . 18
K. Comparison of National Survey of Family Growth (Cycle 5) estimates of the number of births with vital statistics by year, race, and Hispanic origin . . . 19
L. Joint frequency distribution of the education of mother and father of Cycle 5 respondents . . . 21
M. Number and percent of pregnancy-intervals requiring an imputed wantedness . . . 22
N. Availability of family income from the 1993 National Health Interview Survey and National Survey of Family Growth, Cycle 5 . . . 24
O. Generalized standard errors for estimated percentages and corresponding sample sizes from the respondent file: National Survey of Family Growth, Cycle 5 . . . 27
P. Generalized standard errors for estimated percentages and corresponding sample sizes from the pregnancy-interval file: National Survey of Family Growth, Cycle 5 . . . 28

Q. Comparison of three ways of estimating the standard errors for the percent currently using oral contraceptive pills: (1) assuming a simple random sample, (2) using SUDAAN, and (3) using generalized standard errors: 1995 National Survey of Family Growth . . . 29

Figures

1. National Survey of Family Growth Cycle 5 weight assignment process . . . 9
2. Final segmentation of the 1995 National Survey of Family Growth sample members by percent located . . . 14
3. Final segmentation of National Survey of Family Growth – 5 eligible sample members by percent responded . . . 16
4. Receiver operating characteristics curve of the combined location and response models . . . 17
5. Segmentation of the respondent's wantedness of the pregnancy . . . 23

Detailed Tables

1. Distribution of sample women selected for the 1995 National Survey of Family Growth, weighted location and response rates by characteristics of women and their households as measured in the 1993 National Health Interview Survey . . . 31
2. Current Population Survey totals, relative standard errors, and sample sizes for the poststratification adjustment variables, by age, race, parity, and marital status, National Survey of Family Growth, Cycle 5 . . . 35

Appendix Tables

I. Tracing steps in the 1995 National Survey of Family Growth . . . 43
II. National Health Interview Survey predictor variables used in the location model in table III . . . 44
III. Factors measured in the 1993 National Health Interview Survey affecting location rates in the 1995 National Survey of Family Growth . . . 45
IV. Median design effects for nine respondent file, by race/ethnicity and demographic characteristics: 1995 National Survey of Family Growth, Cycle 5 . . . 61
V. Median design effects for seven pregnancy-interval file variables, by race/ethnicity and demographic characteristics: 1995 National Survey of Family Growth, Cycle 5 . . . 62

Appendix Figure

I. Cumulative bias potential associated with National Health Interview Survey nonresponse and subsequent National Survey of Family Growth nonlocation and nonresponse . . . 47

Abstract

Objectives

Cycle 5 of the National Survey of Family Growth (NSFG) was conducted by the National Center for Health Statistics (NCHS) in 1995. The NSFG collects data on pregnancy, childbearing, and women's health from a national sample of women 15–44 years of age. This report describes how the sample was designed, shows response rates for various subgroups of women, describes how the sampling weights were computed to make national estimates possible, shows how missing data were imputed for a limited set of key variables, and describes the proper ways to estimate sampling errors from the NSFG. The report includes both nontechnical summaries for readers who need only general information and more technical detail for readers who need an in-depth understanding of these topics.

Methods

The 1995 NSFG was based on a national probability sample of women 15–44 years of age in the United States and was drawn from 14,000 households interviewed in the 1993 National Health Interview Survey (NHIS). Of the 13,795 women eligible for the NSFG, 10,847 (79 percent) gave complete interviews.

Results

This report recommends using weighted data for analysis and a software package that will estimate sampling errors from complex samples (for example, SUDAAN or comparable software).

The rate of missing data in the 1995 NSFG was very low. However, missing data were imputed for 315 key variables, called "recodes." Of the 315 recodes defined for Cycle 5, 271 variables had missing data on less than 1 percent of the cases; only 44 had 1 percent or more with missing data. These missing values were imputed for all of these 315 variables. The imputation procedures are described in this report.

Keywords: *survey methodology • response rates • imputation • variance estimation*

Sample Design, Sampling Weights, Imputation, and Variance Estimation in the 1995 National Survey of Family Growth

by Frank J. Potter, Ph.D., formerly with the Research Triangle Institute, Vincent G. Iannacchione, M.S., Research Triangle Institute, William D. Mosher, Ph.D., National Center for Health Statistics, Robert E. Mason, Ph.D., Research Triangle Institute, and Jill D. Kavee, M.S., Research Triangle Institute

Introduction

This report describes the procedures used in the 1995 National Survey of Family Growth to select the sample, develop the sampling weights, impute missing data, and estimate sampling errors. These procedures help to ensure that the sample data can be used to make valid national estimates. They also allow researchers to draw statistically sound conclusions from the data. Parts of this report contain a great deal of technical detail. For readers seeking a general understanding of the survey procedures, this introduction provides a brief and less technical summary of the procedures used.

The National Survey of Family Growth (NSFG) is designed and administered by the National Center for Health Statistics (NCHS) of the U.S. Centers for Disease Control and Prevention (CDC), a part of the U.S. Department of Health and Human Services. The purpose of the survey is to produce national estimates of factors affecting pregnancy—including sexual activity, contraceptive use, and infertility—and the health of women and infants. For Cycle 5, interviewing and data processing were conducted by the Research Triangle Institute (RTI), under a contract with NCHS.

A national probability sample of 10,847 civilian noninstitutionalized women ages 15–44 years of age were interviewed between January and October 1995. The interviews were conducted in person by trained female interviewers using laptop, or notebook, computers. This procedure is called computer-assisted personal interviewing (CAPI). The interview, which lasted an average of 103 minutes, collected data on each pregnancy (if any); contraceptive use by the woman and her partner; her ability to bear children; the use of medical services for contraception, infertility, and prenatal care; a history of her marriages, cohabitations, and childhood living situations; her education history, work history, and a variety of demographic and economic characteristics.

The 1995 National Survey of Family Growth was jointly planned and funded primarily by the National Center for Health Statistics, the National Institute for Child Health and Human Development, and the Office of Population Affairs, with additional support from the Administration for Children and Families. Other agencies and individuals also provided helpful advice and assistance. Steven Machlin, M.S., of the Agency for Health Care Policy and Research (AHCPR), served as peer reviewer for this report and made many useful comments. Dr. Anjani Chandra and Christopher Moriarity, NCHS, also read the manuscript and made a number of helpful suggestions.

Sample Design

A total of 10,847 women were interviewed from a national probability sample of households that responded to the 1993 National Health Interview Survey (NHIS). The NHIS is a continuous household survey conducted by NCHS that covers the U.S. civilian noninstitutionalized population. Women for the NSFG sample were selected from all 198 NHIS primary sampling units (PSU's). A PSU is a metropolitan statistical area (MSA), a county, or a group of adjacent counties. PSU's were located in nearly every State and included all of the largest metropolitan areas in the United States. Sample women who moved since the NHIS interview were traced to their new address, and an interviewer conducted the interview with the woman at the new address.

Hispanic and non-Hispanic black women were selected with higher probability than other women so that more reliable statistics for Hispanic and non-Hispanic black women could be produced. All NHIS households containing Hispanic or non-Hispanic black women were included in the NSFG sample; one woman was selected randomly if more than one woman was eligible for the NSFG. Only some households of other race or ethnic identification were selected to be in the sample. Households were selected with probability proportional to the number of eligible women in the household.

Appendix I defines some of the technical terms used in this report. Appendix II discusses some statistical aspects of the linkage of the NSFG to the NHIS.

Sampling Weights

A simple random sample in which response rates and coverage were the same in *every* subgroup would be a ''scale model'' of the population from which it was drawn. However, a survey sample is almost never a scale model in that sense. Groups are often *selected* at different rates and often have different *response* rates. For example, in the NSFG, non-Hispanic black women account for 23 percent of all respondents in the sample but only about 14 percent of the population. ''Sampling weights'' adjust for these different sampling rates, response rates, and coverage rates so that accurate national estimates can be made from the sample.

A respondent's sampling weight can be interpreted as the number of women in the population that she represents. For example, if a woman's sampling weight is 5,000, then she represents an estimated 5,000 eligible women in the population. For the NSFG, the fully adjusted sampling weights were assigned to each respondent and consisted of the NHIS sampling weight and four adjustment factors. The NHIS sampling weight is the inverse of a sample member's probability of selection into the NHIS sample. For example, if the probability of selection is 1 in 4,000, then the initial sampling weight is 4,000. The *first adjustment factor* applied to the NHIS sampling weight was the *inverse of the subsampling rates* used to select the NSFG sample of 14,000 from the 25,534 NHIS women ages 15–44 in households responding to the 1993 NHIS.

Between the 1993 NHIS and the 1995 NSFG, many women in the sample moved, and substantial effort was made to identify their new addresses. The percent located is called the ''location rate.'' The percent who participated in the survey is called the ''participation rate.'' The overall location rate among the 14,000 sample women was approximately 95 percent. Because more young women move, they were generally harder to locate than older women. Among located sample members, non-Hispanic women were more likely to refuse to participate than were Hispanics. To compensate for these different effects, the *second and third adjustment factors* adjusted for the proportion of women who could be *located* and the proportion of those located who *participated*. The *fourth adjustment* was to make the weights match independent estimates of the number of women by age, race, marital status, and parity (the number of live births) obtained from the U.S. Bureau of the Census. This process is called poststratification.

Item Imputation

In any survey, not every question is answered by every person interviewed. Sometimes a respondent cannot remember the fact asked for in a question; sometimes she may refuse to answer. Such missing data create inconsistencies in estimates, which may be confusing for some users of the data. Assigning values to these missing items is called ''imputation''; imputation makes the data complete, more consistent, and easier to use.

In Cycle 5 of the NSFG, there are thousands of variables in the data file. Of these many variables, about 315 recoded variables (called recodes) were selected because they are used frequently in analysis. Missing data for these recodes could create inconsistencies among survey estimates and confusion among data users, so these variables were imputed. The frequency of missing values for the recoded variables in Cycle 5 was very low, in part because CAPI requires the interviewer to enter an acceptable response and then automatically goes to the next appropriate question. The program also performs range and consistency checks to rule out logically impossible answers. The imputation techniques used in Cycle 5 (appendix III) were:

- Logical imputation
- Unweighted hot-deck imputation
- Weighted hot-deck imputation
- (Regression) model-based imputation

Some of the recodes are actually a set of several repetitions of a variable. For example, data were collected on up to 10 periods of working, up to 12 living situations, and up to 15 pregnancies. Counting all these repetitions, about 488 variables were recoded—and if missing, imputed.

Item imputation usually involves assigning a value from a person with reported data for an item (a donor) to a person with missing data for that item (a receptor). The donor cases were selected so that they were as similar as possible to the receptor cases on age, race, and other variables. Except when it was obviously incorrect, actual reported information was *never* replaced by an imputed value. For each recoded variable in the database, an imputation flag identifies whether the value of that variable was imputed. Using the imputation flag, a researcher can identify the observations with an imputed value and the specific type of imputation procedure used for each specific recoded variable.

Variance Estimation

The sampling variance is a measure of the variation of a statistic (such as a proportion or a mean) caused by having taken a sample instead of interviewing the full population. It measures the variation of the estimated statistic over repeated samples. The sampling variance is zero when the full population is observed, as in a census.

For the NSFG, the sampling variance estimate is a function of the sampling design and the population parameter being estimated, and it is called the *design-based sampling variance*. The NSFG data file contains a final weight and information necessary to estimate the sampling variance for a statistic. Most common statistical software (such as SAS and SPSS) will attempt to compute ''population'' variances, which may severely underestimate or overestimate the sampling variances. Special software is required to accurately estimate sampling errors in a complex sample such as the NSFG. Appendix IV describes some of the statistical theory of variance estimation used in the 1995 NSFG. Appendix V shows two sample programs—one for a cross-tabulation and one for a logistic regression—that show how to estimate NSFG sampling variances using a variance estimation program called SUDAAN. A shortcut method to estimate sampling errors for numbers and percentages has also been developed. Appendix VI describes how these ''generalized standard error estimates'' were made.

Conclusion

The rest of this report provides more information about the sample design, the linkage of NSFG to NHIS, weighting, item imputation, and variance estimation. Each major section begins with a ''Summary'' or ''Overview,'' which gives the reader a shorter and less technical review of the topic. The remainder of the text and the appendixes supply full details.

Background

The NSFG was established in 1971 by NCHS. The purpose of the survey is to produce national estimates of factors related to pregnancy and birth rates, such as sexual activity, contraceptive use, and infertility; use of family planning and other medical services; and maternal and infant health. Interviewing for the first cycle of NSFG was conducted in 1973; other cycles were conducted in 1976 (Cycle 2), 1982 (Cycle 3), 1988 (Cycle 4), and in 1995 (Cycle 5). A major function of the successive cycles of the survey is to produce comparable time trend data.

Data for all five cycles were collected from probability samples of women by means of personal interviews. The sample for Cycle 5 was drawn from households interviewed in another survey conducted by NCHS, the 1993 National Health Interview Survey (NHIS).

The sample design and data collection for Cycle 5 were completed under a contract with the Research Triangle Institute (RTI). Cycle 5 is based on interviews with 10,847 women. The interviews were conducted between mid-January and October 1995.

In general, Cycle 5 was modeled after Cycle 4. However, several major aspects of the survey changed between Cycle 4 and Cycle 5 (1). The first change was that the interviews were conducted with laptop, or notebook, computers instead of paper and pencil interviewing. The use of computer-assisted personal interviewing (CAPI) made it possible to collect more detailed data and use more complex question sequences and still improve data quality. However, computer-assisted interviewing requires substantial effort to translate the ordinary logic of questions into computer programming language and to ensure that the specifications are accurate (2).

The second change in the 1995 NSFG, compared with the 1988 NSFG, is that much of the 1995 questionnaire consisted of event histories. An event history is simply a complete list of all occurrences of some event, including the beginning and ending dates of each occurrence, and other important details. In Cycle 5 of the NSFG, the following event histories were collected:

1. regular, vocational, and general equivalency diploma (GED) education
2. periods of living with a mother, father, and grandparents during childhood
3. work
4. marriage, separation, and divorce
5. cohabitation
6. sexual partners in the last 5 years
7. contraceptive use
8. pregnancy

These event histories dramatically improved the usefulness of the NSFG for academic and policy research, but also increased the length of the interview. For the first three cycles, the interviews lasted an average of 60 minutes, and for Cycle 4, 70 minutes. The average interview length for Cycle 5 was nearly 50 percent longer than in Cycle 4 (approximately 103 minutes). These longer, more difficult interviews made it necessary to pay a $20 incentive to each respondent in order to maintain the NSFG's traditionally high response rates. Pretesting showed that incentives increased response rates, reduced costs, and improved the reporting of sensitive items (3,4).

The third change in Cycle 5 compared with Cycle 4 was an increase in sample size, from 8,450 to 10,847 women. For the first time in the NSFG, Hispanic as well as black women were oversampled. The number of Hispanic women interviewed increased from 641 in Cycle 4 to 1,553 in Cycle 5. Questionnaires were administered in Spanish by bilingual interviewers when necessary. For a more detailed discussion of how the survey was planned, how the questionnaire was programmed, and how the fieldwork (interviewing) was done, see reference 1.

Design Specifications

The sample design was developed to achieve the primary survey objectives within the practical constraints of the available funds, time and schedule requirements, and the size and characteristics of the population under study. The principal sample design features for the NSFG were predetermined by the use of the NHIS as the source of the sample of women. The additional specifications for the NSFG were:

- The target population was civilian noninstitutionalized women 15–44 years of age who were living in households or group quarters in the United States, including Alaska and Hawaii. Women in the military and those confined to institutions such as prisons and mental hospitals were specifically excluded.
- The intended sample size was approximately 10,500 completed interviews, selected from households previously interviewed for the 1993 NHIS. Hispanic and non-Hispanic black women would be oversampled to produce more precise estimates for these populations. In addition, only one eligible woman was to be randomly selected from a household for interview.
- Data were to be collected from the sample women by means of CAPI technology. No proxy interviews were accepted.
- All interviewers had to be female.
- The interviewer would collect information on fertility, sexual experience and contraceptive use, sources and types of family planning services, and related aspects of maternal and infant health by using a highly structured interview instrument programmed into a laptop, or notebook, computer.
- The target interview completion rate was 80 percent among those who had already completed the NHIS. This response rate was to be achieved for Hispanic women, non-Hispanic black women, and women of other races.
- The interviewing should be completed in approximately 6 months.
- The contractor, in cooperation with NCHS, was required to design and implement procedures to measure and control the quality of data collection and data preparation.

Sample Design

Summary

For Cycle 5 of the NSFG, a national probability sample of 14,000 women 15–44 years of age was selected from among households that responded to the 1993 NHIS. The NHIS is a continuous multistage household survey conducted by NCHS that covers the U.S. civilian noninstitutionalized population. Data are collected for each household member on health conditions, doctor visits, hospitalizations, disabilities, and other health-related topics, as well as demographic and economic data for the household and household members. The 1993 NHIS was conducted in 198 primary sampling units, or PSU's, where a PSU is a MSA, county, or group of adjacent counties. PSU's were located in nearly every State and included all of the largest metropolitan areas in the United States.

NCHS provided RTI with data files containing household-level and person-level data for all persons in households responding to the 1993 NHIS. All households in the 1993 NHIS containing Hispanic or non-Hispanic black women 15–44 years of age were included in the NSFG sample, along with about 55 percent of NHIS households containing white and other (nonblack, non-Hispanic) women. Thus black and Hispanic women were sampled at a higher rate than other women for the NSFG. If there were more than one eligible woman in the household, one was selected for the NSFG.

Sampled women were drawn from all 198 NHIS PSU's, and women who moved since the NHIS interview were traced to their new address. An interviewer conducted the interview at the new address. Because of the complex design and the unequal sampling rates, the sampling weights must be used to calculate accurate numbers, percents, and other statistics. The sample design must be taken into account to compute accurate sampling errors. The design-based variance assumes the use of the fully adjusted sampling weight. The fully adjusted sampling weight is derived from the sampling design with adjustments to compensate for nonresponse and for adjusting the sample data to independent population estimates by age, race, marital status, and parity from the U.S. Bureau of the Census.

Sample Design of the National Health Interview Survey

The NHIS is a stratified multistage household survey that covers the civilian noninstitutionalized population of the United States. The NHIS is redesigned each decade using data from the most recent census (5). Cycle 5 of the NSFG used the NHIS sample based on the design developed for the period 1985–94. A complete description of the NHIS design is given in reference 5.

For the NHIS, the geographic area of the United States was divided into approximately 1,900 PSU's. A PSU consists of a MSA, an individual county, or a small group of adjacent counties. The 1,900 PSU's were stratified using socioeconomic and demographic variables. The sample was selected with probability proportional to the population size (PPS) within a stratum. The 1985–94 NHIS sample contained 198 PSU's. Under the PPS design, the largest PSU's were selected into the sample with certainty (that is, with probability of 1). These PSU's are called self-representing (SR) PSU's. A total of 52 PSU's was designated as SR PSU's. The remainder of the PSU's were stratified into 73 strata, and 2 PSU's were selected from each stratum. That is, the final NHIS sample of 198 PSU's consisted of 52 SR PSU's and 146 nonself-representing (NSR) PSU's.

Within each sample PSU, a sample of census blocks (or small groups of blocks) was selected. In PSU's with 5–49 percent of their population black persons, blocks in enumeration districts with high concentrations of black persons were selected with a higher probability than other blocks. Within each block (or group of blocks), a cluster of approximately eight housing units was selected. These housing units were spread as evenly throughout the block as possible.

To gain better control over the size of the sample, housing units constructed since the 1980 census were selected through a sample of building permits rather than through area sampling. These units were selected in clusters of four instead of eight.

The NHIS sample is divided into 51 (or sometimes 52) weekly interviewer assignment samples such that each weekly sample represents a national probability sample of housing units. NCHS can then form national samples by combining weekly samples. NCHS processes the weekly samples in batches of 12 or 13 for each quarter of the calendar year. For each quarterly sample, the respondent data are processed and edited, and a fully adjusted sampling weight that allows for national estimates is computed.

In 1993, budget restrictions required NCHS to field only 7 of the 13 weekly samples in the second quarter (April to June); hence the 1993 NHIS sample contained 46 weekly samples (that is, 6 of the 52 weekly samples were dropped). The 1993 NHIS respondent sample included data for 109,671 persons in 43,007 households. In addition, because of budgetary constraints, the households interviewed during the first two quarters of the 1993 NHIS were administered only the core NHIS questionnaire.

National Survey of Family Growth Sample Design

The NSFG sample design required at least 10,500 completed interviews. If a combined location and response rate of 75 percent was obtained, a sample of 14,000 women would be sufficient to achieve the 10,500 completed interviews. In total, 14,000 women were selected. Only one woman per household was selected. All NHIS households containing Hispanic or non-Hispanic black women were included (that is, ''selected with certainty'') in the NSFG subsample. The remaining NHIS households (those containing white women or women of other races) were subsampled with probabilities proportional to the weighted number of women in the household so that each of the sampled women would have the same probability of selection for the NSFG as women in households with more or fewer eligible women.

The 1993 NHIS Sample and 1995 NSFG Sampling Frame

The 1993 NHIS consisted of 46 of the 52 weekly samples. Based on these 46 weekly samples, the 1995 NSFG sampling frame included 25,534 women 15–44 years of age in 21,168 households (1.21 women per household). The ages are based on the estimated midpoint of the data collection period, which was April 1, 1995. Therefore, a woman was included in the sampling frame if she was born between April 1, 1950, and March 31, 1980, inclusive.

A total of 2,135 households contained one or more Hispanic women (called an Hispanic household); 3,206 contained one or more non-Hispanic black women, but no Hispanic women (called a non-Hispanic black household); and 15,827 contained only women of other race/ethnicities (called other households).

The sampling weight for frame members was computed from the NHIS household basic weight before age-sex-race poststratification adjustments (NHIS person file tape, positions 164–169). Based on this weight, the weighted total number of women represented by the sampling frame was 53,587,840. The weighted total number of Hispanic women was 5,709,751; non-Hispanic black women, 6,853,684; and other women, 41,024,225. These weighted counts were computed using the NHIS weight before it was poststratified, so they undercount the actual population by approximately 10 percent, which is consistent with undercounts in other household surveys. The poststratification adjustment of the weight raised the weighted totals to approximately 60,201,000 women, the estimate from the Current Population Survey of the number of women 15–44 years of age in the civilian noninstitutionalized population of the United States in 1995.

Sampling Procedure and Allocations

For the sample selection, the sampling frame of women was stratified by the characteristics of the households in the NHIS clusters. An NHIS cluster represents the sample of households selected for the NHIS in an area or a permit segment. These strata are:

- The minority stratum—1,015 clusters containing only households with Hispanic or non-Hispanic black women.
- The mixed stratum—1,518 clusters containing households with Hispanic or non-Hispanic black women and households with other women.
- The high-density stratum—2,250 clusters containing three or more households with other women.
- The low-density stratum—2,160 clusters containing only one or two-households with other women.

See table A for the number and classification of households in these strata.

The sampling design for Cycle 5 specified that all households with Hispanic or non-Hispanic black women should be included in the NSFG. Therefore, field interviewers had to go to all NHIS clusters in the minority stratum and the mixed stratum (areas containing black or Hispanic households). In clusters containing no black or Hispanic households, about 55 percent of the households were selected. One household was expected to be selected from each cluster in the high-density nonminority stratum. For the low-density nonminority stratum, approximately one-half of the clusters were selected and approximately one household would be selected in the cluster.

Table A. Distribution of National Health Interview Survey clusters and households by cluster composition strata

Cluster composition strata	NHIS[1] clusters	Households			Households per cluster
		Total	Minority[2]	Other	
Total	6,943	21,168	5,341	15,827	3.05
Minority only	1,015	2,920	2,920	–	2.88
Mixed	1,518	5,844	2,421	3,423	3.85
Nonminority only	4,410	12,404	–	12,404	2.81
High density	2,250	9,114	–	9,114	4.05
Low density	2,160	3,290	–	3,290	1.52

– Quantity zero.
[1]NHIS is National Health Interview Survey.
[2]Minority households are households with Hispanic or non-Hispanic black women.

As a source of potential cost reduction in data collection, households in the low-density stratum were undersampled to reduce the number of NHIS clusters with only one household. So the households in the mixed stratum and the high-density stratum were oversampled by about 10 percent. This design results in an increase in the sampling variance of less than 5 percent.

Allocation of the Sample to PSU's

The probabilities of selection of some households in the nonminority strata were large enough that more than one woman was expected to be selected. All these households (1,420) were selected for the sample (''selected with certainty''). The remaining sample of ''other'' women (white and other—not black or Hispanic) not selected with certainty was allocated to the PSU's based on the weighted count of women in each PSU in the three cluster strata.

Sample Selection Within PSU's

For the sample selection of households in each PSU, Chromy's probability minimal replacement sequential selection procedure (6) was used with the weighted number of women in a household as the size measure. The sampling frame in each PSU was stratified by the cluster type (mixed and high-density clusters and low-density clusters). The sampling frame within each stratum was then ordered so that households with similar geographic information (in MSA, central city; in MSA, not in the central city; and not in MSA) were close together. After the household was selected, one woman was randomly selected from each household.

Characteristics of the Sample

Designated Sample Sizes and Probabilities of Selection

Table B shows the number of cases selected from the 1993 National Health Interview Survey (NHIS) and the average selection probability, average sampling weight, and the relative variance of the sampling weights. The average selection probabilities for Hispanic and non-Hispanic black women are 70 to 120 percent higher than for the other race/ethnicity group, since the Hispanic and non-Hispanic black women were oversampled. In addition, selecting only one woman per household resulted in lower selection probabilities (and larger sampling weights) for women in larger

Table B. Designated sample sizes, probability of selection, and average and relative variance of weights, by race/ethnicity and number of eligible women: 1995 National Survey of Family Growth

Race/ethnicity and number of eligible women in household[1]	Sample sizes	Probability of selection	Average weight[2]	Relative variance
All women 15–44 years old	14,000	0.00026124	3,828	0.20
Race/ethnicity				
Hispanic	2,097	0.00036788	2,718	0.30
Non-Hispanic black	3,205	0.00046559	2,148	0.41
Other	8,698	0.00021211	4,715	0.06
Number of eligible women in household				
1	10,546	0.00028619	3,494	0.20
2	2,841	0.00022330	4,478	0.11
3	526	0.00015977	6,259	0.08
4	73	0.00012545	7,971	0.07
5 or more	14	0.00009648	10,365	0.14

[1]Race/ethnicity based on data from the National Health Interview Survey.
[2]The weight is the full-sample sampling weight before any nonresponse or poststratification adjustments.

households than women in smaller households. This occurred almost exclusively in the households with Hispanic or non-Hispanic black women, because all of these households were selected in the sample with certainty. The race/ethnicity classification in table B is based on the NHIS-reported data.

The *relative variance* of the sampling weights in table B is a measure of the potential increase, or decrease, in the sampling variances (the design effect) attributable to unequal sampling weights from over or undersampling portions of the survey population. The relative variance is computed as the ratio of the variance of the full-sample sampling weight to the square of the average weight. The relative variance for the full sample is 0.20, the relative variance for Hispanic women is 0.30, for non-Hispanic black women it is 0.41, and for other women it is 0.06. In addition, selecting only one NSFG sample woman in a household with several women eligible for the NSFG results in a smaller within-household sampling rate (and sometimes a larger full-sample weight) than that experienced in a household with only one woman eligible for the NSFG.

By taking into account the number of women eligible for the NSFG in sampling households, almost all the variability in the overall sampling rates for non-Hispanic, nonblack women was eliminated. The NSFG selected all NHIS households containing only Hispanic or black women, so whatever variations in sampling rates were in the NHIS also appeared in the NSFG. However for white and other women, the sampling rates could be adjusted because not every white woman was selected for the NSFG. As a result, the overall NSFG Cycle 5 sampling rates for Hispanic and black women vary more. This increases their variances somewhat, but that is acceptable because the sample sizes for Hispanic and black women are large enough to be useful for analysis.

Interview Rates

The location and response rates measure the operational performance and the potential for bias in survey estimates. The location rate is the percentage of cases in the sample that were located, where located means obtaining a valid address. The response rate is the percentage of eligible cases for which a completed interview was obtained. Eligible cases include those who completed an interview and those who refused, were not home, or were never located or contacted. Location and response rates for Cycle 5 are shown in table C. The location and response rates are based on the actual count of located cases, completed interviews, and sample cases (14,000 in total)—205 (1.5 percent) were determined to be ineligible and 757 unlocated cases were assumed to be eligible. Subtracting the 757 unlocated and the 205 ineligible from 14,000, 13,038 cases (93.1 percent) were located and eligible. Of the 13,795 eligible cases, 10,847 were completed interviews (78.6 percent, unweighted and 78.7 percent weighted).

All three race/ethnicity groups had similar overall response rates. For the Hispanic and black women, the location rates were lower—about 92 percent for Hispanic or black and 96 percent for non-Hispanic, nonblack women. But the participation rates among Hispanic and black located eligible cases were higher. Sample women under 24 years of age had the highest overall response rates (82 percent) and women 25–29 years had the lowest overall response rate (74.5 percent). Sample women 25–29 years of age had both the lowest location rate (91.7 percent) and the lowest overall response rate. Race/ethnicity and age (as of April 1, 1995) are based on NHIS data.

Sample Sizes, Clustering, and Variation in the Probability of Selection

Cluster size has important effects on survey costs and variances. Larger cluster size tends to reduce costs because interviewers spend less time and money traveling between sample households. But larger cluster sizes also tend to increase variances because

Table C. Response rates among completed cases in the National Health Interview Survey, by race/ethnicity and age: 1995 National Survey of Family Growth

Race/ethnicity[1] and age	Sample sizes	Located cases	Location rate	Eligible women[2]	Completed interviews	Response rate
All women 15–44 years	14,000	13,243	94.6	13,795	10,847	78.6
Race/ethnicity[1]						
Hispanic	2,097	1,926	91.8	2,030	1,613	79.5
Non-Hispanic black	3,205	2,939	91.7	3,169	2,464	77.8
Other	8,698	8,378	96.3	8,596	6,770	78.8
Age[3]						
15–17 years	1,040	1,001	96.3	1,020	841	82.5
18–24 years	2,622	2,452	93.5	2,586	2,122	82.1
25–29 years	2,339	2,146	91.7	2,310	1,722	74.5
30–34 years	2,815	2,656	94.4	2,783	2,172	78.0
35–39 years	2,751	2,632	95.7	2,723	2,127	78.1
40–44 years	2,433	2,356	96.8	2,373	1,863	78.5

[1]Race/ethnicity and age are based on data from the National Health Iinterview Survey.
[2]Unlocatables were assumed to be eligible.
[3]Age as of April 1, 1995.

people who live near each other tend to be similar.

Table D shows the average cluster size and the variation in the weights among completed interviews. The average cluster size is the number of interviews of the indicated type that were obtained, on average, from essentially the same neighborhood (that is, the NHIS segment). The number of clusters with one or more completed interviews was 5,377 (compared with 3,143 in Cycle 4), and the average number of completed interviews per cluster was 2.01 (compared with 2.69 in Cycle 4) (7). This increase in the absolute number of NSFG clusters was necessary to avoid variation in the sampling rates for ‘‘other'' women (not black or Hispanic). This strategy increased the reliability of survey estimates including these women; still, this increased number of survey clusters probably increased survey costs compared with a design with substantially fewer clusters, such as that used for Cycles 1–4. Using the race/ethnicity from the NSFG interview, non-Hispanic black women were more clustered (an average of 1.86 per cluster) than the Hispanic and non-Hispanic, nonblack women (1.52 and 1.69, respectively).

All reports published by NCHS from the NSFG have the data weighted appropriately; that is, cases from underrepresented groups have a larger weight than cases from overrepresented groups. Users of the detailed data file are cautioned that *analyzing the data without weights understates the variances and exaggerates biases that are corrected by the weights.* The last column of table D is the relative variance of the unbiased weights. In a simple random sample, the relative variance of unbiased weights is zero because all sample cases have the same probability of selection. The larger the value of the relative variance of the unbiased weights, the more the probabilities of selection vary. This variation in the probabilities increases the sampling error of the estimates for all women but increases the reliability of the data for the group being oversampled, primarily black or Hispanic women. Thus, ignoring the weights in analysis and significance testing may lead analysts to claim that differences are significant when, in fact, they are not.

Table D. Clustering and weight variation among completed cases in the 1995 National Survey of Family Growth, by race/ethnicity and age

Race/ethnicity and age	Completed interviews	Number of clusters with 1 or more completes	Average number of completes per cluster	Relative variance of fully adjusted weights
All women 15–44 years old	10,847	5,377	2.01	0.23
Race/ethnicity[1]				
Hispanic	1,553	1,020	1.52	0.34
Non-Hispanic black	2,446	1,316	1.86	0.45
Other	6,848	4,064	1.69	0.11
Age[1,2]				
15–17 years	828	786	1.05	0.19
18–19 years	580	561	1.03	0.20
20–24 years	1,526	1,238	1.23	0.24
25–29 years	1,716	1,473	1.16	0.28
30–34 years	2,165	1,875	1.15	0.23
35–39 years	2,125	1,814	1.17	0.20
40–44 years	1,907	1,655	1.15	0.20

[1]Race/ethnicity and age are based on the National Survey of Family Growth interview.
[2]Age as of April 1, 1995.

Sampling Weights

Summary

Data from the 10,847 women in the NSFG sample are used to estimate percents, averages, and other measures for the entire population of 60.2 million women 15–44 years of age in the United States in 1995, and for subgroups such as Hispanic women, black women, teenagers, and others. This is done by attaching a ‘‘sampling weight'' to each sample case to denote the number of women in the population that she represents. The weight for each sample case is then summed to make an estimate for the total population. The average weight for NSFG cases is 60,201,000/10,847 = 5,550. But the weights vary considerably. For example, for other women (non-Hispanic, nonblack) the average weight is 6,559; for Hispanic women, 4,316; and for non-Hispanic black women, 3,357. Given the importance of the NSFG data, considerable effort was spent to construct weights that would produce unbiased, accurate national estimates.

The following section describes in detail how the weights were derived and reiterates the importance of using the weights in analysis of NSFG data.

The weight assignment process for Cycle 5 of the NSFG consisted of four adjustment factors. These were applied to the NHIS weights of the 10,847 Cycle 5 respondents to adjust for the *subsampling nonlocation*, *nonresponse*, and *noncoverage* of sample women:

$$W_{0i} \bullet A_{1i} \bullet A_{2i} \bullet A_{3i} \bullet A_{4i} = W_{4i}$$

where

W_{0i} = NHIS weight assigned to sample woman *i*

A_{1i} = adjustment for differing sampling rates for Hispanic, non-Hispanic black, and other women in Cycle 5

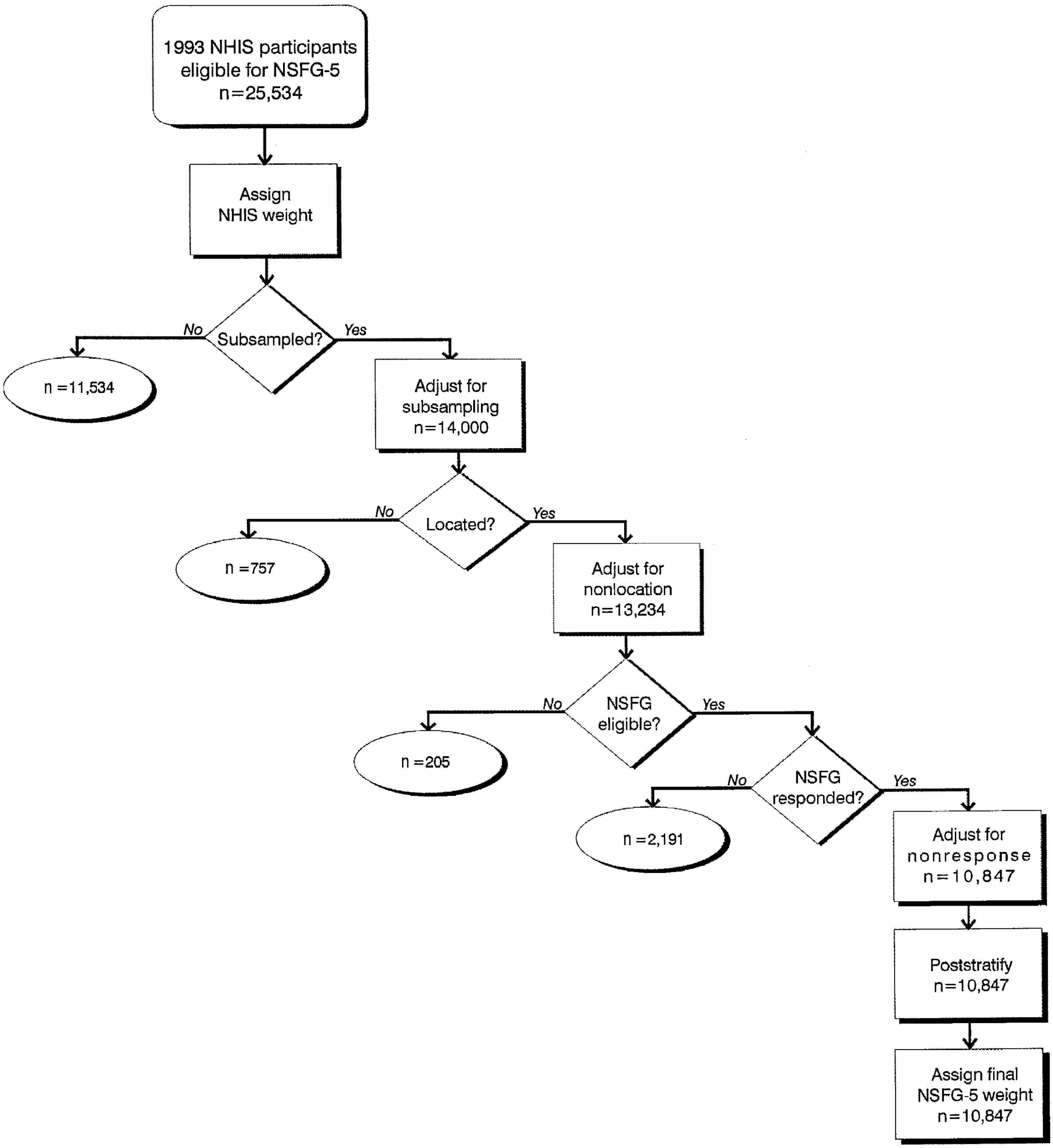
A_{2i} = adjustment for inability to locate some women

A_{3i} = adjustment for nonresponse among those located

A_{4i} = poststratification or adjustment for noncoverage

W_{4i} = fully adjusted Cycle 5 weight

The fully adjusted Cycle 5 weight is called POST_WT in the NSFG data file (locations 12,350–12,359 on the respondent file). A flowchart of the weight assignment process for Cycle 5 is shown in figure 1.



NOTE: NHIS is National Health Interview Survey and NSFG is National Survey of Family Growth.

Figure 1. National Survey of Family Growth Cycle 5 weight assignment process

Weight Adjustments for Location and Response Propensity

Weighting-class adjustments are made to the sampling weights to reduce potential nonresponse bias. The Cycle 4 weighting classes were formed with *segmentation modeling,* which was implemented with the CHAID (Chi-squared Automatic Interaction Detection) software (7,8). This technique was used to divide the Cycle 4 sample into segments, or weighting classes, that had different response rates. Segmentation modeling successively splits a sample into smaller subgroups based on their response rates. The splitting process continues until no more statistically significant predictors can be found or until the subgroups are too small to continue. Weighting-class adjustments are based on the assumption that sample women can be partitioned into cells or weighting classes within which the responses of nonrespondents, had they been obtained, would be similar to those of respondents. Within each weighting class, the inverse of the weighted response rate is applied to the sampling weights of responding sample women so that the sums of the adjusted weights for respondents reproduce the sums of the unadjusted weight for respondents and nonrespondents.

A noticeable distinction between the weight adjustments for Cycles 4 and 5 was the use of *response propensity modeling* through logistic regression in Cycle 5 (9). This technique models the functional relationship between a set of response predictors and a (dichotomous) response outcome. *If the relationship is significant, the model-based adjustment factors greatly reduce the potential for nonresponse bias.* In addition, response propensity modeling provides a formal statistical setting for evaluating variables believed to be related to response. This was particularly useful for evaluating the large number of potential predictor variables available from the NHIS data.

Hispanic and non-Hispanic black sample women were harder to locate than were other sample women. Once found, however, Hispanic and black women were more likely to participate in the survey than other sample women. Separate adjustments were needed to reflect the distinct patterns of availability, including change of address, lack of some or all contact information, and resistance to participation. Mobility was an expected artifact of the Cycle 5 sampling design because of the linkage to the 1993 NHIS and the long time period between the two surveys (13–34 months). The lack of contact information may be generally considered as an indicator of resistance to future survey participation in some cases, and a failure on the part of the interviewer to collect accurate and complete contact information during the NHIS interview in other cases.

The adjustment factors for location and response propensity were calculated separately for each case because each case's values were inserted into a logistic regression equation. As a result, response propensity adjustments can help to reduce nonresponse bias by following the actual response pattern of individual sample women more closely than weighting class adjustments. However, these gains in accuracy may be offset if the variation among the weights causes an excessive increase in variances (10). The unequal weighting effects are shown by race/ethnicity in table E along with the mean and maximum adjustment factors. Note that the mean adjustment factor for subsampling the ‘‘other’’ race category (2.16) is higher than for the Hispanic or the non-Hispanic black categories. This reflects the lower sampling rates that were applied to women in the ‘‘other’’ race category. The subsequent adjustments for location and response propensity and poststratification had little additional effect on unequal weighting. Since no weight was excessively large, weight trimming was unnecessary.

Table E. Weight adjustment summary by race/ethnicity and type of adjustment

Race/ethnicity	Number of sample women	Mean adjustment	Maximum adjustment	Unequal weight effect[1]
NHIS subsampling adjustment: (A1)[2]				
Hispanic	2,097	1.28	7.00	1.30
Black, non-Hispanic	3,205	1.27	6.00	1.41
Other	8,698	2.16	14.24	1.06
Total	14,000	1.83	14.24	1.20
Location-propensity adjustment: (A2)[3]				
Hispanic	1,926	1.08	2.54	1.31
Black, non-Hispanic	2,939	1.08	4.02	1.42
Other	8,378	1.04	2.90	1.07
Total	13,243	1.06	4.02	1.20
Response-propensity adjustment: (A3)[3]				
Hispanic	1,612	1.15	1.89	1.32
Black, non-Hispanic	2,465	1.19	2.66	1.42
Other	6,770	1.23	3.04	1.09
Total	10,847	1.20	3.04	1.23
Poststratification: (A4)[4]				
Hispanic	1,553	1.24	1.88	1.34
Black, non-Hispanic	2,446	1.21	2.52	1.46
Other	6,848	1.12	1.62	1.11
Total	10,847	1.16	2.52	1.23

[1]The unequal weighting effect measures the relative increase in sampling variance attributable to unequal weighting, assuming that equal weighting is optimal.
[2]NHIS is National Health Interview Survey.
[3]Race/ethnicity based on NHIS classification.
[4]Race/ethnicity based on Cycle 5 classification.

National Health Interview Survey Sampling Weights and Adjustments for Cycle 5 Subsampling

The Cycle 5 sampling design permits estimates for civilian noninstitutionalized women between the ages of 15 and 44 living in the United States during 1995. The stratification, clustering, and disproportionate sampling in the sample design imposed a design effect. The design effect is the increase or decrease in the sampling variance attributable to the sampling design relative to the sampling variance of a simple random sample of the same size from the same population.

NHIS Sampling Weight

The NHIS is a stratified multistage household survey that covers the civilian noninstitutionalized population of the United States. A complete description of the NHIS design is given in reference 5. The 1993 NHIS respondent sample included data for 109,671 persons in 43,007 households. The sampling weights for these respondents were computed from the NHIS household basic quarter weight before age-sex-race adjustments (1993 NHIS person file, tape positions 164–169). The NHIS basic quarter weight is designed to make national estimates using data collected in just one quarter. When data are used from a full year of NHIS, the weight is adjusted (divided by four). Because the NHIS sample for the second quarter of 1993 contained only 7 weeks (rather than 13 weeks), the annual NHIS household sampling weight for women in the *i*-th household (HH_i) was computed as:

$$W_{0i} = \begin{cases} 7/46 \bullet \text{basic quarter weight for } HH_i, \\ \quad \text{if sampled in 2d quarter,} \\ 13/46 \bullet \text{basic quarter weight for } HH_i \\ \quad \text{if sampled in other quarters.} \end{cases}$$

Adjustments for Cycle 5 Subsampling

All households containing Hispanic or non-Hispanic black women were selected for the Cycle 5 sample; one woman per household was selected. In other households, the subsampling rates were set at levels designed to achieve equal overall sampling weights. The only exceptions to this equal weighting design were a slight oversampling of households in high-density clusters to reduce data collection costs, and selecting households that represent a large number of women in the population.

The initial adjustment factor applied to the NHIS annual household weight of the 14,000 selected sample women was the inverse of the subsampling rate. That is,

$$W_{1i} = A_{1i} \bullet W_{0i}$$

where A_{1i} = the inverse of the subsampling rate applied to the *i*th sampled woman (SW_i).

As shown in table E, the adjustment factors for subsampling ranged from 1 to 14.24 and nearly eliminated the unequal weighting effect among sample women in the ‘‘other’’ race/ethnicity stratum—it was only 1.06 (table E). The mean adjustment factor of 2.16 in the ‘‘other’’ stratum reflects the lower sampling rate for women in this stratum compared with Hispanic and non-Hispanic black women.

Response Probability Modeling

For Cycle 5, response probability modeling was used to extend the group-level adjustments of the weighting-class approach that was used in Cycle 4 to sample woman-level adjustments derived from the predicted response propensities of a logistic regression model. The primary advantage of response probability modeling over the approach used in Cycle 4 is that a larger number of main effect variables can be used in the adjustment procedure.

With the logistic modeling approach, the marginal totals and any interaction effects between variables will be preserved by including the corresponding main effect and interaction variables in the model. In addition, regression modeling allows valid statistical tests (using SUDAAN) of the significance of the regression coefficients.

Instead of examining ‘‘all possible interactions’’ to detect interaction effects, segmentation modeling (CHAID) was used (8). The significant interactions found using the CHAID model were included in the logistic regression procedure in SUDAAN (11).

Model Development

The variables used in the model are shown in table F. Some of these variables may be viewed as indicators of unavailability, while others may be viewed as indicators of resistance or hostility to surveys. For example, residence in temporary quarters, or multiple jobs may be indicators of unavailability, while refusing to give one’s Social Security number or the telephone number of a contact person may indicate resistance. Distinctive patterns were found, suggesting that the locating process should be treated as a different outcome variable than the cooperation process among those who were located. Tabulations of the NHIS variables used in the development of the location and response propensity models are presented in table 1.

The overall response propensity for a sample woman may be subdivided into the following components. A zero-one indicator L_i for the *i*-th sample woman (SW_i) may be defined as follows: 1 if she was located and 0 if she was not located. Among the 13,038 sample women who were located and found to be eligible for the Cycle 5, R_i was set to 1 if the sampled woman responded (that is, completed the interview) and 0 if she did not.

Then, the overall probability that SW_i responds may be written as

$$\begin{aligned} P[R_i = 1] &= P[L_i = 1] \bullet P[R_i = 1 \mid L_i = 1] \\ &= \lambda_i \bullet \rho_i \end{aligned}$$

That is, the overall probability that a sample woman participates in the survey is equal to the probability that she is located times the probability that she agrees to participate once located.

This approach led to two logistic regression models. The first model for location propensity was applied to the entire sample of 14,000 sample women. The second model for response propensity was applied to the 13,038 sample women who were located and eligible for the study. To simplify the response propensity modeling procedures, an available general model and computer software was used to estimate the model parameters (12).

Location Propensity Model

The following logistic model was developed for the probability that sampled woman i was located:

$$\lambda_i = P\,[L_i = 1 \mid X_i, \beta] = [1 + \exp(-X_i^T \beta)]^{-1}$$

where

X_i = $(1, X_{1i}, \ldots, X_{21i})$, a 22-element vector with a 1 as the first element (the intercept) followed by 21 predictor variables

β = a vector of logistic regression coefficients (for notational convenience the intercept term was included in the vector)

The logistic regression coefficients $\hat{\beta}$ were estimated iteratively by solving the following estimation equations:

$$\sum_{i \in S} (W_{1i} \div \hat{\lambda}_i) X_i^T \hat{\lambda}_i = \sum_{i \in S} (W_{1i} \div \hat{\lambda}_i) X_i^T L_i,$$

where

S = sample of 14,000 women, and

$\hat{\lambda}_i$ = $[1 + \exp(-X_i^T \hat{\beta})]^{-1}$

The location adjusted weight is

$$W_{2i} = W_{1i} \bullet A_{2i}$$

where

$$A_{2i} = \begin{cases} \hat{\lambda}_i^{-1} & \text{if } L_i = 1 \\ 0 & \text{otherwise} \end{cases}$$

Because the first element of X_i is uniformly 1, the adjusted W_{2i} weight sums for located sample women ($L_i = 1$) equal the corresponding W_{1i} weight sum across all sample members. In addition, the weight sum equality constraint holds for any sample subset identified by any zero-one indicator in X_i.

Response Propensity Model

The following logistic model was developed for the probability of participation given that SW_i was located and eligible:

$$\rho_i = P\,[R_i = 1 \mid L_i = 1, Z_i^T \theta] = [1 + \exp(-Z_i^T \theta)]^{-1}$$

Table F. National Health Interview Survey variables used to predict location rates and response rates for the 1995 National Survey of Family Growth

NHIS candidate predictor[1,2]	Significant predictor[3]: Location model	Significant predictor[3]: Response model
Demographic variables		
Family income	X	X
Age	X	X
Marital status	X	X
Poverty level	X	X
Region	X	X
Metropolitan statistical area status	X	X
Education	X	X
Health status	X	X
Race	. . .	X
Ethnicity	. . .	X
Major activity	. . .	X
Class of worker	. . .	X
Employment status	. . .	X
Number of children in household	. . .	X
Number of doctor visits in past year	. . .	X
Time since last doctor visit	. . .	X
Number of conditions	. . .	X
Predominantly black area	X	. . .
Family relationship of sample woman	X	. . .
Urban/rural	. . .	. . .
Living quarters	. . .	. . .
Number of families in household	. . .	. . .
Relationship to NHIS reference person	. . .	. . .
NHIS contact variables		
Name not provided	X	X
Record of calls	X	X
Contact name provided	X	X
NHIS respondent status	. . .	X
Telephone number refused	X	. . .
Social Security number refused	X	. . .
Number of calls	. . .	. . .
Number of additional contacts	. . .	. . .
Number of callbacks for Social Security number	. . .	. . .
Number of callbacks for immunizations	. . .	. . .
Refused height, weight, or health status	. . .	. . .

. . . Variable not significant.
[1]NHIS is National Health Interview Survey.
[2]The levels of each NHIS candidate predictor as well as the location and response rates are presented in table 1.
[3]Predictors included as main effects or segments in the final logistic regression models (= 0.10).
NOTES: The list of NHIS candidate variables comprises all variables on the NHIS public-use files believed to be potentially related to location and/or response propensity. X = signficant at 0.10.

where

Z_i = $(1, Z_{1i}, \ldots, Z_{27i})$, a 28-element vector with a 1 as the first element (the intercept) followed by 27 predictor variables

θ = a vector of logistic regression coefficients (for notational convenience the intercept term is included in this vector)

Analogous to the location propensity model, the logistic regression coefficients $\hat{\theta}$ were estimated iteratively by solving the following estimation equations:

$$\sum_{i \in \xi} (W_{2i} \div \hat{\rho}_i) Z_i^T \hat{\rho}_i = \sum_{i \in \xi} (W_{2i} \div \hat{\rho}_i) Z_i^T R_i$$

where

ξ = Sample of 13,038 located, eligible sample women

$\hat{\rho}_i$ = $[1 + \exp(-Z_i^T \hat{\theta})]^{-1}$

The response-adjusted weight is

$$W_{3i} = W_{2i} \bullet A_{3i}$$

where

$$A_{3i} = \begin{cases} \hat{\rho}_i^{-1} & \text{if } R_i = 1 \\ 0 & \text{otherwise} \end{cases}$$

Table G. Final location propensity logistic model for the 1995 National Survey of Family Growth

Location predictors from the 1993 NHIS[1]	Odds ratio[2]	Beta coefficient	SE beta	Design effect	*t*-Test beta = 0	*p*-Value beta = 0[3]
Intercept	. . .	3.81	0.25	1.01	15.34	<0.0001
Sample woman is 25–29 years old	0.37	–0.99	0.13	1.09	–7.56	<0.0001
Segment–2	0.23	–1.46	0.20	1.04	–7.40	<0.0001
Segment–8	0.13	–2.03	0.28	1.01	–7.14	<0.0001
Segment–3	0.31	–1.18	0.22	1.11	–5.39	<0.0001
Name not provided in NHIS	0.27	–1.30	0.24	1.28	–5.34	<0.0001
Segment–5	0.36	–1.01	0.20	1.27	–5.11	<0.0001
NHIS contact person unknown	0.60	–0.52	0.11	1.11	–4.52	<0.0001
Not married, over 14 years old	0.60	–0.51	0.12	1.04	–4.11	0.0001
Segment–4	0.56	–0.59	0.15	1.15	–3.97	0.0002
Segment–10	0.40	–0.93	0.24	1.14	–3.92	0.0002
Sample woman is 30–34 years old	0.60	–0.51	0.13	1.06	–3.87	0.0002
Sample woman is 15–24 years old	0.62	–0.48	0.13	1.09	–3.81	0.0003
Telephone number given	1.89	0.64	0.19	1.10	3.29	0.0016
Segment–9	0.39	–0.95	0.29	0.97	–3.25	0.0018
Midwest	1.55	0.44	0.15	1.41	3.02	0.0035
Some high school or high school graduate	0.73	–0.31	0.10	1.07	–2.96	0.0041
Record of calls 0–3	1.31	0.27	0.09	1.14	2.88	0.0052
Education less than high school/unknown	0.69	–0.38	0.15	1.12	–2.46	0.0165
Segment–13	0.44	–0.82	0.34	1.22	–2.42	0.0182
Above poverty level	1.26	0.23	0.11	1.16	2.09	0.0401
Health status: excellent	1.19	0.17	0.10	1.20	1.69	0.0948

. . . Category not applicable.

[1]NHIS is National Health Interview Survey; MSA is Metropolitan Statistical Area; Pov Lev is poverty level; and SSN is Social Security number; Unk is unknown, and SW is sample woman.

[2]The ratio of location propensity of the predictor to location propensity of the reference cell of the predictor. For example, women 25–29 years old were 37 percent as likely to be located as women 35–44 years old (the reference cell for age category) after adjusting for the other predictors in the model.

[3]Cutoff level of significance: 0.10.

Segment definitions:

Seg–2: Family income unknown or less than $20,000 and telephone number given and MSA greater than or equal to 1 million and contact person Unk and Pov Lev below or Unk.

Seg–3: Family income unknown or less than $20,000 and telephone number given and MSA greater than or equal to 1 million and contact person known.

Seg–4: Family income unknown or less than $20,000 and telephone number given and MSA less than 1 million and Northeast, South, or West and SSN given.

Seg–5: Family income unknown or less than $20,000 and telephone number given and MSA less than 1 million and Northeast, South, or West and SSN refused.

Seg–8: Family income unknown or less than $20,000 and telephone number refused, Unk or no number and *SW* name not missing and MSA greater than or equal to 1 million living alone or with sample woman.

Seg–9: Family income unknown or less than $20,000 and telephone number refused, Unk or no number and *SW* name not missing and MSA greater than or equal to 1 million living with other relative.

Seg–10: Family income unknown or less than $20,000 and telephone number refused, Unk or no number and *SW* name not missing and MSA less than 1 million and not black area.

Seg–13: Family income unknown or less than $20,000 and telephone number refused, Unk or no number and *SW* name missing.

Because the first element of Z_i is uniformly 1, the adjusted W_{3i} weight sums for responding sample women (that is, $R_i = 1$) equal the corresponding W_{2i} weight sum across all eligible sample members. In addition, the weight sum equality constraint holds for any sample subset identified by any zero-one indicator in Z_i.

The components of Z_i, the vector of predictors for the response propensity model, and X_i, the vector of predictors for the location propensity model, are described in the next sections.

Factors Affecting the Proportion Located

Table F lists the variables on the NHIS that were available for estimating location propensity (that is, the proportion located). All variables were entered in the model, but not all were significant. As expected, predictors indicating the presence or absence of NHIS contact data were significant factors in the final location propensity model shown in table G. The segmentation of the sample shown in figure 2 suggests that the effect of these predictors was affected by a number of demographic factors, especially family income. For example, figure 2 shows that among women with low or unknown family incomes, 94 percent of women whose telephone number was known were located compared with 84 percent of low-income sample women whose phone number was not known. In fact, the lowest segment-level location rate (63.4 percent) occurred among sample women with low or unknown family incomes who either refused to report or did not have a telephone number and who did not provide her name.

In contrast, only one significant predictor of location rates was found among women with family incomes of $20,000 or more (ages 15–39 versus ages 40–44). The lack of segmentation among sample women with a known family income of $20,000 or more suggests that sample women who were willing to provide income (a traditionally sensitive item) were also likely to provide contact information. Other notable NHIS contact indicators included providing the name of a contact person and willingness to provide a Social Security number, which was used for locating a sample member.

Among demographic characteristics, the age and family income of sample women were the most statistically significant factors affecting location propensity. In general, sample women with a low or unknown family income were harder to locate than those with a family income of $20,000 or more, although the large number of segments created within this group implies several exceptions. Other notable demographic characteristics in the model included

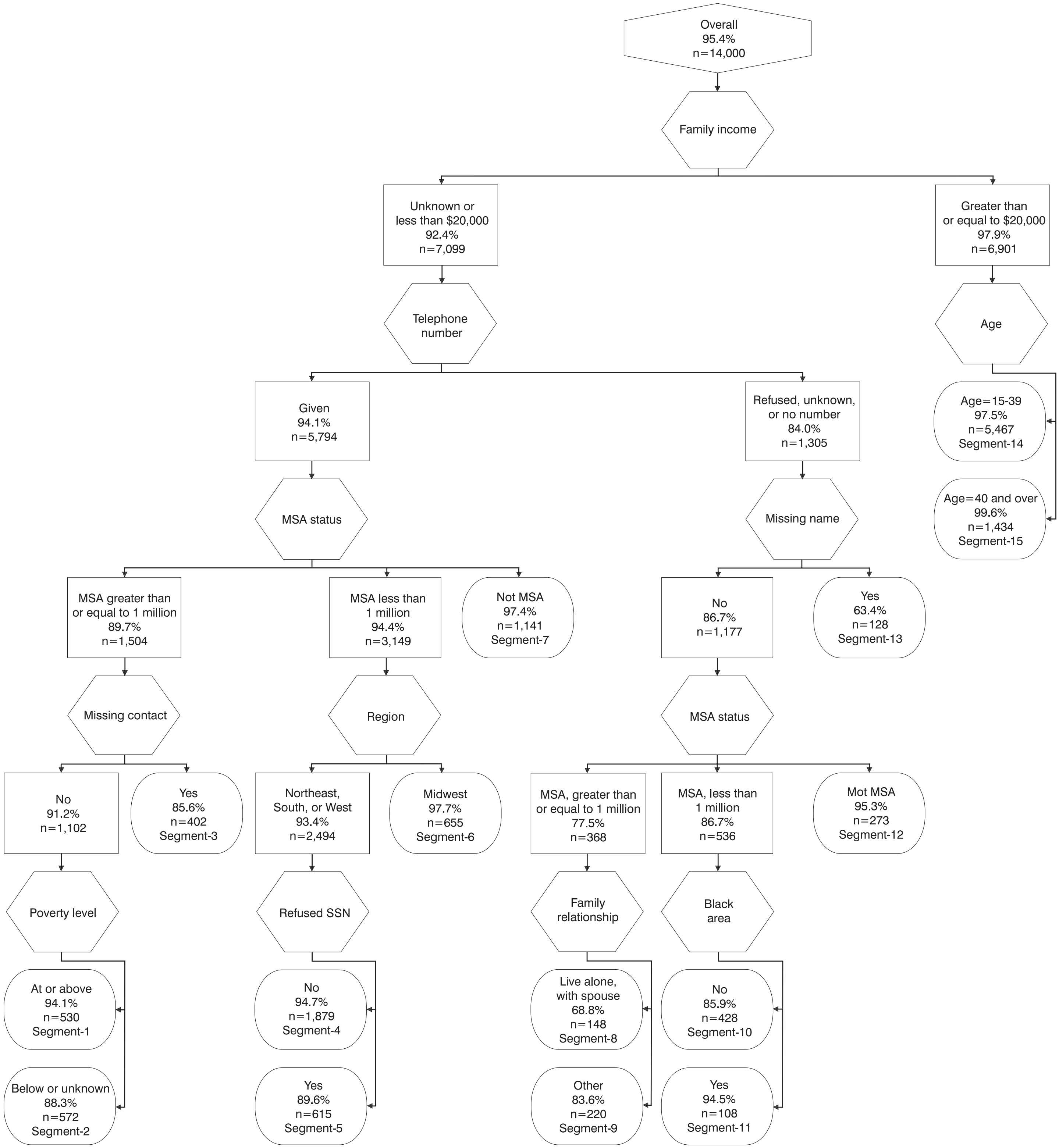


Figure 2. Final segmentation of the 1995 National Survey of Family Growth sample members by percent located

Table H. Final response propensity logistic model for the 1995 National Survey of Family Growth

Response predictors from the 1993 NHIS[1]	Odds ratio[2]	Beta coefficient	SE beta	Design effect	T-Test beta = 0	*p*-Value beta = 0[3]
Intercept	. . .	1.53	0.24	1.39	6.30	<0.0001
Segment–9	0.37	–0.99	0.16	1.46	–6.08	<0.0001
Segment–15	0.50	–0.69	0.12	1.14	–5.76	<0.0001
NHIS contact person unknown	0.65	–0.43	0.07	1.23	–5.76	<0.0001
Hispanic or non-Hispanic black	1.38	0.32	0.06	1.01	5.08	<0.0001
One or no children	0.75	–0.29	0.06	1.13	–4.86	<0.0001
Sample woman is 15–24 years old	1.60	0.47	0.10	1.25	4.74	<0.0001
Self-respondent	1.35	0.30	0.06	1.30	4.79	<0.0001
Going to school less than 18 years	1.79	0.58	0.17	1.21	3.47	<0.0009
Income provided: less than $50,000	1.32	0.27	0.08	1.25	3.46	0.0009
Segment–4	0.46	–0.78	0.23	1.39	–3.39	0.0011
Segment–5	1.56	0.45	0.14	1.26	3.25	0.0017
Northeast or South	0.81	–0.21	0.07	1.74	–3.07	0.0030
Telephone number refused or unknown	0.56	–0.59	0.19	1.48	–3.07	0.0030
Name not provided in NHIS	0.60	–0.52	0.17	1.52	–2.98	0.0039
Segment–2	1.42	0.35	0.12	1.37	2.81	0.0063
Income provided: $50,000 or more	1.28	0.25	0.09	1.25	2.77	0.0072
Segment–14	0.70	–0.35	0.13	1.22	–2.72	0.0083
Number of health conditions: zero	0.85	–0.17	0.06	1.50	–2.59	0.0116
Segment–13	2.24	0.81	0.32	1.08	2.49	0.0150
Two years or less since last doctor visit	0.85	–0.16	0.07	1.41	–2.39	0.0192
SW is employed	1.16	0.15	0.06	1.21	2.38	0.0198
Segment–17	0.61	–0.49	0.21	1.29	–2.31	0.0240
MSA/central city or not central city	0.83	–0.18	0.09	2.01	–2.03	0.0459
Record of calls, two or less	1.11	0.11	0.05	1.19	1.96	0.0540
Working or keeping house	1.30	0.26	0.15	1.21	1.78	0.0797
Segment–7	1.63	0.49	0.27	1.69	1.77	0.0814
Telephone number given	0.79	–0.24	0.14	1.41	–1.71	0.0910

. . . Category not applicable.

[1]*SW* = sample woman; NHIS is National Health Interview Survey; Hlth Stat is health status; VG is very good; Doc is doctor; PI is Pacific Islander.

[2]The ratio of response propensity of the predictor to the response propensity of the reference cell of the predictor. For example, women in households with one or no children were 75 percent as likely to respond as women in households with more than one child (the reference cell for number of children in a household) after adjusting for the other predictors in the model.

[3]Cutoff level of significance: 0.10.

Segment definitions:

Seg–2: SSN given and racial background is white or other and two or less doctor visits and at or above or below poverty level and region is West.

Seg–4: SSN given and racial background is white or other and two or less doctor visits and poverty level unknown and record of calls three or more.

Seg–5: SSN given and racial background is white or other and three or more doctor visits and contact person known and record of calls one or less.

Seg–7: SSN given and racial background is white or other and two or more doctor visits and contact person unknown and family income under $50,000.

Seg–9: SSN given and racial background is Asian or PI and marital status is married/separated/less than 14 years or unknown.

Seg–13: SSN refused and Hispanic origin and major activity is not working.

Seg–14: SSN refused and not Hispanic origin and contact person known and MSA central city or not MSA.

Seg–15: SSN refused and not Hispanic origin and contact person known and MSA not central city.

Seg–17: SSN refused and not Hispanic origin and contact person unknown and hlth stat excellent or vg and last doc visit is one or more years.

marital status, education, region, and MSA status.

Factors Affecting Response Propensity

As in the location propensity model, several of the predictors related to the presence of contact data also were significant in the response propensity model shown in table H and figure 3. For example, more than 2,000 sample women refused to provide a Social Security number during the NHIS but were subsequently located and found eligible for Cycle 5. Once located, however, these sample women were significantly less likely to participate (72.8 percent, figure 3) than the 11,009 who provided their Social Security number (84.2 percent). Similar patterns can be seen for refusal to provide a telephone number or the name of a contact person in the segmentation modeling of located eligibles shown in figure 3.

Among the 11,009 women who gave their Social Security numbers to the NHIS interviewer (figure 3), the 335 Asian or Pacific Islanders had a lower response rate (69.3 percent) than the white and black women. Among the 2,029 women who refused to give their Social Security numbers to the NHIS interviewer, the 341 Hispanic women had a higher response rate (86.1 percent) than the 1,688 non-Hispanic women (70.8 percent).

Evaluation of the Combined Location and Response Models

Generalized Wald statistics, adjusted for design effects, were used to test the goodness-of-fit of the location and response propensity models. However, the overall predicted probability of response was not amenable to conventional regression analysis because of the lack of independence between the models. Therefore, a receiver operating characteristics (ROC) curve was used to assess the overall predictive ability of the combined model. An example of the use of ROC curves to evaluate response propensity models is described by

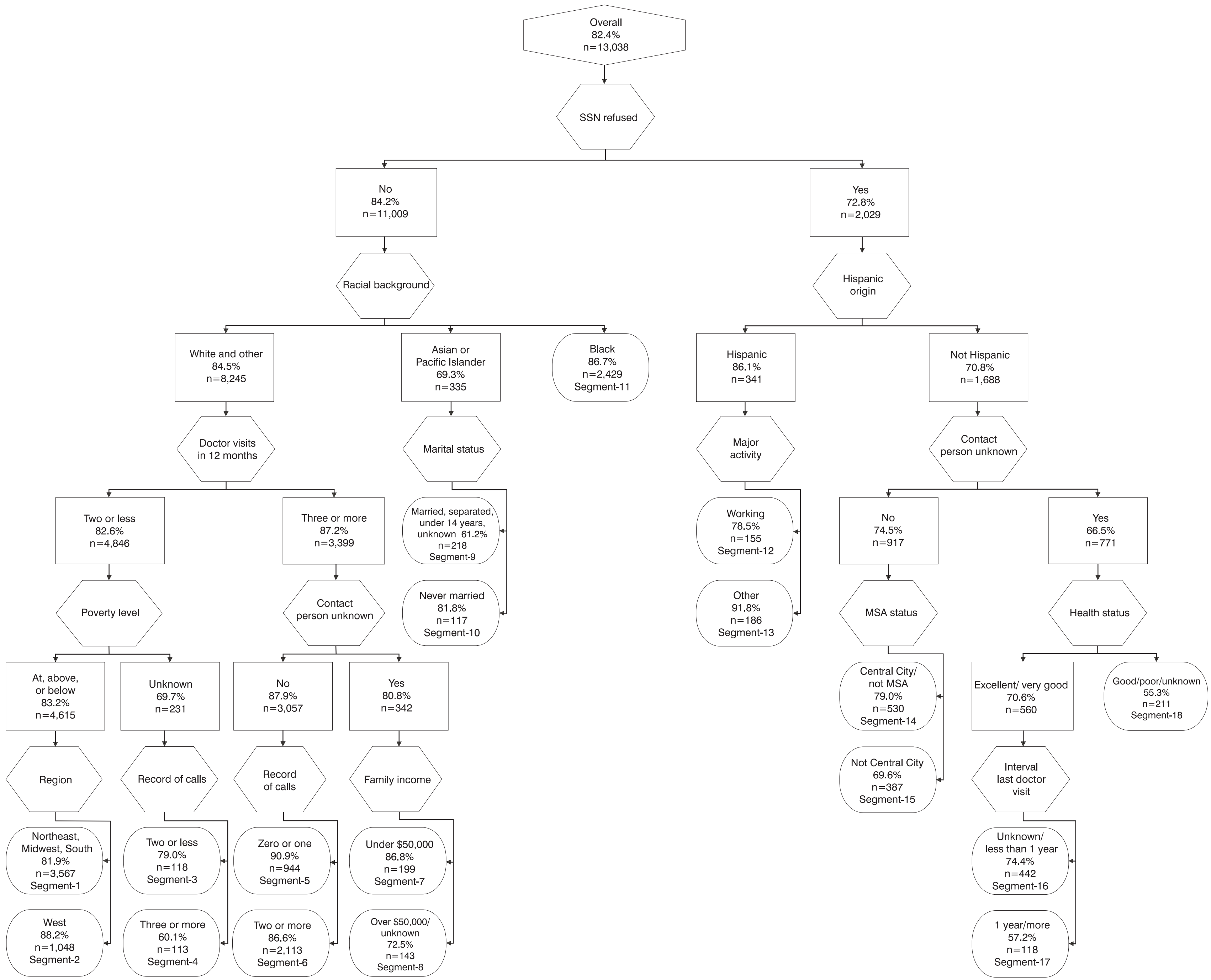


Figure 3. Final segmentation of National Survey of Family Growth Cycle 5 eligible sample members by percent responded

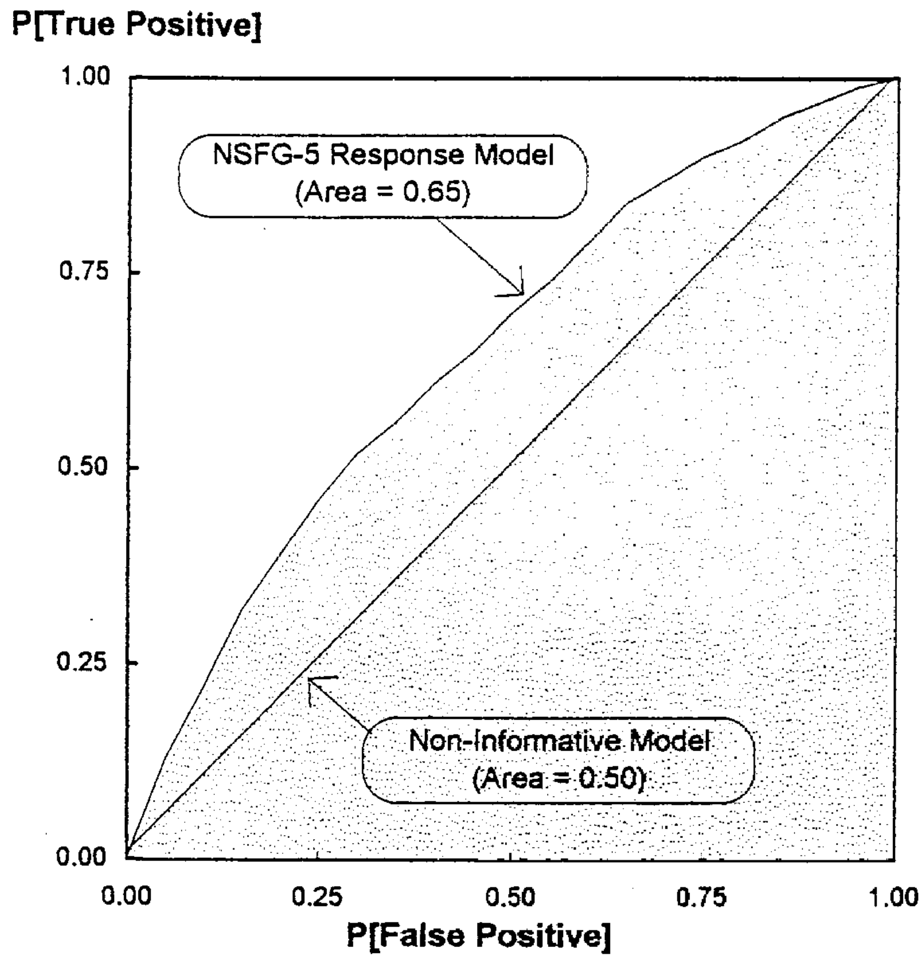


Figure 4. Receiver operating characteristics curve of the combined location and response models

Iannacchione, Milne, and Folsom (13). ROC curves are defined in appendix I.

The area under an ROC curve measures the probability that a randomly chosen pair of observations, one respondent and one nonrespondent, will be correctly ranked (14). This probability of a correct pair-wise ranking is the same quantity that is estimated by the nonparametric Wilcoxon statistic, which tests whether the levels of a quantitative variable in one population tend to be greater than in a second population. No assumptions about how the variable is distributed in the populations are required for the test.

For the combined model, the null hypothesis associated with the Wilcoxon test is that the overall predicted response propensity (that is, $\hat{\lambda}_i \bullet \hat{\rho}_i$) is not a useful discriminator between the responding and nonresponding populations. If the null hypothesis is true, the ROC curve will be a diagonal line with an area of 0.5 that reflects the equally likely chance of making a correct or incorrect decision. If the null hypothesis is not true, the ROC curve will rise above the diagonal and the area under the curve will be significantly greater than 0.5.

As shown in figure 4, the area under the ROC curve developed for the overall predicted response propensity was 0.65 and corresponded to a highly significant Wilcoxon test statistic. The curve indicates that in two of every three randomly chosen pairs of sample women, one responding and the other nonresponding, the predicted overall response propensity of the respondent will be greater than that of the nonrespondent. This level of discrimination implies that the NHIS variables used in the two models are informative but not definitive predictors of a sample woman's overall response propensity.

Poststratification Adjustments

The first four steps in constructing the sampling weight for the NSFG were discussed previously:

- The first step was to use the NHIS weight.
- The second step was to adjust the NHIS weight for the different sampling rates in the NSFG for Hispanic, black, and other women, and for women in households with more than one eligible woman.
- The third step was to adjust the weights for inability to locate the case, using the logistic regression approach described previously.
- The fourth step was to adjust the weights for NSFG nonresponse among located cases, again using a logistic regression approach.

The fifth step, described in the following paragraphs, is to adjust the weighted numbers to independent control totals provided by the U.S. Bureau of the Census. This adjustment to independent control totals is called "poststratification."

Forming poststrata—To "poststratify" an estimate means to make it conform to an independent control total by some mathematical technique. These techniques may be simple or complex, but they are designed to correct for noncoverage—the fact that a survey does not cover a certain portion of the population. Household surveys tend to undercount the population. The poststratification process adjusts the weighted data to match independent estimates of the population. This makes the data consistent with other, more comprehensive sources. Secondly, poststratification can improve precision.

The categories in which the NSFG survey estimate and the independent control total (from the U.S. Bureau of the Census) were made to agree are called "cells." In Cycles 1, 2, and 3, these cells were defined by age, marital status, and race (black and nonblack women). In Cycle 5, these cells were defined by age, race/ethnicity (Hispanic, black, and other), marital status, and parity. As in Cycle 4, estimated totals from the Current Population Survey (CPS) were used to poststratify by marital status, parity, age, and race/ethnicity. The only change in the cell definitions was that three levels of race/ethnicity (Hispanic, non-Hispanic black, and other) were used in Cycle 5 instead of the two (black and other) used in Cycle 4. This made the cells consistent with the stratification used to select the Cycle 5 sample. Ideally, the estimated totals for the poststrata should reflect the April 1, 1995, (the midpoint of data collection) national distribution of women between the ages of 15 and 44. However, because these estimates were not available for all cross-classifications, the June 1994 CPS totals were adjusted to May 1995 marginals for age and race/ethnicity.

Adding a third race/ethnicity level resulted in more potential poststrata than in Cycle 4. As a result, some of the cross-classifications represented small subpopulations that had relatively unstable CPS population estimates. To identify such cells, the percent relative standard error (RSE) of each CPS estimate was calculated. Estimates with an RSE of more than 10 percent were deemed potentially unstable and collapsed with adjacent cells. In general, a 10-percent RSE corresponded to a CPS estimate based on a sample size of about 100 cases.

To maximize the number of multiway cross-classifications, all marginal and submarginal estimates that satisfied the 10-percent RSE criterion were included in the post-stratification model. For example, in the 15–17 year age group, race/ethnicity by parity cross-classifications were too unstable to be used as separate poststrata. Therefore, the submarginal totals for race/ethnicity and parity were controlled separately. The population totals used in the 108 poststrata are presented in table 2.

A poststratification adjustment can be either upward (greater than 1.0) or downward (less than 1.0) for some portions of a sample. For classical poststratification, control totals are required for all possible combinations of poststratification factors. For the NSFG, poststratification was desired on marital status, parity, age, and race/ethnicity. Because control totals were not available for all combinations of these factors, the NSFG weight was poststratified to control totals for the remaining factors. This was implemented repeatedly (a process called "raking") until weighted counts matched all control totals. For Cycle 5, a generalized raking procedure was used.

The actual adjustments were calculated using an exponential model analogous to a generalized raking procedure (9). The exponential model preserves totals of main-effect explanatory variables without necessarily preserving the multiway cross-classification totals of the main effects. Which multiway cross-classification totals are controlled depend on which interaction terms are included in the model.

After the poststratification adjustment factors were calculated, the final Cycle 5 adjusted weight (POST_WT location 12,350–12,359 in the data file) was computed for each respondent as follows:

$$W_{4i} = W_{3i} \bullet A_{4i}$$

where

A_{4i} = the poststratification adjustment factor for SW_i obtained from the generalized raking procedure

The mean poststratification adjustments by age category, race/ethnicity, parity, and marital status are summarized in table J.

The beneficial effects of nonresponse propensity adjustments and poststratification may be offset if the variation in the adjusted weights is excessive. However, the combined weight adjustments for location and response propensity and for poststratification only increased the overall unequal weighting effect from 1.20 to 1.23 (table E). This indicates that the increase in the sampling variances of Cycle 5 estimates was marginal and makes weight trimming unnecessary.

The marginal increase in the sampling variances attributable to the weight adjustments is likely to be more than offset by the reduction in the overall bias of survey estimates produced with the adjusted weights. The estimated number

Table J. Poststratification adjustment summary, by selected characteristics

Characteristic	Current Population Survey total (thousands)	Mean adjustment factor[1]
Age[2]		
15–17 years	5,452	1.21
18–19 years	3,508	1.16
20–24 years	9,051	1.30
25–29 years	9,693	1.19
30–34 years	11,056	1.11
35–39 years	11,211	1.10
40–44 years	10,230	1.09
Race/ethnicity[2]		
Hispanic	6,702	1.24
Black, non-Hispanic	8,210	1.21
Other	45,288	1.12
Parity (number of live births)[3]		
0	25,244	1.24
1	10,704	1.12
2	13,875	1.11
3	6,961	1.12
4 or more	3,416	1.07
Marital status[3]		
Ever married	37,521	1.13
Never married	22,679	1.20
Overall[2]	60,201	1.16

[1]The average poststratification adjustment factor (A_{4i}) applied to the nonresponse-adjusted weights (W_{3i}) of Cycle 5 participants.
[2]May 1995 Current Population Survey totals.
[3]June 1994 Current Population Survey totals adjusted to May 1995 marginals for age and race/ethnicity.

of births using the adjusted weights are compared with vital statistics compiled by NCHS from birth certificates in table K. Except for the small ''other race'' category, the adjusted NSFG estimates are not significantly different from the vital statistics.

Table K. Comparison of National Survey of Family Growth (Cycle 5) estimates of the number of births with vital statistics by year, race, and Hispanic origin

Year, race, and Hispanic origin	NSFG estimates[1] (thousands)	0.95 confidence interval	Vital statistics (thousands)	Ratio of NSFG to vital statistics
Total	15,932	(14,935–16,929)	16,129	0.99
Year				
1991	4,030	(3,665–4,395)	4,111	0.98
1992	4,160	(3,771–4,550)	4,065	1.02
1993	3,909	(3,556–4,261)	4,000	0.98
1994	3,833	(3,489–4,176)	3,953	0.97
Race				
White	12,494	(11,614–13,374)	12,714	0.98
Black	2,363	(2,074–2,652)	2,652	0.89
Other	1,075	(862–1,288)	763	1.41
Hispanic origin				
Hispanic	2,489	(2,040–2,938)	2,585	0.96
Other	13,443	(12,629–14,257)	13,544	0.99

[1]NSFG is National Survey of Family Growth.

Item Imputation

Overview

In any survey, not every question is answered by every person interviewed. Sometimes a respondent cannot remember the fact asked for in a question; sometimes she may refuse to answer. Such missing data create inconsistencies in estimates, which are confusing for some users of the data. Missing data may also introduce bias, because cases with missing data are often not a representative subset of all cases. Assigning values to these missing answers is called ''imputation''; imputation makes the data complete, more consistent, easier to use, and often corrects nonreporting biases. As long as the percent of cases with missing data is low, imputation is a good solution to the problems that missing data can create. Imputation can, however, be labor-intensive and costly. Cycle 5 of the NSFG has thousands of variables, but resources were limited, so it was necessary to select a small percentage of variables to be imputed. Only about 315 recoded variables, or ''recodes,'' were imputed because they were likely to be the most frequently used variables for national estimates. A list of the recodes is shown in appendix III.

Some of the 315 recodes were computed several times; for example, recodes were defined for up to 10 periods of employment, 12 living situations, and 15 pregnancies.

In general, a value was imputed for a woman using data from similar women with complete data for that item. The process of matching a person with reported data (a donor) to a person with missing data (a recipient) used information from the NSFG database and, for some variables, data from the National Health Interview Survey (NHIS) database. Data from a matching donor were assigned to the recipient. Actual reported information was *never* replaced by an imputed value unless the reported information was obviously incorrect. An imputation flag was associated with each imputed variable to enable a researcher to identify which cases had imputed values and what kind of imputation was used.

The NSFG database consists of two files: the respondent file and the pregnancy-interval file. The respondent file contains one record for each of the 10,847 NSFG respondents. The pregnancy-interval file contains one record for up to 15 pregnancy intervals experienced by the respondents. Data for 21,332 pregnancy intervals were collected. Respondents who reported never being pregnant are not represented in this file. Selected items recorded in the pregnancy-interval file are included on the respondent file and labeled with numbers 1–15 to indicate the pregnancy to which they apply.

The frequency of missing values for the data items on which the recodes are based was very low, in part because of the use of computer-assisted personal interviewing (CAPI), which required a response before proceeding to the next question. The CAPI program controlled the flow of the questions (identifying questions to be asked or skipped) and included range and consistency checks for data. As a result, only 11 recoded variables had missing data for more than 3 percent of the observations. Of these, four were pregnancy interval variables measuring the wantedness of the pregnancy. Family income had the most missing data with 1,233 observations (11.4 percent) and accounted for 23.6 percent of all cases needing imputation. The extent of incomplete data for the income variable was expected, and a more sophisticated imputation procedure was used for income.

The following sections provide details on the imputation procedures used on the NSFG Cycle 5 recodes. For reference, the respondent file recodes are listed alphabetically within the questionnaire section in appendix III; the interval file recodes are also shown.

Imputation Procedures for Cycle 5

Four methods of imputation were used for the recodes in NSFG Cycle 5. The methods differed based on the level of sophistication of the imputation procedure and the availability of data for the imputation. An overview of these methods is provided in this section. After each imputation procedure, the imputed values were evaluated. If the initial imputed value was out of range, or inconsistent with other data for that case, the imputation was repeated until the imputed value was acceptable.

Method 1: Logical imputation—The first step in the imputation process was to determine whether the missing answer could be either deduced or guessed from answers to other questions. For example, in the 10th variable in Section H in appendix III, eight cases with missing data for whether the respondent's husband or partner has ever had an infertility test (INFERTH) were imputed using a logical imputation. The imputed values were determined using the variable that measures whether or not the respondent had ever had help getting pregnant (ANYPRGHP). The logical imputations were generally limited to variables with fewer than 10 cases with missing values. If the data were not consistent or if ambiguity existed, then the value was imputed by a ''hot-deck'' procedure.

Method 2: Unweighted hot-deck imputation—Imputation using the hot-deck procedure requires identifying a pool of donors (observations with complete data) with characteristics similar to those of the receptor (the observation with a missing value). A donor is then selected from the pool randomly either with equal probability (unweighted) or with probability proportional to the sampling weight of the donor (weighted). The cases that could donate a value to a case without data are called donor pools, or imputation classes. An imputation class should be sufficiently large so that the number of times a single donor provides a value is minimized, but also sufficiently small so that the donors and receptors are adequately comparable. By creating a group of respondents with similar characteristics for variables believed to be correlated with the missing recode, imputed values are generally more consistent with the life-history information.

For constructing the donor pools, two types of variables were considered: screening variables and classing variables. Screening variables defined the subgroup of the data set that contained values used for imputation. For example, if the recoded variable for ever having cohabited equaled ''yes'' (COHEVER=1), then a date representing the recoded variable for the century month of first cohabitation (COHAB1) was required. In this case, COHEVER is the screening variable for the imputation of COHAB1. The classing variables defined the characteristics with which cases with reported data were matched with cases with missing data; complete agreement between the classing variables of the donor and receptor was not essential. For example, to match donors to receptors by the respondent's age, instead of using the single year of age (AGER), donors and receptors were classified by age in seven categories: 15–17, 18–19, 20–24, 25–29, 30–34, 35–39, and 40–44. Thus, not every matched pair of cases had exactly the same value for age in single years.

When a small number of cases had missing data, identifying the screening and classing variables was relatively straightforward. In general, the primary classing variables were the geographic characteristics of residence (that is, geographic region, metropolitan status, and urban/rural status) and demographic variables (race/ethnicity, educational level, and age). When a recoded variable had a relatively large number of cases with missing data, identifying the classing variables and imputation classes was more important, and more complex ways of choosing them were used.

Because many of the variables had only a few missing values, a structured hot-deck approach was used. In this procedure, one or more screening and classing variables was used to define a group of relatively similar cases—some with and some without data. A random number generator was used to randomly select one of the donors to provide a value for the observation without data.

Returning to the example with the date of first cohabitation (COHAB1), the first table in appendix III shows that 61 cases required data for COHAB1 after determining that these respondents had cohabited at least once (COHEVER=1). Values for COHAB1 were imputed using an unweighted hot-deck procedure from respondents who reported at least one period of cohabitation. (COHEVER=1). Thus, the file was screened for only those cases with COHEVER=1 prior to the imputation procedure. The following classing variables were used to create the imputation classes for COHAB1:

COHSTAT	Cohabitation status relative to the first marriage
FMARNO	Number of marriages
RMARITAL	Informal marital status
RACE	Race
EDUCAT2	Four-level classification of education (0, 1–11, 12, and 13 and over)
AGECAT	Seven-level classification of age (15–17, 18–19, 20–24, 25–29, 30–34, 35–39, and 40–44)

Imputed values were examined against other relevant data to determine their consistency.

Method 3: Weighted hot-deck imputation—Weighted hot-deck imputation (15) replicates the *weighted* distribution of the reported data in the imputed data by using the sampling weights of the item respondents and nonrespondents. The weighted hot-deck procedure takes into account the unequal probabilities of selection in the original sample by using the sampling weight to specify the expected number of times a particular respondent's answer will be used to replace a missing item. These expected selection frequencies are specified so that, over repeated applications of the algorithm, the expected value of the weighted distribution of the imputed values will equal the weighted distribution of the reported answers. Weighted hot-deck procedures were used for variables in which 2.7 to 6.7 percent of the cases required imputation. These variables were:

- Mother's or mother-figure's education (EDUCMOM) (4.3 percent)
- Father's or father-figure's education (EDUCDAD) (6.7 percent)
- Had sexual intercourse in the 3 months before the NSFG Cycle 5 interview, or not (SEXP3MO) (2.8 percent)
- Number of months of no sexual intercourse in the 12 months before the NSFG Cycle 5 interview (NOSEX12) (2.8 percent)
- Number of months of no sexual intercourse in the 36 months before

the NSFG Cycle 5 interview (NOSEX36) (2.8 percent)
- Measure (comparable to Cycle 4) of respondent's wantedness status of pregnancies 1–15 (OLDWR01–15) (5.2 percent)
- New measure of wantedness status of pregnancies 1–15 (WANTRP01–15) (5.2 percent)
- Measure (comparable to Cycle 4) of partner's wantedness status of pregnancies 1–15 (OLDWP01–15) (4.9 percent)
- New measure of partner's wantedness status of pregnancies 1–15 (WANTPT01–15) (5.1 percent)

Details of the imputation methodology are described in more detail in the following text.

Method 4: Regression imputation for family income—Among the recodes that were subject to imputation, family income was missing the most frequently (11.4 percent). The relatively large proportion of women with missing data, combined with the importance of the family income variable for policy analysis, warranted a special model-based approach for assigning the imputed values. Unlike most other recoded variables, family income has a direct counterpart from the 1993 NHIS data. A regression model was used to modify the 1993 family income (when available) based on changes in the respondent's marital status, family size, employment status, and other associated factors. Details of the imputation methodology are described in the section on "Imputation of family income."

Weighted hot-deck imputation—Except for family income, the recodes with the largest number of missing values requiring imputation were parents' education, intercourse in the months before interview, and pregnancy wantedness. The weighted hot-deck procedure was chosen for these recodes to ensure that, within each imputation class, the weighted mean of the imputed values would be the same as the weighted mean of the data directly obtained from the respondents. The following sections describe the steps used to impute the values.

Parental education variables—The weighted hot-deck imputation procedure was used to impute the education of the respondent's father (EDUCDAD) and mother (EDUCMOM). The donor pool consisted of 9,279 respondents who had reported a value for the educational attainment of both parents. Recipient cases fell into three categories:

1. Mother's education reported, father's education missing (450 missing)
2. Mother's education missing, father's education reported (141 missing)
3. Mother's education missing, father's education missing (271 missing)

Imputation classes were based on the respondent's age and race/ethnicity. Within each class, when the education of one parent was known, donors and recipients were sorted by the known education so that the donors and recipients tended to share similar values. When both parents' education was missing, donors and recipients were sorted by the education of the respondent (the daughter) with a single donor chosen for both parental education values. The rationale behind this strategy was the presumption that the education levels of parents and, to a lesser extent, their children, would be correlated. Table L shows the joint frequency distribution of the education levels of the mothers and fathers of Cycle 5 respondents before imputation. The concentration of values on or next to the diagonal provides some evidence of the tendency for couples to have similar levels of education.

Table L. Joint frequency distribution of the education of mother and father of Cycle 5 respondents

		Father's education				
Mother's education	Total	Less than high school	High school	Some college	College degree	Missing
Total	10,141	3,105	3,381	1,032	1,902	721
Less than high school	2,923	1,839 (17.0%)	631 (5.8%)	116 (1.1%)	131 (1.2%)	206 (1.9%)
High school	4,097	914 (8.4%)	2,037 (18.8%)	452 (4.2%)	505 (4.7%)	189 (1.7%)
Some college	1,383	178 (1.6%)	443 (4.1%)	294 (2.7%)	436 (4.0%)	32 (0.3%)
College degree	1,326	96 (0.9%)	221 (2.0%)	167 (1.5%)	819 (7.6%)	23 (0.2%)
Missing	412	78 (0.7%)	49 (0.5%)	3 (0.0%)	11 (0.1%)	271 (2.5%)

0.0 Quantity more than zero but less than 0.01.

NOTE: This table excludes 706 sample women who reported that they did not know the education of either parent. These cases were not imputed and were not used as donor cases.

Sexual intercourse variables—The same weighted hot-deck procedure was also used for the variable measuring sexual intercourse in the 3 months, 12 months, and 36 months prior to the Cycle 5 interview (SEXP3MO), (NOSEX12), and (NOSEX36), respectively.

Imputation classes were defined by cross-classifications of the respondent's

- formal marital status (FMARITAL)
- race/ethnicity (RACE)
- age categorized into seven levels (AGEAPR1)

For respondents missing more than one variable, the same donor was used to maintain the internal consistency among the imputed values. For example, if the donor variable indicated intercourse in the 3 months prior to the interview month (SEXP3MO=1), then the other two variables (NOSEX12, NOSEX36) were imputed by the same donor to ensure consistent values.

Imputing the wantedness of pregnancy—Four recoded variables related to the wantedness of the pregnancy were computed using data from the pregnancy-interval file. Of the 21,332 pregnancies reported by respondents, those with missing values for these recoded variables accounted for 5.5 percent of the pregnancy-interval data, leaving 94.5 percent for the donor pool. The number of pregnancy-intervals requiring an imputed wantedness value are shown in table M. Of the 21,332

Table M. Number and percent of pregnancy-intervals requiring an imputed wantedness

Variable name	Description	Number imputed	Percent imputed
OLDWANTP	Partner's wantedness—Cycle 4 measure	1,050	4.9
OLDWANTR	Respondent's wantedness—Cycle 4 measure	1,114	5.2
WANTPART	Partner's wantedness—new Cycle 5 measure	1,092	5.1
WANTRESP	Respondent's wantedness—new Cycle 5 measure	1,111	5.2

intervals, a total of 1,174 pregnancies (5.5 percent) required one or more imputations, while 1,022 intervals required imputed values for all four wantedness variables. The ''Cycle 4 measure'' was based on the series of questions used in cycles 1–4. The ''new Cycle 5 measure'' improves on the old measure as explained in previous reports (16,17).

Given the importance of data on wantedness for understanding fertility differences, and for public policy, a careful imputation process for the wantedness variables was devised. The procedure included:

- a segmentation (cross-tabulation) analysis on the pregnancy interval data to identify the classing variables using CHAID (7,8)
- a weighted sequential hot-deck procedure
- review of the imputed values
- reimputation of any imputed values that were inconsistent

For the segmentation analysis, OLDWANTR was used as the dependent variable and 21 variables were used as predictor variables. This analysis resulted in 25 segments (shown as ovals in figure 5) made up of combinations of 8 of the 21 predictor variables. Both pregnancy interval and respondent variables were included; however, seven of the eight variables that defined the segments were pregnancy interval variables. Figure 5 displays the segmentation results for OLDWANTR. The numbers in each box are the weighted segment means of the levels of wantedness shown in the legend of figure 5. As a result, low segment means indicate wantedness while high segment means indicate unwantedness.

The weighted sequential hot-deck procedure was used with the 25 segments defining the imputation classes. Within each class the data were grouped by OLDWANTR, WANTRESP, OLDWANTP, and WANTPART. The final step was to search for, and rectify, any inconsistent imputed data based on other variable values defined for that respondent's interval. For example, if HPWNOLD=1 (partner wanted respondent to have a baby at some time), then a value of ''unwanted'' was not assigned to either OLDWANTP or WANTPART.

Imputation of Family Income

The recoded variables of respondent's family income (TOTINCR) required data on the total family income, and the poverty level variable (POVERTY) required total income and the number of family members. The total income recoded variable was missing for 1,233, or 11.4 percent, of the respondents. The relatively large amount of missing data combined with the importance of this variable for analysis warranted a special model-based approach for the assignment of the imputed values. Unlike most other recoded variables, the value of family income in 1993 was known from the 1993 NHIS data. Therefore, a regression model was used to modify the 1993 family income (when available) based on changes in the respondent's marital status, family size, employment status, and other associated factors.

The following steps were taken in the imputation process for family income:

1. Predictors of income were selected from the NSFG Cycle 5 and the NHIS.
2. Weighted segmentation analysis was done using CHAID (7,8) with donor status (donor versus receptor) as the dependent variable, to select variables for logistic regression analysis.
3. A weighted logistic regression analysis was done to model the item-response propensity.
4. The data were reweighted using predicted values from the logistic regression analysis in item 3 so that donors who shared the same characteristics of nonrespondents were given prominence in the assignment of imputed values.
5. A weighted segmentation analysis was performed with TOTINCR as the dependent variable to select main effect and interaction candidates for the linear regression analysis.
6. A weighted linear regression analysis was done to model TOTINCR among respondents.
7. The predicted values obtained in item 6 were used to impute the missing family income values (TOTINCR). Using TOTINCR, poverty level was computed.

These steps are explained further in the following text.

Model Development for Imputing Income

Two separate models were developed for the imputation process: a logistic regression model for whether income was reported or not and a linear regression model for family income. This approach is based on experience that has shown that it is easier to predict whether an item is reported or not (''item response propensity'') than to predict the value of the question's missing values. In addition, fewer variables and simpler models are required to predict the item response propensity than to model the answers themselves.

The first step in imputing the poverty-level income was to select a list of potential predictor variables. Eleven NHIS variables and nine Cycle 5 variables were believed to be good predictors of either family income or the likelihood to report family income. The most important NHIS variable was the 1993 family income. The second line of table N shows that, of the 1,233 cases in the NSFG with missing family income, 78 percent had reported incomes in the 1993 NHIS. Twenty variables were used in a segmentation analysis to choose predictor variables for a subsequent logistic regression analysis. The zero/one variable indicating whether income was reported

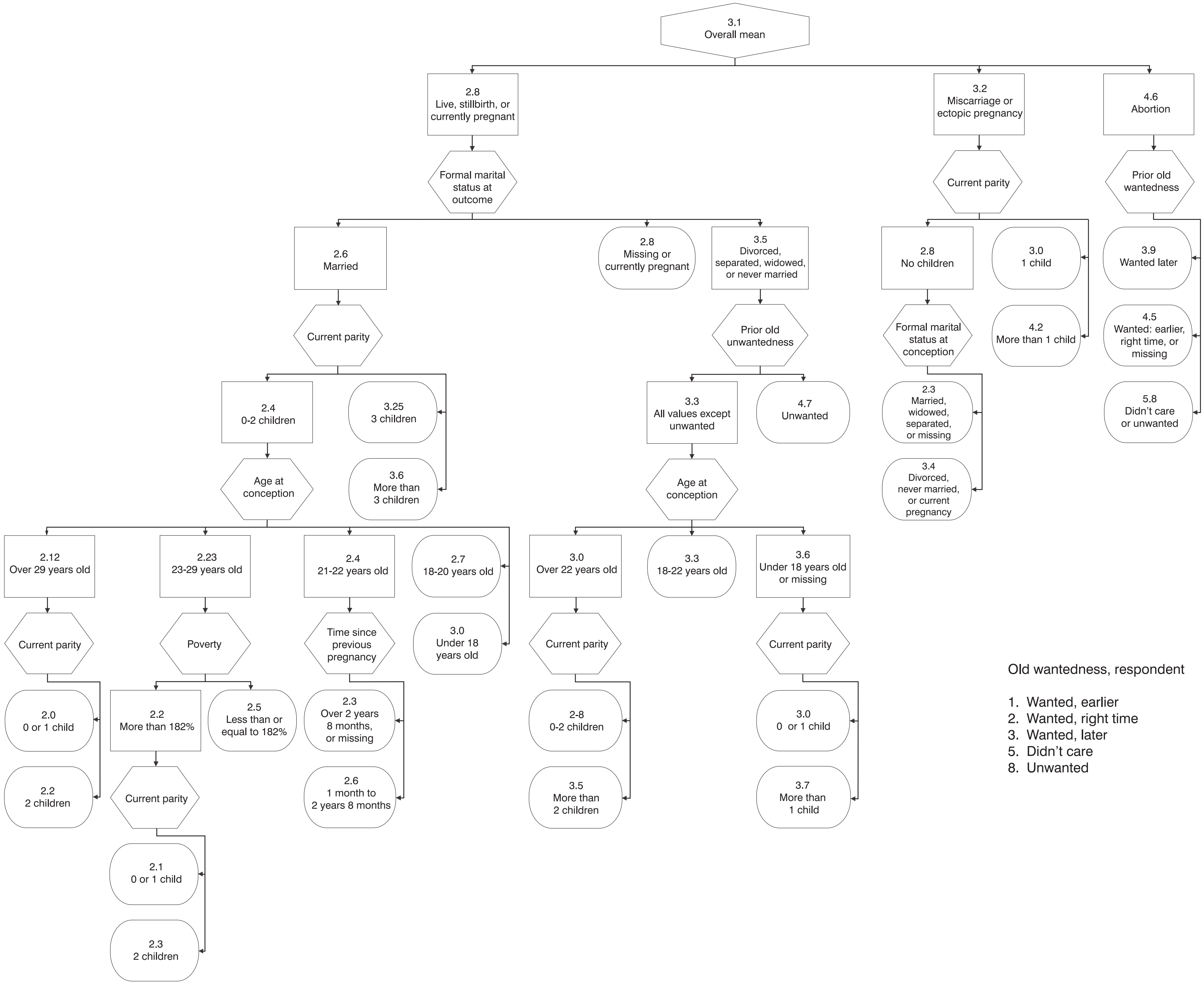


Figure 5. Segmentation of the respondent's wantedness of the pregnancy

Table N. Availability of family income from the 1993 National Health Interview Survey and National Survey of Family Growth, Cycle 5

National Survey of Family Growth, Cycle 5	1993 National Health Interview Survey		
	Available	Missing	Total
Available	8,464 (88%)	1,151 (12%)	9,614 (100%)
Missing	963 (78%)	270 (22%)	1,233 (100%)
Total	9,427 (87%)	1,420 (13%)	10,847 (100%)

was set to one for the 9,614 respondents with a reported value of TOTINCR and to zero for the remaining 1,233 women. The segmentation analysis identified 16 significant main effects and 21 interactions.

Whether income was reported or not (item response propensity) was modeled using a weighted logistic regression with the same zero/one poverty variable described previously. Using the 37 variables identified in CHAID and a backwards elimination selection process, 6 main effects and 12 interactions were identified as significant. The predicted values from this model were used to reweight the data for fitting the subsequent linear regression model. Reweighting increases the weights of respondents with small response propensities (that is, respondents expected to share many of the characteristics of nonrespondents) relative to the weights of respondents with large response propensities.

After reweighting the data, a second segmentation analysis was performed to select variables for main effects and interactions for the linear regression model for predicting family income among respondents. The same 20 variables used in the first segmentation analysis were also used in the family income model. This resulted in 19 main effects and 67 interaction candidates. A backwards elimination linear regression analysis (see appendix I) was used to parse the model and resulted in 15 main effects and 50 interactions. The regression coefficients were used to impute 1995 family income. The resulting imputations for TOTINCR and number of family members (NUMFMHH) were used to impute 1,251 values of poverty level income (POVERTY).

Variance Estimation

Background

The sampling variance is a measure of the variation of an estimator (for example, percent, mean, or regression coefficient) because a sample was used instead of the full population. The sampling variance represents the average squared differences of the observations from their expected value over all possible samples of the same size and using the same sampling design. The classical ‘‘population’’ variance is a measure of the variation among the *individuals* in the population, whereas a sampling variance is a measure of the variation of the *estimate* of a population parameter (for example, a population mean or proportion) over repeated samples. The population variance is different from the sampling variance in the sense that the population variance is a constant, independent of any sampling issues, while the sampling variance becomes smaller as the sample size increases. The sampling variance is zero when the full population is observed, as in a census.

Based on the sampling variance, a series of measures of reliability can be computed for a statistic such as a proportion or mean. The standard error (SE) is the square root of the sampling variance. Over repeated samples of the same size and using the same sampling design, we expect that the true value of the statistic would differ from the sample estimate by less than twice the SE in approximately 95 percent of the samples. The degree of approximation depends on how the data are distributed. The relative standard error (RSE) is the SE divided by the sample estimate and is usually presented as a percentage.

For the National Survey of Family Growth (NSFG), the sampling variance estimate is a function of the sampling design and the population parameter being estimated, and it is called the design-based sampling variance. The design-based variance assumes the use of the fully adjusted sampling weight. The ‘‘fully adjusted sampling weight’’ is derived from the sampling design with adjustments to compensate for nonresponse and for adjusting the ratio of the sampling totals to external totals, such as those by age, race/ethnicity, marital status, and parity from the Bureau of the Census.

For Cycle 5 of the NSFG, the data files include a single fully adjusted sampling weight and information necessary to estimate the sampling variance for a statistic. The weight is called POST_WT in the data file and is in locations 12,350–12,359 in the respondent file. The other variables needed are the collapsed strata (12,347–12,348) and the panel identifier (12,349). Because the NSFG sampling design is complex (that is, a stratified, multistage design with individual sampling rates), both the sampling weight and the sampling design must be taken into account to compute unbiased estimates of population parameters and sampling variances.

Estimating the sampling variance requires using survey data analysis software or specially developed programs designed to accommodate the population parameter being estimated and the sampling design. Several methods are available to compute sampling variances for complex samples. They include ‘‘Taylor Series Approximation’’ techniques, and several pseudo-replication approaches, like balanced repeated replication and the ‘‘Jackknife’’ technique (18). The 1995 NSFG data file includes survey design variables that facilitate using the Taylor Series Approximation approach.

SUDAAN is one of the software packages that use the Taylor Series Approximation approach.

SUDAAN can produce accurate variance estimates for a wide variety of statistics: means, percents, and estimated numbers, as well as coefficients and odds ratios for regression, logistic regression, and proportional hazards models. For these reasons and others, SUDAAN has been used to estimate variances for several NCHS surveys, including the NHIS and the 1995 NSFG. Therefore, this report includes examples of SUDAAN programs.

However, there are several other software packages available that will produce estimates of variances that take complex sample designs into account. Data users are encouraged to use any appropriate software that meets their needs and takes the complex sample design into account. Since NCHS does not endorse any commercial product, the authors of this report wish to alert readers to the fact that other software for variance estimation from complex samples is available.

The information given below was obtained in October 1997 from the following site on the World Wide Web: http://www.fas.harvard.edu/~stats/survey-soft/survey-soft.html. Three of these packages are also discussed in reference 19.

The other available software for variance estimation includes:

1. CENVAR and VPLX, both from the U.S. Bureau of the Census. Both packages are described further on the Bureau's site on the World Wide Web and both may be downloaded free. CENVAR uses the Taylor Series approach to variance estimation, while VPLX uses replication techniques. Both can estimate variances for statistics such as estimated numbers, means, and proportions. For CENVAR: http://www.census.gov/ftp/pub/ipc/www/imps.html. For VPLX: http://www.census.gov/sdms/www/vwelcome.html.
2. The CSAMPLE procedure in the CDC's ''Epi-Info'' software. Epi Info uses the Taylor Series approach and can compute means, proportions, odds ratios, and other simple statistics. Further information is available from CDC's web site at: www.cdc.gov/epo/epi/epi.html or by e-mail at: epiinfo@cdc1.cdc.gov.
3. PC CARP is a package available from the Iowa State University Statistical Laboratory, 219 Snedecor Hall, Ames, IA 50011. PC CARP uses a Taylor Series approach to estimate variances. According to its authors, it will compute variances for estimated numbers, percents, means, and weighted regression.
4. STATA is a commercial general-purpose statistical package that includes variance estimation software. According to its makers, it uses a Taylor Series approach and will compute variances for means, estimated numbers, proportions, and linear, logistic, and probit regression. (Stata Corporation, 702 University Drive East, College Station, TX 77840.)
5. WesVarPC was produced by Westat, Inc., 1650 Research Blvd., Rockville, MD, 20850, and is available free from Westat's web site at http://www.westat.com/wesvarpc/wesvarpc.html. WesVar uses a replication technique and, according to its distributors, estimates variances from estimated numbers, means, percentages, linear regression, and logistic regression.

While the characteristics and capabilities of these packages will change as time passes, the important point is that software is now available to compute valid estimates of sampling errors for statistics from complex samples. Data users are urged to use any appropriate software that is available to them to obtain accurate estimates of sampling errors for the NSFG.

A shortcut method for estimating sampling variances using a generalized variance algorithm has also been developed. The shortcut estimation methods, called generalized standard error (GSE) estimates, are formulated based on a regression model, showing the relationship between a parameter estimate and its sampling variance. The coefficients of the model are estimated using the actual survey estimates of population parameters and direct estimates of their sampling variances. The GSE estimates are easy and quick to compute using the survey estimates and estimated model coefficients. The obvious disadvantage of generalized variance estimates is that they may not accurately reflect the true sampling variance for a given statistic.

Summary of Variance Estimation

Because of the linkage between the NSFG and the NHIS, estimating variance for survey estimates from the NSFG derives partly from the NHIS design. The NHIS uses a highly complex sample design to increase statistical and operational efficiency. Similarly, computing the sampling variance estimate for NHIS estimates is complex. NCHS has developed two estimation algorithms that approximate the sampling variances. The variance estimation procedure for NHIS is described in detail by Massey et al. (5). The procedures for computing the NSFG sampling variance estimates parallel the NHIS procedures, and the sampling variance estimates for NSFG can be computed using adaptations of the NHIS variance estimation procedures.

Survey estimators fall into two general classes: linear and nonlinear estimators. Linear estimators are weighted totals of the persons with an attribute, or means and proportions if the denominators are known (for example, when the denominator is a poststratum total or a sum of poststrata totals). Nonlinear estimators include proportions and means when the denominators are unknown and are estimated from the survey, as well as ratios, and correlation and regression coefficients. The variances of nonlinear statistics cannot, in general, be expressed exactly. Woodruff (20) suggested a procedure in which a nonlinear estimator is linearized by a Taylor Series Approximation. The sampling variance equation is then used on this linear form (called a linearized variate) to produce a variance approximation for the original nonlinear estimator.

NCHS has decided to use Taylor Series linearization in Cycle 5 of the NSFG (alternatives include balanced repeated replication or the Jackknife procedure (5,21). The Taylor Series Approximation can take into account the effect of the nonresponse adjustment, the ratio adjustment, and the poststratification adjustment and only requires the computation of one fully adjusted sampling weight.

Software, such as SUDAAN (11), that uses the Taylor Series linearization procedure or a valid replication procedure, can handle the multistage design and the components of variance in the NHIS design and the NSFG design. Variance estimation using the SUDAAN software is described further in appendix IV. Two sample SUDAAN programs are shown in appendix V.

Generalized Variance Estimation

Generalized standard error (GSE) estimates are shortcut methods that approximate the sampling variance of a population estimate when a specific or direct estimate of the variance is not available. GSE estimates are based on a modeled relationship between a set of key parameter estimates (for example, percents) and their associated direct variances (calculated by, for example, SUDAAN). The direct estimates of variances are considered the most accurate because they are calculated directly from the data. The GSE estimates are made from a regression equation in which the sampling variances are predicted from the survey estimates (for example, percents—see appendix VI). Like any regression equation, the GSE estimates have error around the regression line. The GSE estimates, however, should be better (more accurate) than Simple Random Sample (SRS) estimates. For this reason, the direct variance estimates obtained from SUDAAN or a similar survey analysis package are usually better than the GSE estimates described in the following text. See appendix VI for further details on how the GSE estimates were calculated.

The model developed for the GSE algorithm is for estimating the sampling variance of a proportion and assumes that the subpopulations of interest (for example, the denominator of the proportion) will be approximate combinations of the 108 poststratification cells used in computing the fully adjusted analysis weights (see table 2 of this report for a listing of poststrata). The poststratification cells are cross-classifications of race/ethnicity, age, marital status, and parity. The sampling variance for any survey estimate that is a combination of the poststratification cells (the denominator) is zero because these counts are assumed to be known without error. Since the denominator for a proportion is a combination of the poststratification cells (for example, Hispanic women 25–29 years of age), then only the numerator contributes to the sampling variance. The poststratification totals are estimates from the Current Population Survey. While some of these totals have a non-zero sampling variance, for present purposes it is assumed that the poststratification cell totals are known without error.

The most commonly used model for GSE for subpopulation proportions relates the relative sampling variance for an estimate to the inverse of the survey estimate (18). The model is of the form

$$V^2(P) = S^2 / P^2 = \alpha + \beta / X \qquad (1)$$

where

$V^2(P)$ = the relative sampling variance of the estimated proportion P
S^2 = the sampling variance of the estimated proportion P
α, β = the model coefficients to be estimated
X = the survey estimate of the numerator of the proportion P

Alternatively, the relative variance may be expressed in terms of a design effect (18), which is the ratio of the variance of a survey estimate to the variance that would have been obtained from a simple random sample (SRS) of the same sample size. The design effect (deff) for an estimated proportion P is defined as:

$$\text{deff} = S^2 / [P\,(1–P) / n] \qquad (2)$$

where

S^2 = the sampling variance of the estimated proportion P
n = the sample size

The denominator of equation (2) is the estimated sampling variance of a proportion when simple random sampling is used. The use of design effects allows the model in (1) to be recast as

$$\begin{aligned} V^2(P) &= (1–P)\ \text{deff} / (Pn) \\ &= –\text{deff} / n + (N\ \text{deff} / n) / X \\ &= \alpha + \beta / X \qquad (3) \end{aligned}$$

where

n = the sample size
X = the survey estimate of the numerator of the proportion P
N = the population size or denominator of the proportion P

Design effects provide a summary measure of the combined effects of stratification, clustering, and unequal weighting on the variance of a survey estimate. The design effects are particularly useful for estimating GSE's because they identify the subpopulations that are most affected by the sample design. For example, the design effects for Hispanic and non-Hispanic black women are generally larger than those obtained for other women because of the oversampling of minorities in the NSFG sample design. Therefore, separate variance estimates were made for Hispanic and black women. The procedure for deriving the GSE estimates is described in appendix VI.

For example, to obtain GSE's for NSFG percentage estimates, first determine whether the estimate is a respondent characteristic (for example, percent of women using the oral contraceptive pill in 1995) or a pregnancy interval characteristic (for example, percent of pregnancies that were wanted). Table O presents GSE's for respondent characteristics, and table P presents GSE's for pregnancy interval characteristics. Each table provides GSE's for Hispanic women, non-Hispanic black women, and all women and white women. Determine the appropriate race category and then obtain the GSE that corresponds to the sample size (row entry) and the percentage estimate (column entry). For example, the generalized standard error for a percentage estimate of 15 percent that is based on a sample size of 1,000 Hispanic women is 1.25 percentage points, as shown in table O.

Table O. Generalized standard errors for estimated percentages and corresponding sample sizes from the respondent file: National Survey of Family Growth, Cycle 5

Sample size	Weighted size (000's)	Percentage									
		50	45 or 55	40 or 60	35 or 65	30 or 70	25 or 75	20 or 80	15 or 85	10 or 90	5 or 95
		Hispanic women									
100	431	5.42	5.51	5.50	5.39	5.17	4.83	4.35	3.73	2.91	1.82
200	863	3.90	3.96	3.95	3.87	3.72	3.47	3.13	2.68	2.09	1.31
300	1,295	3.21	3.27	3.26	3.19	3.06	2.86	2.58	2.21	1.72	1.08
400	1,726	2.80	2.85	2.84	2.79	2.67	2.50	2.25	1.93	1.50	0.94
500	2,158	2.52	2.56	2.56	2.50	2.40	2.24	2.02	1.73	1.35	0.85
600	2,589	2.31	2.35	2.34	2.30	2.20	2.06	1.86	1.59	1.24	0.78
700	3,021	2.15	2.18	2.18	2.13	2.05	1.91	1.72	1.48	1.15	0.72
800	3,452	2.01	2.05	2.04	2	1.92	1.79	1.62	1.38	1.08	0.68
900	3,884	1.90	1.94	1.93	1.89	1.82	1.70	1.53	1.31	1.02	0.64
1,000	4,315	1.81	1.84	1.84	1.80	1.73	1.61	1.46	1.25	0.97	0.61
1,100	4,747	1.73	1.76	1.76	1.72	1.65	1.54	1.39	1.19	0.93	0.58
1,200	5,179	1.66	1.69	1.69	1.65	1.58	1.48	1.33	1.14	0.89	0.56
1,300	5,610	1.60	1.62	1.62	1.59	1.52	1.42	1.28	1.10	0.86	0.54
1,400	6,042	1.54	1.57	1.57	1.53	1.47	1.37	1.24	1.06	0.83	0.52
1,500	6,473	1.49	1.52	1.52	1.48	1.42	1.33	1.20	1.03	0.80	0.50
1,553	6,702	1.47	1.49	1.49	1.46	1.40	1.31	1.18	1.01	0.79	0.49

For the above table, the coefficients in equation (2) of appendix VI are: b0 = 0.4279; b1 = 117; b2 = 0.5047; b3 = 0.0479.

Sample size	Weighted size (000's)	50	45 or 55	40 or 60	35 or 65	30 or 70	25 or 75	20 or 80	15 or 85	10 or 90	5 or 95
		Non-Hispanic black women									
100	336	5.75	5.76	5.71	5.60	5.40	5.13	4.75	4.25	3.57	2.59
200	671	4.13	4.15	4.11	4.03	3.89	3.69	3.42	3.06	2.57	1.86
300	1,007	3.41	3.42	3.39	3.32	3.21	3.04	2.82	2.52	2.12	1.53
400	1,343	2.97	2.98	2.96	2.90	2.80	2.65	2.46	2.20	1.85	1.34
500	1,678	2.68	2.68	2.66	2.60	2.52	2.39	2.21	1.98	1.66	1.20
600	2,014	2.45	2.46	2.44	2.39	2.31	2.19	2.03	1.81	1.52	1.10
700	2,350	2.28	2.29	2.27	2.22	2.14	2.03	1.89	1.69	1.42	1.03
800	2,685	2.14	2.15	2.13	2.08	2.01	1.91	1.77	1.58	1.33	0.96
900	3,021	2.02	2.03	2.01	1.97	1.90	1.81	1.67	1.50	1.26	0.91
1,000	3,356	1.92	1.93	1.91	1.87	1.81	1.72	1.59	1.42	1.20	0.87
1,100	3,692	1.84	1.84	1.83	1.79	1.73	1.64	1.52	1.36	1.14	0.83
1,200	4,028	1.76	1.77	1.75	1.72	1.66	1.57	1.46	1.31	1.10	0.79
1,300	4,363	1.70	1.70	1.69	1.65	1.60	1.52	1.40	1.26	1.06	0.76
1,400	4,699	1.64	1.64	1.63	1.60	1.54	1.46	1.36	1.21	1.02	0.74
1,600	5,370	1.54	1.54	1.53	1.50	1.45	1.37	1.27	1.14	0.96	0.69
1,800	6,042	1.46	1.46	1.45	1.42	1.37	1.30	1.20	1.08	0.90	0.65
2,000	6,713	1.38	1.39	1.38	1.35	1.30	1.24	1.14	1.02	0.86	0.62
2,200	7,384	1.32	1.33	1.32	1.29	1.24	1.18	1.09	0.98	0.82	0.59
2,400	8,056	1.27	1.27	1.26	1.24	1.19	1.13	1.05	0.94	0.79	0.57
2,446	8,210	1.26	1.26	1.25	1.22	1.18	1.12	1.04	0.93	0.78	0.57

For the above table, the coefficients in equation (2) of appendix VI are: b0 = 0.0876; b1 = 0.1915; b2 = 0.0262; b3 = 0.0495.

Sample size	Weighted size (000's)	50	45 or 55	40 or 60	35 or 65	30 or 70	25 or 75	20 or 80	15 or 85	10 or 90	5 or 95
		All women and white women									
100	555	4.95	4.93	4.86	4.73	4.54	4.28	3.94	3.51	2.92	2.092
200	1,110	3.61	3.60	3.54	3.45	3.31	3.13	2.88	2.56	2.13	1.53
300	1,665	30	2.99	2.95	2.87	2.76	2.60	2.39	2.13	1.77	1.27
400	2,220	2.64	2.63	2.59	2.52	2.42	2.28	2.10	1.87	1.56	1.11
500	2,775	2.38	2.37	2.34	2.28	2.18	2.06	1.90	1.69	1.41	1.01
600	3,330	2.19	2.18	2.15	2.09	2.01	1.90	1.75	1.55	1.29	0.93
700	3,885	2.04	2.04	2.01	1.95	1.87	1.77	1.63	1.45	1.21	0.86
800	4,440	1.92	1.92	1.89	1.84	1.76	1.66	1.53	1.36	1.14	0.81
900	4,995	1.82	1.82	1.79	1.74	1.67	1.58	1.45	1.29	1.08	0.77
1,000	5,550	1.74	1.73	1.71	1.66	1.59	1.50	1.38	1.23	1.03	0.73
1,200	6,660	1.60	1.59	1.57	1.53	1.47	1.38	1.27	1.13	0.94	0.68
1,600	8,880	1.40	1.40	1.38	1.34	1.29	1.21	1.12	0.99	0.83	0.59
2,000	11,100	1.27	1.26	1.24	1.21	1.16	1.10	1.01	0.90	0.75	0.54
3,000	16,650	1.05	1.05	1.04	1.01	0.97	0.91	0.84	0.75	0.62	0.45
4,000	22,200	0.93	0.92	0.91	0.88	0.85	0.80	0.74	0.66	0.55	0.39
5,000	27,750	0.84	0.83	0.82	0.80	0.77	0.72	0.67	0.59	0.49	0.35
6,000	33,300	0.77	0.77	0.76	0.74	0.71	0.67	0.61	0.54	0.45	0.33
8,000	44,400	0.68	0.67	0.66	0.65	0.62	0.58	0.54	0.48	0.40	0.29
10,000	55,500	0.61	0.61	0.60	0.58	0.56	0.53	0.49	0.43	0.36	0.26
10,847	60,201	0.59	0.59	0.58	0.56	0.54	0.51	0.47	0.42	0.35	0.25

For the above table, the coefficients in equation (2) of appendix VI are: b0 = -0.1513; b1 = 0.0810; b2 = 0.0493; b3 = 0.0908.

Table P. Generalized standard errors for estimated percentages and corresponding sample sizes from the pregnancy-interval file: National Survey of Family Growth, Cycle 5

Sample size	Weighted size (000s)	Percentage									
		50	45 or 55	40 or 60	35 or 65	30 or 70	25 or 75	20 or 80	15 or 85	10 or 90	5 or 95
		Hispanic women									
100	407	6.26	6.20	6.08	5.92	5.70	5.42	5.05	4.58	3.96	3.03
200	813	4.59	4.55	4.46	4.34	4.18	3.97	3.71	3.36	2.90	2.22
300	1,221	3.83	3.79	3.72	3.62	3.49	3.31	3.09	2.80	2.42	1.85
400	1,628	3.37	3.33	3.27	3.19	3.07	2.91	2.72	2.47	2.13	1.63
500	2,034	3.05	3.02	2.96	2.88	2.78	2.64	2.46	2.23	1.93	1.48
600	2,441	2.81	2.78	2.73	2.66	2.56	2.43	2.27	2.06	1.78	1.36
700	2,848	2.62	2.60	2.55	2.48	2.39	2.27	2.12	1.92	1.66	1.27
800	3,255	2.47	2.45	2.40	2.34	2.25	2.14	1.99	1.81	1.56	1.20
900	3,662	2.35	2.32	2.28	2.22	2.14	2.03	1.89	1.72	1.48	1.13
1,000	4,069	2.24	2.21	2.17	2.12	2.04	1.94	1.81	1.64	1.41	1.08
1,100	4,476	2.14	2.12	2.08	2.03	1.95	1.85	1.73	1.57	1.35	1.04
1,300	5,290	1.99	1.97	1.93	1.88	1.81	1.72	1.61	1.46	1.26	0.96
1,500	6,103	1.87	1.85	1.81	1.76	1.70	1.61	1.51	1.37	1.18	0.90
1,900	7,731	1.68	1.66	1.63	1.59	1.53	1.45	1.35	1.23	1.06	0.81
2,300	9,358	1.54	1.53	1.50	1.46	1.40	1.33	1.24	1.13	0.97	0.75
2,700	10,986	1.44	1.42	1.39	1.36	1.31	1.24	1.16	1.05	0.91	0.69
3,100	12,614	1.35	1.34	1.31	1.28	1.23	1.17	1.09	0.99	0.85	0.65
3,500	14,241	1.28	1.26	1.24	1.21	1.16	1.11	1.03	0.94	0.81	0.62
3,900	15,869	1.22	1.21	1.18	1.15	1.11	1.05	0.98	0.89	0.77	0.59
3,942	16,040	1.21	1.20	1.18	1.15	1.10	1.05	0.98	0.89	0.77	0.59

For the above table, the coefficients in equation (2) of appendix VI are: b0 = -0.1616; b1 = -0.3059; b2 = -0.1756; b3 = 0.1060.

Sample size	Weighted size (000s)	50	45 or 55	40 or 60	35 or 65	30 or 70	25 or 75	20 or 80	15 or 85	10 or 90	5 or 95
		Non-Hispanic black									
100	314	6.15	6.25	6.27	6.22	6.07	5.83	5.46	4.94	4.20	3.08
200	628	4.49	4.56	4.58	4.54	4.43	4.25	3.98	3.60	3.06	2.25
300	942	3.73	3.79	3.81	3.77	3.69	3.54	3.31	3	2.55	1.87
400	1,256	3.28	3.33	3.34	3.31	3.23	3.10	2.91	2.63	2.24	1.64
500	1,570	2.96	3.01	3.02	2.99	2.92	2.80	2.63	2.38	2.02	1.48
600	1,884	2.72	2.77	2.78	2.76	2.69	2.58	2.42	2.19	1.86	1.37
700	2,199	2.54	2.58	2.59	2.57	2.51	2.41	2.25	2.04	1.73	1.27
800	2,513	2.39	2.43	2.44	2.42	2.36	2.26	2.12	1.92	1.63	1.20
900	2,827	2.27	2.30	2.31	2.29	2.24	2.15	2.01	1.82	1.55	1.14
1,000	3,141	2.16	2.20	2.20	2.18	2.13	2.05	1.92	1.73	1.47	1.08
1,100	3,455	2.07	2.10	2.11	2.09	2.04	1.96	1.84	1.66	1.41	1.04
1,500	4,711	1.80	1.83	1.83	1.82	1.77	1.70	1.59	1.44	1.23	0.90
1,900	5,968	1.61	1.64	1.65	1.63	1.59	1.53	1.43	1.30	1.10	0.81
2,300	7,224	1.48	1.50	1.51	1.50	1.46	1.40	1.31	1.19	1.01	0.74
2,700	8,480	1.38	1.40	1.40	1.39	1.36	1.30	1.22	1.10	0.94	0.69
3,100	9,737	1.29	1.31	1.32	1.31	1.28	1.22	1.15	1.04	0.88	0.65
3,900	12,250	1.16	1.18	1.19	1.18	1.15	1.10	1.03	0.93	0.79	0.58
4,700	14,762	1.07	1.09	1.09	1.08	1.06	1.01	0.95	0.86	0.73	0.54
5,500	17,275	1.00	1.01	1.02	1.01	0.98	0.94	0.88	0.80	0.68	0.50
6,135	19,272	0.95	0.96	0.97	0.96	0.94	0.90	0.84	0.76	0.65	0.47

For the above table, the coefficients in equation (2) of appendix VI are: b0 = 0.1335; b1 = 0.4491; b2 = 0.00418; b3 = 0.09139.

Sample size	Weighted size (000s)	50	45 or 55	40 or 60	35 or 65	30 or 70	25 or 75	20 or 80	15 or 85	10 or 90	5 or 95
		All women and white women									
100	519	5.36	5.51	5.59	5.60	5.52	5.35	5.06	4.62	3.97	2.95
200	1,038	3.90	4.01	4.07	4.08	4.02	3.89	3.68	3.36	2.89	2.15
300	1,556	3.24	3.33	3.38	3.38	3.34	3.23	3.06	2.79	2.40	1.78
400	2,075	2.84	2.92	2.96	2.97	2.93	2.83	2.68	2.45	2.10	1.56
500	2,594	2.56	2.64	2.67	2.68	2.64	2.56	2.42	2.21	1.90	1.41
600	3,113	2.36	2.42	2.46	2.46	2.43	2.35	2.23	2.03	1.74	1.30
700	3,631	2.20	2.26	2.29	2.30	2.26	2.19	2.07	1.89	1.63	1.21
800	4,150	2.07	2.12	2.16	2.16	2.13	2.06	1.95	1.78	1.53	1.14
900	4,669	1.96	2.01	2.04	2.05	2.02	1.95	1.85	1.69	1.45	1.08
1,000	5,188	1.86	1.92	1.95	1.95	1.92	1.86	1.76	1.61	1.38	1.03
1,600	8,300	1.50	1.55	1.57	1.57	1.55	1.50	1.42	1.30	1.11	0.83
2,000	10,375	1.36	1.40	1.42	1.42	1.40	1.36	1.28	1.17	1.00	0.75
2,400	12,450	1.25	1.28	1.30	1.31	1.29	1.25	1.18	1.08	0.92	0.69
3,000	15,563	1.13	1.16	1.18	1.18	1.16	1.13	1.06	0.97	0.83	0.62
3,600	18,675	1.04	1.07	1.08	1.08	1.07	1.04	0.98	0.89	0.77	0.57
5,000	25,938	0.89	0.92	0.93	0.93	0.92	0.89	0.84	0.77	0.66	0.49
8,000	41,500	0.72	0.74	0.75	0.75	0.74	0.72	0.68	0.62	0.53	0.40
12,000	62,250	0.60	0.61	0.62	0.62	0.62	0.60	0.56	0.51	0.44	0.33
17,000	88,188	0.51	0.52	0.53	0.53	0.52	0.51	0.48	0.44	0.38	0.28
24,418	126,667	0.43	0.44	0.45	0.45	0.44	0.43	0.41	0.37	0.32	0.24

For the above table, the coefficients in equation (2) of appendix VI are: b0 = 0.0942; b1 = 0.6815; b2 = -0.0123; b3 = 0.0836.

Interpolation may be used for sample sizes and/or percentage estimates that are not shown in the tables. Alternatively, the GSE's may be generated with spreadsheet programs by using the coefficients that are shown below each table.

GSE's for an estimated total can be obtained by multiplying the GSE for the corresponding percentage found in Table O or P by the weighted sample count and then dividing by 100. That is,

$$GSE(Y) = GSE(P,n) * (W/100)$$

where

Y = estimated population total for domain of interest

P = estimated population percentage for domain of interest

n = sample size for domain of interest

$GSE(P,n)$ = GSE from table O or P for estimated proportion and sample size n

W = weighted sample size for domain of interest

The weighted sample sizes for each domain of interest are shown next to the sample sizes in tables O and P.

Continuing the examples shown previously, the GSE of the estimated total number of Hispanic women with the characteristic of interest is 53,938 and is obtained by multiplying the GSE of the percentage (1.25) by the weighted size (4,315,000) found in table O and then dividing by 100.

Comparison of Generalized Standard Error Estimates to Direct Estimates

Three methods for calculating standard errors have been discussed to this point—the SRS method, the direct design-based method (SUDAAN, Wes Var, PC CARP et al.), and the generalized standard error estimate, or GSE. Table Q shows the proportion of women currently using the oral contraceptive pill in 1995, and the standard errors as calculated by the SRS, SUDAAN, and GSE methods—by age, race, Hispanic origin, and religion.

As suggested earlier, the SUDAAN estimates are normally the most accurate. The SRS estimates are always lower than the SUDAAN estimates—sometimes a little lower and sometimes substantially lower. Thus, if an analyst uses the SRS standard errors, he or she will find more significant differences than actually exist. Note also that the sum of the differences between the GSE and SUDAAN estimates in table Q (0.67) is lower than the sum of the differences between the SRS and SUDAAN estimates (1.07). Thus, if the SUDAAN estimates are the most accurate, the GSE estimates are, on average, more accurate than SRS estimates. It is recommended that analysts use estimates from SUDAAN or another design-based procedure when they are available. Use the GSE estimates when SUDAAN or other design-based estimates are not available.

Table Q. Comparison of three ways of estimating the standard errors for the percent currently using oral contraceptive pills: (1) assuming a simple random sample, (2) using SUDAAN, and (3) using generalized standard errors: 1995 National Survey of Family Growth

Characteristic	Sample size	Percent using the pill	Standard errors (percent)[1]		
			SRS[2]	SUDAAN[3]	GSE[4]
Total[5]	10,847	17.3	0.36	0.43	0.44
Age					
15–19 years	1,416	13.1	0.89	0.92	0.99
20–24 years	1,519	33.5	1.21	1.52	1.43
25–29 years	1,739	26.8	1.06	1.3	1.29
30–34 years	2,148	20.5	0.87	1.03	1.01
35–39 years	2,144	8.2	0.59	0.61	0.75
40–44 years	1,881	4.3	0.47	0.57	0.59
Race, origin, and religion					
White Protestant	3,503	18.6	0.66	0.69	0.74
White Catholic	1,802	19.3	0.93	0.96	1.01
Hispanic	1,553	13.6	0.87	0.95	1.01
Non-Hispanic black	2,446	14.8	0.72	0.79	0.93

[1]Approximate 95 percent confidence intervals (CI's) may be obtained by adding and subtracting 1.96 times the standard error of interest. For example, the 95 CI for the 17.3% estimate across all sample women is +/–0.72% if SRS is assumed. This compares to +/– 0.86% for using a direct SUDAAN estimate of the standard error.

[2]SRS Is simple random sample.

[3]Computed using Proc Descript in SUDAAN. Program code is shown in appendix V.

[4]GSE is generalized standard errors.

[5]Total includes white women and women of other races with other religions or no religion, not shown separately.

Hypothesis Tests

An estimate of the standard error of the difference, $X-Y$, between any two aggregates or percents is given by

$$SE\ (X-Y) = \sqrt{[SE(X)]^2 + [SE(Y)]^2}$$

This expression provides a good estimate of the standard error for uncorrelated statistics, but it can be considered only a rough approximation otherwise.

Because estimates from the 1995 NSFG are based on a large sample of women and because the variance estimates were based on 198 PSU's, the test statistics

$$t = \frac{X-Y}{SE\ (X-Y)}$$

will be approximately normally distributed unless the sample size is very small. Therefore, individual two-tailed significance tests of differences between statistics from Cycle 5 data can be performed with an approximate significance level of alpha by computing t and comparing it with the two-tailed 1-α critical value for the normal distribution.

Example: From table Q, the estimated percentage of white Protestants using the pill was 18.6 percent compared with 19.3 percent for white Catholics. The corresponding standard errors (SUDAAN estimates) are 0.7 percent and 1.0 percent, respectively. To test whether this difference is significant at the 0.05 level of significance, compute

$$t = \frac{19.3 - 18.6}{\sqrt{(0.7)^2 + (1.0)^2}}$$

$$= 0.57$$

The two-tailed critical value for a normal test statistic at the 0.05 level of significance is 1.96. Therefore, the 0.7 percent difference between white Protestants and white Catholics is not significant at the 0.05 level.

References

1. Kelly JE, Mosher WD, Duffer AP, Kinsey SH. Plan and operation of the 1995 National Survey of Family Growth. National Center for Health Statistics. Vital Health Stat 1(36). 1997.
2. Gardenier JS. Problems, trade-offs, and solutions for CAPI surveys. In: American Statistical Association, 1994 Proceedings of the Section on Survey Research Methods, pp. 857–60. 1994.
3. Mosher WD, Pratt WF, Duffer AP. CAPI, event histories, and incentives in the NSFG Cycle 5 Pretest. In: American Statistical Association, 1994 Proceedings of the Section on Survey Research Methods, pp. 59–63. 1994.
4. Duffer AP, Lessler J, Weeks MF, Mosher W. Effects of incentive payments on response rates and field costs in a pretest of a national CAPI survey. In: American Statistical Association, 1994 Proceedings of the Section on Survey Research Methods, pp. 1386–91. 1994.
5. Massey JT, Moore TF, Parsons VL, Tadros W. Design and estimation for the National Health Interview Survey, 1985–94. National Center for Health Statistics. Vital Health Stat 2(101). 1989.
6. Chromy JR. Sequential sample selection methods. In: American Statistical Association, 1979 Proceedings of the Section on Survey Research Methods, pp. 401–6. 1979.
7. Judkins DR, Mosher WD, Botman S. National Survey of Family Growth: Design, estimation, and inference, National Center for Health Statistics, Vital and Health Stat (2)109. 1991.
8. Magidson J. SPSS for Windows: CHAID, release 6.0. Belmont, MA: Statistical Innovations, Inc., 1993.
9. Folsom RE. Exponential and logistic weight adjustments for sampling and nonresponse error reduction. In: American Statistical Association, 1991 Proceedings of the Section on Survey Research Methods, pp. 197–201. 1991.
10. Potter FJ. A study of procedures to identify and trim extreme sampling weights. In: American Statistical Association, Proceedings of the Section on Survey Research Methods, pp. 225–230. 1990.
11. Shah BV, Barnswell BG, Bieler GS. SUDAAN user’s manual: Software for the analysis of correlated data, release 6.40. Research Triangle Institute: Research Triangle Park, North Carolina. 1995.
12. Folsom RE, Witt MB. Testing a new attrition nonresponse adjustment method for SIPP. In: American Statistical Association, Proceedings of the Section on Survey Research Methods, pp. 428–33. 1994.
13. Iannacchione VG, Milne JG, Folsom RE. Response probability weight adjustments using logistic regression. In: American Statistical Association, Proceedings of the Section on Survey Research Methods, pp. 637–42. 1991.
14. Manley JA, McNeil BJ. The meaning and use of the area under a receiver-operating characteristic (ROC) curve. Diagnostic Radiology, 143:29–36. 1982.
15. Cox BG. The weighted sequential hot-deck imputation procedure. In: 1980 American Statistical Association, Proceedings of the Section on Survey Research Methods, pp. 721–26. 1980.
16. Abma J, Chandra C, Mosher W, Peterson L, Piccinino L. Fertility, family planning, and women’s health: New data from the 1995 National Survey of Family Growth. National Center for Health Statistics. Vital Health Stat 23(19). 1997.
17. Brown SS, Eisenberg L, eds. The best intentions: Unintended pregnancy and the well-being of children and families. Washington DC: National Academy Press. 1995. Especially pp. 21–82 and 286–95.
18. Wolter KM. Introduction to variance estimation. New York: Springer-Verlag. 1985.
19. Cohen S. An evaluation of alternative PC-based software packages developed for the analysis of complex survey data. The American Statistician 51(3) 285–92. 1997.
20. Woodruff RS. A simple method for approximating the variance of a complicated estimate. JASA 66:411–4. 1971.
21. Casady RJ, Parsons VL, and Snowden CB. Simplified variance estimation for the National Health Interview Survey. In: American Statistical Association, 1987 Proceedings of the Section on Survey Research Methods, pp. 412–7. 1987.
22. Bachrach C, Horn M, Mosher W, Shimizu I. National Survey of Family Growth, Cycle 3: Sample design, weighting, and variance estimation. National Center for Health Statistics. Vital Health Stat (2)98. 1985.
23. Waksberg J, Sperry S, Judkins D, Smith V. National Survey of Family Growth, Cycle IV, evaluation of linked design. National Center for Health Statistics. Vital Health Stat (2)117. 1993.
24. Botman SL, Waksberg J, Sperry S, Pratt WF. Integration of two surveys. In: American Statistical Association, 1989 Proceedings of the Section on Survey Research Methods, pp. 572–7. 1989.
25. Waksberg J, Northrup D. Integration of the design of the National Survey of Family Growth with the National Health Interview Survey. National Center for Health Statistics. Vital Health Stat (2)96. 1985.
26. Mathiowetz N, Northrup D, Sperry S, Waksberg J. Linking the National Survey of Family Growth with the National Health Interview Survey: Analysis of field trials. National Center for Health Statistics. Vital Health Stat 2(103). 1987.
27. Mason RE, Pate DK, Potter FJ, Weeks MF, George BJ, Duffer AP Jr. National Survey of Family Growth, Cycle V: Comparison of NHIS-linked person and non-linked area household design alternatives, rep no 5442–02–01S. Research Triangle Institute: Research Triangle Park, North Carolina. 1994.
28. Binder DA. On the variances of asymptotically normal estimators from complex surveys, International Statistical Rev (51)279–92. 1983.
29. Binder DA. On the variances of asymptotically normal estimators from complex surveys, Survey Methodology, Vol. 7, No. 2, pp. 157–70. 1981.

Table 1. Distribution of sample women selected for the 1995 National Survey of Family Growth, weighted location and response rates by characteristics of women and their households as measured in the 1993 National Health Interview Survey

National Health Interview Survey characteristic	Sample count	Number located	Weighted percent located	Number eligible[1]	Weighted percent responded[2]
All sample women	14,000	13,243	95.4	13,795	78.7
Age					
15–17 years	1,040	1,001	96.1	1,020	81.4
18–24 years	2,622	2,452	94.4	2,586	82.1
25–29 years	2,339	2,146	92.7	2,310	74.5
30–34 years	2,815	2,656	95.1	2,783	77.2
35–39 years	2,751	2,632	96.4	2,723	78.6
40 years and over	2,433	2,356	97.4	2,373	78.9
Race					
White	9,634	9,236	96.2	9,498	79.7
Black	3,264	2,991	91.7	3,227	77.8
Asian or Pacific Islander	452	424	94.1	434	62.3
Other	650	592	91.7	636	77.6
Hispanic origin					
Puerto Rican	266	237	90.6	256	79.0
Mexican	1,183	1,091	92.2	1,153	81.1
Hispanic other	553	508	92.8	527	78.0
Non-Hispanic	11,998	11,407	95.7	11,859	78.5
Marital status					
Married	7,391	7,065	96.3	7,290	79.2
Less than 14 years	229	222	96.6	226	84.6
Widowed, divorced, or separated	1,439	1,322	92.4	1,419	77.0
Never married	4,827	4,535	94.8	4,748	78.3
Unknown	114	99	88.0	112	64.5
Number of children in household					
0	6,742	6,350	94.7	6,626	77.3
1	2,533	2,400	95.8	2,496	78.7
2	2,781	2,663	96.7	2,747	80.5
3	1,346	1,278	95.8	1,331	81.2
4 or more	598	552	93.9	595	80.7
Education					
8th grade or less	1,250	1,150	92.9	1,219	76.2
Some high school	2,276	2,120	94.2	2,238	79.9
High school diploma	4,835	4,548	95.0	4,763	77.4
Some college	3,122	3,003	96.6	3,092	80.4
College graduate	2,421	2,345	97.2	2,392	80.1
Unknown	96	77	80.2	91	45.9
Education of responding adult					
Less than high school	1,657	1,487	90.7	1,620	74.5
High school	4,813	4,515	94.7	4,750	78.3
Some college	3,624	3,466	96.0	3,582	79.5
Bachelor's degree or higher	3,831	3,716	97.3	3,773	80.4
Family income					
Under $10,000	1,805	1,627	90.7	1,781	80.1
$10,000–19,999	2,149	1,974	92.2	2,107	77.5
$20,000–29,999	1,937	1,865	96.7	1,911	81.8
$30,000–39,999	1,783	1,740	98.0	1,769	81.0
$40,000–49,999	1,488	1,455	98.0	1,470	82.5
$50,000 or more	2,708	2,661	98.3	2,682	80.3
Unknown	2,130	1,921	91.5	2,075	68.0
Poverty index					
At or above poverty line	10,419	10,041	96.8	10,288	80.0
Below poverty line	2,375	2,148	91.0	2,338	78.6
Unknown	1,206	1,054	88.7	1,169	65.0

See footnotes at end of table.

Table 1. Distribution of sample women selected for the 1995 National Survey of Family Growth, weighted location and response rates by characteristics of women and their households as measured in the 1993 National Health Interview Survey—Con.

National Health Interview Survey characteristic	Sample count	Number located	Weighted percent located	Number eligible[1]	Weighted percent responded[2]
Employment status in past 2 weeks					
Employed	8,256	7,873	95.9	8,164	78.6
Unemployed	603	551	92.3	600	78.2
Not in labor force	3,516	3,255	93.8	3,434	76.5
Under 18 years, not applicable	1,625	1,564	96.3	1,597	83.2
Class of worker					
Private	6,689	6,354	95.6	6,625	78.2
Government	1,494	1,449	97.5	1,481	82.0
Self-employed	450	426	95.8	441	77.7
Not in labor force	3,541	3,278	93.8	3,459	76.5
Under 18 years	1,625	1,564	96.3	1,597	83.2
Unknown	201	172	86.9	192	67.2
Major activity					
Working	7,693	7,329	95.9	7,604	77.9
Keeping house	2,957	2,746	94.1	2,901	77.9
Going to school	1,266	1,192	94.7	1,247	81.2
Under 18 years	1,625	1,564	96.3	1,597	83.2
Other, unknown	459	412	91.3	446	69.3
Living quarters					
House, apartment, or flat	12,940	12,246	95.4	12,747	78.6
Mobile home	678	645	95.3	671	79.3
Hotel or group quarters	382	352	93.4	377	79.7
Region					
Northeast	2,732	2,573	95.0	2,679	76.1
Midwest	3,234	3,117	97.2	3,199	81.4
South	4,859	4,595	95.1	4,806	77.4
West	3,175	2,958	94.1	3,111	79.8
Metropolitan statistical area status					
MSA, central city, population 1 million or more	3,065	2,769	90.9	2,996	74.5
MSA, central city, less than 1 million	2,185	2,067	95.5	2,152	79.7
MSA, not central city	6,122	5,853	96.0	6,041	77.8
Not MSA	2,628	2,554	97.5	2,606	83.5
Urban/rural residence					
Urban	10,887	10,221	94.7	10,710	77.6
Rural	3,113	3,022	97.4	3,085	81.7
Area oversampled for black persons					
No	12,270	11,652	95.5	12,097	78.0
Yes	1,730	1,591	92.6	1,698	77.5
Number of calls in NHIS[3]					
0–1	4,010	3,811	96.1	3,959	81.4
2	3,223	3,075	95.9	3,179	80.2
3	2,238	2,139	95.9	2,198	77.9
4 or more	4,529	4,218	94.0	4,459	75.5
Number of contacts in the NHIS					
0	567	92	18.7	567	0.0
1	2,214	2,071	94.1	2,062	76.2
2	4,783	4,730	99.1	4,759	89.5
3	2,342	2,315	99.1	2,328	82.0
4	1,467	1,453	99.1	1,462	81.4
5	813	797	98.5	809	74.8
6	540	531	98.4	537	72.3
7	372	368	98.9	370	71.4
8	258	250	97.6	258	68.3
9 or more	644	636	98.8	643	64.6

See footnotes at end of table.

Table 1. Distribution of sample women selected for the 1995 National Survey of Family Growth, weighted location and response rates by characteristics of women and their households as measured in the 1993 National Health Interview Survey—Con.

National Health Interview Survey characteristic	Sample count	Number located	Weighted percent located	Number eligible[1]	Weighted percent responded[2]
Number of tracing attempts					
0	11,266	11,199	99.4	11,096	83.7
1	1,009	843	85.0	994	61.8
2	573	444	78.7	562	59.5
3	379	279	78.0	374	57.9
4	220	147	72.8	219	53.1
5 or more	553	331	62.5	550	44.0
Number of callbacks required for SSN[4]					
0	11,365	10,738	95.2	11,201	78.7
1 or more	1,722	1,659	96.8	1,695	80.7
Unknown	913	846	93.7	899	74.4
Number of callbacks required for immunizations					
0	11,039	10,435	95.3	10,876	78.6
1 or more	2,088	1,999	96.3	2,060	81.1
Unknown	873	809	93.5	859	73.4
Number of additional contacts					
0	11,471	10,806	95.0	11,300	77.9
1–8	1,399	1,355	97.2	1,379	82.7
9 or more	1,130	1,082	96.2	1,116	80.8
Health status of sample woman					
Excellent	4,949	4,718	96.1	4,882	79.1
Very good	4,307	4,080	95.2	4,248	78.9
Good	3,541	3,329	95.0	3,483	79.0
Fair or poor	1,154	1,072	93.5	1,136	75.3
Unknown	49	44	92.0	46	62.9
Number of health conditions					
0	8,575	8,079	95.0	8,422	76.8
1	3,288	3,129	95.8	3,258	81.3
2	1,257	1,201	96.6	1,246	83.5
3	501	475	95.8	495	79.9
4	209	201	96.4	207	82.6
5 or more	170	158	93.0	167	77.5
Doctor visits in past 12 months					
None	2,675	2,485	94.0	2,613	74.7
1	3,519	3,342	95.5	3,462	77.4
2	2,270	2,156	95.5	2,239	78.0
3–10	4,013	3,811	95.9	3,974	81.3
11 or more	1,523	1,449	95.8	1,507	82.5
Interval since last doctor visit					
Less than 1 year	11,399	10,821	95.6	11,254	79.6
1–2 years	2,158	2,016	94.5	2,112	75.4
2 years or more	246	228	93.7	238	73.6
Never or unknown	197	178	91.3	191	67.8
Refusal on height, weight, or health status					
Did not refuse	13,858	13,125	95.5	13,660	78.9
Refused 1 or more	142	118	84.4	135	60.6
Name of sample woman					
Reported	13,680	13,008	95.9	13,486	79.5
Refused	320	235	73.6	309	44.1
NHIS contact person					
Known and reported	11,402	10,933	96.5	11,252	81.2
Not reported	2,598	2,310	90.1	2,543	66.9

See footnotes at end of table.

Table 1. Distribution of sample women selected for the 1995 National Survey of Family Growth, weighted location and response rates by characteristics of women and their households as measured in the 1993 National Health Interview Survey—Con.

National Health Interview Survey characteristic	Sample count	Number located	Weighted percent located	Number eligible[1]	Weighted percent responded[2]
Sample woman refused to give SSN					
No	11,646	11,159	96.4	11,496	81.1
Yes	2,354	2,084	90.0	2,299	65.5
Telephone status					
Number given	12,373	11,856	96.4	12,210	79.8
Number refused	601	516	86.9	584	60.1
No number	1,026	871	84.8	1,001	73.1
Type of family recode					
Primary family	12,691	12,048	95.7	12,509	79.0
Secondary family	36	32	91.3	36	71.4
Primary individual	1,018	933	92.5	1,001	74.9
Secondary individual	255	230	91.9	249	79.3
Family relationship of sample woman					
Living alone	976	895	92.7	962	75.4
Living with nonrelative	297	268	91.4	288	77.6
Living with spouse	7,258	6,953	96.5	7,160	79.4
Living with relative	5,469	5,127	94.5	5,385	78.4
Number of families in the household					
1	13,616	12,890	95.4	13,422	78.6
2 or more	384	353	93.4	373	80.3
Relationship to NHIS reference person					
Reference person	4,590	4,227	93.3	4,519	78.4
Spouse	5,664	5,453	96.8	5,599	79.2
Other relative	3,746	3,563	95.4	3,677	78.2
Number of persons in household					
1	1,020	929	92.1	1,005	74.1
2	2,902	2,713	94.5	2,855	77.0
3	3,360	3,193	95.5	3,320	78.7
4	3,611	3,465	96.7	3,559	80.0
5	1,927	1,838	95.9	1,897	80.6
6	712	671	95.1	701	81.1
7 or more	468	434	93.8	458	74.6
NHIS respondent status					
Self	8,733	8,237	95.3	8,627	79.7
Part self, part proxy	816	771	95.1	807	77.5
Proxy	4,326	4,125	95.7	4,240	77.2
Unknown	125	110	90.9	121	70.4
Quarter interviewed for NHIS					
First	4,034	3,808	95.2	3,977	78.2
Second	2,192	2,057	94.7	2,150	77.0
Third	3,952	3,751	95.4	3,892	77.8
Fourth	3,822	3,627	95.8	3,776	81.0

0.0 Quantity more than zero but less than 0.05.

[1]Sample women not located were assumed to be eligible.

[2]Weighted number of responding sample women divided by the weighted number of eligible sample women.

[3]NHIS is National Health Interview Survey.

[4]SSN is Social Security number.

Table 2. Current Population Survey totals, relative standard errors, and sample sizes for the poststratification adjustment variables, by age, race, parity, and marital status: National Survey of Family Growth, Cycle 5

Age, race, parity, and marital status	CPS total[1]	RSE in CPS[2]	Sample size in CPS
Female, 15–29 years			
Parity:			
0	18,340,992	0.9	9,611
1	4,757,967	2.0	2,358
2	2,929,214	2.6	1,506
3	1,212,316	4.1	621
4 or more	464,053	6.7	231
Marital status:			
Ever married	9,738,274	1.3	4,943
Never married	17,966,268	0.9	9,384
Female, 15–17 years			
Race:			
Black, non-Hispanic	853,463	4.6	432
Hispanic	687,909	7.3	258
Other	3,911,014	2.2	2,247
Parity:			
0	5,244,906	1.9	2,831
1 or more	207,480	10.1	106
Female, 18–19 years			
Race:			
Black, non-Hispanic	538,133	6.0	257
Hispanic	461,876	8.7	175
Other	2,508,431	2.8	1,311
Black and Hispanic by parity:			
0	725,699	5.1	317
1 or more	274,310	8.5	115
Other race by parity:			
0	2,237,803	3.0	1,175
1 or more	270,628	8.8	136
Marital status:			
Ever married	325,705	8.0	145
Never married	3,182,735	2.5	1,598
Female, 20–24 years			
Black, non-Hispanic by parity:			
0	658,451	5.3	287
1	352,262	7.3	154
2 or more	317,707	7.7	135
Hispanic by parity:			
0	584,043	7.5	255
1	579,232	7.6	255
Other race by parity:			
0	4,671,559	2.0	2,546
1	1,198,161	4.1	620
2 or more	689,407	5.4	379
Black, non-Hispanic by marital status:			
Ever married	239,960	9.0	94
Never married	1,088,460	4.0	482
Hispanic by marital status:			
Ever married	505,144	8.1	222
Never married	658,131	7.0	288
Other race by marital status:			
Ever married	2,359,263	2.9	1,221
Never married	4,201,864	2.1	2,324

See footnotes at end of table.

Table 2. Current Population Survey totals, relative standard errors, and sample sizes for the poststratification adjustment variables, by age, race, parity, and marital status: National Survey of Family Growth, Cycle 5—Con.

Age, race, parity, and marital status	CPS total[1]	RSE in CPS[2]	Sample size in CPS
Female, 25–29 years			
Race:			
Black, non-Hispanic	1,345,589	3.5	607
Hispanic	1,216,905	5.2	522
Other	7,130,400	1.6	3,887
Black and Hispanic by parity:			
0	806,752	4.8	357
1	610,660	5.6	263
2	572,024	5.8	254
3 or more	573,058	5.8	255
Other race by parity:			
0	3,411,779	2.4	1,843
1	1,639,785	3.5	881
2	1,388,193	3.8	769
3 or more	690,643	5.4	394
Black and Hispanic by marital status:			
Ever married	1,401,619	3.5	597
Never married	1,160,875	3.9	532
Other race by marital status:			
Ever married	4,835,491	2.0	2,630
Never married	2,294,909	2.9	1,257
Female, 30–44 years			
Parity:			
0	6,901,403	1.6	3,744
1	5,947,721	1.8	3,208
2	10,946,270	1.2	5,858
3	5,748,869	1.8	3,082
4 or more	2,951,799	2.6	1,589
Marital status:			
Ever married	27,784,197	0.6	14,978
Never married	4,711,865	2.0	2,503
Female, 30–34 years			
Race:			
Black, non-Hispanic	1,455,690	3.4	691
Hispanic	1,232,970	5.0	473
Other	8,367,164	1.5	4,854
Black and Hispanic by parity:			
0	535,958	6.0	241
1	517,364	6.1	230
2	694,033	5.2	304
3	543,736	6.0	233
4 or more	397,569	7.0	156
Other race by parity:			
0	2,374,586	2.9	1,355
1	1,805,520	3.3	1,033
2	2,664,582	2.7	1,540
3	1,132,174	4.2	676
4 or more	390,302	7.2	250
Black and Hispanic by marital status:			
Ever married	1,772,563	3.0	739
Never married	916,097	4.5	425
Other race by marital status:			
Ever married	7,073,109	1.6	4,088
Never married	1,294,055	3.9	766

See footnotes at end of table.

Table 2. Current Population Survey totals, relative standard errors, and sample sizes for the poststratification adjustment variables, by age, race, parity, and marital status: National Survey of Family Growth, Cycle 5—Con.

Age, race, parity, and marital status	CPS total[1]	RSE in CPS[2]	Sample size in CPS
Female, 35–39 years			
Race:			
Black, non-Hispanic	1,439,489	3.4	674
Hispanic	1,066,959	5.5	415
Other	8,704,138	1.4	4,875
Black and Hispanic by parity:			
0	375,134	7.3	166
1	411,839	6.9	195
2	770,116	5.0	321
3	553,049	5.9	234
4 or more	396,310	7.1	173
Other race by parity:			
0	1,824,245	3.3	1,020
1	1,471,949	3.7	801
2	3,212,169	2.5	1,765
3	1,577,233	3.6	902
4 or more	618,542	5.8	387
Black and Hispanic by marital status:			
Ever married	1,851,996	3.0	784
Never married	654,452	5.4	305
Other race by marital status:			
Ever married	7,841,620	1.5	4,375
Never married	862,518	4.9	500
Female, 40–44 years			
Race:			
Black, non-Hispanic	1,249,222	3.8	619
Hispanic	872,549	6.6	318
Other	8,107,881	1.5	4,562
Black and Hispanic by parity:			
0	292,741	8.5	127
1	383,104	7.3	182
2	542,184	6.0	257
3	453,120	6.8	190
4 or more	420,622	7.0	181
Other race by parity:			
0	1,498,739	3.7	835
1	1,357,945	3.9	767
2	3,033,186	2.6	1,671
3	1,489,557	3.7	847
4 or more	728,454	5.4	442
Black and Hispanic by marital status:			
Ever married	1,768,559	3.1	767
Never married	353,212	7.6	170
Other race by marital status:			
Ever married	7,475,330	1.6	4,225
Never married	632,551	5.8	337

[1]CPS is Current Population Survey.

[2]RSE is relative standard error.

NOTE: May 1, 1995, Current Population Survey estimates for age group and race combinations. June 1, 1994, relative distribution of marital status and parity within each age group and race combination.

Appendix I

Definitions of Terms

Backwards elimination—Backwards elimination is a model-selection method often used in linear or logistic regression analysis. The method begins with the fitting of a model that uses all potential predictor (independent) variables. Then, the least significant variable that does not meet a specified level of significance is removed. (Once a variable is removed from the model, it remains excluded.) The process is repeated, deleting the weakest variable one at a time, until all variables left in the model are significantly related to the dependent variable at or beyond the specified level (for example, 0.05 or 0.10).

Clustering—A method of sampling in which the sampling unit contains more than one population element. Large surveys like the NHIS and NSFG often are geographically clustered because the cost per sample member is lower than a geographically dispersed sample. Clusters for the NSFG consisted of counties or MSA's (that is, PSU's) at the first stage of selection, groups of households or area segments within selected PSU's at the second stage, and households within selected area segments at the third stage.

Computer-Assisted Personal Interviewing (CAPI)—In the 1995 NSFG, a CAPI program was installed on a laptop computer to perform a personal interview of the respondent in her own home. A CAPI program selects the questions that are appropriate for a given respondent, selects the appropriate wording for each question, and determines whether an answer is valid or not. The result is higher quality data. The CAPI program was written in Blaise, a CAPI system developed by the Netherlands Central Bureau of Statistics. Version 2.5 of Blaise was used. Later versions of the program are now available.

Design effect—The ratio of the sampling variance for the sampling design compared with the sampling variance of a simple random sample of the same size from the same population. Design effects provide a summary measure of the combined effects of stratification, clustering, and unequal weighting on the variance of a survey estimate.

Donor case—In hot-deck imputation, a case with complete data that ''donates'' its value on a variable to a case with missing data.

Eligible woman—In the NSFG, a woman who was eligible to be selected for the NSFG sample. Specifically, she was born between April 1, 1950, and March 31, 1980, and in 1995 was in the civilian noninstitutionalized population of the United States. Thus, women born before April 1, 1950, or after March 31, 1980, or in the military, in prison, or some other institution, women who had left the United States, or women who died since the NHIS interview, were not eligible.

Item imputation—The process of assigning answers to cases with missing data (don't know, refused, or not ascertained). For example, if respondent 00123's education was missing and a value of 14 years (high school graduate plus 2 years of college) was assigned for her, then her education was imputed. Imputation was done in one of four ways in the 1995 NSFG: logical, unweighted hot-deck, weighted hot-deck, and regression. The purposes of imputation are to make the data complete, more consistent, and easier to use; and to reduce or eliminate bias caused by differential failure to respond. For most of the variables for which imputation was done in the NSFG, less than 1 percent of the cases received an imputed value.

Location rate—In this report, the location rate is the percent of sample women that were located, where located means that a correct telephone number or address was obtained. The correctness of the address was confirmed by talking by telephone or in person to a member of the woman's household. This was often, but not always, the woman selected for the NSFG.

National Center for Health Statistics (NCHS)—NCHS is the Nation's principal health statistics agency. It designs, develops, and maintains a number of systems that produce data related to demographic and health concerns. These include data on registered births and deaths, the National Health Interview Survey (NHIS), the National Health and Nutrition Examination Survey (NHANES), the National Health Care Survey, and the National Survey of Family Growth (NSFG), among others. NCHS has conducted the NSFG since 1973. NCHS is one of the ''Centers'' for Disease Control and Prevention (CDC), which is part of the U.S. Department of Health and Human Services.

National Health Interview Survey (NHIS)—The NHIS is a principal source of information on the health of the civilian noninstitutionalized population. The survey, conducted continuously since 1957, collects information from approximately 110,000 people in 43,000 households each year on health status, access to care and insurance, health services utilization, health behaviors, and other topics. The survey consists of a set of core data items that are repeated each year and a set of supplements that can change each year to address current health topics. Households interviewed in the 1993 NHIS were used as the sampling frame for the 1995 NSFG.

Participation rate—The percent of those located who participate or respond to the survey. In Cycle 5 of the NSFG, the overall unweighted participation rate was 10,847 divided by 13,243 cases located, or 81.9 percent.

Primary sampling unit (PSU)—A unit that is used for the first, or primary, stage of sampling. (Secondary units are parts of primary sampling units, and tertiary units are parts of secondary units.) In the NHIS and NSFG a PSU is a Metropolitan Statistical Area, a county, or a group of contiguous counties. The 1993 NHIS had 198 PSU's and the 1995 NSFG used all of these areas.

Race/ethnicity—Race/ethnicity, as reported in the 1993 NHIS, was used to design and select the NSFG sample. Three categories were used for purposes of sample design: Hispanic, non-Hispanic black, and other. Hispanic women and non-Hispanic black women were sampled at higher rates than others in the 1995 NSFG in order to obtain adequate numbers of Hispanic and black women for analysis. Thus, when this report contains tables showing ''race/ethnicity,'' the three categories are *those used to design and select the sample*. In reports that are designed to present substantive results, the ''other'' category is split into ''non-Hispanic white'' and ''non-Hispanic other race'' categories.

Raking procedure—Raking, also known as iterative proportional fitting, is a technique used to adjust the frequencies of a multidimensional table to conform to new marginal totals while preserving the internal structure of the table. A generalized raking procedure in the form of an exponential model was used to poststratify the NSFG sampling weights to estimated totals obtained from the Current Population Survey. The exponential model preserves totals of main-effect explanatory variables as well as totals associated with interaction terms.

Receiver operating characteristic (ROC) curves—ROC curves are used to evaluate the ability of statistical methods to predict an outcome. For the NSFG, an ROC curve was used to assess the overall predictive ability of the response propensity model that was used to adjust the sampling weights.

The ROC curve was constructed by considering a range of cutoff points for predicting whether a sample woman was a respondent (completed an interview) or a nonrespondent (did not complete an interview). The possible cutoff points ranged from always predicting response (that is, a cutoff less than the lowest predicted probability) to never predicting response (that is, a cutoff greater than the highest predicted probability).

Then, a point on the ROC curve was obtained by plotting the weighted proportion of respondents with a predicted probability greater than a specified cutoff (that is, the proportion of true positives) versus the weighted proportion of nonrespondents with a predicted probability greater than the cutoff (the proportion of false positives).

The ROC curve was obtained by computing the proportion of true and false positives for the entire range of possible cutoff points. A ROC curve that rises noticeably above the diagonal (where the proportion of true and false positives is equal) indicates that the statistical model is likely to correctly classify most sample women as either respondents or nonrespondents across the range of cutoff points.

Receptor case—In imputation, a case with missing data that receives the imputed value from another (donor) case. If Case A has missing data and Case A receives an imputed value from Case B, then Case A is the receptor case and Case B is the donor case.

Recodes or recoded variables—Variables constructed from other variables in the NSFG. These were the only variables in the NSFG data file for which missing data were imputed.

Response rate—Respondents divided by the number of eligible persons in the sample. In this report, the response rate is the number of respondents divided by the number of women in the sample (excluding ineligibles), times 100. Response rates can be calculated based on weighted or unweighted data. The overall *unweighted* response rate was 10,847 divided by 13,795, times 100, or 78.6 percent. The corresponding *weighted* response rate (shown in table 1 of this report) was 78.7 percent.

Respondents—Persons who answer, or respond to, a survey. In the 1995 NSFG, the ''respondents'' were the 10,847 women who completed an NSFG interview.

Research Triangle Institute (RTI)—RTI, located in Research Triangle Park, North Carolina, was the contractor selected to conduct the 1995 NSFG. RTI was established in 1958 and is an independent, not-for-profit organization with a staff of 1,450. RTI conducts survey research, as well as medical and laboratory research in the natural sciences.

Sampling variance—The sampling variance is a measure of the variation of a statistic, such as a proportion or a mean, that is due to having taken a sample instead of interviewing the full population. It measures the variation of the estimated proportion or mean over repeated samples. The sampling variance is zero when the full population is observed, as in a census. For the NSFG, the sampling variance estimate is a function of the sampling design and the population parameter being estimated (for example, a proportion or mean). Most common statistical software will attempt to compute ''population'' variances, which may under- or over-estimate the sampling variance. Estimating the sampling variance requires special software such as a replication technique, or the Taylor Series Approximation, or an adjustment to the standard variance formulas (see the ''Variance Estimation'' section).

Sampling weight—The number of women in the population that a woman in the sample represents. For example, if a woman's sampling weight is 5,000, then she represents an estimated 5,000 women in the population. Similarly, on the pregnancy interval file, the weight is the estimated number of pregnancies in the population that is represented by that sample pregnancy interval. The NSFG sampling weights adjust for different *sampling rates*, *response rates*, and *coverage rates* among sample women so that accurate national estimates can be made from the sample. Because it adjusts for all these factors, it is sometimes called a ''fully adjusted'' sampling weight.

Selected with certainty—When PSU's, households, or other units in a sampling frame are ''selected with certainty,'' it means that all of them are included in the sample. For the NHIS, 52 of the 198 PSU's were selected with certainty and are referred to as ''self-representing'' PSU's. All NHIS households with Hispanic or non-Hispanic black women were selected with certainty for the NSFG.

Simple random sample—A sample in which all members of the population are selected directly and have an equal chance to be selected for the sample. The NSFG sample is not a simple random sample because it was stratified, because it was selected in stages, and because the selection probability of women varied by their race and ethnicity and by the number of eligible women in their household.

Strata; Stratification—Stratification is the partitioning of a population of sampling units into mutually exclusive subpopulations (strata). Typically, stratification is used to increase the precision of survey estimates for subpopulations important to the survey's analytic objectives. The NSFG sample was stratified at each stage of selection: PSU's were stratified using socioeconomic and demographic variables; area segments within PSU's were stratified by concentration of black population; and women were stratified by race and ethnicity.

SUDAAN—SUDAAN is a statistical software package developed by the Research Triangle Institute to analyze data from complex sample surveys like the NSFG, as well as other observational and experimental studies involving clustered data. A complex sample may be multistage, stratified, and/or clustered. Information about SUDAAN can be obtained by phone: (919) 541–6236 or by e-mail: sudaan@rti.org.

Wald statistics—The Wald Chi-square and Wald F statistics are used for hypothesis testing in SUDAAN. The Wald chi-square statistic is the weighted analog of the conventional Pearson chi-square statistic, which assumes an equally-weighted sample. The Wald F statistic assumes that a finite number of denominator degrees of freedom (ddf) are available for testing. In SUDAAN, the number of ddf is assumed to be the number of PSU's minus the number of first-stage strata used for variance estimation. For the NSFG, there are 186 ddf available for testing using the with-replacement design option in SUDAAN.

Weight—See ''Sampling Weight.''

Wilcoxon statistic—The nonparametric Wilcoxon statistic is used to test whether the levels of a quantitative variable in one population tend to be greater than in a second population. It is nonparametric because no assumptions about how the variable is distributed in the populations are required for the test. Wilcoxon Statistics are defined in more detail in textbooks on nonparametric statistics, such as M. Hollander and D. Wolf, *Nonparametric Statistical Methods*, NY: John Wiley & Sons, 1972.

Appendix II

Linkage of the National Survey of Family Growth to the National Health Interview Survey

Previous Research on Linked Samples

Cycles 1, 2, and 3 of the National Survey of Family Growth (NSFG), conducted in 1973, 1976, and 1982, respectively, were designed as stand-alone area household surveys. In Cycle 4 (1988) and Cycle 5 (1995) the NSFG sample consisted of a subsample of women from the National Health Interview Survey (NHIS).

The linkage of the NSFG to the NHIS is a complex topic. The effects of linkage include:

1. Linkage makes new kinds of analysis possible, by allowing analysts to add a number of variables from the NHIS to the NSFG data. In the 1995 NSFG public-use data file, about 75 variables from the 1993 NHIS screener and core questionnaire are included (locations 12,938–13,056). This allows analyses combining NHIS and NSFG data.
2. Linkage permits the use of the NHIS to adjust for nonresponse to the NSFG. Thus, the nonresponse adjustments described in this report for Cycle 5 are far more sophisticated than those for Cycle 3 (22).
3. Linkage eliminates the expense of listing households and screening household members (in area-frame samples). These savings are offset by expenses on tracing persons who move (in linked samples).
4. Given fixed goals for oversampling certain populations, a linked sample may require more PSU's and area segments to obtain the desired sample size (for example, of black or Hispanic women in the 1995 NSFG). A larger number of PSU's and segments raises costs for interviewer time and travel, but reduces variances.
5. Nonresponse associated with screening (in area-frame samples) is eliminated, but nonresponse associated with tracing (inability to locate people who move) is incurred in linked samples.

The net effects of factors 3, 4, and 5 are likely to vary from survey to survey. In the 1995 NSFG, nonresponse associated with tracing was relatively large (5.5 percent), because of the long interval between the 1993 NHIS and the 1995 NSFG (13–34 months) and the incompleteness of the locator data. Nonresponse from tracing may increase the risk of nonresponse bias, but some or all of this bias may be eliminated by the more sophisticated nonresponse adjustments that are made possible by linking the NSFG with the NHIS data.

In this appendix, one limited aspect of this complex topic will be examined: nonresponse associated with tracing and nonparticipation (for example, refusals), and their potential effects on bias. This is intended to move the study of linked samples one step forward.

One objective of linking the two surveys, according to Waksberg et al. (23), was to ''reduce NSFG costs while keeping sampling error constant.'' Since the credibility and quality of the data rest in part on the response rates, the linkage should not reduce response rates significantly and should not increase nonresponse bias significantly.

Botman et al. (24) described research on the redesign of the NHIS that would allow linkage with other surveys, including the NSFG. Massey et al. (5) described the complete set of objectives adopted for the redesign of the NHIS, the research undertaken to develop the design, and the resulting sample design and estimation procedures.

Waksberg and Northrup (25) studied the design choices affecting the linkage of the NSFG to the NHIS. They compared the cost and variance implications of eight designs, drawing on Cycle 3 experience to quantify their cost and variance models. Their approach considered

- using all 200 of the NHIS PSU's versus a subsample of 100 PSU's
- using a person-linked versus an address-linked design
- fielding the NSFG only after enough NHIS interviews are accumulated versus continuous interviewing as the NHIS cases become available

For surveys having comparable variances, Waksberg and Northrup (25) estimated cost savings of between 28 and 35 percent for a linked design relative to a stand-alone area household survey, depending on other features of the designs. Based on the proportion of persons who might be expected to move between the two surveys, they projected an increase in the nonresponse rate of

- 0.5 percent, given continuous interviewing as the NHIS information becomes available
- 1 to 1.5 percent, using all of the NHIS PSU's
- 2 to 3 percent, using a half-sample of the NHIS PSU's (requiring a longer lag time to accumulate a sufficient number of NHIS observations)

Based on Cycle 3 experience, Waksberg and Northrup projected a response rate of about 84 percent for a stand-alone area household sample, and between 81 and 83 percent for a linked design, depending on the elapsed time between the two surveys. In these designs, they considered NHIS nonrespondents to be ineligible for the NSFG.

Mathiowetz et al. (26) reported on field trials conducted to determine:

- What are the effects of using a person-linked sample on response rates and level of effort compared with an address-linked sample?
- What are the effects on cost and response rates of initial contacts by telephone versus in person?
- How does the elapsed time between the two surveys affect the tracking effort and the willingness of sample persons to participate in the second study?
- Are the above effects the same in the various race and marital status categories?

To evaluate these issues, 1,315 NHIS households were selected for a "Reproductive Health Survey" and allocated to groups testing:

- a person-linked versus address-linked design
- initial contacts by phone versus in person
- varying elapsed times between the two surveys

The outcome variables were response rates, levels of effort, and costs. These outcomes were reported separately by race (black versus nonblack) and marital status (never married versus ever married).

The authors cautioned that the response rates obtained in the field trials are not directly comparable with rates that might be expected for the NSFG Main Study because of differences in the lengths of the questionnaires and other operational features. Further, interpretation of some of the treatment main effects is complex. Mathiowetz et al. (26) concluded that

- Somewhat higher response rates would be achieved with an address-linked sample (although the difference was not statistically significant in the "Reproductive Health Survey").
- The cost of data collection in a person-linked sample was 10–15 percent lower than in an address-linked sample.
- The mode of initial contact, telephone versus in-person, did not appear to affect response rates. However, the telephone contacts produced important reductions in the overall level of effort and costs.
- There was no clear effect of elapsed time between the two surveys.

In retrospect, it appears that the results of the experiment did not predict the results of Cycle 5 for at least four reasons:

1. The lag time between the NHIS and the Reproductive Health Survey ranged from 1 to 15 months, but the median lag time was only 5 months. In Cycle 5 of the NSFG, the lag time was much longer—it ranged from 13 to 34 months and the median lag time was 22 months. The original intent in Cycle 5 was to have a median lag time of just 11 months, but delays caused by computerizing the questionnaire and delays in obtaining clearances resulted in longer lag times than anticipated.
2. The questionnaire for the Reproductive Health Survey was only 10 to 15 minutes long compared with an average of 103 minutes for the 1995 NSFG interview. The hypothesis in the Reproductive Health Survey was that it might be easier for sample persons to refuse to do an interview over the phone, so initial telephone contact might increase refusal rates. However, a valid test of the hypothesis would have to use an interview in which the interview length was in the usual range of NCHS interviews conducted in the 1990's—1 to 2 hours. An effect of initial telephone contact on refusal rates may well be found if the interview was as long as the NSFG, but not found if the interview were only 10 to 15 minutes.
3. Respondents to the Reproductive Health Survey had signed waivers permitting the Census Bureau to release their information to the NHIS. This suggests that they were a more cooperative subsample, or that they felt committed to responding because they had signed the waiver. NSFG respondents had not signed any such waiver.
4. The number of PSU's in the Reproductive Health Survey was only 10—and most of the cases were in just 2 PSU's. Thus, the sample was far more concentrated (132 cases per PSU) than in the 1995 NSFG, where there were only 55 cases per PSU (10,847 respondents in 198 PSU's). This meant that mobility and travel costs were much higher in the NSFG than in the Reproductive Health Survey.
5. The comparisons assumed fixed variances and the rate of oversampling of black women used in Cycle 3. Cycles 4 and 5 used a lower rate of oversampling than in Cycle 3.

In short, the results of the Reproductive Health Survey did not predict the results of the 1995 NSFG because of a number of key differences between the surveys. All of these differences tended to raise costs and lower response rates in the 1995 NSFG compared with what was expected based on the Reproductive Health Survey.

Cycle 4 of the NSFG, conducted in 1988, made the first use of a person-linked design. The survey was fielded only after sufficient NHIS information was available to provide the necessary number of NSFG observations. This required information from the fourth quarter of 1985 through the first quarter of 1987. Interviews were conducted between January and August 1988. A complex system of subsampling both the NHIS first-stage units and the sample households at different rates was used to achieve desired oversampling rates for black women while preserving the NSFG requirement of sampling only one woman per household.

The Cycle 4 design and estimation is described by Judkins et al. (7). Waksberg et al. (23) evaluated the Cycle 4 design, recalibrating the cost and variance models used by Waksberg and Northrup (25) to reflect the actual Cycle 4 experience, and extrapolating this experience to estimate the cost of an area household design that would provide comparable variances to the linked design. Waksberg et al also assumed that a linked design would have a shorter interval between accumulating sufficient NHIS information and fielding the NSFG sample than was actually experienced in Cycle 4. Relative to an area household design having comparable variances, the linked design was estimated to have produced a 22 percent cost saving, approximately $900,000. If Cycle 4 had been fielded without delays, they projected a cost savings of 27–28 percent (23).

The Cycle 5 design differed from the Cycle 4 design in three major ways:

1. Oversampling of minorities: only black women were oversampled in Cycle 4; both Hispanic and black women were oversampled in Cycle 5.
2. For a number of reasons, the number of PSU's increased from 156 to 198 and the number of segments increased from 3,143 to 5,377—an increase of 71 percent. This difference made the sample much more dispersed in Cycle 5 than in Cycle 4.
3. Because of long-delayed data needs that had to be addressed by the NSFG, the interview length increased by almost 50 percent, from 70 minutes in Cycle 4 to 103 minutes in Cycle 5. This longer interview increased refusal rates (from 8 to 12 percent) and increased the costs of fieldwork.

The Cycle 5 sampling frame was constructed using NHIS information from all of 1993. Enough NHIS cases were accumulated to provide the target number of interviews for black and Hispanic women. This increased the absolute number of NSFG clusters necessary to avoid the variation in the sampling rates for women neither black nor Hispanic. This strategy increased the reliability of survey estimates for these women, but the increased number of clusters probably increased the survey costs compared with a design with substantially fewer clusters, such as that used for Cycle 4.

Based on cost estimates made following the Cycle 5 pretest, the cost of an unlinked area probability sample of the same number and distribution of women (10,500 total, 1,800 Hispanic and 3,000 non-Hispanic black) was estimated (27). Composite size measures were used to develop the area design so that the only contribution to the unequal weighting effect for the minorities was due to selecting a single sample woman per household. The estimated overall unequal weighting effect for the area design was comparable to the effect experienced in the person-linked Cycle 4 design and about 56 percent of the effect experienced in the Cycle 3 area sample. However, the estimated cost of the linked design remained lower than the estimated cost of an area design. Ultimately, the actual cost of the linked design proved to be higher than the estimate, primarily because of the large amount of effort required to (a) locate sample women who moved, (b) convert initial refusals, and (c) work a more dispersed sample.

Cycle 5 Tracing Activities

Cycle 5 tracing activities were divided into two parts: advance tracing and field tracing. Before the main study data collection, a thorough multistep ''advance tracing'' procedure was used to secure and confirm a current address and telephone number for each sample woman. The advance tracing activities were designed to give the field interviewers the correct address before they tried to contact each sample woman. These activities were conducted from July 1994 through November 1994. Then, as problems with the information sent to the field were identified, a second set of tracing activities was carried on during the data collection period. These field tracing activities were conducted from May 1995 through October 1995 (1).

The tracing activities in the aggregate were successful in locating 13,243, or 94.6 percent, of the total sample of 14,000 women.

NHIS Tracing Data

The NHIS information used in the tracing procedures (called ''locator information'' in reference 1) consisted of the following items:

- The sample woman's name, address, telephone number, Social Security number (SSN), date of birth, race, and marital status.
- The NHIS reference person's name, address, telephone number, SSN, and relationship to the sample woman. Typically, the NHIS reference person was the identified head of household (or householder).
- The NHIS contact person's name, address, telephone number, and relationship to the sample woman. Contact persons were identified by the NHIS respondent at the time of interview and were generally

Table I. Tracing steps in the 1995 National Survey of Family Growth

Tracing step	Date	Number of cases sent	Number of of cases updated/located
1. NCOA submission[1]	07/01/94	33,521	5,537
2. Mailing to postmasters for Rural Route addresses	07/22/94	NA[2]	NA[2]
3. Telematch submission[3]	07/22/94	32,876	4,608
4. Telephone tracing	08/01/94–10/31/94	14,000	11,787
5. Tracing contractor submission	08/29/94–11/21/94	1,599	863
6. NCOA resubmission	12/01/94	33,704	1,287
7. Postcard mailing	12/19/94	14,000	NA[2]
8. Field tracing by field interviewers	01/14/95–10/21/95	14,000	13,273
9. DMV requests[4]	01/14/95–10/31/95	952	545
10. Database searches	01/15/95–10/31/95	2,459	1,512
11. U.S. Bureau of the Census tracing	08/01/95–09/30/95	641	149

[1]NCOA is national change of address.
[2]NA is not ascertained.
[3]Telematch is a computerized database of residential telephone numbers. At the time of this study, it contained names and addresses for 65 million phone numbers in the United States.
[4]State Department of Motor Vehicles (DMV) listings show the most recent address of persons in that State, based on their name and date of birth.

relatives, neighbors, or friends who knew the respondent well.

- Whether the NSFG sampled woman was also the NHIS respondent.
- Date of the NHIS interview.

The objective of the tracing procedure was to provide current addresses and telephone numbers. About 38 percent of sample women had one or more pieces of tracing information missing: 31 percent had missing SSN's, 25 percent had no contact person listed, and 2 percent had completed the NHIS interview but had refused to give their names. Table I shows the major steps in tracing and the yield of each step. The tracing process is described in detail in another report (1) and summarized here.

Advance Tracing

The first step in advance tracing was to use the U. S. Postal Service's National Change of Address (NCOA) system to update address information for the sample women, for reference persons in the households, and for contact persons they named in the NHIS. Lists of rural route addresses by ZIP + 4 Code areas were also sent to appropriate postmasters with the request for street addresses. Out of the total of 33,521 individual records submitted, new (that is, different) addresses and forwarding addresses were obtained for 5,537 (17 percent).

Following the NCOA submission, the updated list was sent to a commercial service that uses names, street addresses, and ZIP Codes as search criteria to locate telephone numbers. RTI (the primary contractor), and its tracing subcontractors were required to protect the confidentiality of sample women and their families by the same laws that require NCHS employees to protect confidentiality. Note also that tracing information, or locator information, was used only to find a new address and telephone number to ask the woman for an interview. No other use was made of the tracing, or locator, information.

The next step in the advance tracing operation was to telephone all sample women to verify the telephone numbers and addresses obtained to this point. If the telephone number was confirmed by someone in the household, the case was classified as ''located'' and ready for assignment to a field interviewer. Otherwise, additional attempts to obtain and verify a number for the sample women were made (1).

At the end of these steps, 11,859 sample women (84.7 percent of the total sample of 14,000) had been located, including those with and without telephone numbers and a few who were ineligible or had died since the NHIS.

Further address searches for cases not yet located were conducted using credit and driver's license files (1). New addresses obtained for either the sample woman or the secondary sources were confirmed by telephone contact. This step yielded an additional 863 (6.2 percent) confirmed cases.

Two final steps were taken just before interviewing to identify cases who might have moved during the months of advance tracing: first, to resubmit the address to the U.S. Postal Service's NCOA for a final check; second, postcards preprinted with the request ''Do Not Forward–Address Correction Requested'' were sent to each address. New addresses were given to the field interviewers.

When interviewing began (January 1995), 977 (7.0 percent) of the 14,000 sample women had not yet been located. These cases were assigned to interviewers for an in-person followup and to the contractor's central office staff to check databases (1).

Toward the end of the data collection period, NCHS arranged to have the Bureau of the Census (which does the interviewing for the NHIS) trace the remaining unlocatable cases. Information on 641 sample women was sent to the Census regional offices. This effort located 149 sample women and resulted in 69 completed interviews.

Characteristics of Unlocated Sample Women

To better understand factors influencing location rates, a statistical model was developed to assess the association between the experienced location rates and

- the items comprising the NHIS tracing information
- selected characteristics of the sample women

The objective of the modeling exercise was to assess the effect on the success of the tracing activities when various pieces of the NHIS tracing data are missing. The outcome variable in the model takes on the value 1 if the sample woman was successfully located, and 0 if she was not located.

Logistic regression is preferable to linear regression for dichotomous outcomes when the predictor variables are not normally distributed with a common covariance matrix or when predicted values are generated. In this case however, interest centered only on identifying those factors that significantly influenced location rates. Therefore, a linear regression model was adequate to determine which factors were good predictors of location rates. The model was developed and evaluated using design-consistent variance estimation in SUDAAN. The results of this model are described in the following text.

A separate logistic regression model was used to predict the location probabilities of various NSFG subpopulations. These probabilities were used to adjust the weights for nonresponse. The development and use of that model is described in the section on ''Sampling Weights.''

Factors A through L in table II are derived from the NHIS information used in the tracing procedures. Factor M is the calendar quarter in which the NHIS interview occurred. Factors N and O (education and urban/rural residence) are characteristics of the sample woman thought to have an influence on the success of the tracing operation. Table III shows the test of the statistical significance of the association between the factors and the experienced location rate. A 0.05 level of significance was used, but many of the variables were significant at much lower levels.

The model has a multiple correlation coefficient (multiple R) of 0.26 and an R^2 value of 0.07. With a 0–1 dependent variable, a multiple R of 0.26 and an R^2 of 0.07 are not uncommon. (The location propensity model included interaction effects and additional demographic characteristics such as family income, which are highly significant but made only marginal improvement in the multiple correlation coefficient.) This low R^2 suggests that factors other than the items making up the NHIS tracing information and the characteristics of the sample women included in the model are more important determinants of success in locating the sample women. For example, some women in the sample may not wish to be located, for reasons such as that the respondent or her spouse is an illegal immigrant, or that the spouse is in prison, or owes child support, or has credit problems, or the woman or her spouse is working long hours and is not home much. There may also be other important factors affecting location rates.

The NHIS items that related directly to the sample woman and were used in the tracing process were the sample woman's name, address, telephone number, Social Security number, date of birth, race, and marital status. Each of

Table II. National Health Interview Survey Predictor variables used in the location model in table III

	Factor	Categories
A.	Name	Provided Refused
B.	P.O. box only, no street address	Yes No
C.	ZIP + 4 code	Provided Not available
D.	Telephone	Has telephone, number provided Has telephone, number not provided Telephone status unknown
E.	Social Security number	Provided Refused
F.	Age	15–17 years 18–24 years 25–29 years 30–34 years 35–39 years 40–44 years
G.	Imputed month of birth	Imputed Reported
H.	Race/ethnicity	Hispanic Non-Hispanic black Other
I.	Marital status	Married, spouse present Married, spouse absent Widowed Divorced Separated Never married Unknown
J.	Reference person	Sample woman is the reference person Other
K.	Contact person	Complete or partial information provided (at least one of name, address, telephone number, and relationship to sample woman) No information provided
L.	Respondent information	The sample woman was the NHIS respondent A proxy for the sample woman was the NHIS respondent Some NHIS information was obtained from the sample woman and some from a proxy unknown respondent
M.	NHIS quarter	January–March 1993 April–June 1993 July–September 1993 October–December 1993
N.	Educational attainment	Less than high school High school graduate Some college College graduate
O.	Urban/rural residence	Urban Rural

NOTE: 1 = located; 0 = not located; NHIS National Health Interview Survey.

Table III. Factors measured in the 1993 National Health Interview Survey affecting location rates in the 1995 National Survey of Family Growth

Source of variation	Degrees of freedom	Wald F	Probability
Model	32	11.36	<0.0001
Name (provided versus missing)	1	17.77	0.0001
Post Office box only (versus street address)	1	0.78	0.3792
ZIP + 4 code	1	8.70	0.0043
Telephone	2	17.66	<0.0001
Has telephone, number provided versus number not provided	1	2.75	0.1018
Has telephone, number provided versus telephone status unknown	1	33.27	<0.0001
Social Security number	1	13.47	0.0005
Age	5	13.47	<0.0001
15–17 years versus 18–24 years	1	16.06	0.0001
15–17 years versus 25–29 years	1	23.71	<0.0001
15–17 years versus 30–34 years	1	9.17	0.0034
15–17 years versus 35–39 years	1	4.28	0.0421
15–17 years versus 40–44 years	1	1.81	0.1828
Imputed month of birth	1	14.46	0.0003
Race/ethnicity	2	9.41	0.0002
Other versus Hispanic	1	1.81	0.0058
Other versus non-Hispanic black	1	8.08	0.0004
Marital status	6	1.94	0.0862
Married, spouse present versus married, spouse absent	1	3.76	0.0565
Married, spouse present versus widowed	1	1.03	0.3127
Married, spouse present versus divorced	1	5.05	0.0277
Married, spouse present versus separated	1	0.61	0.4381
Married, spouse present versus never married	1	0.77	0.3831
Married, spouse present versus marital status unknown	1	0.03	0.8570
Is sample woman the reference person?	1	12.43	0.0007
Contact person (named versus not named)	1	18.10	0.0001
National Health Interview Survey respondent	3	0.39	0.7595
Sample woman versus proxy	1	0.33	0.5681
Sample woman versus sample woman and proxy	1	0.09	0.7692
Sample woman versus unknown respondent	1	0.65	0.4235
National Health Interview Survey quarter	3	1.42	0.2432
January–March versus April–June	1	0.51	0.4767
January–March versus July–September	1	0.39	0.5371
January–March versus October–December	1	3.96	0.0504
Educational attainment	3	8.99	<0.0001
Less than high school versus high school graduate	1	5.10	0.0270
Less than high school versus some college	1	17.95	0.0001
Less than high school versus college graduate	1	17.23	0.0001
Urban/rural residence	1	24.17	<0.0001

these items was included in the model. Address was replaced by two variables: one measuring whether the address was only a Post Office box or a street address (factor B), and another measuring whether a complete ZIP Code was available (factor C). Date of birth was replaced by the age categories listed in factor F.

Based on the size of the Wald F test statistic, having the sample woman's name (Wald F = 17.77) and telephone number (Wald F = 33.27) were among the most important pieces of tracing information affecting location rates. The estimated location rate when the sample woman's name was provided was 95.9 percent compared with 73.6 percent when her name was not provided. The rates cited appear in table 1.

With respect to the telephone number, the important fact was whether the sample woman had a telephone number and reported it to the NHIS interviewer. The location rate for women with telephones who provided the number was 96.4 percent versus 84.8 percent for women with unknown telephone status.

Next in importance were the age of the woman and having her Social Security number. The estimated location rates for each of the age categories used in the model was 15–17 years, 96.1 percent; 18–24 years, 94.4 percent; 25–29 years, 92.7 percent; 30–34 years, 95.1 percent; 35–39 years, 96.4 percent; and 40–44 years, 97.4 percent. The most difficult age groups to locate are women in their twenties, probably because they tend to change addresses more frequently than other age groups.

Whether or not the age (actually the month of birth) of the woman was known as opposed to imputed was also a significant factor (F = 14.46). The estimated overall location rate for women whose ages were known was 95.9 percent compared with 76.6 percent for women whose ages were imputed.

The location rate for sample women whose Social Security number was reported was 96.4 percent compared with 92.2 percent for women whose Social Security number was not provided (Wald F = 13.47).

The race of the sample woman also contributed significantly to differences in location rates. The location rate for Hispanic women (92.3 percent) was not significantly different from that for others (96.4 percent), but the location rate for non-Hispanic black women (91.7 percent) was significantly lower.

With respect to marital status, sample women who were married with spouse present were compared with each of the other categories. Although the overall or average contribution of the marital status variable was not significant, the comparison between married women with spouse present (96.4 percent) and divorced women (92.8 percent located) was significant. The location rates for married women with spouse absent, widows, and separated women, were lower than for divorced women, but not significantly so, because of smaller sample sizes. The estimated percent located was 87.6 percent for married women with spouse absent, 90.3 percent for widows, and 91.8 percent for separated women.

Having reported an NHIS reference person and an NHIS contact person were also significantly associated with the location rates. When reference persons were reported, the estimated location rate was 96.2 percent compared with 93.3 percent when reference persons were not. Similarly, for the contact person, the rates were 96.5 and 90.1 percent, respectively. On the other hand, whether the sample woman responded for herself in the NHIS interview made no difference. The quarter during which the NHIS data collection took place had a marginally significant effect (0.0504) on location rates, after controlling for other factors. This suggests that short delays in data collection are not a significant problem in locating respondents but long delays may be.

Educational attainment was significantly associated with location rates. The rates themselves were 93.7 percent for those with less than a high school education; 95.0 percent for high school graduates; 96.6 percent for those with some college; and 97.7 percent for college graduates. Compared with women with less than a high school education, each of the above differences in location rates is significant.

Finally, differences between urban and rural residents were highly significant. Estimated location rates for urban residents were 94.7 percent compared with 97.4 percent for rural residents.

Effect of Unlocated Sample Women

A distinction is often made between the terms ‘‘noncoverage’’ and ‘‘undercoverage.’’ Noncoverage refers to any failure of the sampling frame to include the totality of the inferential population. Undercoverage refers to any failure to obtain information for every unit of observation selected into the sample. This section attempts to quantify the bias potential due to undercoverage associated with the NSFG and noncoverage associated with the use of NHIS respondents as the source information for constructing the NSFG frame.

Because response variable values are necessarily missing for the nonrespondents and for unlocated individuals, the actual biases associated with the parameter estimates are, of course, unknown. However, given the rates at which missing data occurs, the *potential* for bias can be quantified for sample estimates of population proportions. The bias associated with an estimated proportion can be bounded above and below for any value of the proportion.

The bounds show the worst case because the procedures used to compensate for missing data (described in the ‘‘Sampling Weights’’ and ‘‘Item Imputation’’ sections) reduce biases to much less than the extremes indicated by the bounds. Examining the bounds is however a useful way to assess the *relative* contributions of the components of the missing data problem:

1. Nonresponse to the NHIS
2. Inability to locate women in the NSFG
3. Nonresponse to the NSFG

In what follows, let N_R denote the respondent set (that is, all respondents to the NSFG). The complement nonrespondent set is denoted by

$$N_R^- = N - N_R$$

Then the minimum (that is, most negative) and maximum bias that can occur in association with the sample estimate of a population proportion, P, are given by

$$\min\{\text{bias}\{P\}\} = \frac{N_R^-}{N_R}(P-1) \quad \text{if} \quad \frac{N_R^-}{N} \le P \le 1$$

$$= -P \quad \text{if} \quad 0 \le P \le \frac{N_R^-}{N}$$

$$\max\{\text{bias}\{\hat{P}\}\} = \frac{N_R^-}{N_R} P \quad \text{if} \quad 0 \le P \le \frac{N_R^-}{N}$$

$$= 1 - P \quad \text{if} \quad \frac{N_R}{N} \le P \le 1$$

That is, the minimum bias is equal to the ratio of nonrespondents’ to respondents’ times (P−1) if the value of the population proportion is greater than the nonresponse rate in the population. If the value of the population proportion is less than the nonresponse rate, then the minimum bias is simply equal to the negative of the proportion. Similarly, the maximum bias is equal to the ratio of nonrespondents’ to respondents’ times the value of the population proportion if the proportion is less than the response rate, and it is equal to (1−P) if the proportion is greater than the response rate. Note that the proportion and the rates in the above expressions are the population parameters.

Figure I illustrates the bias bounds associated with

- the NHIS nonresponse rate
- the cumulative effect of the NHIS nonresponse rate and the subsequent NSFG nonlocation rate
- the cumulative effect of the NHIS nonresponse rate, the NSFG nonlocation rate, and the NSFG nonresponse rate

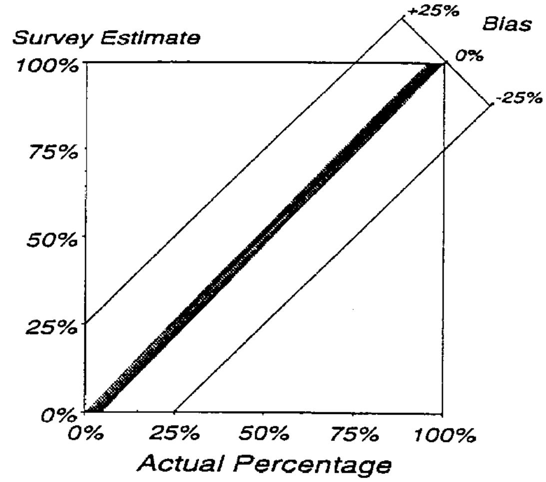
For the NHIS, Massey et al. (5) remark that ‘‘Historically, usually less than 5 percent of all eligible . . . sampled households do not respond.’’ Figure 1 assumes an NHIS response rate of 0.95. The Cycle 5 (weighted) location rate of 0.954 makes the cumulative rate for these two components equal to 0.906. Finally, the conditional NSFG response rate among located cases is 0.811, making the cumulative rate over all three components in figure 1 equal to 0.735.

Figure I answers the question, ‘‘What is the *potential* bias that can occur in a sample estimate of a population proportion given the response rates in the NHIS and NSFG?’’ The value of the proportion is entered on the y-axis. Then the expected value of the sample estimate of the proportion lies in the interval along the x-axis bounded by the points at which the value of the proportion intersects the shaded area in the figure. The shaded area above the diagonal line on each of the graphs in the figure is the potential positive bias, and the shaded area below the diagonal line is the potential negative bias.

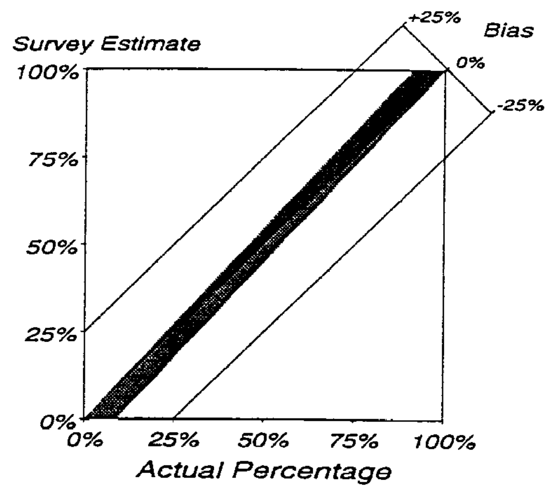
For example, consider a population proportion of 0.25. The intervals will not be symmetric and using a proportion

Shaded areas depict bias potential attributable to:

1. NHIS nonresponse:



2. NHIS nonresponse and NSFG nonlocation:



3. NHIS nonresponse and NSFG nonlocation and NSFG Nonresponse:

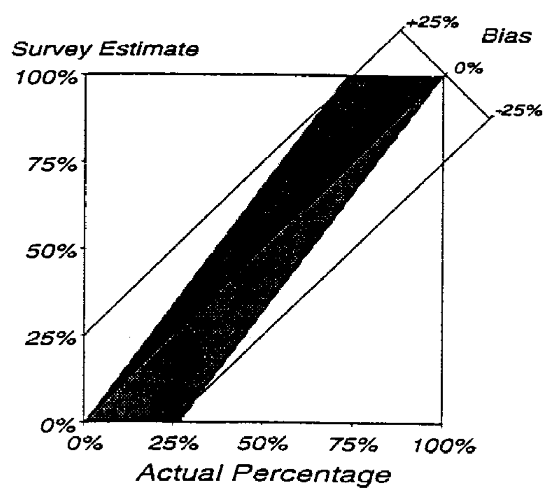


Figure I. Cumulative bias potential associated with National Health Interview Survey nonresponse and subsequent National Survey of Family Growth nonlocation and nonresponse

of 0.25 provides more opportunity for negative than for positive bias (see figure I, graph 1). Given the bias potential due to NHIS nonresponse, the expected value of the sample estimate lies in the range from 0.211 to 0.263. Adding the NSFG nonlocation component to this (figure I, graph 2) widens the range to between 0.172 and 0.276.

Finally, adding the NSFG nonresponse component (figure I, graph 3) produces a range from 0 to 0.340. If the NSFG was not linked to the NHIS, then applying only the NSFG conditional nonresponse rate yields an interval from 0.075 to 0.308 for the same population proportion.

Whether or not this difference can be considered important depends on a number of factors. Perhaps foremost among these are 1) the effectiveness of the nonresponse adjustments and other bias-reduction techniques—and the nonresponse adjustments shown in this report (which were made possible by linkage to the NHIS) are very detailed; and 2) the policy and program implications associated with using estimates that carry the larger potential bias. As shown in figure I, graph 3, for proportions smaller than the nonresponse rate, there is a greater chance for negative bias (underestimating the proportion) than for positive bias. This suggests that estimated statistics based on small proportions could be underestimated when the survey's response rate is low. For proportions larger than the response rate, survey estimates may be overreported. In both cases, the bias potential diminishes as the response rate increases.

It was noted earlier that *the potential bias overstates the actual bias: it is a worst-case scenario.* On the other hand, expecting a missing data compensation procedure to adjust totally for the missing data biases may be overly optimistic. The receiver operating characteristics curve associated with the response propensity model used for Cycle 5 (see the section on ''Sampling Weights'') while yielding a highly significant result, does not suggest a definitive prediction of every sample woman's response propensity. If, for example, the model was successful in reducing the potential bias by 65 percent (the area under the receiver operating characteristic curve), then the interval around the expected value of a sample estimate of a population proportion of 0.25 for the linked design is reduced (becoming 0.162 to 0.282, assuming the model is equally successful in reducing biases in both directions) but not eliminated. Hence, the bias issue should be considered along with the other advantages and disadvantages of a linked design.

Discussion

Some statistical issues in developing the design for an NSFG linked to the NHIS are listed in the following text.

1. The current practice of sampling one woman per household arises out of privacy/confidentiality concerns, but given the level of the NSFG's oversampling of black and Hispanic women, this feature reduces the statistical efficiency of the black and Hispanic estimates. There is, therefore, a trade-off between the one-per-household requirement and the variances for black and Hispanic women.
2. If the variation in sampling rates for black or Hispanic women could be reduced, the reliability of the statistics for these women would be increased.
3. If other methods cannot be used to increase the statistical efficiency of the sample for black and Hispanic women, the cost-effectiveness of the current level of NSFG oversampling of black and Hispanic women should be evaluated. The current level of oversampling of these populations increases the design effects (variances) significantly.
4. The design and development of the NSFG should be planned so that it can be fielded as soon as possible after the needed frame is available from the NHIS. This will increase the proportion located and reduce the cost of tracing.

Number of imputations and corresponding imputation specifications for the recoded variables from the respondent file ordered alphabetically within questionnaire section: National Survey of Family Growth, Cycle 5

Variable Name	Variable Number	Variable Labels	Number Imputed	Type of Imputation	Class Variables / General Notes	Screening Variables
Section A						
AGEAPR1	102	Age as of April 1, 1995	0	none	No imputations	
AGEFDTH	122	R Age at Father's Death	53	Logical	Based on Y_CHAN,DATECH,SEE_DA	
AGEFMAR	124	R Age at Natural Father's Marriage	214	Logical	Based on FEMPAR,MALPAR,DATECH	
AGELSTED	109	Age Last Enrolled in Regular School	37	Logical	Based on HIGRADE	
AGELSTVC	110	Age Last in Vocational Training Program	16	Hot Deck	VOCTEC,AGECAT	
AGEMDTH	121	R Age at Mother's Death	26	Logical	Based on Y_CHAN,DATECH,SEE_MO	
AGEMMAR	123	R Age at Natural Mother's Marriage	334	Logical	Based on FEMPAR,MALPAR,DATECH	
AGEMOMB1	135	Age of Mother(Moth-Figure) at 1st Birth	228	Hot Deck	EDUCMOM,REGION	
AGEPARDS	120	R Age at Parents' 1st Sep/Divorce	102	Logical	Based on FEMPAR,MALPAR,DATECH	
AGER	101	Age at Interview	0	none	No imputations	
CMALONE	137	Date R 1st Lived Alone	16	Hot Deck	REGION,RACE,NUMLIVST,EDUCAT2,MAR_FLAG, BAB_FLAG,AGECAT	AGER<30
CMBGWK01	114.01	Date Began 1st Work (or 18th B-Day)	24	Hot Deck	WORKPDS,REGION,RACE,EDUCAT2,AGECAT	WORKPDS>0
CMBGWK02	114.02	Date Began 2nd Period of Work	43	Hot Deck	WORKPDS,REGION,RACE,EDUCAT2,AGECAT	WORKPDS>0
CMBGWK03	114.03	Date Began 3rd Period of Work	28	Hot Deck	WORKPDS,JOBON18,CHANGE,REGION,RACE, EDUCAT2,AGECAT	WORKPDS>0
CMBGWK04	114.04	Date Began 4th Period of Work	17	Hot Deck	WORKPDS,JOBON18,CHANGE,REGION,RACE, EDUCAT2,AGECAT	WORKPDS>0
CMBGWK05	114.05	Date Began 5th Period of Work	13	Hot Deck	Same donor as CMBGWK01	
CMBGWK06	114.06	Date Began 6th Period of Work	13	Hot Deck	Same donor as CMBGWK01	
CMBGWK07	114.07	Date Began 7th Period of Work	14	Hot Deck	WORKPDS,JOBON18,CHANGE,REGION,RACE, EDUCAT2,AGECAT	WORKPDS>0
CMBGWK08	114.08	Date Began 8th Period of Work	13	Hot Deck	Same donor as CMBGWK01	
CMBGWK09	114.09	Date Began 9th Period of Work	13	Hot Deck	Same donor as CMBGWK01	
CMBGWK10	114.10	Date Began 10th Period of Work	13	Hot Deck	Same donor as CMBGWK01	
CMCHFM01	129.01	Date Change 1st Parental Living Situatn	110	Hot Deck	NUMLIVST,FMARNO,RACE,EDUCAT2,AGECAT	
CMCHFM02	129.02	Date Change 2nd Parental Living Situatn	32	Hot Deck	NUMLIVST,FMARNO,RACE,EDUCAT2,AGECAT	
CMCHFM03	129.03	Date Change 3rd Parental Living Situatn	11	Hot Deck	NUMLIVST,FMARNO,RACE,EDUCAT2,AGECAT	
CMCHFM04	129.04	Date Change 4th Parental Living Situatn	5	Hot Deck	NUMLIVST,FMARNO,RACE,EDUCAT2,AGECAT	
CMCHFM05	129.05	Date Change 5th Parental Living Situatn	5	Hot Deck	NUMLIVST,FMARNO,RACE,AGECAT	
CMCHFM06	129.06	Date Change 6th Parental Living Situatn	3	Hot Deck	Same donor as CMCHFM01,CMCHFM05	
CMCHFM07	129.07	Date Change 7th Parental Living Situatn	2	Hot Deck	Same donor as CMCHFM01,CMCHFM05	
CMCHFM08	129.08	Date Change 8th Parental Living Situatn	0	none	No imputations	
CMCHFM09	129.09	Date Change 9th Parental Living Situatn	0	none	No imputations	
CMCHFM10	129.10	Date Change 10 Parental Living Situatn	0	none	No imputations	
CMCHFM11	129.11	Date Change 11 Parental Living Situatn	0	none	No imputations	
CMCHFM12	129.12	Date Change 12 Parental Living Situatn	0	none	No imputations	
CMENWK01	115.01	Date Ended 1st Period of Work	124	Hot Deck	WORKPDS,JOBON18,CHANGE,REGION,RACE, EDUCAT2,AGECAT	WORKPDS>0
CMENWK02	115.02	Date Ended 2nd Period of Work	36	Hot Deck	WORKPDS,JOBON18,CHANGE,REGION,RACE, EDUCAT2,AGECAT	WORKPDS>0
CMENWK03	115.03	Date Ended 3rd Period of Work	25	Hot Deck	WORKPDS,JOBON18,CHANGE,REGION,RACE, EDUCAT2,AGECAT	WORKPDS>0

Variable Name	Variable Number	Variable Labels	Number Imputed	Type of Imputation	Class Variables / General Notes	Screening Variables
Section B						
OUTCOM01-15	215	Outcome of 1st-15th Pregnancy	6	Hot Deck	See Interval Level Appendix	
PARITY	211	Total Number of Live Births	0	none	No imputations	
PREGNUM	203	Total Number of Pregnancies	0	none	No imputations	
RCURPREG	202	Pregnant at Time of Interview	3	Hot Deck	REGION,RACE,RMARITAL,AGER	INTENT=1
STLLBRTH	208	# Completed Pregs Ending in Stillbirth	0	Recalculated	Based on OUTCOME	
YRPREG01-15	217	Year of 1st-15th Pregnancy Outcome	233	Recalculated	See Interval Level Appendix	
Section C						
AGEDD1	320	Age at Divorce or Death, 1st Marriage	54	Recalculated	Based on MARDIS01,MAREND01	
AGEDISS1	319	Age at Dissolution of 1st Marriage	43	Recalculated	Based on MARDIS01,MAREND01	
COHAB1	328	Date of 1st Cohabitation	61	Hot Deck	COHSTAT,FMARNO,RMARITAL,RACE,EDUCAT2,AGECAT	COHEVER=1
COHEVER	327	Ever Cohabited (Outside of Marriage)	0	none	No imputations	
COHOUT	330	Outcome of First Cohabitation	30	Hot Deck	Same donor as COHOUT	
COHSTAT	329	Cohab Status Relative to 1st Marriage	58	Hot Deck	FMARNO,RMARITAL,RACE,EDUCAT2,AGECAT	COHEVER=1
CON1MAR1	326	Mos btw/1st Conception & 1st Marriage	56	Recalculated	Based on MARDAT01,DATCON01	
DATESEX1	339	Date of 1st (Voluntary) Sex After Menarc	114	Hot Deck	RAPE,COHEVER,MAR_FLAG,RACE,PREGCNT, AGECAT,MENARCHR	SEXEVER=1
DATEVOL1	338	Date of 1st Voluntary Sex	98	Hot Deck	RAPE,COHEVER,MAR_FLAG,RACE,PREGCNT, AGECAT,MENARCHR	SEXEVER=1
DD1REMAR	322	Mos btw Divorce/Death & Remarriage/Intv	47	Recalculated	Based on MARDIS01,MARDAT02	
EVVOLSEX	332	Ever Had Voluntary Sex	8	Hot Deck	RMARITAL,HADSEX,SEXEVER,RACE,AGER	
FMAR1AGE	318	Age at First Marriage	22	Recalculated	Based on MARDAT01	
FMARNO	302	Number Of Marriages	3	Hot Deck	RACE,RMARITAL,EDUCAT2,AGECAT,COMPREG	
HADSEX	331	Ever Had Sex at All	5	Hot Deck	Same donor as SEXEVER	
MAR1_NOW	323	Years Since First Marriage	22	Recalculated	Based on MARDAT01	
MAR1BIR1	324	Mos btw/1st Marriage & 1st Birth	29	Recalculated	Based on MARDAT01,BABY1MO	
MAR1CON1	325	Mos btw/1st Marriage & 1st Conceptn/Intv	49	Recalculated	Based on MARDAT01,DATCON01	
MAR1DISS	321	Mos btw/1st Marriage & Diss/Intv	57	Hot Deck	FMARNO,RMARITAL,RACE,EDUCAT2,AGER,COMPREG	
MARDAT01	303	Date of 1st Marriage	22	Hot Deck	FMARNO,RMARITAL,RACE,EDUCAT2,AGECAT	
MARDAT02	304	Date of 2nd Marriage	13	Hot Deck	FMARNO,RMARITAL,RACE,EDUCAT2,AGECAT	
MARDAT03	305	Date of 3rd Marriage	7	Hot Deck	FMARNO,RACE,EDUCAT2,AGECAT	
MARDAT04	306	Date of 4th Marriage	3	Hot Deck	Same donor as MARDAT01	
MARDAT05	307	Date of 5th Marriage	3	Hot Deck	Same donor as MARDAT01	
MARDIS01	308	Date of Dissolution of 1st Marriage	43	Hot Deck	FMARNO,MAREND01,RACE,EDUCAT2,FMAR1AGE	
MARDIS02	309	Date of Dissolution of 2nd Marriage	22	Hot Deck	FMARNO,MAREND02,RACE,EDUCAT2	
MARDIS03	310	Date of Dissolution of 3rd Marriage	10	Hot Deck	FMARNO,MAREND03,RACE,EDUCAT2,AGECAT	
MARDIS04	311	Date of Dissolution of 4th Marriage	3	Hot Deck	Same donor as MARDAT01	
MARDIS05	312	Date of Dissolution of 5th Marriage	3	Hot Deck	Same donor as MARDAT01	
MAREND01	313	How 1st Marriage Ended	6	Hot Deck	FMARNO,RMARITAL,RACE,EDUCAT,AGER	
MAREND02	314	How 2nd Marriage Ended	3	Hot Deck	Same donor as MARDAT01	
MAREND03	315	How 3rd Marriage Ended	3	Hot Deck	Same donor as MARDAT01	
MAREND04	316	How 4th Marriage Ended	3	Hot Deck	Same donor as MARDAT01	
MAREND05	317	How 5th Marriage Ended	3	Hot Deck	Same donor as MARDAT01	

Variable Name	Variable Number	Variable Labels	Number Imputed	Type of Imputation	Class Variables / General Notes	Screening Variables
Section E						
OLDWP01-15	426	Wantedness of 1-15 Preg-Partner (Old)	1050	Weighted Hot Deck	See text	PREGNUM>0
OLDWR01-15	427	Wantedness of 1-15 Preg-Respondent (Old)	1114	Weighted Hot Deck	See text	PREGNUM>0
PILLR	417	Ever Used Pills For Any Reason	3	Hot Deck	FMARITAL,RACE,HISPANIC,AGER	
SEX1MTHD	420	Birth Control Method Used at 1st Sex	65	Hot Deck	FMARITAL,RACE,HISPANIC,AGER	0<=SEX1MTHD<=20
SEXP3MO	410	Sex in the 3Mths Prior To Interview	306	Weighted Hot Deck	FMARITAL RACE RAGEAPR1 (See Text)	NOSEX36^=95
SOURCEM1	425	Source of Methd in Mos Befr Intv:1st Rep	13	Hot Deck	FMARITAL,RACE,HISPANIC,AGER	
SOURCEM2	425	Source of Methd in Mos Befr Intv:2nd Rep	3	Hot Deck	FMARITAL,RACE,HISPANIC,AGER	
SOURCEM3	425	Source of Methd in Mos Befr Intv:3rd Rep	0	none	No imputations	
SOURCEM4	425	Source of Methd in Mos Befr Intv:4th Rep	0	none	No imputations	
WANTP5	430	Number of Wanted Pregnancies In Last 5yr	0	none	No imputations	
WANTPT01-15	429	Wantedness of 1-15 Preg-Partner (New)	1092	Weighted Hot Deck	See text	PREGNUM>0
WANTRP01-15	428	Wantedness of 1-15 Preg-Respondent (New)	1111	Weighted Hot Deck	See text	PREGNUM>0
Section F						
FP1CHECK	502	Services 1st FP Visit: Chk-Up For BC	2	Hot Deck	RACE,RMARITAL,AGER	
FP1COUBC	503	Services 1st FP Visit: BC Counseling	2	Hot Deck	RACE,RMARITAL,AGER	
FP1COUST	504	Services 1st FP Visit: Sterility Counsel	1	Hot Deck	RACE,RMARITAL,AGER	
FP1MTHD	500	Services 1st FP Visit: BC Meth/Prescript	2	Hot Deck	RACE,RMARITAL,AGER	
FP1STER	501	Services 1st FP Visit: Sterilizing Oper	0	none	No imputations	
FPABOR12	514	Did R Have Abortion Last 12Mos	0	none	No imputations	
FPCHEC12	509	Did R Have Check-Up Last 12Mos	12	Hot Deck	Same donor as FPCHEC12,FPTITCHK,FPPAYCHK	FPCHEC12^=.
FPCLIEXM	549	Services at 1st Clin Visit: Medical Exam	6	Hot Deck	REGION,RMARITAL,RACE (Sorted by AGECAT)	
FPCLIFP	547	Services 1st Clin Visit: Family Planning	3	Hot Deck	REGION,RMARITAL,RACE (Sorted by AGECAT)	
FPCLIINF	550	Services 1st Clin Visit: Test/Trt Infect	6	Hot Deck	REGION,RMARITAL,RACE (Sorted by AGECAT)	
FPCLIOTH	551	Services 1st Clin Visit: Any Other Srv	4	Hot Deck	REGION,RMARITAL,RACE (Sorted by AGECAT)	
FPCLIPRE	548	Services 1st Clin Visit: Preg Test/Abort	4	Hot Deck	REGION,RMARITAL,RACE (Sorted by AGECAT)	
FPCNBC12	510	Did R Have BC Counseling Last 12Mos	11	Hot Deck	Same donor as FPCNBC12,FPTITCBC,FPPAYCBC	FPCNBC12^=.
FPCNST12	511	Did R Have Sterile Counseling Last 12Mos	5	Hot Deck	Same donor as FPCNST12,FPTITCST,FPPAYCST	FPCNST12^=.
FPTITMED	517	Type Clinic Used for Med Srv Last 12Mos	0	none	No imputations	
FPMOCLVT	552	Date 1st Clinic Visit After Menarche	0	none	No imputations	
FPMTHD12	508	Did R Have BC Meth/Presc Last 12Mos	42	Hot Deck	Same donor as FPMTHD12,FPTITBC,FPPAYBC	FPMTHD12^=.
FPPART12	516	Did R Have Post-Preg Care Last 12Mos	0	none	No imputations	
FPPAYABO	537	Pay Method(Last 12Mos): Abortion	0	none	No imputations	
FPPAYBC	532	Pay Method(Last 12Mos): BC Meth/presc	49	Hot Deck	RACE,RMARITAL,AGECAT	
FPPAYCBC	534	Pay Method(Last 12Mos): BC Counseling	15	Hot Deck	RACE,RMARITAL,AGECAT	
FPPAYCHK	533	Pay Method(Last 12Mos): Chk-Up for BC	15	Hot Deck	RACE,RMARITAL,AGECAT	
FPPAYCST	535	Pay Method(Last 12Mos): Steril Counsel	7	Hot Deck	Same donor as FPCNST12 (n=5)	

Variable Name	Variable Number	Variable Labels	Number Imputed	Type of Imputation	Class Variables / General Notes	Screening Variables
Section H						
ADVICE	607	Infertility Svcs: Advice	8	Logical	Based on ANYPRGHP	
ANYHIV	628	Ever Had an HIV Test	41	Hot Deck	REGION,RMARITAL,RACE (Sorted by AGECAT)	
ANYMSCHP	601	Any Medical Help to Prevent Miscarriage	0	none	No imputations	
ANYPRGHP	600	Any Medical Help to Become Pregnant	8	Hot Deck	REGION,RMARITAL,RACE (Sorted by AGECAT)	SEXEVER=1
BADSPERM	625	Diagnosis: Semen or Sperm Problem	8	Logical	Based on ANYPRGHP	
BEDREST	614	Miscarriage Svcs: Complete Bed Rest	7	Logical	Based on ANYMSCHP	
ENDOMET	610	Infertility Svcs: Endometriosis Surgery	8	Logical	Based on ANYPRGHP	
ENDOPROB	624	Diagnosis: Endometriosis	8	Logical	Based on ANYPRGHP	
FIBROIDS	611	Infertility Svcs: Fibroids Surgery	8	Logical	Based on ANYPRGHP	
INFERTH	606	Infertility Svcs: Testing on H/P	8	Logical	Based on ANYPRGHP	
INFERTR	605	Infertility Svcs: Testing on R	8	Logical	Based on ANYPRGHP	
INFEVER	602	Ever Used Infertility Services	8	Logical	Based on ANYPRGHP	
INFSRC	612	Source of Most Infertility Services	10	Hot Deck	REGION (Sorted by AGECAT)	SEXEVER=1, ANYPRGHP=1
INSEM	608	Infertility Svcs: Artificial Insemination	8	Logical	Based on ANYPRGHP	
INVITRO	609	Infertility Svcs: In Vitro Fertilization	8	Logical	Based on ANYPRGHP	
LIMACTIV	615	Miscarriage Svcs: Limit Physic. Activity	7	Logical	Based on ANYMSCHP	
MISCDRUG	617	Miscarriage Svcs: Drugs to Prevent	7	Logical	Based on ANYMSCHP	
MISCTEST	616	Miscarriage Svcs: Diagnostic Tests	7	Logical	Based on ANYMSCHP	
MSCSRC	619	Source of Most Miscarriage Services	0	none	No imputations	
OPROBINF	626	Diagnosis: Other Infertility Problems	8	Logical	Based on ANYPRGHP	
OVULATE	603	Infertility Svcs: Ovulation Drugs	8	Logical	Based on ANYPRGHP	
OVULPROB	621	Diagnosis: Problems with Ovulation	8	Logical	Based on ANYPRGHP	
PIDTREAT	627	Ever Been Treated for PID	3	Hot Deck	REGION,RMARITAL,RACE (Sorted by AGECAT)	
PRIVINSM	620	Private Insurance for Miscarriage Svcs	0	none	No imputations	
PRIVINSP	613	Private Insurance for Help Getting Preg	12	Hot Deck	REGION,RMARITAL,RACE (Sorted by AGECAT)	SEXEVER=1, ANYPRGHP=1
PURSTRNG	618	Miscarriage Svcs: Stitches in Cervix	7	Logical	Based on ANYMSCHP	
TUBEBLOK	622	Diagnosis: Blocked Tubes	8	Logical	Based on ANYPRGHP	
TUBES	604	Infertility Svcs: Blocked Tubes Surgery	8	Logical	Based on ANYPRGHP	
TUBLPELV	623	Diagnosis: Other Tubal or Pelvic Problem	8	Logical	Based on ANYPRGHP	
WHYCONDM	629	Reasons for Using Condoms	20	Hot Deck	REGION,RMARITAL,RACE (Sorted by AGECAT)	SEXEVER=1, CONDOMR=1
Section I						
BRTHPLCE	705	Geographic Region of R's Birthplace	48	Logical	Based on RSTATE	
GADSCOR	710	Score for Generalized Anxiety Disorder	55	Logical	Based on anxiety variables	
GENANX	709	R Ever Experienced Generalized Anxiety	55	Logical	Based on anxiety variables	
HISPANIC	707	Hispanic Origin	7	Logical	Based on HISPGRP,BACKBCKG,RACE_BNB	
LABORFOR	711	Labor Force Status	40	Hot Deck	REGION,RMARITAL,RACE (Sorted by AGECAT)	TOTINCR>0
METRO	701	1=MSA,Central City/2=MSA,Other/3=Not MSA	0	none	Geocoded variable	
POVERTY	712	Poverty Level of Income	1,251	Recalculated	Based on TOTINCR and/or NUMFMHH; See text	
RACE	708	Race	3	Logical	Based on HISPGRP,BACKBCKG,RACE_BNB	
REGION	703	Geographic Region of R's Residence	17	Logical	Based on NHIS region	

^= Not equal to.
>= Greater than or equal to.
R is Respondent.

NOTE: See National Survey of Family Growth file documentation for explanation of variable name.

Number of imputations and corresponding imputation specifications for the recoded variables from the pregnancy-interval file ordered alphabetically within questionnaire section: National Survey of Family Growth, Cycle 5

Variable Name	Variable Number	Variable Labels	Number Imputed	Type of Imputation	Class Variables	Screening Variables
Section B						
AGECON	233	Age at Conception	248	Recalculated	Based on DATEND, PRGLNGTH	
AGEPREG	229	Age at Pregnancy Outcome	233	Recalculated	Based on DATEND	
BFEEDWKS	245	Duration of Breast Feeding in Weeks	7	Hot Deck	RACE,EDUCAT2,AGEPREGC	OUTCOME=1
DATECON	232	Date of Conception	248	Recalculated	Based on DATEND, PRGLNGTH	
DATEND	227	Date Pregnancy Ended	233	Hot Deck	RACE,RMARITAL,OUTCOME	OUTCOME^=6
DELIVERY	236	Mode of Delivery	0	none	No imputations	
FMARCON2	235	Marital Status at Conception-2 Cats	235	Recalculated	Based on DATEND, Marriage dates	
FMARCON5	234	Formal Marital Status at Conception	380	Recalculated	Based on DATEND, Marriage dates	
FMAROUT2	231	Marital Status at Preg Outcome-2 Cats	229	Recalculated	Based on DATEND, Marriage dates	
FMAROUT5	230	Formal Marital Status at Preg Outcome	371	Recalculated	Based on DATEND, Marriage dates	
LBW1	241	Low Birthweight - Baby #1	12	Hot Deck	RACE,EDUCAT2,AGEPREGC	OUTCOME=1
LBW2	242	Low Birthweight - Baby #2	0	none	No imputations	
LEARNPRG	237	Weeks Pregnant When R Learned of Preg	5	Hot Deck	RACE,EDUCAT2,AGEPREGC	OUTCOME^=, YRPREG>=91
MATERNLV	246	Use of Maternity Leave	37	Hot Deck	RACE,EDUCAT2,AGEPREGC	OUTCOME=1, WORKBORN=2 OUTCOME=1, WORKBORN=7,8
OUTCOME	225	Pregnancy Outcome	6	Hot Deck	RACE,RMARITAL,NUM_PREG,AGER	2<=OUTCOME<=5
PAYDELIV	240	Payment for Delivery	35	Hot Deck	RACE,EDUCAT2,AGEPREGC	OUTCOME=1
PAYPNC	239	Payment for Prenatal Care	3	Hot Deck	RACE,EDUCAT2,AGEPREGC	OUTCOME^=6, YRPREG>=91, GETPRENA NE 2
PNCAREWK	238	Weeks Pregnant at 1st Prenatal Care	1	Hot Deck	RACE,EDUCAT2,AGEPREGC	OUTCOME^=(2,6), GETPRENA=1,YRPREG>=91
PRGLNGTH	226	Duration of Completed Pregnancy in Weeks	31	Hot Deck	RACE,RMARITAL,OUTCOME,AGEPREGC(4)	OUTCOME^=6
SEX1	243	Sex of Baby #1	3	Hot Deck	RACE,EDUCAT2,AGEPREGC	OUTCOME=1
SEX2	244	Sex of Baby #2	0	none	No imputations	
YRPREG	228	Year Pregnancy Ended	233	Recalculated	Based on DATEND	
Section E						
OLDWANTP	426	Preg Wantedness Cycle IV Version - Prtnr	1,050	Weighted Hot Deck	See text	
OLDWANTR	427	Preg Wantedness Cycle IV Version - Resp	1,114	Weighted Hot Deck	See text	
WANTPART	429	Wantedness of Pregnancy - Partner	1,092	Weighted Hot Deck	See text	
WANTRESP	428	Wantedness of Preg - Respondent	1,111	Weighted Hot Deck	See text	

^– Not equal to.
>= Greater than or equal to.

Appendix IV

Variance Estimation Using Taylor Series Approaches

The NSFG sample is obtained using a complex multistage sampling design with unequal selection probabilities and the clustering of respondents, so it is *not* based on a simple random sample. Accurate variance estimation must take into account the complexity of the sampling design. A variance estimate based on a simple random sample assumption (the estimate available from most statistical software packages) usually will NOT be accurate and will likely underestimate the actual sampling variance. For some data items (or populations), the naive variance estimate may be more inaccurate than for others. This appendix gives a more formal mathematical explanation of how Taylor series linearization approaches, such as those used in SUDAAN software, can be used. Appendix V shows an example SUDAAN program.

The Taylor series linearization method for variance estimation (18) is illustrated here for statistics that can be defined explicitly as functions of linear statistics estimated from the survey sample—including means, totals, proportions, ratios of the form $\Sigma wx/\Sigma wy$, and linear regression coefficients. A linearized variable, Z_i, is defined based on the Taylor series expansion of the function, and then substituted into the variance formula appropriate under the specified design for any linear statistic estimated from the sample.

The technique will be illustrated for a statistic which is a function of two linear statistics, although it extends to any number of linear statistics and to statistics that are vectors. Let $\hat{\theta}$ be an estimate of the population parameter θ, with $\hat{\theta} = F(X,Y)$ where X and Y are two linear sample statistics. Let $\mu_x = E(X)$ and $\mu_y = E(Y)$ where the expectation operator E denotes averaging over repeated sampling from the target population. $\hat{\theta}$ can be expanded, assuming usual regularity conditions, in a Taylor series about μ_x and μ_y, so that

$$\begin{aligned}\hat{\theta} = {} & F(\mu_x,\mu_y) + \partial F_x(\mu_x,\mu_y)(X-\mu_x) \\ & + \partial F_y(\mu_x,\mu_y)(Y-\mu_y) \\ & + \text{higher order terms}\end{aligned}$$

where the $\partial F_x(\mu_x,\mu_y)$ and $\partial F_y(\mu_x,\mu_y)$ functions are first-order partial derivatives of F with respect to X and Y evaluated at their respective expectations μ_x and μ_x. If the higher order terms are negligible, then

$$\begin{aligned}\text{Var}[\hat{\theta}] &\doteq E[\hat{\theta} - F(\mu_x,\mu_y)]^2 \\ &= \{(\partial F_x)^2 E(X-\mu_x)^2 + (\partial F_y)^2 E(Y-\mu_y)^2 \\ &\quad + 2(\partial F_x)(\partial F_y) E[(X-\mu_x)(Y-\mu_y)]\} \\ &= \{(\partial F_x)^2 \text{Var}(X) + (\partial F_y)^2 \text{Var}(Y) \\ &\quad + 2(\partial F_x)(\partial F_y) Cov(X,Y)\} \qquad (1)\end{aligned}$$

where

$$\partial F_x = \partial F_x(\mu_x,\mu_y) \quad \text{and} \quad \partial F_y = \partial F_y(\mu_x,\mu_y)$$

An equivalent computational procedure for producing the Taylor series variance estimate suggested by Woodruff (20) recognizes that the variable portion of the linearization in equation 1 is

$$Z = (\partial F_x)X + (\partial F_y)Y$$

and therefore,

$$\begin{aligned}\text{Var}[\hat{\theta}] &\doteq \text{Var}[(\partial F_x)X + (\partial F_y)Y] \\ &= \text{Var}(Z) \qquad (2)\end{aligned}$$

Noting that X and Y are linear statistics formed from the corresponding response variates x_i and y_i, measured on the ith sample unit, the variance approximation in equation 2 can be produced by substituting the linearized variable

$$Z_i = (\partial F_x)x_i + (\partial F_y)y_i$$

for x_i or y_i in the variance formula appropriate for computing Var(X) or Var(Y) under the specified sample design. To obtain a sample estimate for the Taylor series variance approximation, one replaces the population-evaluated derivative functions in Z_i with the corresponding sample analogies, i.e.

$$Z_i = [\partial F_x(X,Y)]x_i + [\partial F_y(X,Y)]y_i$$

Binder (28,29) proposed and justified using an implicit differentiation method for estimating the variance for a vector of survey statistics. Binder's results are particularly useful when the parameters are implicitly defined—such as for logistic regression coefficients and survival models.

Appendix V

Example SUDAAN Program Code and Output

The following example SUDAAN program uses two procedures—PROC DESCRIPT and PROC LOGISTIC—to analyze the NSFG data. (PROC LOGISTIC is referred to as PROC RLOGIST in SAS-Callable SUDAAN.) In this example, PROC DESCRIPT estimates the number, percentage, and associated standard errors of women 15–44 years of age who were currently using the oral contraceptive pill, by age group (AGECAT) and current religious affiliation (WRELIG).

PROC LOGISTIC (or RLOGIST) fits a logistic regression model of the effect of age (AGECAT), religious affiliation (WRELIG), and parity (PARITY) on the proportion of women who currently use the pill.

This example was run on a VAX computer, but the same program code (except for file names) should run on most PC or mainframe computers. The sample SUDAAN programs use two variables, COL_STR and PANEL, locations 12,347–12,349 in the NSFG respondent file) to identify strata and clusters for variance estimation. The use of the WR option in the sample SUDAAN program indicates that finite population correction factors can be omitted. There are four values of PANEL for each COL_STR.

Example SUDAAN Program

SUDAAN Program Code

1. Procedure Statements: Include DESIGN=WR

2. SUDAAN Design Parameters Statement
   NEST Statement:
   a. COL_STR NHIS collapsed strata
   b. PANEL NHIS national panel identifier (values=1-4)

   TOTCNT Statement No statement required

   WEIGHT Statement
   POST_WT Final NSFG analysis weight

3. Categorical variables:
   SUBGROUP statement includes classing variables and LEVEL statement identifies the number of valid levels (AGECAT values: 1-6, WRELIG values: 1-3)

4. Non-categorical variables:
   PARITY: Number of live births
   CONPILL: 1 if currently using the Pill (CONSTAT1=5), 0 otherwise

```
                              S U D A A N
          Software for the Statistical Analysis of Correlated Data
         Copyright      Research Triangle Institute       June 1996
                              Release 7.01

1   PROC DESCRIPT DATA="Y:\\DATA\\PILLGVAR" FILETYPE=SAS DESIGN=WR;

2   NEST COL_STR PANEL;

3   WEIGHT POST_WT;

4   VAR CONPILL;

5   CATLEVEL 1;

6   SUBGROUP AGECAT WRELIG;

7   LEVELS   6     3;

8   TABLES AGECAT WRELIG;

9   SETENV LINESIZE=100 PAGESIZE=60;

10  PRINT NSUM TOTAL PERCENT SEPERCENT
        / STYLE=NCHS TOTALFMT=F8.0 PERCENTFMT=F6.2 SEPERCENTFMT=F6.3;

11  TITLE "NSFG CYCLE V - Standard Error Table for Pill Use (DESIGN=WR)";

Number of observations read    :  10847    Weighted count : 60200604
Number of observations skipped :      0
(WEIGHT variable nonpositive)
Denominator degrees of freedom :    186
```

Date: 09-30-96 Research Triangle Institute Page : 1
Time: 09:22:31 The DESCRIPT Procedure Table : 1

NSFG CYCLE V - Standard Error Table for Pill Use (DESIGN=WR)
by: Variable, AGECAT.

Variable AGECAT	Sample Size	Total	Percent	SE Percent
CONPILL: 1				
Total	10847	10419124	17.31	0.428
15-19	1416	1186119	13.08	0.915
20-24	1519	2998231	33.48	1.522
25-29	1739	2620084	26.75	1.296
30-34	2148	2251128	20.52	1.034
35-39	2144	929504	8.23	0.612
40-44	1881	434057	4.29	0.567

NSFG CYCLE V - Standard Error Table for Pill Use (DESIGN=WR)
by: Variable, WRELIG.

Variable WRELIG	Sample Size	Total	Percent	SE Percent
CONPILL: 1				
Total	10847	10419124	17.31	0.428
White Protestants	3503	4266496	18.62	0.692
White Catholics	1802	2288672	19.34	0.961
Others	5542	3863956	15.18	0.530

```
PROC LOGISTIC DATA="Y:\\DATA\\PILLGVAR" FILETYPE=SAS DESIGN=WR;
NEST COL_STR PANEL;
WEIGHT POST_WT;
SUBGROUP AGECAT WRELIG;
LEVELS   6     3;
MODEL CONPILL = AGECAT WRELIG PARITY;
TITLE "NSFG CYCLE V - Logistic Regression Model for Pill Use (DESIGN=WR)";
```

```
Number of observations read         :  10847    Weighted count: 60200604
Number of observations skipped      :      0
(WEIGHT variable nonpositive)
Observations used in the analysis   :  10847    Weighted count: 60200604
Observations with missing values    :      0    Weighted count:        0
Denominator degrees of freedom      :    186

Number of non-zero responses: 1807
Number of zero responses    : 9040

LOGISTIC has converged in 4 iterations

Multiple R-Square for the dependent variable CONPILL: 0.089977

-2 * Normalized Log-Likelihood with Intercepts Only :  9994.99
-2 * Normalized Log-Likelihood Full Model           :  9010.93
Approximate Chi-Square (-2 * Log-L Ratio)           :   984.06
Degrees of Freedom                                  :        8
Approximate P-Value                                 :     0.00

Note: The approximate Chi-Square is not adjusted for clustering.
      Refer to hypothesis test table for adjusted test.
```

Date: 09-30-96 Research Triangle Institute Page : 1
Time: 09:22:31 The LOGISTIC Procedure Table : 1

For response variable CONPILL
NSFG CYCLE V - Logistic Regression Model for Pill Use (DESIGN=WR)

Independent Variables and Effects	Beta Coeff.	SE Beta	T-Test B=0	P-value T-Test B=0
Intercept	-2.84	0.15	-18.39	0.0000
AGECAT				
15-19	0.77	0.17	4.60	0.0000
20-24	2.11	0.16	13.39	0.0000
25-29	1.91	0.15	12.37	0.0000
30-34	1.69	0.16	10.61	0.0000
35-39	0.68	0.17	4.07	0.0001
40-44	0.00	0.00	.	.
WRELIG				
White Protestants	0.33	0.06	5.27	0.0000
White Catholics	0.36	0.08	4.49	0.0000
Others	0.00	0.00	.	.
PARITY	-0.29	0.03	-11.04	0.0000

For response variable CONPILL
NSFG CYCLE V - Logistic Regression Model for Pill Use (DESIGN=WR)

Contrast	Degrees of Freedom	Wald F	P-value Wald F
OVERALL MODEL	9	359.98	0.0000
MODEL MINUS INTERCEPT	8	78.46	0.0000
INTERCEPT	.	.	.
AGECAT	5	92.68	0.0000
WRELIG	2	17.21	0.0000
PARITY	1	121.90	0.0000

Date: 09-30-96 Research Triangle Institute Page : 3
Time: 09:22:31 The LOGISTIC Procedure Table : 1

For response variable CONPILL

NSFG CYCLE V - Logistic Regression Model for Pill Use (DESIGN=WR)

Independent Variables and Effects	Odds Ratio	Lower 95% Limit	Upper 95% Limit
Intercept	0.06	0.04	0.08
AGECAT			
15-19	2.16	1.55	3.01
20-24	8.28	6.06	11.30
25-29	6.79	5.00	9.21
30-34	5.40	3.95	7.39
35-39	1.98	1.42	2.76
40-44	1.00	1.00	1.00
WRELIG			
White Protestants	1.39	1.23	1.58
White Catholics	1.43	1.22	1.67
Others	1.00	1.00	1.00
PARITY	0.75	0.71	0.79

Appendix VI

How the Generalized Standard Error Estimates Were Made

Two formulas were used for the Generalized Standard Error (GSE) estimation procedures, one for the respondent data and one for the pregnancy interval data. Median design effects were used in the formulas instead of mean design effects because extreme values can distort measurements based on means. Median design effects for the respondent data, displayed in table IV, were based on the proportion of women:

1. Whose first menstrual period was before age 13
2. Who had had at least one completed pregnancy
3. Who had had at least one live birth
4. Who were fecund
5. Whose current contraceptive method was either the pill or a male condom
6. Who had ever used the pill
7. Who had ever used a male condom
8. Whose first method of contraception was either the pill or a male condom
9. Who intended to have additional children

Median design effects for the pregnancy-interval data, displayed in Table V, were based on the proportion of babies:

1. Who were not breastfed
2. Who were delivered vaginally
3. Whose prenatal care was paid for by the mother's personal income and/or private insurance

Table IV. Median design effects for nine respondent file variables, by race/ethnicity and demographic characteristics: 1995 National Survey of Family Growth, Cycle 5

Characteristic	Race/ethnicity: Hispanic	Race/ethnicity: Black non-Hispanic	Race/ethnicity: Other	Total
Total	1.36	1.67	1.30	1.46
Age				
15–17 years	1.21	1.24	1.01	1.10
18–19 years	1.31	1.43	1.08	1.14
20–24 years	1.27	1.52	1.23	1.39
25–29 years	1.44	1.36	1.26	1.47
30–34 years	1.16	1.37	1.16	1.30
35–39 years	1.37	1.47	1.09	1.19
40–44 years	1.29	1.51	1.05	1.17
Marital status				
Married	1.38	1.46	1.13	1.27
Wid/div/sep[1]	1.31	1.60	1.16	1.32
Never married	1.21	1.51	1.19	1.37
Education				
Less than high school	1.37	1.46	1.05	1.24
High school diploma	1.33	1.53	1.19	1.32
Some college	1.33	1.55	1.46	1.57
College graduate	1.43	1.39	1.15	1.24
Poverty level				
0–100%	1.45	1.67	1.26	1.53
101–200%	1.35	1.58	1.48	1.54
201–399%	1.23	1.56	1.14	1.22
400% or more	1.30	1.44	1.13	1.21
Metropolitan residence				
Metropolitan	1.41	1.61	1.41	1.55
Nonmetropolitan	1.19	1.68	1.11	1.20
Rural/urban residence in 1995				
Urban	1.37	1.60	1.42	1.54
Rural	1.06	1.44	0.98	1.05
Labor force status in 1995				
Full-time work	1.42	1.48	1.15	1.31
Part-time work	1.39	1.69	1.10	1.25
In school	0.97	1.28	1.02	1.10
Other	1.43	1.54	1.21	1.41

[1]Widowed/divorced/separated.

Table V. Median design effects for seven pregnancy-interval file variables, by race/ethnicity and demographic characteristics: 1995 National Survey of Family Growth, Cycle 5

Characteristic	Race/ethnicity			Total
	Hispanic	Black, non-Hispanic	Other	
Total	2.19	2.46	1.77	2.23
Age at outcome[1]				
15–17 years	1.54	1.66	1.13	1.46
18–19 years	1.45	1.43	1.17	1.39
20–24 years	1.69	2.20	1.47	1.76
25–29 years	1.66	1.81	1.38	1.56
30–34 years	1.54	1.43	1.19	1.31
35–39 years	1.21	1.54	1.40	1.33
40 years and over	0.95	1.39	1.25	1.26
Marital status at outcome				
Married	2.04	1.95	1.78	1.90
Wid/div/sep[2]	1.61	1.69	1.25	1.56
Never married	2.03	2.72	1.61	2.38
Education[3]				
Less than high school	2.51	2.92	1.74	2.57
High school diploma	2.10	2.44	1.95	2.20
Some college	2.25	1.98	1.72	1.85
College graduate or higher	1.68	1.53	1.61	1.58
Poverty level[3]				
0–100%	2.08	3.02	1.71	2.60
101–200%	1.94	2.22	1.83	2.00
201–400%	1.75	1.85	1.72	1.88
400% or more	1.71	2.32	1.52	1.66
Metropolitan residence[3]				
Metropolitan	2.27	2.60	1.61	2.12
Nonmetropolitan	1.27	1.87	2.21	2.35
Rural/urban residence[3]				
Urban	2.23	2.59	1.66	2.05
Rural	2.00	2.05	2.04	1.88
Labor force status[3]				
Full-time work	2.37	2.47	1.54	1.88
Part-time work	1.88	2.85	1.60	1.86
In school	1.65	1.87	1.37	1.70
Other	2.09	2.71	1.85	2.37

[1]Status at the end of the pregnancy interval.
[2]Widowed/divorced/separated.
[3]Measured at time of interview.

4. Whose delivery was paid for through personal income and/or private insurance Three additional variables include the proportion of *pregnancies* in which:
5. The length of the pregnancy exceeded 38 weeks
6. The first baby was male
7. The outcome was a live birth

The direct estimates of the sampling variances were computed for each of these outcomes using the ''with-replacement'' variance estimator (that is, DESIGN = WR) in the SUDAAN procedure DESCRIPT. The parameter estimates and the direct variance estimates were computed for the respondent data file (one record for each of the 10,847 responding women) and for the pregnancy-interval database (one or more records for each of the 7,761 responding women with at least one pregnancy). For each data file, the parameter estimates and their sampling variances were computed for all women and for each of the three race/ethnicity categories (Hispanic, non-Hispanic black, and other women).

Generalized standard errors (GSE's) were obtained from a prediction equation involving the design effect (deff) estimates. The model was initially based on the design effect for an estimated proportion. The resulting prediction equation was based on the following log (base 10) linear relationship between the design effect (deff), the proportion p, and the sample size n:

$$\log(\text{deff}) = \beta_0 + \beta_1 \log(P) + \beta_2 \log(1 - P) + \beta_3 \log(n) \quad (1)$$

where

$\beta_0, \beta_1, \beta_2, \beta_3$ = regression coefficients for the intercept, $\log(P)$, $\log(1-P)$, and $\log(n)$, respectively.

Separate models were fit for the respondent and pregnancy-interval data within three race/ethnicity categories: Hispanic, non-Hispanic black, and overall. By substituting the fitted model in equation 1 back into the definition of the design effect, a prediction equation for the GSE is

$$\text{GSE}_{ij}(P) = \frac{10^{(b_{0ij})/2} \bullet P_{ij}^{(1+b_{0ij})/2} \bullet (1-P_{ij})^{(1+b_{2ij})/2}}{n_{ij}^{(1-b_{3ij})/2}}$$

$$i = 1,2 \qquad j = 1,2 \qquad (2)$$

where

b_0, b_1, b_2, b_3 = estimated regression coefficients for the intercept, $\log(P)$, $\log(1-P)$, and $\log(n)$, respectively.

The i-index depicts whether the standard error approximation is for a respondent proportion or a pregnancy-interval proportion. The j-index identifies the three race/ethnicity categories.

Vital and Health Statistics series descriptions

SERIES 1. **Programs and Collection Procedures**—These reports describe the data collection programs of the National Center for Health Statistics. They include descriptions of the methods used to collect and process the data, definitions, and other material necessary for understanding the data.

SERIES 2. **Data Evaluation and Methods Research**—These reports are studies of new statistical methods and include analytical techniques, objective evaluations of reliability of collected data, and contributions to statistical theory. These studies also include experimental tests of new survey methods and comparisons of U.S. methodology with those of other countries.

SERIES 3. **Analytical and Epidemiological Studies**—These reports present analytical or interpretive studies based on vital and health statistics. These reports carry the analyses further than the expository types of reports in the other series.

SERIES 4. **Documents and Committee Reports**—These are final reports of major committees concerned with vital and health statistics and documents such as recommended model vital registration laws and revised birth and death certificates.

SERIES 5. **International Vital and Health Statistics Reports**—These reports are analytical or descriptive reports that compare U.S. vital and health statistics with those of other countries or present other international data of relevance to the health statistics system of the United States.

SERIES 6. **Cognition and Survey Measurement**—These reports are from the National Laboratory for Collaborative Research in Cognition and Survey Measurement. They use methods of cognitive science to design, evaluate, and test survey instruments.

SERIES 10. **Data From the National Health Interview Survey**—These reports contain statistics on illness; unintentional injuries; disability; use of hospital, medical, and other health services; and a wide range of special current health topics covering many aspects of health behaviors, health status, and health care utilization. They are based on data collected in a continuing national household interview survey.

SERIES 11. **Data From the National Health Examination Survey, the National Health and Nutrition Examination Surveys, and the Hispanic Health and Nutrition Examination Survey**—Data from direct examination, testing, and measurement on representative samples of the civilian noninstitutionalized population provide the basis for (1) medically defined total prevalence of specific diseases or conditions in the United States and the distributions of the population with respect to physical, physiological, and psychological characteristics, and (2) analyses of trends and relationships among various measurements and between survey periods.

SERIES 12. **Data From the Institutionalized Population Surveys**—Discontinued in 1975. Reports from these surveys are included in Series 13.

SERIES 13. **Data From the National Health Care Survey**—These reports contain statistics on health resources and the public's use of health care resources including ambulatory, hospital, and long-term care services based on data collected directly from health care providers and provider records.

SERIES 14. **Data on Health Resources: Manpower and Facilities**—Discontinued in 1990. Reports on the numbers, geographic distribution, and characteristics of health resources are now included in Series 13.

SERIES 15. **Data From Special Surveys**—These reports contain statistics on health and health-related topics collected in special surveys that are not part of the continuing data systems of the National Center for Health Statistics.

SERIES 16. **Compilations of Advance Data From Vital and Health Statistics**—Advance Data Reports provide early release of information from the National Center for Health Statistics' health and demographic surveys. They are compiled in the order in which they are published. Some of these releases may be followed by detailed reports in Series 10–13.

SERIES 20. **Data on Mortality**—These reports contain statistics on mortality that are not included in regular, annual, or monthly reports. Special analyses by cause of death, age, other demographic variables, and geographic and trend analyses are included.

SERIES 21. **Data on Natality, Marriage, and Divorce**—These reports contain statistics on natality, marriage, and divorce that are not included in regular, annual, or monthly reports. Special analyses by health and demographic variables and geographic and trend analyses are included.

SERIES 22. **Data From the National Mortality and Natality Surveys**—Discontinued in 1975. Reports from these sample surveys, based on vital records, are now published in Series 20 or 21.

SERIES 23. **Data From the National Survey of Family Growth**—These reports contain statistics on factors that affect birth rates, including contraception, infertility, cohabitation, marriage, divorce, and remarriage; adoption; use of medical care for family planning and infertility; and related maternal and infant health topics. These statistics are based on national surveys of women of childbearing age.

SERIES 24. **Compilations of Data on Natality, Mortality, Marriage, Divorce, and Induced Terminations of Pregnancy**—These include advance reports of births, deaths, marriages, and divorces based on final data from the National Vital Statistics System that were published as supplements to the *Monthly Vital Statistics Report* (MVSR). These reports provide highlights and summaries of detailed data subsequently published in *Vital Statistics of the United States*. Other supplements to the MVSR published here provide selected findings based on final data from the National Vital Statistics System and may be followed by detailed reports in Series 20 or 21.

For answers to questions about this report or for a list of reports published in these series, contact:

Data Dissemination Branch
National Center for Health Statistics
Centers for Disease Control and Prevention
6525 Belcrest Road, Room 1064
Hyattsville, MD 20782-2003

(301) 436–8500
E-mail: nchsquery@cdc.gov
Internet: www.cdc.gov/nchswww

DEPARTMENT OF
HEALTH & HUMAN SERVICES

Centers for Disease Control and Prevention
National Center for Health Statistics
6525 Belcrest Road
Hyattsville, Maryland 20782-2003

STANDARD MAIL (A)
POSTAGE & FEES PAID
PHS/NCHS
PERMIT NO. G-281

DHHS Publication No. (PHS) 98-1398, Series 2, No. 124
7-0762 (2/98)