U.S. Department of Education
Institute of Education Sciences
NCES 2006–179

2004 National Study of Postsecondary Faculty (NSOPF:04)

Methodology Report

Technical Report

U.S. Department of Education
Institute of Education Sciences
NCES 2006–179

2004 National Study of Postsecondary Faculty (NSOPF:04)

Methodology Report

Technical Report

May 2006

Ruth Heuer
Brian Kuhr
Mansour Fahimi
T.R. Curtin
Marjorie Hinsdale
Lisa Carley-Baxter
Pat Green
RTI International

Linda Zimbler
Project Officer
National Center for
Education Statistics

U.S. Department of Education
Margaret Spellings
Secretary

Institute of Education Sciences
Grover J. Whitehurst
Director

National Center for Education Statistics
Mark Schneider
Commissioner

The National Center for Education Statistics (NCES) is the primary federal entity for collecting, analyzing, and reporting data related to education in the United States and other nations. It fulfills a congressional mandate to collect, collate, analyze, and report full and complete statistics on the condition of education in the United States; conduct and publish reports and specialized analyses of the meaning and significance of such statistics; assist state and local education agencies in improving their statistical systems; and review and report on education activities in foreign countries.

NCES activities are designed to address high priority education data needs; provide consistent, reliable, complete, and accurate indicators of education status and trends; and report timely, useful, and high quality data to the U.S. Department of Education, the Congress, the states, other education policymakers, practitioners, data users, and the general public.

We strive to make our products available in a variety of formats and in language that is appropriate to a variety of audiences. You, as our customer, are the best judge of our success in communicating information effectively. If you have any comments or suggestions about this or any other NCES product or report, we would like to hear from you. Please direct your comments to:

National Center for Education Statistics
Institute of Education Sciences
U.S. Department of Education
1990 K Street NW
Washington, DC 20006–5651

May 2006

The NCES World Wide Web Home Page address is http://nces.ed.gov
The NCES World Wide Web Electronic Catalog is: http://nces.ed.gov/pubsearch

Suggested Citation
Heuer, R., Kuhr, B., Fahimi, M., Curtin, T.R., Hinsdale, M., Carley-Baxter, L., and Green, P. (2005). *2004 National Study of Postsecondary Faculty (NSOPF:04) Methodology Report* (NCES 2006-179). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

For ordering information on this report, write:
U.S. Department of Education
ED Pubs
P.O. Box 1398
Jessup, MD 20794–1398

Or call toll free 1–877–4ED–Pubs

Content Contact:
Aurora D'Amico
(202) 502–7334
aurora.d'amico@ed.gov

Executive Summary

The 2004 National Study of Postsecondary Faculty (NSOPF:04), conducted by RTI International (RTI) and sponsored by the U.S. Department of Education's National Center for Education Statistics (NCES), is a nationally representative study that collects data regarding the characteristics, workload, and career paths of full- and part-time postsecondary faculty and instructional staff at public and private not-for-profit 2- and 4-year institutions in the United States. Conducted previously in 1988, 1993, and 1999, it serves a continuing need for data on faculty and instructional staff.

For the first time, NSOPF:04 is being conducted as a component study of the 2004 National Study of Faculty and Students (NSoFaS:04). The student component—the 2004 National Postsecondary Student Aid Study (NPSAS:04)—is a nationally representative study of students enrolled in all levels of postsecondary education. Historically, there has been considerable overlap in the institutions selected for participation in NSOPF and NPSAS; therefore, institution sampling and contacting activities for both studies were coordinated to help minimize response burden on institutions and to improve data collection efficiency.

This report describes the methodology and findings of NSOPF:04, which took place during the 2003–04 academic year. A field test, conducted in the 2002–03 academic year, was used to plan, implement, and evaluate methodological procedures, instruments, and systems proposed for use in the full-scale study. The 2004 National Study of Postsecondary Faculty Field Test Methodology Report (Heuer et al. 2004) is available from NCES.

This methodology report is designed to report solely for NSOPF:04. NPSAS:04 procedures and results—provided in a separate report—are discussed here only as they impact or overlap with those outlined for NSOPF:04.

Target Population and Sample Design

The NSOPF:04 sample consists of postsecondary institutions and their full- and part-time faculty and instructional staff. The sampled institutions represent all public and private not-for-profit Title IV-participating, degree-granting institutions in the 50 states and the District of Columbia, as reported in the *2002 Integrated Postsecondary Education Data System (IPEDS*) data files. Stratified, systematic samples of institutions and faculty were designed to allow detailed comparisons and high levels of precision. A customized cost/variance optimization program was implemented to efficiently secure targeted levels of precision for key estimates.

A two-stage sampling methodology was utilized. In the first stage, the institution sample was drawn based on a probability proportional to size (PPS) selection methodology, where each institution was assigned a composite measure of size (MOS) that reflected the number of eligible faculty and instructional staff in each of six strata. A sample of 1,080 postsecondary institutions was selected for participation; 1,070[*] of these were eligible. Each institution was asked to provide a list of all of the full- and part-time faculty and instructional staff that the institution employed during the fall 2003 term. Institutions were asked to include all employees with faculty

[*] Throughout this report, faculty and institution counts are rounded to the nearest 10 to protect the confidentiality of faculty and institutions. However, percentages cited are based on the original unrounded numbers.

status (both instructional and non-instructional) and all others with instructional responsibilities, regardless of faculty status. A total of 980 institutions provided a list suitable for sampling.

In the second stage of sampling, full- and part-time faculty and instructional staff employed by participating institutions as of November 1, 2003 were selected. Sampling was conducted on a flow basis, as lists were received, checked for accuracy, and processed. A total of 35,630 faculty were sampled from participating institutions. Of these, 34,330 were eligible.

Instrumentation

The NSOPF:04 *institution* questionnaire was designed to be self-administered via the Internet; the NSoFaS:04 website for institutional participation provided secure access to the questionnaire and information about each component of the study. To expedite completion, it could also be administered as a computer-assisted telephone interview (CATI), if necessary. The instrument was divided into major sections that collected information on the number of faculty and instructional staff employed at the target institution, the policies and practices that affected full-time faculty and instructional staff, the policies and practices that affected part-time faculty and instructional staff, and the percentage of undergraduate instruction assigned to various instructional personnel.

The NSOPF:04 *faculty* instrument was also designed as a web-based instrument for self-administration via the Internet and by CATI for nonresponse follow-up. The faculty website, like the institution website, provided secure access to the self-administered questionnaire as well as additional information about the study.

Both instruments were designed to accommodate the mixed-mode data collection approach and to ensure the collection of high-quality data. Design considerations included appropriate question wording for both self-administered and telephone interviews, and checks for out-of-range or inconsistent values. The *faculty* instrument consisted of the following eight sections grouped by topic:

- employment during the fall 2003 term (including academic rank, tenure status, and field of teaching);
- academic and professional background (including highest degree earned and employment history);
- institutional responsibilities and workload (including instructional activities and other work responsibilities performed in a typical week);
- scholarly activities (including productivity, funding of scholarly activities, and field of research);
- job satisfaction and retirement plans;
- monetary compensation (including income from the institution and other sources, structure of the employment contract, and household income);
- sociodemographic information (including gender, race, date of birth, marital status, number of dependent children, and citizenship); and
- opinions about working conditions at the institution.

Institution Contacting

Sampled institutions were contacted by mail, e-mail, and telephone beginning in spring 2003 to allow institutions sufficient time to plan for the study and to resolve any potential roadblocks to participation. Institution contacts were designed to verify institutional eligibility, secure timely participation in each survey component, and identify a staff person at each institution—called the Institution Coordinator—to respond to all NSoFaS:04 data requests. The Institution Coordinator was mailed an introductory letter and accompanying information packet, and then contacted by telephone to confirm the institution's intent and ability to participate within schedule constraints. At this time, each coordinator was asked to complete a *Coordinator Response Form* that confirmed the data items requested for each component of NSoFaS:04 and the projected deadlines for completion of the study. Upon request, project staff prepared additional information packets for Institutional Review Boards (IRBs) and other deliberative bodies within institutions to secure the institution's participation.

Beginning in fall 2003, each Institution Coordinator was mailed a binder containing complete specifications for participation. Institution Coordinators were asked to provide electronic lists of all eligible faculty and instructional staff on November 1, 2003, and to complete the institution questionnaire by December 6, 2003. Follow-up activities continued with the Institution Coordinator until all requested data was supplied.

Of the 1,070 eligible institutions, 980 (91 percent unweighted and weighted) provided faculty lists, and 920 (86 percent unweighted; 84 percent weighted) completed the institution questionnaire.

Help Desk and Interviewer Training

Training programs were developed for help desk operators who would respond to questions of sample members attempting to complete the web-based survey and for telephone interviewers who would conduct the nonresponse follow-up. Help desk operators received specific training in "frequently asked questions" regarding the instrument and technical issues related to completion of the self-administered questionnaire via the Internet. In addition, help desk operators received the same training as telephone interviewers because they were expected to complete the instrument over the telephone if requested by a caller. The telephone interviewer training focused on techniques for successfully locating and interviewing sample members, and covered such topics as administrative procedures required for case management, quality control of interactions with sample members and other contacts, and the organization and operation of the web-based faculty instrument to be used in data collection.

Faculty Locating and Survey Completion

NSOPF:04 data collection procedures were designed to locate sample members, encourage prompt completion of the self-administered questionnaire via the Internet, and conduct telephone interviews with nonrespondents.

Upon receipt of faculty lists, contact information for the sampled faculty and instructional staff was reviewed and assessed for completeness. Incomplete information was supplemented by searches of the institution's website for telephone and address information. Intensive tracing was performed when all telephone numbers for a respondent were exhausted.

Faculty data collection utilized a mixed-mode approach; sample members could participate either by web-based self-administered questionnaire or by an interviewer-administered telephone interview. The participation of sample members was initially requested in a letter, which provided both instructions for completing the web questionnaire and completing the interview via CATI. Periodic reminder letters and e-mail messages were sent to nonrespondents to encourage their participation.

After 4 weeks, interviewers began calling the sample members directly to attempt a CATI interview. An early-response incentive was provided to encourage prompt completion of the instrument. Incentives were also offered to sample members who refused or were unresponsive.

Of the 34,330 eligible sample members, 26,110 (76 percent, unweighted and weighted) completed the faculty questionnaire during a field period from January to October of 2004. Seventy-six percent of respondents completed the self-administered web questionnaire, and 24 percent were interviewed by telephone. The average time to complete the survey was 30 minutes.

Evaluation of Operations and Data Quality

Evaluations of operations and procedures focused on the joint institution contacting endeavor, the timeline for data collection from institutions (faculty lists and institution questionnaires) and faculty (CATI and self-administered interviews), tracing and locating procedures, refusal conversion efforts, the effectiveness of incentives, and the length of the faculty interview.

Results of the data quality evaluations included the following:

- Eighty-two percent of faculty list counts were within 10 percent of the corresponding institution questionnaire counts. There were greater variances between list counts and IPEDS, which is based on a narrower definition of faculty. Patterns of discrepancies between IPEDS and list data followed expected patterns, with list counts larger than those from IPEDS.
- Item nonresponse was below 15 percent for 87 of the 90 items in the institution questionnaire and for 141 out of the 162 items in the faculty questionnaire.
- Of the 26,550 eligible sample members who started the interview, 570 (2 percent) broke off before completing the interview. Of these, 430 broke off before completing the workload section and were not considered to be partial completes. Of the 140 partial completes, 48 percent broke off in the scholarly activities section; 9 percent broke off in the job satisfaction section; 29 percent in the compensation section; 11 percent in the characteristics section; and 4 percent in the opinions section.
- A new assisted coding system, used to code field of teaching, highest degree field, and principal field of scholarly activity, coded 77 percent of verbatim strings; 23 percent of strings required manual coding.
- A recoding of 10 percent of teaching, research, and highest degree verbatim strings showed 71 percent were coded correctly, 13 percent incorrectly, and the remaining 15 percent were too vague to code. The coding performed by web respondents was more often accepted as correctly coded than that done by CATI interviewers.

- Of the approximately 25,760 postsecondary institutions coded in the faculty instrument, 1,130 (4 percent) were initially deemed uncodeable. Based on the institution information collected, however, 1,030 of these institutions were positively identified and recoded.

NSOPF:04 Data Files and Products

NSOPF:04 data can be accessed both through the NCES Data Analysis System (DAS) for public use and through electronically documented, restricted access data files (with associated Electronic Codebooks). The public-use DAS may be accessed on the NCES website at http://nces.ed.gov/das/.

Using DAS, researchers are able to

- create their own analysis tables;
- view the highlights of report findings, with figures and tables, for various postsecondary topics;
- see a comprehensive listing of analyses regarding postsecondary education and download the reports; and
- view and download DAS table parameter files (TPFs) used to generate report tables.

An ongoing series of descriptive statistical reports may be accessed online or ordered through NCES as they are released. Descriptive reports focus on topics of interest, such as undergraduate teaching, teaching with technology, distance education instruction, gender and racial/ethnic composition of the faculty population, tenure status, work activities and compensation, and characteristics of part-time faculty. Publications available for public use may be downloaded or ordered at http://nces.ed.gov/pubsearch/getpubcats.asp?sid=011.

Foreword

This report describes the methods and procedures used for the data collection effort of the 2004 National Study of Postsecondary Faculty (NSOPF:04). NSOPF:04 serves a continuing need for data on faculty and instructional staff, all of whom directly affect the quality of education in postsecondary institutions.

We hope that the information provided here will be useful to a wide range of interested readers and that the results reported in the forthcoming descriptive summary report will encourage others to use the NSOPF:04 data. We welcome recommendations for improving the format, content, and approach, so that future methodology reports will be more informative and useful.

C. Dennis Carroll
Associate Commissioner
Postsecondary Studies Division

Acknowledgments

The authors gratefully acknowledge the assistance of staff members of the National Center for Education Statistics (NCES) and the Institute of Education Sciences (IES) for their advice, guidance, and review in the design and conduct of the study, and in the preparation of this document. We are especially grateful to Linda Zimbler, the NSOPF:04 Project Officer; James Griffith, Program Director, Postsecondary Longitudinal and Sample Survey Studies (PLSSS); Tracy Hunt-White, Statistician (PLSSS), and C. Dennis Carroll, Associate Commissioner, Postsecondary Studies Division (PSD), for their constructive input and review.

Particular thanks are extended as well to the Technical Review Panel members, who provided considerable insight and guidance in the development of the design and instrumentation of this study. For their careful review of the document, we also extend particular thanks to Ralph Lee, NCES Office of the Deputy Commissioner (ODC); Professor James Palmer, Illinois State University; Michael Cohen; Bureau of Transportation Statistics; and Marilyn Seastrom, Chief Statistician and Program Director, NCES Statistical Standards Program (ODC).

We also extend our thanks to the project staff members of contractors RTI International, MPR Associates, Inc. and Pinkerton Computer Consultants, Inc. A number of staff from these organizations—including statisticians, analysts, survey managers, programmers, data collectors, and interviewers, too numerous to name here—worked long hours on this study. At RTI, we are especially indebted to Sallie Fiore, Sharon Powell, Lynne Hawley, and Craig R. Hollingsworth for their excellent and tireless efforts in preparing the drafts and final version of this report.

Most of all, we are indebted to the many Chief Administrators, Institution Coordinators, institution respondents, and faculty and instructional staff members who participated in the NSOPF:04 study. Their willingness to take the time to share information made this study a success.

Contents

Table **Page**

Figure **Page**

Chapter 1
Overview of NSOPF:04

This document describes the study design, procedures, and outcomes for the 2004 National Study of Postsecondary Faculty (NSOPF:04), which was conducted for the National Center for Education Statistics (NCES) of the U.S. Department of Education, Washington, DC, as authorized by Title I, Section 153, of the Education Sciences Reform Act of 2002 [PL 107-279]. For the 2004 cycle, NSOPF:04 was conducted as a component study of the 2004 National Study of Faculty and Students (NSoFaS:04) under contract by RTI International,[1] with the assistance of MPR Associates, Inc., and Pinkerton Computer Consultants, Inc. Results for the student component, the 2004 National Postsecondary Student Aid Study (NPSAS:04), are provided in a separate methodology report (Cominole et al.).

This introductory chapter provides an overview of NSOPF:04, including a description of the background and purpose of the study, the types of policy-relevant issues addressed, the changes to the study from previous cycles, the data and reports generated from the study, and the schedule of data collection activities.

1.1 Background and Purpose of NSOPF:04

NSOPF:04 was a comprehensive nationwide study of the characteristics, workload, and career paths of postsecondary faculty and instructional staff.[2] The study was based on a nationally representative sample of all full- and part-time faculty and instructional staff at public and private not-for-profit 2- and 4-year degree-granting institutions in the United States. The NSOPF:04 full-scale sample consisted of 35,630 faculty and instructional staff selected from 980 sampled institutions in the 50 states and the District of Columbia.[3]

NSOPF:04 comprises the fourth cycle of the National Study of Postsecondary Faculty. Previous studies, conducted in 1988, 1993, and 1999 (called NSOPF:88, NSOPF:93, and NSOPF:99, respectively), provided national profiles of faculty and instructional staff in postsecondary institutions, national benchmarks for faculty productivity and workload, and information on institutional policies and practices that affect faculty. The fourth cycle of the National Study of Postsecondary Faculty, NSOPF:04, expanded the information about faculty and instructional staff in two ways: (1) it allowed for comparisons to be made over an extended period of time, and (2) it helped examine emerging issues concerning faculty, such as changes related to increased use of the Internet and distance education.

NSOPF:04 was designed to address a variety of policy-relevant issues concerning faculty, instructional staff, and postsecondary institutions. The study included faculty and institution questionnaires covering general policies concerning faculty. Information obtained from these two sources helped address important questions about postsecondary education, such as the following:

[1] RTI International is a trade name of Research Triangle Institute.
[2] References to "faculty" in this report include instructional staff and others (e.g., administrators) with faculty status (who may or may not have instructional duties).
[3] Throughout this report, faculty and institution counts are rounded to the nearest 10 to protect the confidentiality of faculty and institutions. However, percentages cited are based on the original unrounded numbers.

- What are the background characteristics of full- and part-time faculty?
- What are their workloads and how is their time allocated between classroom instruction and other activities?
- What are the current teaching practices and uses of technology among postsecondary faculty and instructional staff?
- How satisfied are they with current working conditions and institutional policies?
- How are faculty and instructional staff compensated by their institutions? How important are other sources of income?
- What are the career and retirement plans of faculty and instructional staff?
- What retirement packages are available to faculty and instructional staff?
- Have institutions changed their policies on granting tenure to faculty members? Are changes anticipated in the future?

1.2 Methodological Issues and Changes for NSOPF:04

1.2.1 Combining NSOPF and NPSAS

NSOPF:04 was, in one respect, unlike any previous cycle of NSOPF, as it was conducted in tandem with another major study, NPSAS:04, under one overarching contract: NSoFaS:04. NCES recognized that, historically, there has been considerable overlap in the institutions selected for participation in NSOPF:04 and NPSAS:04. By combining the two independent studies under one contract, NCES sought to minimize the response burden on institutions and to realize data collection efficiencies. The NSOPF:04 and NPSAS:04 studies retain their separate identities. The purpose of this report is to summarize the methodology of NSOPF:04; sampling and data collection procedures for NPSAS:04 are referred to only as they are combined with, or impact, the parallel procedures for NSOPF:04.

The combination of NSOPF:04 and NPSAS:04 into NSoFaS:04 had important implications for the NSOPF:04 institution sample design and institution contacting procedures. Institutions for the NSOPF:04 sample were selected as a subsample of the NPSAS:04 sample institutions.[4] This combination resulted in a somewhat larger sample of institutions for the full-scale study than previous NSOPF cycles (1,070 eligible institutions compared to 960 in 1999) and created a need to balance the design requirements of both studies in all institution-related study procedures.

1.2.2 Institution Sampling and List Collection

Apart from the changes necessitated by combining NSOPF:04 and NPSAS:04, as noted above, the key change in sampling procedures for NSOPF:04 was its use of a customized cost/variance optimization technique. This procedure was designed to identify the allocation that would accommodate all analytical objectives of this survey while minimizing data collection

[4] The larger NPSAS sample includes about 400 schools not eligible for NSOPF, including less-than-two-year and proprietary schools, and schools located in Puerto Rico. It also includes about 140 institutions that were NSOPF-eligible but not included in the sample because the precision requirements for NSOPF could be met without their inclusion.

costs. As with the institution-level sampling, a customized cost/variance optimization technique was used to determine the optimal allocation of faculty to the sampling strata.

In previous cycles, delays in receiving faculty lists created critical delays in sampling and contacting respondents during the time optimal to reach them (i.e., prior to the close of the regular academic year). Because the perceived burden of NSoFaS:04 would likely be greater than that of the individual studies by themselves, an advance notification and early contacting strategy was developed for this cycle. The purpose of advance notification and early contacting was to provide sufficient time to resolve any roadblocks to participation, allow the Institution Coordinator sufficient time to plan staffing and resources for the study, and to allow sufficient time for the completion of any review process the institution required, thereby facilitating the finish of data collection prior to the deadline.

For faculty list collection, procedures were developed that would encourage institutions to provide lists of faculty and complete related documentation (including the institution questionnaire) online. On the NSoFaS:04 website, a secure tool for uploading lists was provided to eliminate the need for institutions to send data files through conventional mail.

The institution questionnaire was designed as a single integrated web/computer-assisted telephone interview (CATI) instrument; there was no hardcopy instrument, although a facsimile was provided to allow dissemination of questions to different departments.

Table 1 summarizes the data collection schedule for the full-scale study.

Table 1. Schedule of major NSOPF:04 data collection activities: 2004

Activity	Start date[1]	End date[2]
Select institution sample	May 22, 2002	August 25, 2002
Institutional recruitment/early contacting of institution coordinators[3]	March 10, 2003	September 29, 2003
Obtain faculty lists[4]	September 29, 2003	July 11, 2004
Implement institution questionnaire	September 29, 2003	October 22, 2004
Select faculty samples	November 6, 2003	July 12, 2004
Send mail and e-mail to faculty	January 15, 2004	October 1, 2004
Implement faculty web questionnaire	January 15, 2004	October 6, 2004
Implement faculty CATI interviewing	February 12, 2004	October 5, 2004

[1] This is the date on which the activity was initiated for the first applicable institution and/or its associated faculty.
[2] This is the date on which the activity was completed for the last applicable institution and/or its associated faculty.
[3] The Chief Administrator's office at each institution was contacted to appoint an Institution Coordinator, who served as the primary point of contact to deal with specific survey-related questions, correspondence, and follow-up.
[4] Faculty sampling rates were determined based upon frame counts using Integrated Postsecondary Education Data System (IPEDS) information, and selected on a rolling basis as lists were received.
NOTE: CATI = computer assisted telephone interview.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

1.2.3 Faculty Sampling and Data Collection

Precision goals for NSOPF:04 were to secure national-level survey estimates with precisions comparable to or better than those of NSOPF:99 for the overall faculty population. As with institution-level sampling, a customized cost/variance optimization technique was used to allocate the sample faculty to the institution and person strata while minimizing cost and

variance. Further details about faculty sampling may be found in Section 2.1; sample allocation to strata is fully detailed in appendix A."

Sample size was significantly larger than in the previous cycle: 35,630 faculty were sampled for NSOPF:04; of which, 34,330 were eligible. The final eligible sample for NSOPF:99 was 19,210. Criteria for faculty eligibility are discussed in section 2.1.2.

Prior to sampling, faculty counts from all lists provided by participating institutions were checked against both the Integrated Postsecondary Education Data System (IPEDS) and the counts provided by the institution on their institution questionnaire. (In 1999, the IPEDS comparison was used as a quality control check only when institution questionnaire counts were absent). As in NSOPF:99, institutions were contacted to resolve any discrepancies between data sources.

As in past cycles, faculty data collection utilized a mixed-mode approach; however, for NSOPF:04, sample members could participate only by a web-based self-administrated questionnaire or by an interviewer-administered telephone interview—there was no hardcopy version of the questionnaire. The participation of sample faculty members was initially requested in a letter that provided both instructions for completing the web questionnaire and calling to complete the interview via CATI. After 4 weeks, interviewers contacted the sample faculty members who had not completed the questionnaire to attempt a telephone interview. An early-response incentive was provided to encourage prompt completion of the instrument. Refusal or nonresponse incentives were also offered to selected sample members. Incentives are discussed in section 3.2.5.

1.3 NSOPF:04 Products

Data from the full-scale study will be used by researchers and policymakers to examine a wide range of topics, including who faculty are, what they do, and whether and how they are changing over time. NSOPF:04 provides data on each of these topics. The NCES Data Analysis System (DAS) for public release has been constructed from the data and is available to the public at http://nces.ed.gov/das. Electronically documented, restricted access data files with associated Electronic Codebooks (ECBs) are also available to qualified researchers.

The following types of reports are products of NSOPF:04: (1) this methodology report, providing details of sample design and selection procedures, data collection procedures, weighting methodologies, estimation procedures and design effects, and the results of nonresponse analyses; and (2) a series of descriptive statistical reports on key topics of interest. These topics include undergraduate teaching, faculty work activities and compensation, gender and racial/ethnic composition, and characteristics of part-time faculty. NSOPF:04 publications can be accessed electronically through the NCES website at http://nces.ed.gov/pubsearch/getpubcats.asp?sid=011.

Special tabulations are available on a limited basis from the National Education Data Resource Center (NEDRC) upon request. Use of NEDRC services is most appropriate for well defined questions that are likely to yield a few tables. It is recommended that those requiring more extensive research and in-depth analysis apply for direct access to the restricted access data files. Questions regarding NEDRC services may be directed by e-mail to nedrc@pcci.com or to Aurora D'Amico at aurora.d'amico@ed.gov or (202) 502-7334.

The remainder of this report contains the details of various activities. Chapter 2 details the survey design and implementation. Data collection outcomes are reported in chapter 3. Chapter 4 presents evaluations of the quality of data collected from institutions and faculty. Chapter 5 details procedures for data file development and imputation. Chapter 6 reports on procedures for weighting and variance estimation.

Chapter 2
Design and Implementation of NSOPF:04

This chapter provides a detailed summary of the design and implementation of the 2004 National Study of Postsecondary Faculty (NSOPF:04) full-scale study. Sampling of institutions and of faculty and instructional staff is discussed in detail. In addition, instrument design and data collection procedures are described.

A Technical Review Panel (TRP) meeting was held on September 8–9, 2003. The panel, comprised of nationally recognized experts in higher education, reviewed the impact of methodological changes in sampling and data collection, including combining NSOPF:04 with NPSAS:04, the elimination of paper instruments, shortening the data collection period, and revisions to the instruments. The list of panel members is provided in appendix B.

2.1 Sampling Design

NSOPF:04 employed a two-stage sampling methodology for selection of eligible faculty and instructional staff based on a cost/variance optimization process, details of which are provided in appendix A. In the first step, samples of eligible institutions were selected within the following 10 institutional strata:

- public doctoral;
- public master's;
- public baccalaureate;
- public associate;
- public other/unknown;
- private not-for-profit doctoral;
- private not-for-profit master's;
- private not-for-profit baccalaureate;
- private not-for-profit associate; and
- private not-for-profit other/unknown.

In the second step, samples of faculty members were selected within sampled institutions using a stratified systematic sampling where the six strata were defined in the following hierarchical order:

- Hispanic;
- non-Hispanic Black;
- Asian and Pacific Islander;
- full-time female;
- full-time male; and

- all other.

The institution frame was comprised of all 3,380 eligible postsecondary institutions, while the faculty frame included all faculty and instructional staff in the corresponding institutions, which was estimated to include approximately 1.1 million individuals (Zimbler 2001).[5]

The composition and eligibility definitions for these frames are outlined below.

2.1.1 Institution Frame

The institution frame for the NSOPF:04, like previous NSOPF cycles, consisted of all institutions meeting the following criteria:

- located in the 50 states or the District of Columbia;
- classified as participating in Title IV[6] student aid programs;
- public or private not-for-profit;
- 2- or 4-year degree-granting;
- offers educational programs designed for students beyond high school;
- academically, occupationally, or vocationally oriented; and
- makes programs available to the public.

The resulting frame was a subset of that used for the National Postsecondary Student Aid Study (NPSAS:04), in that NSOPF:04 did not include private for-profit less-than-2-year non-degree-granting or Puerto Rican institutions that were included in NPSAS:04.

The institution frame for NSOPF:04 was constructed from the Winter 2001–02 Integrated Postsecondary Education Data System Data Collection (Winter:02 IPEDS) file. To allow precise survey estimates for sectors of interest to the education community, this set of institutions was stratified based on institution control and level of degree offered. Institution control distinguished between public and private not-for-profit institutions, while level of degree offered was based on the 2000 Carnegie classification system[7] for segmentation of institutions. Table 2 summarizes the number of the eligible institutions for each of the resulting 10 primary institutional strata, based on the Winter:02 IPEDS file.

[5] This was used as a preliminary estimate and was adjusted later.
[6] Postsecondary institutions which have signed Title IV federal student aid program participation agreements with the U.S. Department of Education.
[7] The Carnegie Classification is a taxonomy of colleges and universities in the United States according to such variables as degrees awarded, number of fields covered, and specialization.

Table 2. Institution frame for the NSOPF:04, by Carnegie code, institution control, and degree granted: 2004

Degree granting	Total	Carnegie code	Public	Private not-for-profit
Total	3,380	†	1,700	1,680
Doctoral	300	15, 16, and 52	190	110
Master's	590	21 and 22	270	320
Bachelor's	570	31, 32, and 33	90	480
Associate's	1,180	40 and 60	1,030	150
Other/unknown	730	51, 53–59, and unknown	110	620

† Not applicable.
NOTE: For sampling purposes, public baccalaureate, private associate, and other/unknown institutions are collapsed into a single stratum. Definitions of Carnegie codes are available at http://www.carnegiefoundation.org/classification. The institution universe counts include institutions that were added after the sample was selected to account for institutions that became eligible for NSOPF:04 after construction of the institution sampling frame from the Winter:02 IPEDS. Also, the 44 institutions that had an unknown Carnegie code at the time of sample selection have been reassigned to their appropriate strata. Therefore, there are no longer any institutions with unknown Carnegie codes in the sample, but some still remain in the universe. Numbers have been rounded to the nearest 10. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), Fall 2000.

2.1.2 Faculty Frame

The second-stage sampling frame for NSOPF:04 includes faculty and instructional staff in the eligible postsecondary institutions. This includes both instructional faculty and faculty with no instructional responsibilities (e.g., research or administrative faculty) as well as staff with instructional responsibilities regardless of faculty status. In summary, eligible individuals for the NSOPF:04 study included any faculty and instructional staff who

- were permanent, temporary, adjunct, visiting, acting, or postdoctoral appointees;
- were employed full- or part-time by the institution;
- taught credit or noncredit classes;
- were tenured, nontenured but on tenure track, or nontenured and not on tenure track;
- provided individual instruction, served on thesis or dissertation committees, advised, or otherwise interacted with first-professional, graduate, or undergraduate students;
- were in professional schools (e.g., medical, law, dentistry); or
- were on paid sabbatical leave.

Ineligible individuals for NSOPF:04 included staff who:

- were graduate or undergraduate teaching or research assistants;
- had instructional duties outside of the United States, unless on sabbatical leave;
- were on leave without pay;
- were not paid by the institution, e.g., those in the military or part of a religious order;
- were supplied by independent contractors; or
- who otherwise volunteer their services.

2.1.3 Institution Sample Selection

The administration of NSOPF:04 consisted of a sample of 35,630 faculty and instructional staff across a sample of 1,080 institutions in the 50 states and the District of Columbia. This section provides details regarding the composition and construction of the institution sampling frame and methods used for selection of the institution sample.

Institution frame construction

The institution sample was selected using Chromy's sequential probability minimum replacement (PMR) sampling algorithm (Chromy 1979) to select institutions with probabilities proportional to a composite measure of size, details of which are provided in appendix A. For this purpose, each institution was assigned a measure of size (MOS) based on the number of eligible faculty and instructional staff and students in the given institution. Specifically, the composite size measure was the sum of cross products of sampling rates and population sizes for the groups, operating as the expected combined sample size at an institution. This measure was designed to ensure that student and faculty in certain minority strata would have a higher chance of selection. For faculty, these minority strata included:

- Hispanic;
- non-Hispanic Black or African American;
- Asian and Pacific Islander;
- female, full-time employee;
- male, full-time employee; and
- all others.

It should be noted that the MOS for each institution was calculated to reflect the number of students in the given institutions, since for this administration the institution samples for NPSAS:04 and NSOPF:04 were selected jointly. That is, precision requirements for NSoFaS:04 were considered jointly by reflecting both the faculty and student design objectives. Faculty counts needed for MOS calculations were initially obtained from the Fall Staff Survey component of the Winter:02 IPEDS data collection. However, this source could not provide all information necessary to classify faculty members into one of the above sampling strata. For instance, in a number of institutions faculty counts were not reported, while for others reported counts were not indexed by race and ethnicity. As a result, the missing information was imputed in two steps. In the first step, unreported (missing) faculty counts were imputed, while in the second step, faculty reported as unknown race/ethnicity or nonresident aliens were distributed among the known race categories using a special procedure, details of which are provided in appendix A.

Institution sample selection

The institution sampling frame was constructed from the IPEDS-IC files and was partitioned into institutional strata based on institutional control, highest level of offering, and Carnegie classification.[8] As mentioned earlier, the sample of institutions was selected probability

[8] More detailed information about the Carnegie classification can be found in appendix A.

proportional to size (PPS) based on the number of faculty and students at each institution, using Chromy's sampling algorithm. Sample sizes and their corresponding sampling rates were established using a customized cost/variance optimization procedure, which aimed to identify the allocation that would accommodate all analytical objectives of this survey while minimizing data collection costs.

Table 3 summarizes the distribution of the resulting sample of institutions for NSOPF:04. Subsequent to selection of the sample, the resulting institutions were contacted and asked to provide lists of eligible faculty and instructional staff for their institutions.

Table 3. Distribution of NSOPF:04 institution universe and sample, by institution control and degree granted: 2004

	Total		Public		Private not-for-profit	
Degree granting	Universe	Sample	Universe	Sample	Universe	Sample
Total	3,380	1,080	1,700	680	1,680	400
Doctoral	300	300	190	190	110	110
Master's	590	200	270	120	320	80
Bachelor's	570	160	90	30	480	130
Associate's	1,180	350	1,030	340	150	10
Other/unknown	730	70	110	10	620	60

NOTE: The universe and sample counts include institutions that were added after the sample was selected to account for institutions that became eligible for NSOPF:04 since construction of the institution sampling frame from the Winter:02 IPEDS. Also, the 44 sample institutions that had an unknown Carnegie code at the time of sample selection were reassigned to their appropriate strata. Therefore, there are no longer any institutions with unknown Carnegie codes in the sample, but some still remain in the universe. Universe and sample counts are rounded to the nearest 10. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

2.1.4 Faculty Sample Selection

This section provides an overview of the faculty sample selection procedures, which include methods used for frame construction and the technical details of cost/variance optimization process for selection of the initial sample sizes and calculation of needed sampling rates.

Faculty frame construction

The sampling frames for selection of faculty and instructional staff were constructed institution-by-institution. Each sampled institution was asked to provide a complete listing of eligible full- and part-time faculty and instructional staff. The majority of lists were delivered electronically; however, some of these lists were abstracted from online sources such as institution directories or supplied on paper.

Faculty sample selection

The sample of faculty was selected using an equal probability stratified systematic sampling, within cells indexed by institutional and faculty strata. As detailed in the next section, a customized cost/variance optimization program was utilized.

Determining initial faculty sample sizes and sample allocation

A special cost/variance optimization program was used to determine the desired allocation of respondents to institution-by-person strata, the goal of which was to secure at least the same level of precision for key estimates as those achieved during the previous administration of the survey. This optimization process, which is detailed in appendix A, consisted of the following steps:

- establishing precision requirements for key estimates;
- constructing a cost model specific to the structure of the NSOPF:04 sample;
- developing a relative variance model; and
- determining the optimum sample allocation.

Faculty sample selection

Faculty members were sampled as faculty lists were received from participating institutions. Prior to selecting the faculty sample for a given institution, expected sample sizes for each faculty stratum were calculated using the institution-specific faculty list counts and sampling rates. These sampling rates were then modified, as necessary, for the reasons given below.

- Rates were increased across all faculty strata to ensure that at least ten faculty members were selected from each institution, if possible.
- Rates were increased within faculty strata to guarantee that at least one faculty member was selected per stratum within each institution, if possible.
- The sample yield was monitored throughout the months during which faculty lists were received, and the faculty sampling rates were adjusted periodically for institutions for which sample selection had not yet been performed to ensure that the desired faculty sample sizes were achieved.

Stratified systematic sampling was used to select faculty members from the faculty lists. Specifically, from each list (institution) sample faculty were selected within each faculty stratum defined by race/ethnicity, gender, and employment status using the corresponding rate for the given institution-faculty stratum, with academic field serving as an implicit sort variable. Whenever a list contained insufficient data to identify faculty strata, a systematic sample of faculty was selected using the overall sampling rate for the institution. For hard copy lists, the resulting sample was then keyed to create an electronic file. The following table 4 provides a summary of the required sample sizes, which were determined based on the cost/variance optimization process and the resulting completed interviews by faculty stratum.

Table 4. Distribution of NSOPF:04 faculty sample sizes and completed interviews by faculty stratum: 2004

Faculty stratum	Required sample size	Completed interviews
Total	24,500	26,100
Non-Hispanic Black	1,600	2,060
Hispanic	1,300	1,700
Asian and Pacific Islander	900	1,610
Other full-time female	4,600	5,850
Other full-time male	8,300	8,500
Other part-time	7,800	6,380

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

The information supplied for each sampled faculty member (e.g., name, academic field, residence) was checked against that of faculty previously selected from other institutions to identify and eliminate respondents sampled twice. Duplicates were eliminated from the sample of the current institution. Once the de-duplication process was complete and the institution's final sample file was created, the institution's final sample file was added to the master dataset. The master dataset contained all sampled faculty members and their relevant sampling information.

2.2 Instrumentation

This section describes the institution and faculty instruments that were developed for the NSOPF:04 full-scale study conducted during the 2003–04 academic year with a national sample of postsecondary institutions and faculty and instructional staff. Data collection for the study was by self-administered questionnaires on the Internet or computer-assisted telephone interviews (CATIs) with web nonrespondents. In contrast to the data collection approach for NSOPF:99, no paper-and-pencil questionnaire options were provided.[9] Facsimiles of the electronic instruments, which provide item wording, response options, and information on respondent groups, are included in appendix C.

2.2.1 Development of Instrumentation

Project staff from RTI and MPR Associates were responsible, respectively, for developing and implementing study instrumentation for NSOPF:04 and for ensuring that the instruments, where possible, retained analytic comparability with earlier data collection rounds of the study. Revisions to the institution and faculty/instructional staff instruments built upon the NSOPF:99 instruments, and included the comments and suggestions of the Technical Review Panel (TRP), sample respondents contacted after the study for additional information, and other government officials and postsecondary researchers. (Copies of the NSOPF:99 data collection instruments for postsecondary institutions and faculty/instructional staff are included as appendixes A and B, respectively, in Abraham et al. 2002.) In May 2002, meetings with the TRP were conducted to review the relevance of policy issues examined in NSOPF:99, the importance of emerging issues (such as increased use of the Internet and distance education) not included in

[9] A "facsimile" of the institution questionnaire—what the electronic instrument might have looked like if it was rendered as a hard-copy document—was included with the binder materials distributed to Institution Coordinators. However, this 12-page document was marked "informational copy only" and was not used for data collection.

the 1999 instruments, and the consequences of adding, revising, or deleting items from the NSOPF:99 instruments.[10]

Following contract award for NSOPF:04, project staff developed and tested multiple versions of the institution and faculty/instructional staff instruments. A field test version of the instrumentation was developed at the start of the 2002–03 academic year and closely reviewed by members of the study TRP, government officials, postsecondary researchers, and other interested individuals. Then during the fall and spring terms of 2002–03, field test data collection for NSOPF:04 permitted the evaluation of the revised institution and faculty/staff instrumentation under conditions comparable to those to be employed during the NSOPF:04 full-scale study.[11]

Several policy, methodological, and practical concerns guided the development of instrumentation for NSOPF:04. To ensure the comparability of data elements from earlier rounds of the postsecondary faculty study in 1988, 1993, and 1999, one of the primary objectives of instrumentation was to maintain the trend analyses for this national, cross-sectional study. However, this goal was balanced by the importance of adequately considering emerging issues, while at the same time developing instruments that could be completed quickly and efficiently by sample members. For example, almost 70 percent of the institution responses for the 1999 study were obtained via paper-and-pencil questionnaire, and the average time to complete the institution questionnaire was 90 minutes. For the NSOPF:99 faculty questionnaire, over one-half (54 percent) of the respondents completed hardcopy instruments, with an average web and paper questionnaire completion time of 51 minutes; the average CATI completion time was 55 minutes.

Based on these considerations, the goals for the NSOPF:04 instrumentation included several elements:

- All data collection would be completed electronically, using web-based self-administered questionnaires, with telephone interviews for those who did not respond to the web self-administered questionnaires. No paper and pencil instruments would be received.
- All data collection instruments for the study would be shorter than the NSOPF:99 instruments, thus simultaneously increasing response rates while reducing the potential for bias and the need for costly refusal conversion efforts. The targets for average time to complete the instruments were set at 45 minutes for the institution questionnaire and 30 minutes for the faculty/instructional staff questionnaire.
- Consistent with the transition to all-electronic data collection, the NSOPF:04 instrumentation was designed to be easier for sample members to complete, to be easier for the study team to process, and to provide higher quality data.
- Finally, the instrumentation team sought to address emerging issues as well as to maintain comparability with earlier rounds of the study.

[10] One important element in this process was a consideration of recent literature in the field; for example, *Developing the 2004 Faculty Survey: Themes from the Literature on Postsecondary Education*, developed by the American Institutes for Research (Berger et al. 2002).

[11] Field test data collection for the institution questionnaire took place from September 2002 through June 2003; faculty/instructional staff field test data collection lasted from January 2003 through June 2003.

With these goals established, planning and design for the NSOPF:04 institution and faculty/instructional staff questionnaires began. Specification for both instruments was in RTI's Instrument Development and Documentation System (IDADS), a tool developed specifically for the design of complex electronic data collection instruments (see also section 2.5.1). Using IDADS, instrument designers entered information about each instrument item, including the variable data definition, formatting, and the desired on-screen presentation.[12] For each of the NSOPF:04 instruments, designers specified the variable names and labels, values and value labels, "applies to" fields, and variable definitions (e.g., numeric, continuous, maximum and minimum values, field size, etc.).

2.2.2 Instrument Programming

Despite the different data collection modes for NSOPF:04, the self-administered web instruments for the institution and faculty/instructional staff respondents were identical to their corresponding CATI instruments. Both instruments were web-based products, located on U.S. Department of Education servers. The instruments were developed using Microsoft Corporation's Active Server Pages (ASP) web programming language.[13] This approach resulted in a computer-assisted data collection program that facilitated the preloading of full-screen data entry and editing of "matrix-type" responses. The web and CATI system presented interviewers with screens of questions to be completed, with the software guiding the respondent through the interview. Inapplicable questions were skipped automatically based on prior response patterns. On-screen clarification was available for all items.[14] The instrument also provided real-time error checking for inconsistent or out-of-range responses and minimized the potential for inadvertently skipped items.

2.2.3 Institution Questionnaire

Instrumentation activities for the NSOPF:04 institution questionnaire began in May 2002 with revisions to the NSOPF:99 instrument. Project staff began working with a revised version of the NSOPF:99 instrument that incorporated the lessons learned from the NSOPF:99 data collection, including the comments and suggestions for instrumentation provided by both the NSOPF TRP and a small number of study respondents who were contacted for additional information after the completion of NSOPF:99 data collection.

This information formed the input for the NSOPF:04 field test institution questionnaire that was administered to a purposive sample of 150 postsecondary institutions during the 2002–03 academic year. The interpretation of responses from the field sample members that completed the instrument (77 percent of the sample of institutions that were eligible to participate), results

[12] In addition to instrument development, IDADS also provides a reference system for instrument reviewers and testers and serves as the data documentation system for the data products developed.

[13] Active Server Pages (ASP) dynamically produce hypertext markup language (html) pages designed to facilitate information retrieval across the Internet. ASP code includes small embedded programs or scripts that are processed on a web server when accessed by users employing browser programs such as Netscape or Internet Explorer. Before responses are returned to a user, the request typically accesses databases and develops a customized response.

[14] Each data collection screen or form for the NSOPF:04 field test faculty instrumentation included a link to a page of "help text" prepared specifically for the item and including key definitions, descriptions of respondents to whom the item applied, and other useful information. In an attempt to shorten the administration time for the full-scale instrument, the help text was shortened and appeared on the same form as the question wording and response options. This reduced the need for loading a separate web page for help. A separate help text web page was available for the institution questionnaire for both the field test and full-scale versions of the instrument.

of debriefing sessions with institution contact personnel for the field test who were responsible for encouraging response from the institutions, and data collection timing information for the field test also served to inform revisions to the full-scale study institution questionnaire.

After careful consideration of this input and examination of the data collected during the 1998–99 academic year—including the patterns of responses and missing data, as well as time to complete estimates—instrument revisions were implemented. Like the NSOPF:99 institution questionnaire, the NSOPF:04 instrument was divided into major sections that collected information on the number of faculty and instructional staff employed at the target institution; the policies and practices that affected, respectively, full-time and part-time faculty and instructional staff; and the percentage of undergraduate instruction assigned to various instructional personnel. Descriptions of the information included in these sections follow (see also the instrument facsimile in appendix C):

- The first section (items 1A and 1B) collected information on the number of faculty and instructional staff employed either full time or part time at the target postsecondary institution during the fall term of the target academic year (2003–04). For NSOPF:04, institution personnel were requested to provide these counts "as of November 1, 2003 (or during the fall term of the 2003–04 academic year when your faculty lists are considered complete)."
- Institution instrument items 2 through 13 defined the second section of the questionnaire and collected information on the employment of the target institution's *full-time* faculty and instructional staff. After first collecting information on the numbers of these personnel who entered or exited full-time employment during the previous academic year (2002–03 school year), this section examined the characteristics and policies of the target institution's tenure system, employee benefits, union representation (if any), and personnel evaluation, as applied to full-time faculty and instructional staff.
- The third section of the institution questionnaire (items 14 through 18) examined the employment of the target institution's *part-time* faculty and instructional staff. This section used items similar to those for full-time faculty and instructional staff in the previous section. These items included the availability of retirement plans to part-time faculty, the availability of and institution-level support for various types of employee benefits, and the characteristics of the institution's personnel evaluation system.
- The fourth instrument section included a single question (19) that collected information on the percentage of the target institution's undergraduate instructional activities assigned to various instructional groups, including full-time faculty and instructional staff, part-time faculty and instructional staff, teaching assistants such as graduate students, and others individuals.
- The last section of the NSOPF:04 institution questionnaire (item 20) collected respondent contact information and feedback on data collection. This section attributed the item responses for the entire institution questionnaire to individual respondents at the institution, which allowed data collection staff to recontact respondents for clarification of responses. These data elements—respondent name, job title, telephone number, and e-mail address—were not maintained after data collection was completed.

Appendix D provides a crosswalk of NSOPF:04 institution questionnaire items to the institution questionnaires from NSOPF:88, NSOPF:93, and NSOPF:99. Table 5 notes how the NSOPF:04 questionnaire differs from the NSOPF:99 questionnaire. As noted in this table, nine items from the NSOPF:99 questionnaire were eliminated from the NSOPF:04 institution questionnaire, 14 items were revised, and three items for NSOPF:99 were repeated without change.

Table 5. Content and formatting changes to the NSOPF:99 institution questionnaire in preparation for the NSOPF:04 instrument: 2004

NSOPF:99			NSOPF:04	
Item	Content	Action	Item	Changes
1	Numbers full/part-time faculty and instructional staff	Revised	1	Slight wording and instruction changes
2	Change in total number of full-time faculty and instruction staff over the past 5 years	Deleted		
3	Policies to decrease the number of full-time faculty and instructional staff	Deleted		
4	Availability of tenure system	Unchanged	3	
5	Changes in full-time faculty and instructional staff between fall terms	Revised	2	One response option added (item 2f), slight wording change throughout, distinction among tenured, tenure track, and not tenure track eliminated
6	Number of staff considered for/granted tenure	Revised	4/5	Asked as two questions with first as gate item.
7	Maximum number of years on tenure track	Unchanged	6	
8	Changes in tenure policy in past 5 years	Revised	7/8/ 7sp	Broken into three items; response options revised (Option E, discontinued tenure, asked only of respondents who answered “no” to tenure availability)
9	Other actions to reduce tenured faculty	Deleted		
10	Number of full-time positions sought to hire	Unchanged	9	
11	Retirement plans available to full-time staff	Deleted		
12	Employee benefits available to full-time faculty and instructional staff	Revised		Broken into two items, part 10A serves as gate question
			10A	Response categories for benefits were changed to All, Some, None, Don’t know
			10B	Fully and partially subsidized categories combined
13	Additional employee benefits available to full-time faculty and instructional staff	Revised	11	Response categories for benefits changed to All, Some, None, and Don’t know; Slight wording change
14	Percentage of salary contributed by institution to benefits	Deleted		
15	Collective bargaining for full-time faculty and instructional staff	Revised	12	Percentage of faculty represented by union eliminated
16	Teacher assessment with full-time faculty and instructional staff	Revised	13	Response options changed to Yes, No, Don’t Know; “Other, specify” option was eliminated
17	Availability of retirement plans for part-time faculty and instructional staff	Revised	14	Item wording revised for web data collection
18	Types of retirement plans available for part-time faculty and instructional staff	Deleted		
19	Criteria for eligibility for retirement plans for part-time faculty and instructional staff	Deleted		

See notes at end of table.

Table 5. Content and formatting changes to the NSOPF:99 institution questionnaire in preparation for the NSOPF:04 instrument: 2004—Continued

NSOPF:99			NSOPF:04	
Item	Content	Action	Item	Changes
20	Employee benefits available to part-time faculty and instructional staff	Revised		Broken into two items, part 15A serves as gate question
			15A	Response categories for benefits were changed to All, Some, None, Don't know
			15B	Fully and partially subsidized categories combined
21	Additional employee benefits available to part-time faculty and staff	Revised	16	Response categories for benefits changed to All, Some, None, and Don't know; Slight wording change
22	Benefit eligibility criteria for part-time faculty and instructional staff	Deleted		
23	Percentage of salary contributed by institution to benefits	Deleted		
24	Collective bargaining for part-time faculty and instructional staff	Revised	17	Percentage of faculty represented by union eliminated
25	Teacher assessment with part-time faculty and instructional staff	Revised	18	Response options changed to Yes, No, Don't Know; "Other, specify" option was eliminated
26	Undergraduate instruction by instructional staff type	Revised	19	Response options changed

NOTE: Numbers in table correspond with the question number in the instrument.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

2.2.4 Faculty Questionnaire

The NSOPF:04 questionnaire for faculty and instructional staff was divided into several sections that described the study and respondents' rights (informed consent); nature of employment; academic and professional background; instructional responsibilities and workload; scholarly activities; job satisfaction; compensation; background characteristics; and opinions. Included within the final section, where applicable, were items that collected address information for sample members who were eligible for response incentives. (See section 3.2.5 for additional information about the early-response and refusal conversion incentives.) Table 6 describes the instrument sections, including the number of forms (or screens) and data elements in each. Like the instrumentation for the study waves in 1988, 1993, and 1999, the NSOPF:04 faculty and instructional staff questionnaire emphasized descriptive and behavioral attributes rather than attitudinal measures.

The design of the faculty and instructional staff questionnaire included input from members of the NSOPF:99 TRP and representatives of offices of the U.S. Department of Education, as well as an analysis of the data collected during the 1999 study. Because the NSOPF:99 instrument took 55 minutes to complete, designers made a concerted effort to shorten the instrument and make it more efficient.[15] Several questions were eliminated, and other questions were shortened or otherwise simplified. The instrument was then evaluated in a field

[15] Efficiency for the NSOPF:04 instrument was gained by developing a shorter, tighter, and more focused interview that used state-of-the-art technology and design techniques. The sections and items were rearranged, coding procedures revised considerably to be interactive, skip patterns were employed, range checks were inserted, and other changes were implemented to make the instrument operate more efficiently.

test carried out during the 2002–03 academic year under conditions similar to those employed during the full-scale study in 2003–04.

Table 6. Overview of the NSOPF:04 questionnaire for faculty and instructional staff: 2004

Section	Forms/items[1]	Examples of content
Total	81/183[2]	
Informed consent	6/0	Description of the NSOPF:04 study and respondents' rights as participants.
A. Nature of employment	17/18	Does the respondent have instructional responsibilities during the Fall 2003 term? Does the respondent have faculty status? When did the person begin working? What are the respondent's rank, tenure status, and teaching field?
B. Academic and professional background	16/23	What is the respondent's highest degree? Where, when, and in what area was it earned? Is this the respondent's first academic job? Where else did the person work? Does the respondent teach? How long has the person been teaching?
C. Instructional responsibilities and workload	13/66	How many hours during an average week does the sample member spend on instruction, research, and other activities? How many classes are taught, and what are their characteristics (e.g., duration, number/type of students, evaluation type)? What level of advising and individual instruction is offered?
D. Scholarly activities	7/20	What scholarly activities have sample members completed in their lifetime and during past 2 years? What is their principal scholarly field? Are scholarly activities funded?
E. Job satisfaction	2/10	How satisfied is the respondent with instructional duties and employment at the target school? What are the person's retirement plans?
F. Compensation	7/12	What is the respondent's compensation from the target institution and all other sources? What is the structure of the employment contract? What is the household income?
G. Sociodemographic characteristics	8/13	What is the respondent's sex, date of birth, race/ethnicity, marital status, citizenship, and disability status? Does the person support dependents?
H. Opinions	2/5	What are the respondent's opinions about the faculty reward system at the target institution? Would the sample member seek an academic career again?
I. Incentive information	3/16	Where applicable, these forms also collected address information from sample members qualified for nonresponse incentives.

[1] The faculty/instructional staff questionnaire was divided into forms (screens) and items. Each form was structured to include related items. The first number is the number of forms in the section, and the second number is the number of items included on those forms. Forms and items were often skipped based on the responses to earlier forms and items.

[2] The number of items in the faculty questionnaire (183) differs from the number of faculty items reported elsewhere in this document (e.g., 162 analysis variables and 144 stochastically imputed variables) because some items were for internal use only (e.g., verbatim text strings used to code field of teaching [see section 4.3.3], school name and city to code IPEDS [see section 4.3.4], and contact information for sending incentives).

NOTE: IPEDS = Integrated Postsecondary Education Data System.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Following the field test, additional items were modified and eliminated to reach the desired 30-minute interview. The average CATI and web interview for the NSOPF:04 field test took 42 minutes, considerably longer than anticipated. The results of the NSOPF:04 field test reliability reinterview (Heuer et al. 2004), the policy relevance of each instrument item, and the input received from responding sample members, telephone interviewers/help desk staff, and members of the NSOPF:04 TRP were used to identify 27 forms from the full-scale study for elimination. Table 7 describes the NSOPF:04 field test items on these forms that were eliminated

following the field test. The table also provides the average time to complete each item during the field test.[16] After adjusting this average time to complete by the proportion of the overall respondent population that reached the item (i.e., complicated and time-consuming items will have little impact on the average time to complete the entire interview if most respondents do not attempt the item), these item reductions were expected to reduce the average time to complete the full-scale instrument by approximately 7 minutes.

Table 7. Items removed from the NSOPF:04 faculty/instructional staff questionnaire following the NSOPF:04 field test and estimated time savings: 2004

Instrument change	Estimated time savings (in minutes)
Total	7.483
Q7: Part-time faculty: years employed part-time	0.163
Q17B: Holds Ph.D. in addition to professional degree	0.005
Q17C: Year received doctoral degree	< 0.001
Q17C2VS: Doctoral field: verbatim	< 0.001
Q17C2CD: Online coding: doctoral field	< 0.001
Q17C3: Online coding: doctoral degree institution	0.003
Q17D2: Online coding: bachelor's degree institution	0.642
Q19C: Number classes taught at other postsecondary institution	0.065
Q20: Non-postsecondary education jobs related to teaching field	0.102
Q22: Total number of postsecondary institutions employed as faculty	0.222
Q25: First postsecondary faculty position—academic rank	0.115
Q29: Previous job related to teaching field	0.178
Q30: Years teaching in postsecondary institutions	0.152
Q34A–Q34D: Percent allotment of other time	1.273
Q40A–Q40G: Uses of website	0.410
Q43A–Q43D: Plan/develop instruction/employment opportunities	0.865
Q44A–Q44F: Training opportunities	0.933
Q45: Hours professional training in 2003	0.452
$Q52A_i$cat: Categorical items for Q52AA–AG nonrespondents	0.105
Q58: Primary funding source	0.115
Q59: Number of grants/contracts	0.130
Q60A: Total funding grants/contracts	0.058
Q60B: Range total funding grants/contracts	0.012
Q63: Age expecting to stop working at postsecondary institution	0.338
Q76A–Q76E: Type of disability	0.015
Q78: Number of dependents	0.235
Q84: Respondent comments and suggestions	0.895

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

It should be noted that approximately 7 minutes of the overall time to complete the field test interview were associated with "transit" time—in other words, the time involved to transmit information to each respondent, to "write" the form and related text onto each sample member's screen, to transmit the responses back for storage, and to begin the transmission of the next item. Interview transit times are dependent on many factors such as server bandwidth, processing

[16] The average time to complete each item is based on the time each respondent took to answer each item, as well as the time required to transmit the data collection image to the respondent and to transmit and write the information when the response was completed.

efficiency, and instrument content (e.g., other things being equal, the transit time for a form with little text and graphic information will be less than another form with more text and graphics). Notably, transit times are often dependent on factors beyond the control of instrument designers. For example, the type of internet connection used by the sample member (telephone dial-up modem versus direct Ethernet connection with fiber optic lines) and the number of other users on the respondent's internet service provider at the time will affect transit time. Section 3.3.1 describes the interview completion time and transit times for the NSOPF:04 full-scale faculty and instructional staff questionnaire.

To reduce transmission time from that experienced in the NSOPF:04 field test, project staff carefully reviewed the code efficiency of the web applications. The project also utilized an outside and independent review of the study procedures and programs. TechSages LLP, a computer consulting firm located in Durham, NC, reviewed the NSOPF:04 field test computer code. The group offered several recommendations for optimizing the code to improve execution speed including changes to database connectivity implementation, code structure, and variable scoping.

In addition to the changes in data processing and the reduction in the number of items included in the questionnaire noted above, instrument designers also implemented several other content related changes for the full-scale study. These included the following:

- Instrument designers eliminated the faculty/instructional staff questionnaire's online help and replaced it with more targeted information placed directly on the form containing the question. For the field test, a callable help screen was available for each form of the faculty interview. By selecting a help button at the bottom of each form, the respondent could review a screen of related definitions, examples, and other information about the item. While these help screens provided useful information, accessing them did require the transmission of an additional form and, consequently, an increase in the interview completion time. While adding text, such as definitions or examples, does increase the transit time of a screen, the increase is negligible relative to the increases in interview time that would be obtained by accessing and transmitting a second web page of help text for the item. Adding definitions and examples to the original form of the interview reduced the need for help screens.
- For the full-scale study's questionnaire, project staff also developed an online assisted-coding routine for respondent's academic area or discipline. Assisted coding provided significant time savings over the online coding in the field test, which used two pull-down boxes for each academic discipline. The assisted-coding procedure developed for the full-scale study eliminated pull-down boxes for common disciplines (e.g., mathematics or English), considerably reducing the time each respondent took to code academic field. The pull-down boxes were available for unusual disciplines or when the sample member was not satisfied with the result of the assisted-coding activity. To use the assisted-coding routing, the sample member entered the name of the relevant academic field, and then confirmed or discarded the results of the matches with an assisted-coding dictionary developed from the Classification of Instructional Programs (U.S. Department of Education 2002).
- Instrument developers also improved item wording, and especially screen fills to reduce item wording.

- Finally, the full-scale study instrument combined a number of instrument screens, thus reducing the number of overall forms and the number of data transmissions. (For example, forms Q65 and Q80 in the full-scale study instrument combined previously independent forms.)

Table 8 compares and contrasts the faculty and instructional staff instruments used for the NSOPF:04 and NSOPF:99 full-scale studies. As noted in this table, 39 items were eliminated from the 1999 instrument, 51 items were simplified or otherwise revised, 1 item was added, and 3 items were unchanged.

Table 8. Content and formatting changes to the NSOPF:99 faculty questionnaire in preparation for the NSOPF:04 instrument: 2004

NSOPF:99			NSOPF:04	
Item	Content	Action	Item	Changes
1	Instructional duties	Revised	1	Wording change to highlight that teaching includes credit and noncredit courses; on screen descriptions of instruction duties
2	Credit status of instructional duties	Revised	2	Response options for instructional duties item changed to Yes/No for credits awarded for classes
3	Principal activity	Revised	4	Pubic service option added; other specify for administration removed
4	Faculty status	Revised	3	Faculty status “defined” as at target institution
5	Full- and part-time status	Revised	5	Response category order changed
		New	6	For part-time faculty/instructional staff, is position your primary employment
6	Reason working in part-time position	Revised	8	Wording for response option modified; reason for holding PT position eliminated
7	Year began job	Revised	9	Stem wording revised
8	Rank	Revised	10	Open-ended specify field eliminated; examples given for “other” response option
9	Year achieved rank	Revised	11	Stem modified to specify at “any institution”; response population subset to professors or associate professors only
10	Tenure status/date of tenure	Revised	12/13	Stem modified to specify tenure at “any institution”
11	Duration of contract	Deleted		
12	Type of appointment	Deleted		
13	Chair of department	Deleted		
14	Principal field of teaching	Revised	16	Assisted coding of teaching field discipline using Classification of Instructional Programs added
15	Principal field of research	Revised	54	Stem wording changed to field of “scholarly activities”; assisted coding of discipline using Classification of Instructional Programs added; response population subset to respondents without specified teaching field
16	Degrees obtained (year received, field, and name, city, state of institution awarding)	Revised		Formatted for web data collection, stem wording changed
			17A1	Highest degree only collected
			17A1B	If highest degree reported is professional degree, does respondent also have PhD
			17A2	Year received highest degree
			17A3	Assisted coding of discipline using Classification of Instructional Programs added

See notes at end of table.

Table 8. Content and formatting changes to the NSOPF:99 faculty questionnaire in preparation for the NSOPF:04 instrument: 2004—Continued

NSOPF:99			NSOPF:04	
Item	Content	Action	Item	Changes
			17A4	Name, city, and state of institution awarding highest degree collected; respondent assisted online coding of institution using Integrated Postsecondary Education Data System added
			17D	Year bachelor's degree awarded (if highest degree above bachelor's);
17	Working toward a degree	Deleted		
18	Degree working toward	Deleted		
19	Primary employment	Deleted		
20	Outside consulting	Deleted		
21	Other professional employment	Revised	18	Stem changed to include "all" positions outside of target institution
22	Number of other professional jobs during fall term	Revised	19A 19B	Formatted to include gate question; number of jobs expanded to include information on "full-time jobs" at other postsecondary institutions
23	Total jobs held in postsecondary education	Deleted		
24	First and most recent jobs in higher education: years held, institution type, primary responsibility, employment status and title	Revised	21/23/ 24/26	NSOPF:04 simplified this question from 18 to 4 data elements. Item for 2004 asks if current job was first postsecondary education position, when position began, employment status, and tenure status of the position
25	Years teaching in higher education	Deleted		
26	Number of positions ever held outside of higher education	Revised	27	Changed to positions ever held outside postsecondary education since highest degree
27	Job status of those positions	Deleted		
28	First and most recent jobs outside of higher ed: Type of employer, and primary responsibility	Revised	28	Item simplified from 10 to 1 data elements. Item now collects only the employment sector of most recent job: first profession position outside higher education eliminated
29	Scholarly activities during career; scholarly activities during past 2 years	Revised	52A/52B	Formatted for web instrument; joint/sole responsibility eliminated; stem wording and item strings revised
30	Average time spent in activities per week	Revised	31	Item strings reworded and revised to include more examples' open-ended specify field eliminated
31	Allocation of working time, preferred allocation of working time	Revised	32	Preferred allocation eliminated; item reformatted for web instrument; response categories combined, reworded, and simplified (e.g., asked only about time at target institution, focus changed to instructional activities, professional growth/ administration/service combined
32	Committee assignments	Revised	48	Reformatted for web; stem wording revised to eliminate student level and number of committees chaired and served
33	Number of classes taught	Revised	35A	Reformatted for web; item expanded to include the number of "classes/sections" taught for credit and not for credit, wording revised to include "taught for credit toward degree"
34	Number of different courses taught	Deleted		
35/36	Number of remedial classes taught; number of remedial classes not creditable towards degree	Revised	35B	Item wording revised to include "remedial or developmental classes"; second item on distance education added
37	Number of continuing education classes taught	Deleted		
38	Number of noncredit continuing education classes taught	Deleted		

See notes at end of table.

Table 8. Content and formatting changes to the NSOPF:99 faculty questionnaire in preparation for the NSOPF:04 instrument: 2004—Continued

NSOPF:99			NSOPF:04	
Item	Content	Action	Item	Changes
39	Number of students in all noncredit classes	Deleted		
40	Number of classes taught for credit	Revised		See question 35A above
41	Details on up to five credit classes, including discipline; description (weeks class met, credit hours, hours class met/week, number teaching assistants, number students, class team taught, hours per week respondent taught, and remedial and/or distance education); level of students, instructional method; and instructional medium	Revised	36/37	Reformatted for web; gate item of teaching/lab assistants added; class description matrix simplified; information collected included number of weeks and hours per week respondent taught class, credits for the class, number of students, and level of students
42	Undergraduate evaluation methods	Revised	38	Stem wording revised, response options added, deleted, and revised; response values reworded
43	Websites	Revised	39	Stem wording changed to include all instructional duties; response population subset to persons with instructional duties
44	How websites used	Deleted		
45	E-mail	Deleted		
46	Student percentage using e-mail	Deleted		
47	Hours spent responding to student e-mail	Revised	41	Stem wording revised to include “communicating with students”; response population subset to persons with instructional duties
48	Internet access available	Deleted		
49	Individual instruction	Revised	46/47/ 47B	Gate question added; stem wording changed; item reformatted for web
50	Contact hours with advisees	Revised	50	Reformatted for web; stem wording revised
51	Office hours	Revised	51	Stem wording expanded to include in-person and online office hours
52	Engaged in research	Revised	53	Stem wording revised; reference period the entire academic year
53	Type of primary research	Revised	56	Stem wording revised to include “principal scholarly activity”; reference period the academic year; open-ended specify field eliminated
54	Engaged in funded research	Revised	55	Stem wording revised to include “scholarly activities at target school” and exclude funding from basic salary; reference period the academic year
55	Principal/co-principal investigator on funded research	Deleted		
56	Number supported by grants	Deleted		
57	Sources of funding	Deleted		
58	Total number of grants	Deleted		
59a	Total funds	Deleted		
59b	How received funds were used	Deleted		
60	Evaluation of facilities and resources	Deleted		
61	Use of institutional funds	Deleted		
62	Number and type of administrative committees	Deleted		
63	Hours spent on administrative committee work	Revised	49	Reformatted for web, stem wording revised to include examples
64	Union membership	Revised	14/15	Item reformatted for web; gate question added
65	Satisfaction with instructional duties	Revised	61	Number of response options reduced; new options added
66	Job satisfaction	Revised	62	Number of response options reduced
67	Likelihood of leaving job	Deleted		

See notes at end of table.

Table 8. Content and formatting changes to the NSOPF:99 faculty questionnaire in preparation for the NSOPF:04 instrument: 2004—Continued

NSOPF:99			NSOPF:04	
Item	Content	Action	Item	Changes
68	Age to stop working at postsecondary institution	Deleted		
69	Factors influencing possible decision to leave	Deleted		
70	Most important factor regarding decision	Deleted		
71	Option to draw on retirement	Deleted		
72	Retired previously	Unchanged	64	
73	Early retirement option	Deleted		
74	Age planning to retire	Unchanged	65	
75a/76	Basic salary for academic year/ Compensation for calendar year	Revised	66/66B	Reformatted for web with follow-up screen for nonrespondents; stem wording revised to stress confidentiality; item wording revised to simplify response categories and provide examples
75b	Basis of basic salary	Revised	67/68/69	Reformatted for web into separate items; item wording revised to collect contract length and other pay arrangements; open-ended specify field eliminated
77	Income of spouse/significant other	Deleted		
78	Number of persons in household	Deleted		
79	Household income	Revised	70A/70B	Reformatted for web with follow-up screen for nonrespondents; stem wording revised to include respondent's salary reported earlier and onscreen definition of household income; follow-up screen for item nonrespondents added
80	Number of dependents	Revised	79	Item changed to number of dependent children
81	Gender	Unchanged	71	
82	Month and year of birth	Revised	72	Birth month eliminated
83	Ethnicity	Revised	73	Reformatted for web instrument
84	Race	Revised	74	Response options reordered to match current federal standards for collecting racial information
85	Disability	Revised	75	Stem wording revised to include additional on-screen definitions
86	Type of disability	Deleted		
87	Marital status	Revised	77	Wording and order of response options modified
88	Employment of spouse/significant other	Deleted		
89	Country of birth	Revised	80	Revised to ask if born in U.S.
90	Citizenship status	Revised	81	Visa status and distinction between native/naturalized citizenship eliminated
91	Parent and spouse education level	Deleted		
92	Opinions about target institution	Revised	82/83	Number of response options reduced; new options added
—	Open-ended comments	Deleted		

NOTE: Numbers in table correspond with the question number in the instruments. Question numbers 7, 20, 22, 25, 29, 30, 33, 34, 40, 42, 43, 44, 45, 57, 58, 59, 60, 63, 76, and 78 in the NSOPF:04 faculty questionnaire were eliminated before data collection, and the instrument was not renumbered.

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

2.3 Institution Data Collection

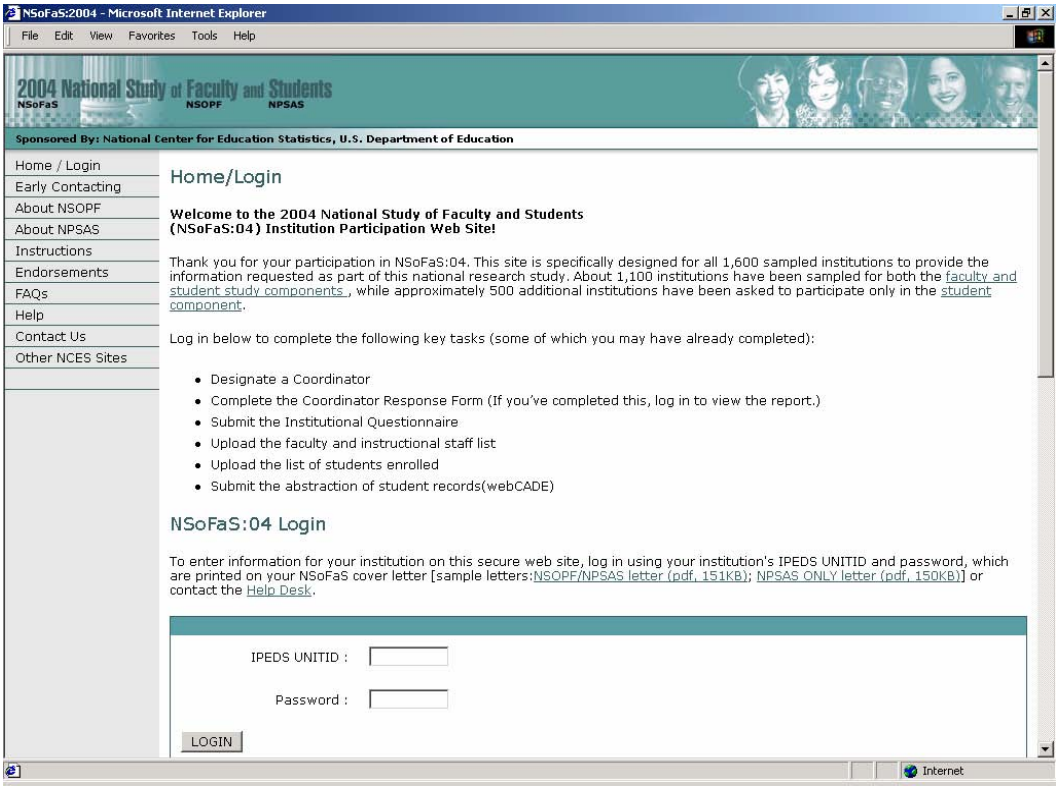
The goals of the institution data collection for the NSOPF:04 study were to collect a list of full- and part-time faculty and instructional staff (referred to as a "faculty list") from each sampled institution and to obtain a completed questionnaire from each sampled institution. As described in section 2.1.4, the faculty list was used for selecting the faculty sample and also

provided the contact information used for faculty data collection activities. The institution questionnaire, detailed in section 2.2.3, collected information on the policies and practices affecting full- and part-time faculty and instructional staff. To facilitate the process of obtaining faculty lists and completing the institution questionnaire, an institution website was developed, and for each sampled institution, the Chief Administrator (CA) was asked to appoint an Institution Coordinator (IC).

2.3.1 Institution Website

The NSoFaS:04 website served a number of functions for both the NSOPF:04 and NPSAS:04 studies. For institutions, it was a central repository for all study documents and instructions. It allowed for the uploading of electronic lists of faculty and instructional staff. In addition, it housed the institution questionnaire for the Institution Coordinator to complete online. Figure 1 presents the home page of the NSoFaS:04 website.

Figure 1. The 2004 National Study of Faculty and Students institution website home page



SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Faculty and Students (NSoFaS:04) website.

Visitors to the website were provided with the following links (see navigation bar on the left side of the screen):

- *Early Contacting* provided information about the early institution contacting process for NSoFaS:04 for the initial stage. Section 2.3.2 provides details of early institution contacting.

- *About NSOPF* (faculty) provided information on the study's mandate and research objectives, with a link to National Center for Education Statistics (NCES) reports from previous study cycles.
- *About NPSAS* (student) provided comparable information (as noted above) for the student component of NSoFaS:04.
- *Instructions* provided links that allowed institution staff to view and print copies of various NSOPF:04 and NPSAS:04 forms (in pdf format).
- *Endorsements* listed the 25 national organizations that endorsed both studies. (The 24 NSOPF:04 endorsements are listed in appendix E; one endorsement was applicable only to proprietary schools that were eligible for NPSAS:04 but ineligible for NSOPF:04).
- *Frequently Asked Questions (FAQs)* included questions and answers concerning all stages of data collection for both components of NSoFaS:04.
- *Help* provided the help desk toll-free number and e-mail address for contacting project staff, along with instructions for logging in.
- *Contact Us* contained address information for RTI International.
- *Other NCES Sites* links to three NCES websites that provided more information about NCES programs and how to order publications.

All data entry applications were protected by Secure Sockets Layer (SSL) encryption. Further security was provided by an automatic "time out" feature, through which a user was automatically logged out of the NSOPF:04 institution questionnaire if the system was idle for 30 minutes or longer. The system did not use any persistent "cookies,"[17] thus adhering to the Department of Education's privacy policy.

A status screen, shown in figure 2, indicated which stages of institution data collection were completed (denoted by a check mark) and allowed institutions to select those stages that were not yet completed. Once a stage was completed, it was no longer accessible via the Web.

[17] A persistent cookie is a piece of information, such as an IPEDS ID, that can be stored in a file on the user's computer. This information could then be used to identify a computer without the user even logging into the application.

Figure 2. The 2004 National Study of Faculty and Students institution website status screen

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Faculty and Students (NSoFaS:04) website.

2.3.2 Institution Contacting

The eligible institution sample for the NSoFaS:04 consisted of 1,630 institutions, of which 1,070 were sampled for NSOPF:04 as well as NPSAS:04. These 1,070 institutions were recruited to participate in both components of NSoFaS:04 (NSOPF:04 and NPSAS:04). The fielding of NSOPF:04 and NPSAS:04 together as the National Study of Faculty and Students was one of three changes made in the institution contacting procedures for this cycle of NSOPF.

The second change was to administer the institution questionnaire as a web or CATI instrument, with no hardcopy equivalent.

The third change was to begin recruiting institutions and initiating coordinator contacts in March 2003—a full 8 months prior to the November reference date for the fall term, and roughly 5 to 6 months earlier than the September start dates of previous cycles. This change was prompted by the need to draw a faculty sample and subsequently contact sampled faculty for participation prior to the 2004 summer break. It was hoped that the additional lead time would allow schools to better plan for the staffing and resources required for participation within the study's schedule constraints, allow institutions additional time to initiate and complete any internal review procedures they felt necessary, and also allow the contractor time to work with institutions to resolve any potential roadblocks to their participation. This advance notification was intended both to speed up receipt of faculty lists, and to positively impact the institution response rate. By sampling and contacting faculty earlier in the academic year, it was hoped that a higher faculty response rate could be achieved.

Prior to the field test, endorsements from organizations that had previously endorsed NSOPF and/or NPSAS were renewed and extended, as appropriate, to both NSoFaS:04 component studies. An effort was also made to solicit new endorsements from other organizations as well. In all, 25 organizations endorsed both components of NSoFaS:04; 24 of these were relevant to NSOPF:04.[18] These endorsements were featured on all project letterhead and pamphlets and on the NSoFaS website. In addition, several of these organizations continued to promote the study throughout the data collection period in newsletters and other communications with their member institutions. See appendix E for a list of the 24 organizations that endorsed NSOPF:04.

For NSOPF:04, data collection proceeded in four stages:

- verification;
- institution recruitment;
- advance notification of the coordinator; and
- faculty list and institution questionnaire data collection procedures.

Procedures for each stage of data collection are outlined below.

Verification

Verification began on January 23, 2003, and was completed prior to the start of institution recruitment on March 10, 2003. Institution contactors were trained to contact the institution at their main number, verify address information and confirm the name and contact information for the CA at the institution. They also confirmed that the school was Title IV eligible and open to the general public during the fall 2003 term.

Institutions flagged as potentially ineligible—including closed institutions and institutions that indicated they were not Title IV eligible or open to the general public—were forwarded to project staff for review. Project staff also reviewed instances of sampled institutions merging with other institutions (sampled or unsampled), possible changes in mission that could affect the institution's sampling strata, and changes in name or address, to confirm the institution was eligible and correctly identified.

Institution recruitment and advance notification of the coordinator

Institution recruitment began on March 10, 2003. The Chief Administrator (CA) at each institution sampled for NSoFaS:04 was sent the following materials (see appendix F for copies of these letters and pamphlets):

- a cover letter, printed on NCES letterhead, providing background information on NSOPF:04 and NPSAS:04;
- an NSoFaS:04 pamphlet summarizing the objectives of both NPSAS:04 and NSOPF:04, and providing background information and selected findings for each component;

[18] One of the 25 organizations, associated with for-profit schools ineligible for NSOPF, was asked only for an endorsement for NPSAS.

- an NSOPF:04 pamphlet, included to show what had been prepared for mailing to the sampled faculty;
- a NPSAS:04 pamphlet, included to show what had been prepared for mailing to sampled students; and
- a project timeline outlining the flow of activities for both component studies of NSoFaS:04 , and the projected schedule for each.

A team of institution contactors followed up with the CA by telephone. The CA was asked to name an Institution Coordinator (IC) by completing the Designation of Coordinator form online, or providing the information over the telephone. Once the IC was identified, they were mailed an identical packet, with a cover letter informing them that they would be mailed complete instructions for their participation in each component in September.

During this advance notification stage of data collection, ICs were asked to complete an online Coordinator Response Form (CRF) which could also be administered by CATI (see appendix F). This instrument confirmed that the institution could supply the items requested for the faculty and student lists within the stated schedule constraints. It also contained items designed to expedite collection of student record information for the student component.

ICs who indicated that a formal review process (such as an Institutional Review Board [IRB] review) was necessary before their institution would agree to participate were forwarded additional project materials as appropriate. A complete IRB approval packet was prepared for this purpose and mailed to the IC upon request. This packet included copies of instruments, as well as complete descriptions of relevant survey procedures (e.g., confidentiality and informed consent).

Faculty list collection procedures

Complete instructions for participation in both NSOPF:04 and NPSAS:04 were sent to all designated ICs on September 29, 2003. Binders continued to be mailed to ICs on a flow basis as they were designated. The mailing, which was packaged in a three-ring binder, included the following materials:

- a cover letter describing the study, the institution's password, IPEDS unit ID,[19] and URL (web address) necessary to access the NSoFaS:04 website (a separate letter was created for NPSAS:04-only sampled institutions);
- a copy of the letter that went to the CA, and a facsimile of the Designation of Coordinator form;
- a complete list of endorsements;
- a project timeline outlining the flow of activities for both component studies of NSoFaS:04 , and the projected schedule for each;
- instructions for preparing the list of faculty and instructional staff, including a list of data elements requested, and a suggested file layout;

[19] Chief Administrators and Institution Coordinators used their institution IPEDS unit ID and a password to authenticate to the institution website. Faculty and instructional staff were assigned a study ID and password to authenticate to the faculty website.

- complete instructions for participation in each phase of NSoFaS:04; and
- a list of transmittal options for sending faculty lists, by mail, e-mail, and direct upload to the NSoFaS:04 website, together with an express courier packet and label for mailing the lists if required.

The instructions directed the ICs to provide a list of full- and part-time faculty and instructional staff, including all personnel who had faculty status or any instructional responsibilities during the fall 2003 term. Institutions were encouraged to submit an electronic list by uploading it to the secure website. The data items requested for each listed faculty or instructional staff member were

- full name;
- academic discipline;
- department/program affiliation;
- full-time/part-time status;
- gender;
- race/ethnicity;
- employee ID number (to eliminate duplicates from sample); and
- contact information (institution and home mailing address, institution and home e-mail address [if available], and home and campus telephone numbers).

Follow-up with ICs was conducted by telephone, mail, and e-mail. Telephone prompts to the ICs were made for institutions that had not provided lists. To minimize the number of contacts made to an IC, prompting for NSOPF:04 was combined with prompting for NPSAS:04. E-mail prompts to ICs, keyed to pending project deadlines, were regularly utilized. E-mail prompts focused on timely completion of requested materials and encouraged review of the instructions for participation. As faculty lists were received, they were reviewed for completeness, readability, and accuracy. Additional follow-up to clarify the information provided or retrieve missing information was conducted by the institution contactors as necessary.

Counts of full- and part-time faculty were collected in both the institution questionnaire and in the faculty lists. For each institution, the counts of full- and part-time faculty were checked against those provided in the institution questionnaire and against 2001 IPEDS Fall Staff Survey data. IPEDS data were used for discrepancy checks whenever institution questionnaire data were unavailable but also served as an additional check to catch inaccuracies in matching questionnaire/list data that otherwise would not have been discovered. For further details regarding quality control checks, see section 4.1.2.

Reimbursement for the time and staff involved in providing the faculty list was offered to institutions indicating a difficulty in complying with the request within schedule constraints. A refusal conversion letter was mailed to institutions that had not responded by November 21, 2003. The letter underscored the offer of reimbursement. Beginning in May 2004 a flat $500 reimbursement was offered to institutions for providing the outstanding faculty and student lists

by the end of June. This offer was extended both to explicit refusals and schools which indicated cooperation but had yet to comply.

For institutions lacking the resources to provide a complete list of full- and part-time faculty despite the offer of reimbursement, list information was, if possible, abstracted from course catalogs, faculty directories, and other publicly available sources. Those institutions for which usable lists were identified were notified of this sampling procedure; institutions which indicated that they did not want their faculty included in the sample were excluded. Faculty lists abstracted in this fashion were reviewed for completeness against IPEDS before being approved for sampling. Faculty list collection continued through July 11.

Institution questionnaire data collection procedures

Institution Coordinators were asked to complete the institution questionnaire (described in section 2.2.3) using the study's institution website. Institution questionnaire follow-up was conducted simultaneously with follow-up for lists of faculty. If an institution was unable to complete the questionnaire online, efforts were made to collect the information over the telephone. This often involved contacting multiple offices within the institution, as questions about benefits and tenure policies could most frequently be completed by human resources and/or the academic affairs office, while questions about faculty counts and turnover were typically answered by institutional research staff.

To expedite data collection, missing questionnaire data was, in some instances, abstracted directly from benefits and policy documentation supplied by the institution, or publicly available on the institution's website. In addition, several large multi-campus systems provided data for their campuses at a system level or indicated that specific policy and benefits information was the same for all related campuses.

Refusal conversion efforts for the institution questionnaire were conducted with institutions regardless of whether they supplied a list of faculty. After August, institutions which had not completed the questionnaire were offered a reimbursement of $50 for providing the questionnaire within schedule constraints. Data collection for the institution questionnaire closed on October 22, 2004.

Administrative systems and procedures

To efficiently track all mail and telephone follow-up (both incoming and outgoing) and processing and sampling activities, the study utilized an Institution Contacting System (ICS) specifically designed to meet the needs of the NSoFaS:04 project. The ICS was accessible to contactors, Call Center[20] supervisors, and project staff. The NSoFaS:04 ICS was designed so that a change in status (e.g., a completed Designation of Coordinator form) automatically generated the next step (e.g., a mailout to the IC and an automatic appointment for telephone follow-up). Electronic call notes documented the outcome of every conversation. The system allowed interviewers to set appointments for future follow-up. Through the ICS, the interviewer had the ability to designate an IC, provide contact information, and access the institution questionnaire and other data collection instruments. The ICS gave interviewers the ability to generate an automatic e-mail to ICs containing the password and IPEDS unit ID required for access. The

[20] RTI's Call Center Services provides telephone, web, and tracing services for a wide variety of projects, and operates two call centers: one in Raleigh, NC, and one in Greenville, NC.

problem report form feature of the ICS allowed institution contactors to immediately forward specific call notes to an e-mail box monitored by project staff. This ensured that refusals, requests for remails, and calls requiring follow-up by project staff were handled promptly.

Quality Circle meetings, attended by interviewers, supervisors, team leaders, and project staff, were held on a weekly basis to share ideas for gaining institutional cooperation and suggestions for improving procedures. Project staff solicited feedback from call center personnel on the ICS, scripts, and handling problems reported by respondents (e.g., difficulties accessing the website).

2.4 Faculty Data Collection

The NSOPF:04 utilized a mixed-mode data collection methodology that allowed sample members to participate either by web-based self-administered questionnaire or via an interviewer-administered telephone interview. At the start of faculty data collection, introductory materials were sent to sample members via first class mail as well as electronic mail (if an e-mail address was available). The initial letter included instructions for completing the self-administered questionnaire on the Internet or by calling a toll-free number to complete a telephone interview. After an initial 4-week period, telephone interviewers began calling sample members. The self-administered web instrument remained available to respondents throughout data collection. An early-response incentive, designed to encourage sample members to complete the self-administered questionnaire prior to outgoing CATI calls, was offered to sample members who completed the questionnaire within 4 weeks of the initial mailing. Incentives were also offered to selected sample members as necessary (i.e., those who refused and other nonrespondents).

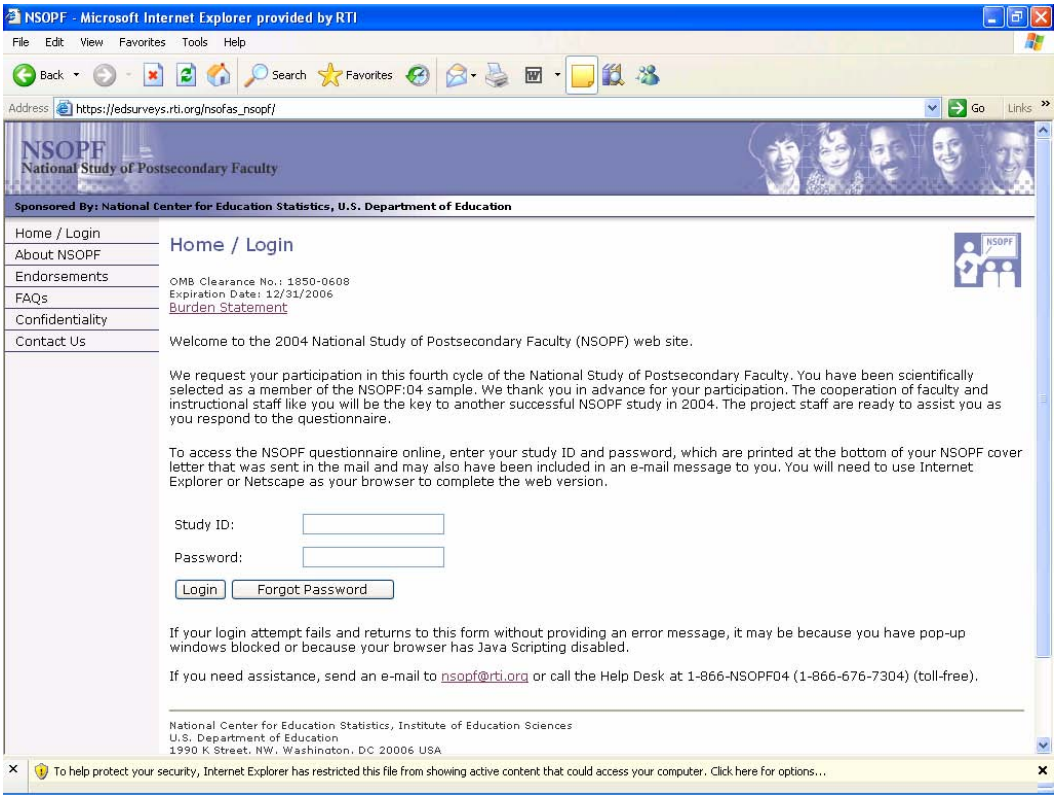
2.4.1 Faculty Website

The website for the NSOPF:04 served a dual purpose. The primary function was to provide access to the web questionnaire for the sampled faculty and instructional staff. The secondary function was to provide information about the study, the selected sample, the sponsor, the contractor, and confidentiality. In addition to the information available on the site, links were provided to other relevant sites (e.g., NCES). The home page of the NSOPF:04 faculty website is depicted in figure 3.

The initial login page provided access to the self-administered questionnaire. The login process involved entering a specific study ID and password, which were provided to the respondent in every letter and e-mail message. Respondents could also obtain their study ID and password by sending an e-mail to the project, or by contacting a help desk agent at the NSOPF:04 toll-free number.

As with the institution application, the web instrument was protected by SSL encryption, an automatic time out feature, and omission of any persistent cookies.

Figure 3. The NSOPF:04 faculty website home page

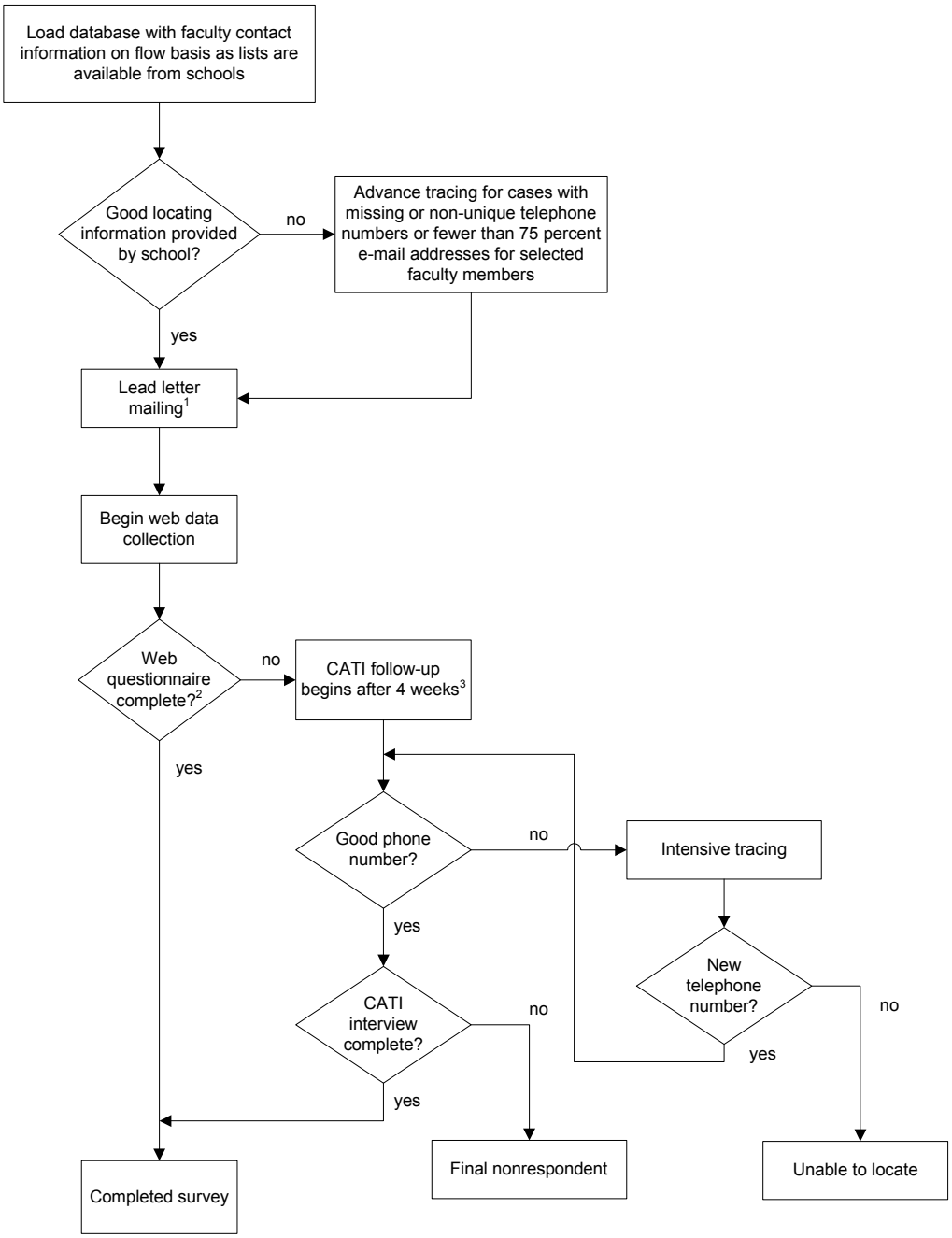


SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

2.4.2 Locating and Interviewing Procedures

The NSOPF:04 faculty data collection design involved locating sample members, providing an opportunity for the faculty or instructional staff to complete the self-administered questionnaire, and following up with web nonrespondents after 4 weeks to conduct a computer-assisted telephone interview. The data collection period lasted approximately 9 months (January 15 through October 6, 2004). Data collection activities for faculty are shown in figure 4.

Figure 4. NSOPF:04 faculty data collection overview



[1] If a home address was available for the sample member, the lead letter package was mailed to the home. If there was no home address, the package was mailed to the school address. If there was no specific school address available, the package was mailed to the main address on file for the school. Sending packages to the home address resulted in a higher response rate compared to sending packages to the school address (78 percent versus 67 percent; $\chi 2 = 565.6$, $p < .0001$).
[2] The web interview option was available throughout data collection, even after telephone follow-up began.
[3] The sample member's office and home telephone numbers were called by computer-assisted telephone interview (CATI) interviewers. If no specific telephone number was available for the sample member, the school's main telephone number was used.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Mailouts

Faculty and instructional staff were sent a lead letter, instructions for accessing the web instrument via the Internet or with the assistance of a telephone interviewer, and a study pamphlet. (Examples of these materials are included in appendix F.) The lead letter introduced the study and listed the organizations that endorsed the study. If an e-mail address was available for a sample member, the introduction to the study was also sent via e-mail.

Periodically throughout the data collection period, reminder letters and e-mail messages were sent to nonrespondents to encourage their participation and to notify them of the incentive, if applicable. Examples of these follow-up contacts are included in appendix F.

Locating

Identifying a valid mailing address and telephone number for all selected faculty and instructional staff sampled from known institutions was critical to the success of the NSOPF:04. Locating activities were conducted in two stages: advance tracing, which took place before data collection began, and intensive tracing conducted during data collection.

Advance tracing. Upon receipt of faculty lists from participating institutions, contact information for the sampled faculty and instructional staff was reviewed and assessed for completeness. Schools for which fewer than 75 percent of the sampled cases had e-mail addresses (n = 430) were selected for tracing before being sent a lead letter. Prior to CATI operations, home contact information was sent to Telematch to obtain the latest telephone numbers.

Initial tracing efforts included searches on the school's website for contact information. When this was not an option, more extensive database searches were employed during intensive tracing. In some cases, the searches confirmed or updated the contact information provided by the institution; in other cases, the searches resulted in new contact information. All locating information obtained as a result of these searches was loaded into the NSOPF:04 database.

Intensive tracing. Intensive tracing was performed on a case if advance tracing did not yield a telephone number for loading in CATI, or if the case was designated as a dead end in CATI (i.e., there were no more telephone numbers to call for the case). The following steps were performed by the tracing unit to locate sample members.

- Check the preloaded information using an online directory assistance search. This step was intended to identify the easy-to-locate cases (e.g., cases with the correct telephone number but the wrong area code).
- Conduct credit bureau database searches. The tracing staff had access to various proprietary databases (TransUnion, Equifax, and Experian) containing current address and phone listings for the majority of consumers with a credit history.
- Conduct additional intensive tracing. This step included (but was not limited to) searches using Lexis-Nexis and FastData, directory assistance calls, and searches of institution websites for campus directories.

Tracing staff checked all new leads procured during their tracing efforts to confirm the addresses and telephone numbers that were obtained. When a telephone number for a sample member was confirmed, telephone interviewing resumed for that case. Cases with new address

information were mailed a lead-letter packet. If the tracing staff located a new e-mail address for a sample member, the information was loaded into the database for future e-mail reminders and other mailings to nonrespondents.

Staff training

The mixed-mode design of the NSOPF:04 data collection required the development of three separate training programs for data collectors: help desk training, CATI interviewer training, and tracing. In addition, separate training sessions were conducted for supervisors and monitors.

Detailed NSOPF:04 interviewer manuals were distributed at the outset of each training session. These manuals served as both an instruction guide for the training lectures, discussions, and practical exercises and as a reference guide for use after completion of training. Supplemental chapters that covered additional duties were provided for supervisors, monitors, and help desk agents. The manual's table of contents and an agenda for telephone interviewer training are included in appendix G.

All training sessions included a study overview, a review of the confidentiality requirements, a demonstration interview, an in-depth review of the instrument, hands-on practice exercises with the instrument, and open-ended coding modules. In addition, the help desk and telephone interviewer training sessions included the following additional topics:

- *Help desk agents* reviewed the "frequently asked questions" in detail, with a focus on responses to technical issues as well as instrument-specific questions, and instructions for documenting each call to the study hotline.
- *Telephone interviewers* were trained in techniques for gaining cooperation of sample members, and of other contacts, as well as techniques for addressing the concerns of reluctant participants and for avoiding refusals.

Self-administered questionnaires

The first phase of data collection, lasting 4 weeks after the lead letters were mailed, provided an opportunity for respondents to complete the self-administered questionnaire via the Internet before the telephone follow-up calls began. The web interview site remained available 24 hours a day, 7 days a week, thereby giving sample members the option to complete the questionnaire online during the entire 9 months of data collection.

Help desk operations

The NSOPF:04 help desk opened on January 15, 2004, in anticipation of the first respondent calls after the lead-letter mailing. The help desk staff were available to assist sample members who had questions or problems accessing and completing the self-administered questionnaire. A toll-free hotline was set up to accept incoming help desk calls. If technical difficulties prevented a sample member from completing the self-administered questionnaire, a help desk staff member, also trained to conduct telephone interviews, would encourage the caller to complete a telephone interview rather than to attempt the self-administered questionnaire.

All incoming calls from sample members were documented using the help desk software. In addition to this primary documentation function, the software provided information needed to

verify a sample member's identity, login information (study ID and password) for the web questionnaire, and a means for tracking calls that could not be resolved immediately.

The help desk software also provided project staff with reports on the types and frequency of problems experienced by sample members, as well as a way to monitor the resolution status of all help desk inquiries.

Telephone interviewing

Telephone prompts to nonrespondents began on February 12, 2004, at the end of the early-response incentive period. CATI procedures included attempts to locate, gain cooperation from, and interview study sample members who had not completed the questionnaire online. Interviewers encouraged respondents to complete the interview by telephone as soon as they made contact. However, if the sample member expressed a preference for completing the self-administered questionnaire via the Internet, a callback was scheduled for 1 week later. During these callbacks, interviewers again prompted the faculty members to complete the questionnaire by telephone.

Refusal conversion procedures were used to gain cooperation from individuals who refused to complete the questionnaire. When a refusal was first encountered, either because the sample member refused or because a "gatekeeper" (secretary or spouse) refused on behalf of the sample member, the case was referred to a refusal conversion specialist. Refusal conversion specialists were selected from among those interviewers most skilled at obtaining cooperation and were given training in refusal conversion techniques tailored to NSOPF:04. The refusal training emphasized ways to gain cooperation, overcome objections, address the concerns of gatekeepers, and encourage participation.

2.5 Data Collection Systems

2.5.1 Instrument Development and Documentation System

The Instrument Development and Documentation System (IDADS) is a controlled web environment in which project staff developed, reviewed, modified, and communicated changes to specifications, code, and documentation for the NSOPF:04 instrument. All information relating to the NSOPF:04 instrument was stored in a Structured Query Language (SQL) Server database and was made accessible through Windows and web interfaces. There are three modules within IDADS: specification, programming, and documentation.

Initial specifications were generated within the IDADS *specification module*. This module enabled access for searching, reviewing, commenting on, updating, exporting, and importing information associated with instrument development. All records were maintained individually for each item, which provided a historical account of all changes requested by both project staff and NCES.

Once specifications were finalized, the *programming module* within IDADS produced hypertext transfer markup language (html), Active Server Pages (ASP), and JavaScript template program code for each screen based on the contents of the SQL Server database. This output included screen wording, response options, and code to write the responses to a database, as well as code to automatically handle such web instrument functions as backing up and moving forward, and recording timer data. For questions that had changed significantly since the field

test, the programming staff edited the automatically generated code to customize screen appearance and program response-based routing. For questions with minor changes, the programming staff simply modified the program code used in the field test.

The *documentation module* contained the finalized version of all instrument items, their screen wording, and variable and value labels. Also included were the more technical descriptions of items such as variable types (alpha or numeric), information regarding to whom the item was administered, and frequency distributions for response categories. The documentation module was used to generate the instrument facsimiles and the Electronic Codebook (ECB) input files.

2.5.2 Integrated Management System

All aspects of the study were under the control of an Integrated Management System (IMS). The IMS was a comprehensive set of desktop tools designed to give project staff and NCES access to a centralized, easily accessible repository for project data and documents. The NSOPF:04 IMS consisted of three components: the management module, the Receipt Control System (RCS), and the Case Management System (CMS).

The *management* module of the IMS contained tools and strategies to assist project staff and the NCES project officer in managing the study. All information pertinent to the study was located there, accessible via the Internet, in a secure desktop environment. Available on the IMS website were the project schedule, monthly progress reports, daily data collection reports and status reports (available through the RCS described below), project plans and specifications, project information and deliverables, instrument specifications, staff contacts, the project bibliography, a document archive, and frequencies for the faculty and institution data. The IMS management module also had a download area from which the client and subcontractors retrieved large files when necessary.

The *Receipt Control System* (RCS) was an integrated set of systems that monitored all activities related to data collection, including tracing and locating. Through the RCS, project staff were able to perform stage-specific activities, track case statuses, identify problems early, and implement solutions effectively. RCS locator data were used for a number of daily tasks related to sample maintenance. Specifically, the mailout program produced mailings to sample members, the query system enabled administrators to review the locator information and status for a particular case, and the mail return system enabled project staff to update the locator database. The RCS also interacted with the Case Management System and tracing unit databases, sending locator data among the three systems as necessary.

The *Case Management System* (CMS) was the technological infrastructure that connected the various components of the CATI system, including the questionnaire, utility screens, databases, call scheduler, report modules, links to outside systems, and other system components. The call scheduler assigned cases to interviewers in a predefined priority order. In addition to delivering appointments to interviewers at the appropriate time, the call scheduler also calculated the priority scores (the order in which cases need to be called based on preprogrammed rules), sorted cases in non-appointment queues, and computed time zone adjustments to ensure that the sampled respondents were not phoned outside the specified calling

hours.[21] The call scheduler also allowed callbacks to be set and assigned status codes to the case. Using an algorithm based on the previous call results, the call scheduler determined which telephone number (e.g., home or work) associated with the case should be called next.

[21] Call Center hours were 9:00 a.m. to 11:00 p.m. Monday through Friday, 9:00 a.m. to 5:00 p.m. Saturday, 1:30 p.m. to 9:30 p.m. Sunday, Eastern Standard Time. The CMS was programmed to account for time zones such that respondents would not be called after 9:00 p.m. local time. Work numbers were only called 8:00 a.m. to 5:00 p.m. Monday through Friday, local time.

Chapter 3
Data Collection Outcomes

The success of the 2004 National Study of Postsecondary Faculty (NSOPF:04) was dependent upon achieving high levels of cooperation at all stages of the data collection process. The data collection results—namely the institution and faculty response rates, along with the results of the efforts that contributed to those rates—are the focus of this chapter.

3.1 Institution Data Collection Results

3.1.1 Institution Participation

Of the 1,080 institutions selected to participate in NSOPF:04, 1,070 were eligible institutions.[22] Of the eligible institutions, 97 percent (unweighted) appointed an Institution Coordinator (IC) to assist with study requirements and 85 percent completed the Coordinator Response Form (CRF), indicating their initial intent to participate in both components of the study and adhere to project timelines. Ultimately, 91 (unweighted) percent of the eligible institutions provided a list of faculty and 86 percent completed institution questionnaires.

Fifty-seven institutions indicated having policies that required the 2004 National Study of Faculty and Students (NSoFaS:04) survey request be submitted to their Institutional Review Board (IRB) for formal approval. One advantage of the advance notification period is that it allowed the contractor sufficient time to prepare customized IRB approval packages for submission to each of these institutions. This procedure expedited the approval process and alleviated the burden on the IC. Of the 60 institutions that were sent IRB approval packages, all but three approved participation in NSOPF:04.

Faculty lists

Two key changes in data collection procedures had the potential to impact faculty list participation rates for NSOPF:04; namely the advance notification initiative begun in March 2003, and the decision to combine the data collection efforts for NSOPF:04 with the 2004 National Postsecondary Student Aid Study (NPSAS:04) under the NSoFaS:04. Table 9 compares the participation rate achieved in NSOPF:04 with previous cycles.

[22] Ineligible institutions included institutions treated as mergers and reported for by other institutions, closed institutions, and institutions that did not meet eligibility requirements. Numbers of institutions are rounded to the nearest 10.

Table 9. Number and percentage of institutions providing faculty lists, by cycle of National Study of Postsecondary Faculty (NSOPF): 1988 to 2004

NSOPF cycle	Number of eligible institutions	Number of institutions providing list	Unweighted percent participation rate[1]
NSOPF:88 field test	110	100	91.4
NSOPF:88 full-scale study	480	450	93.5
NSOPF:93 field test	140	120	89.0
NSOPF:93 full-scale study	960	820	84.9
NSOPF:99 field test	160	150	90.1
NSOPF:99 full-scale study	960	820	85.4
NSOPF:04 field test	150	130	89.9
NSOPF:04 full-scale study	1,070	980	91.3

[1] Percentages are based on the number of eligible institutions within the row under consideration, and are based on original unrounded numbers.
NOTE: Numbers of eligible institutions and numbers of institutions providing lists are rounded to the nearest 10.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

A total of 980 (91 percent, unweighted and weighted) of eligible institutions provided a faculty list, with all institutional strata exceeding a weighted participation rate of 85 percent. The breakdown of institutions providing faculty lists, by institution type, is presented in table 10.

Table 10. Number and percentage of institutions providing faculty lists, by type of institution: 2004

Institution type	Number of institutions		Percent participation rate[1]	
	Eligible	Participating	Unweighted	Weighted
Total	1,070	980		
Public doctoral	190	180	92.7	93.2
Public master's	120	100	89.7	89.1
Public bachelor's	30	30	92.9	88.4
Public associate's	330	290	89.1	87.4
Public other	10	10	100.0	100.0
Private not-for-profit doctoral	110	100	92.0	95.6
Private not-for-profit master's	80	80	92.6	86.8
Private not-for-profit bachelor's	130	120	94.6	93.1
Private not-for-profit associate's	10	10	75.0	96.0
Private not-for-profit other	60	60	93.3	91.8

[1] Percentages are based on the number of eligible institutions within the row under consideration, and are based on original unrounded numbers.
NOTE: Number of eligible and participating institutions are rounded to the nearest 10. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Institution questionnaire

A total of 920 institutions, representing 84 percent of eligible institutions, completed the institution questionnaire. Table 11 provides a breakdown of institution participation by strata.

Table 11. Number and percentage of institutions providing institution questionnaires, by type of institution: 2004

Institution type	Number of institutions		Participation rate percent[1]	
	Eligible	Participating	Unweighted	Weighted
Total	1,070	920	86.1	84.2
Public doctoral	190	170	86.5	84.7
Public master's	120	110	90.5	89.6
Public bachelor's	30	30	100.0	100.0
Public associate's	330	290	89.1	83.6
Public other	10	10	87.5	98.9
Private not-for-profit doctoral	110	90	80.4	83.7
Private not-for-profit master's	80	70	81.5	79.8
Private not-for-profit bachelor's	130	110	83.8	77.7
Private not-for-profit associate's	10	10	75.0	86.0
Private not-for-profit other	60	50	76.7	76.2

[1] Percentages are based on the number of eligible institutions within the row under consideration, and are based on original unrounded numbers.
NOTE: Number of eligible and participating institutions are rounded to the nearest 10. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Table 12 shows the breakdown of completed institution questionnaires by mode of administration. Those completed in computer assisted telephone interview (CATI) mode include instances where the questionnaire was finalized with interviewer assistance (e.g., questionnaires wholly or partially data-entered by project staff from information supplied on hardcopy by the institution) and questionnaires completed, wholly or partly, by CATI. Web completions are defined as those questionnaires transmitted as complete by the institution, although some institutions may have provided some responses in CATI. Nearly 81 percent of institutions completed the institution questionnaire using the Web, and 19 percent completed it with the help of an interviewer. By comparison, in 1999, the only previous cycle in which a web questionnaire was used, 69 percent of the questionnaires were done on paper, and 31 percent were done on the Web. The percentage of interviews completed at least in part by CATI is fairly consistent with the number of questionnaires completed with interviewer assistance in previous cycles.

Table 12. Number and percentage of institutions providing institution questionnaires, by mode of administration: 2004

Mode	Number of participating institutions	Unweighted response rate[1]
Total	920	100.0
Web	740	80.5
CATI	180	19.5

[1] Percentages are based on the total number of participating institutions, and are based on original unrounded numbers.
NOTE: Number of participating institutions are rounded to the nearest 10. Detail may not sum to totals because of rounding. CATI = computer assisted telephone interview.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

3.1.2 Institution Survey Completion Timing

The timing analysis was conducted by embedding time stamps in the programming code for each form (screen) in the survey. From these time stamps, the number of seconds spent on each screen (on-screen time) and the transit time between screens (i.e., the time required to transmit data to the server, for the server to store the data and assemble the next page, and for the page to be transmitted and loaded on the computer) were calculated. A cumulative on-screen time and a cumulative transit time for the institution survey also were calculated from the time stamps. The sum of the cumulative on-screen and transit times was the total instrument time (i.e., the number of minutes it took to administer the institution questionnaire).

Unlike most questionnaires, which require the respondent to complete the survey in sequential order, the institution questionnaire included a status screen that allowed respondents to jump to particular questions they could answer, and skip over ones they could not answer. For most institutions, the questionnaire was completed in multiple internet sessions and, in some cases, by multiple people at the institution.

The target time to complete the institution questionnaire was 45 minutes. Based on the time stamps for each form, the actual time to complete the questionnaire ranged from less than 1 minute to 4 hours and 9 minutes, with an average of 35 minutes.[23] Of these 35 minutes, approximately 31 minutes, on average, were spent answering questions (on-screen time) and 5 minutes, on average, were spent in transit. These numbers may be misleading because some institutions may have completed the sample hardcopy version of the questionnaire in advance, so their time to complete the web questionnaire simply reflected the time it took to key in their responses.

Table 13 reports the average and maximum times (in seconds) to complete each form in the institution instrument. Ten forms (screens) of the institution survey took more than 1 minute to complete, on average. Each of these forms required the respondent to look up information and/or requested several pieces of information, which accounts for the longer times on these screens.

[23] The average time excludes 28 cases with unexplained negative transit times. Some very short survey completion times may be attributed to institutions who answered a small subset of items in the questionnaire. Very long survey times may be the result of the respondent timing out repeatedly while in the questionnaire (e.g., answering the phone).

Table 13. Average and maximum completion time, in seconds, for forms in the institution questionnaire: 2004

Form	Description	Time in seconds		Number of cases
		Average	Maximum	
I1	Number full-/part-time faculty, fall 2003	211	1,319	920
I1B	Have full-/part-time faculty, fall 2003	25	166	200
I2	Changes in number of full-time faculty	252	1,585	920
I2A	Reason for discrepancy, I1A and I2G	84	538	240
I3	Full-time tenure: has tenure system	26	434	920
I4	Full-time tenure: number considered for tenure, 2002-03	62	620	750
I5	Full-time tenure: number granted tenure, 2002-03	16	164	650
I6	Full-time tenure: maximum years on tenure track	39	399	740
I7	Full-time tenure: institution actions, last 5 years	55	436	740
I7SP	Full-time tenure: number early retirees, last 5 years	53	586	400
I8	Full-time tenure: discontinued tenure system, last 5 years	15	125	190
I9	Full-time faculty: positions sought to fill, fall 2003	45	403	910
I10A	Full-time faculty: benefits available	111	1,091	910
I10B	Full-time faculty: benefits subsidized	58	443	900
I11	Full-time faculty: other benefits available	80	427	910
I12	Full-time faculty: union representation	20	408	910
I13	Full-time faculty: teaching assessment	64	457	910
I14	Part-time benefit: retirement plan	38	435	910
I15A	Part-time faculty: benefits available	61	479	910
I15B	Part-time faculty: benefits subsidized	35	319	530
I16	Part-time faculty: other benefits available	52	327	910
I17	Part-time faculty: union representation	14	200	910
I18	Part-time faculty: teaching assessment	41	320	910
I19	Undergraduate instruction: percent assignment	126	941	920
I20	Comments/suggestions	143	785	920

NOTE: The number of cases per form varies due to the interview skip logic. Outliers for each form were top coded (mean + 3 standard deviations). Numbers of cases are rounded to the nearest 10.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

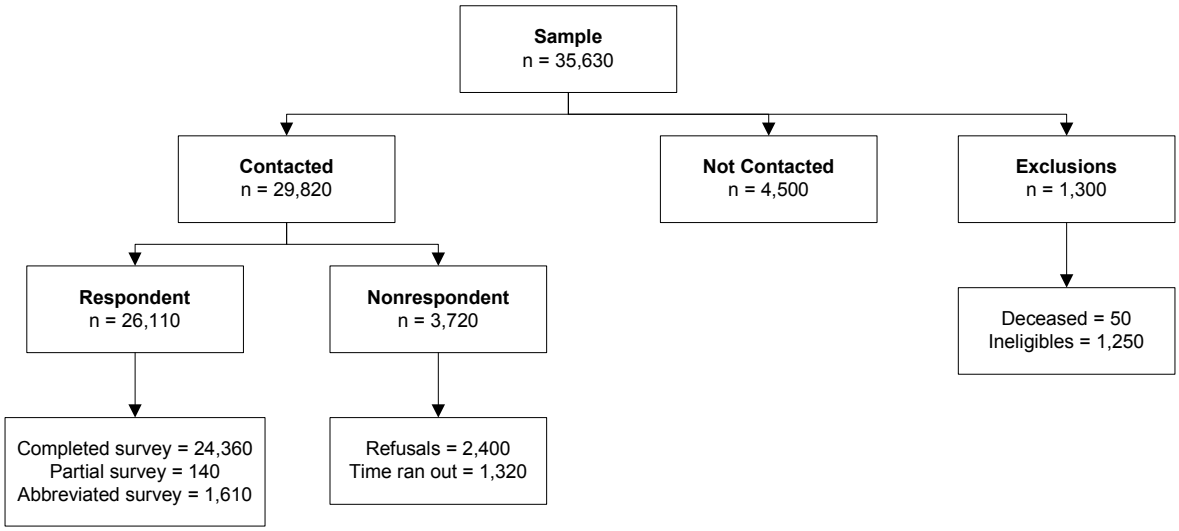
3.2 Faculty Data Collection Results

Faculty data collection efforts for NSOPF:04 consisted of three essential steps: locating (identifying telephone numbers and addresses for sample members), contacting (carrying out the necessary steps to reach the faculty member), and encouraging survey completion by web-based self-administration or CATI. This section describes the results of the NSOPF:04 data collection effort and evaluates the effectiveness of the data collection procedures used in locating, contacting, and interviewing sample members.

3.2.1 Response Rate

Overall contacting and survey completion results for the faculty contact phase of NSOPF:04 are presented in figure 5. Of the 35,630 cases in the original sample, 1,300 (4 percent) were excluded because they were ineligible for the study or deceased. Of the 34,330 eligible sample members, 29,820 (87 percent) were contacted and 26,110 completed the survey, for an unweighted and weighted response rate of 76 percent achieved in the 9-month period from January 15 to October 6, 2004.

Figure 5. Contacting and survey completion outcomes: 2004



NOTE: Numbers rounded to the nearest 10. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Table 14 shows the response rates of faculty by institution type. Response rates range from 67 percent (weighted) of faculty at public bachelor's degree-granting institutions to 91 percent (weighted) of faculty at private not-for-profit associate's degree-granting institutions.

Table 14. Number of sampled, eligible, and responding faculty and response rates, by institution type: 2004

	Faculty			Percent response rate[2]	
Institution type	Sampled	Eligible[1]	Responding[1]	Unweighted	Weighted
Total	35,629	34,330	26,110	76.1	75.6
Institution level					
2-year	9,188	8,830	6,440	73.0	73.7
4-year non-doctorate-granting	8,747	8,430	6,720	79.7	78.6
4-year doctorate-granting	17,694	17,070	12,950	75.8	75.0
Institution control					
Public	23,280	22,450	17,120	76.2	76.0
Private not-for-profit	12,349	11,880	8,990	75.7	74.7
Institution type					
Public doctoral	9,827	9,500	7,460	78.6	78.1
Public master's	3,485	3,350	2,620	78.1	78.5
Public bachelor's	693	680	510	75.4	67.4
Public associate's	9,129	8,770	6,420	73.1	73.7
Public other	146	140	110	73.6	73.3
Private not-for-profit doctoral	4,652	4,470	3,160	70.7	68.2
Private not-for-profit master's	3,020	2,890	2,270	78.6	78.5
Private not-for-profit bachelor's	3,218	3,120	2,520	80.8	78.7
Private not-for-profit associate's	242	240	190	79.8	91.0
Private not-for-profit other	1,217	1,160	850	73.1	70.6

[1] Numbers rounded to the nearest 10.
[2] Percentages are based on the number of eligible faculty within the row under consideration. Percentages are based on original unrounded numbers.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Table 15 presents faculty response rates by when the lead letter package was mailed. Response rates range from 56 percent for one of the later mailings to 80 percent for the first mailing.

Table 15. Faculty response rates, by date lead letter package was mailed: 2004

Date mailed	Number eligible[1]	Percent response rate[2]
Total	34,330	76.1
January 14-February 13, 2004	18,690	80.3
February 14-March 13, 2004	7,910	76.3
March 14-April 13, 2004	1,230	74.6
April 14-May 13, 2004	3,330	67.1
May 14-June13, 2004	1,520	55.9
June 14-July 13, 2004	1,260	62.7
July 14-July 21, 2004	390	67.5

[1] Numbers rounded to the nearest 10.
[2] Percentages are based on the number of eligible faculty within the row under consideration. Percentages are based on original unrounded numbers.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

3.2.2 Locating and Survey Completion

Most of the faculty lists provided by the institutions contained contact information for sample members, including the sample member's name, office telephone number, school name, school address, and department. For some cases, home addresses also were provided. In addition, a number of approaches were used to locate faculty and instructional staff, including advance tracing, the initial mailing to all sample members, follow-up letters and e-mails to nonrespondents, telephone tracing (interviewers calling telephone numbers provided on the faculty lists as well as any additional numbers obtained during the course of making those calls), and intensive tracing (i.e., using consumer databases, internet searches, and criss-cross directories).

Before the start of data collection, schools' faculty lists were assessed for completeness of contact information. As necessary, advance tracing, described in section 2.4.2, was conducted. As shown in table 16, the contact information provided by the school proved effective in contacting faculty and instructional staff; 83 percent of sample members required no intensive tracing, while the remaining 17 percent required intensive tracing. Intensive tracing was required when a case did not have a telephone number associated with it or the CATI calls had exhausted all numbers for the case without reaching the sampled individual. Approximately 52 percent of cases sent to intensive tracing were located, compared to 90 percent of cases not sent to intensive tracing. Further, only 40 percent of cases sent to intensive tracing completed an interview compared with 80 percent of cases not sent to intensive tracing.

Table 16. Locate and interview rates, by intensive tracing efforts: 2004

Intensive tracing status	Total	Located		Completed survey	
		Number[1]	Percent[2]	Number[1]	Percent[3]
Total	35,629	29,820	83.7	26,110	73.3
Intensive tracing required	5,943	3,080	51.8	2,360	39.7
No intensive tracing required	29,686	26,750	90.1	23,750	80.0

[1] Numbers rounded to the nearest 10.
[2] Percentages are based on the total within the row under consideration.
[3] Percentages are based on the number of eligible faculty within the row under consideration. Percentages are based on original unrounded numbers.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Table 17 provides an overview of the primary sources used by tracers during the intensive tracing process. Tracers generally used multiple sources when tracing a case, so no one source can be pinpointed as the one that resulted in the "locate." Among the sources used most frequently for intensive tracing were internet searches, directory assistance, and various consumer database searches.

Table 17. Locate rates, by intensive tracing source: 2004

Tracing source	Total	Located	
		Number	Percent[1]
Internet search	3,726	1,739	46.7
Directory assistance	3,529	1,730	49.0
Consumer database search—Lexis-Nexis	2,911	1,251	43.0
Reverse phone lookup—FastData	2,238	1,105	49.4
Name search—FastData	3,276	1,531	46.7
Address search—FastData	2,279	997	43.7
Neighbor search—FastData	14	4	28.6
Directory Assistance Plus—FastData	526	216	41.1
Consumer database search—TransUnion	2,374	1,131	47.6
Consumer database search—Experian search on Social Security number	1,394	723	51.9
Consumer database search—Experian address search	1,690	762	45.1
Other collateral source	1,714	768	44.8

[1] Percentages are based on the total within the row under consideration.
NOTE: Most cases were traced using multiple sources; therefore, row totals and percents are not mutually exclusive.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

The breakdown of faculty and instructional staff requiring intensive tracing, by faculty status and institution type, is presented in table 18. Thirty-two percent of part-time faculty required intensive tracing, compared to 7 percent for full-time faculty ($\chi^2 = 3806.9$, $p < .0001$). Seventeen percent of faculty at public institutions required intensive tracing compared to 16 percent at private not-for-profit institutions ($\chi^2 = 16.5$, $p < .0001$).

Table 18. Faculty and instructional staff requiring intensive tracing procedures, by employment status and institution type: 2004

Employment status and institution type	Total	Intensive tracing	
		Number	Percent[1]
Total	35,629	5,943	16.7
Employment status			
Full-time	21,891	1,544	7.1
Part-time	13,008	4,210	32.4
Unknown employment status	730	189	25.9
Institution control			
Public	23,280	4,019	17.3
Private not-for-profit	12,349	1,924	15.6
Institution type			
Public doctoral	9,827	951	9.7
Public master's	3,485	455	13.1
Public bachelor's	693	134	19.3
Public associate's	9,129	2,465	27.0
Public other	146	14	9.6
Private not-for-profit doctoral	4,652	733	15.8
Private not-for-profit master's	3,020	473	15.7
Private not-for-profit bachelor's	3,218	462	14.4
Private not-for-profit associate's	242	27	11.2
Private not-for-profit other	1,217	229	18.8

[1] Percentages are based on the number of sample members within the row under consideration.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

The results of faculty and instructional staff locating and survey completion, by faculty status and institution type, are shown in table 19. Ninety percent of full-time faculty members were located, compared with 75 percent of part-time faculty ($\chi^2 = 1414.6$, $p < .0001$). Eighty-one percent of full-time faculty completed the survey, compared with 69 percent of part-time faculty ($\chi^2 = 903.8$, $p < .0001$). When examined by institution type, locate rates ranged from 80 to 88 percent. Survey completion rates ranged from 71 percent for faculty at private not-for-profit doctorate-granting institutions to 81 percent at private not-for-profit baccalaureate-granting institutions.

Table 19. Faculty locating and survey completion results, by employment status and institution type: 2004

Employment status and institution type	Total sample	Located		Number eligible[1]	Completed survey	
		Number[1]	Percent[2]		Number[1]	Percent[3]
Total	35,629	29,820	83.7	34,330	26,110	76.1
Employment status						
Full-time	21,891	19,580	89.5	21,390	17,250	80.6
Part-time	13,008	9,740	74.9	12,270	8,430	68.7
Unknown employment status	730	500	68.8	660	420	64.3
Institution control						
Public	23,280	19,520	83.9	22,450	17,120	76.2
Private not-for-profit	12,349	10,300	83.4	11,880	8,990	75.7
Institution type						
Public doctoral	9,827	8,600	87.5	9,500	7,460	78.6
Public master's	3,485	2,950	84.5	3,350	2,620	78.1
Public bachelor's	693	560	81.7	680	510	75.4
Public associate's	9,129	7,280	79.8	8,770	6,420	73.1
Public other	146	130	85.6	140	110	73.6
Private not-for-profit doctoral	4,652	3,770	81.1	4,470	3,160	70.7
Private not-for-profit master's	3,020	2,540	84.0	2,890	2,270	78.6
Private not-for-profit bachelor's	3,218	2,800	87.0	3,120	2,520	80.8
Private not-for-profit associate's	242	200	84.3	240	190	79.8
Private not-for-profit other	1,217	990	80.9	1,160	850	73.1

[1] Numbers rounded to the nearest 10.
[2] Percentages are based on the number of sample members within the row under consideration. Percentages are based on original unrounded numbers.
[3] Percentages are based on the number of eligible sample members within the row under consideration. Percentages are based on original unrounded numbers.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

The results of faculty and instructional staff survey completion by mode of data collection are presented in table 20. A total of 19,780 respondents (76 percent) completed the self-administered questionnaire and 6,330 respondents (24 percent) completed the CATI interview. (It should be noted that 59.2 percent completed the survey during the early phase, without telephone followup). While NSOPF:04 exceeded the goal of having 50 percent of completes by web, a substantial portion of these web surveys were completed only after having been called by a CATI interviewer.

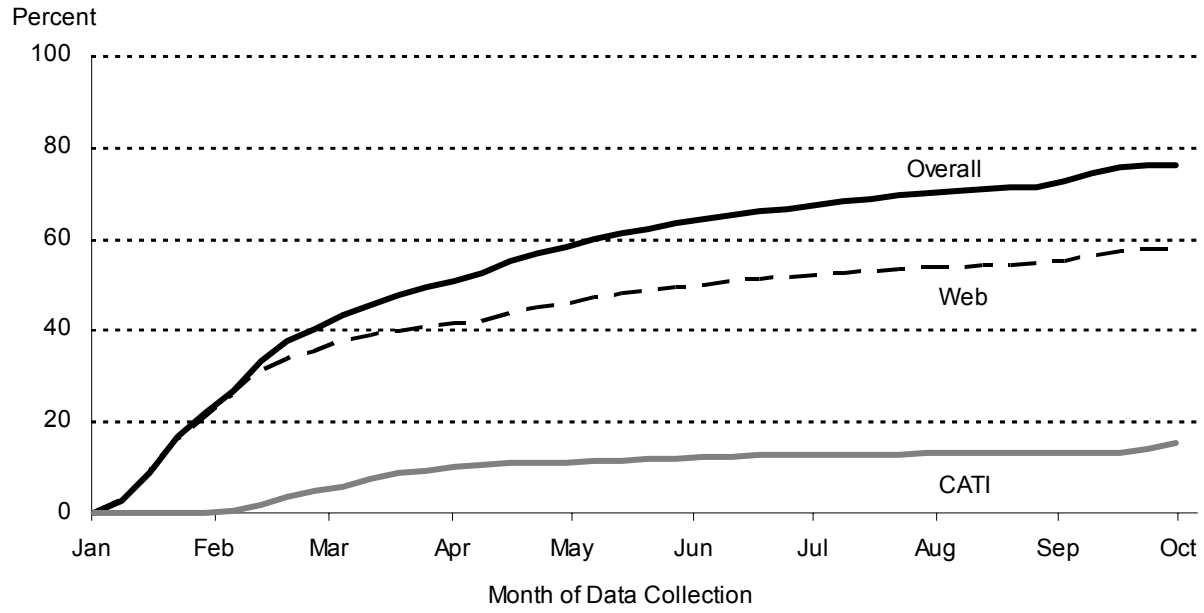
Table 20. Response rates and mode of completion, by employment status and institution type: 2004

Employment status and institution type	Number eligible[1]	Total complete		Mode of completion			
				Self-administered		CATI	
		Number[1]	Percent[2]	Number[1]	Percent[3]	Number[1]	Percent[3]
Total	34,330	26,110	76.1	19,780	75.8	6,330	24.3
Employment status							
Full-time	21,390	17,250	80.6	13,980	81.0	3,280	19.0
Part-time	12,270	8,430	68.7	5,500	65.2	2,940	34.8
Unknown employment status	660	420	64.3	300	71.9	120	28.1
Institution control							
Public	22,450	17,120	76.2	12,850	75.1	4,270	24.9
Private not-for-profit	11,880	8,990	75.7	6,930	77.1	2,060	22.9
Institution type							
Public doctoral	9,500	7,460	78.6	6,090	81.6	1,370	18.4
Public master's	3,350	2,620	78.1	1,980	75.6	640	24.4
Public bachelor's	680	510	75.4	360	69.6	160	30.4
Public associate's	8,770	6,420	73.1	4,350	67.7	2,070	32.3
Public other	140	110	73.6	70	68.9	30	31.1
Private not-for-profit doctoral	4,470	3,160	70.7	2,520	79.9	640	20.2
Private not-for-profit master's	2,890	2,270	78.6	1,700	74.8	570	25.2
Private not-for-profit bachelor's	3,120	2,520	80.8	1,950	77.4	570	22.6
Private not-for-profit associate's	240	190	79.8	160	83.7	30	16.3
Private not-for-profit other	1,160	850	73.1	600	70.6	250	29.5

[1] Numbers rounded to the nearest 10.
[2] Percentages are based on the number of eligible sample members within the row under consideration. Percentages are based on original unrounded numbers.
[3] Percentages are based on the number of completed interviews within the row under consideration. Percentages are based on original unrounded numbers.
NOTE: All percents are unweighted. Reporting excludes 1,300 cases determined to be ineligible for study. CATI = computer assisted telephone interview. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Eighty-one percent of full-time faculty completed the self-administered survey, compared to 65 percent of part-time faculty ($\chi^2 = 776.6$, $p < .0001$). Seventy-seven percent of faculty and instructional staff at private not-for-profit institutions completed the self-administered survey, compared to 75 percent of faculty at public institutions ($\chi^2 = 13.6$, $p < .0002$). Self-administered web survey completion rates by institution type ranged from 68 percent for public associate's degree-granting schools to 84 percent for private not-for-profit associate's degree-granting schools. The cumulative response rate, overall and by mode, is shown in figure 6.

Figure 6. Cumulative response rates, by mode of completion: 2004



NOTE: Mode of completion for respondents who switched modes was determined by the mode at the time of survey completion.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

3.2.3 E-mail Contacting Efforts

E-mail addresses of faculty and instructional staff were requested in the faculty lists. Where e-mail addresses were not provided by the institution, efforts were made through an advance search of the institution's online directory for e-mail addresses of sample members as well as other database searches. In addition, some sample members provided e-mail addresses when contacted by a telephone interviewer. E-mail addresses were available for 27,980 (82 percent) of the 34,330 eligible sample members.

Periodically throughout the data collection period, e-mail messages were sent to nonrespondents to encourage their participation (see appendix F). Sample members who received e-mails were more likely to complete the survey (78 percent) compared to sample members to whom no e-mail reminders were sent (53 percent; $\chi^2 = 1867.2$, $p < .0001$). Respondents with e-mail addresses were more likely to complete the self-administered web questionnaire (79 percent) than were respondents who were not sent e-mail reminders (55 percent; $\chi^2 = 976.1$, $p < 0.0001$).

3.2.4 Refusal Conversion Efforts

Refusal conversion measures were used to gain cooperation from individuals who refused to participate when contacted by telephone interviewers. Refusals came not only from sample members, but also occasionally from other household members or other contacts (such as

secretaries).[24] Whenever a refusal was encountered, unless it was deemed hostile, the case was referred to a specialist trained in refusal conversion techniques. Refusal conversion specialists were chosen based on their performance as interviewers, with those who were the most skilled in obtaining cooperation given additional training in converting refusals. This training was tailored to the concerns of faculty members and gatekeepers regarding participation, and focused on gaining cooperation and encouraging participation.

Ten percent of contacted cases refused to participate at some point during data collection. However, 18 percent of these cases were successfully converted and eventually completed the survey. Fifty-nine percent of the converted cases completed the web self-administered questionnaire, and 41 percent completed a telephone interview. An abbreviated instrument, consisting of sections A (nature of employment), B (academic/professional background), and G (sociodemographic characteristics) from the faculty instrument, was developed to convert nonrespondents by offering a shorter (10 minute) interview. The abbreviated instrument, used only in the final 3 weeks of data collection, yielded 1,610 interviews.

3.2.5 Incentives

For the NSOPF:04 full-scale data collection, three types of incentives were offered to eligible sample members. In accordance with findings from the NSOPF:04 field test incentive experiment[25] (Heuer et al. 2004), incentives were offered during two phases of data collection: an initial early-response incentive period and a nonresponse incentive period. In addition to those periods, refusal incentives were made available following initial refusals. During each incentive phase, respondents were offered the choice of a $30 check or a $30 gift certificate to Amazon.com.

The initial early-response incentive was offered to all sample members for completion of the questionnaire within the first 4 weeks of data collection. The early-response incentive was designed to increase the response rate during the initial phase of data collection and promote a higher rate of web self-administered responses and reduce costs associated with telephone interviewing. Following the initial 4-week period, CATI telephone prompting began. During this second phase of the study, no incentive was offered to respondents for completing the interview. All nonrespondents from the first phase were contacted by telephone and asked to complete the survey, either on the phone or via the Web at their convenience.

Any sample member who refused to participate in the study was flagged for the refusal incentive. A refusal conversion letter was sent out to explain the study and request that the sample member reconsider the decision not to participate and to announce the reinstitution of the $30 incentive for participating.

[24] Nearly 77 percent of all refusals were made by sample members, while the remaining 23 percent were made by other household members or other contacts. Of the sample members who initially refused, 17 percent eventually completed an interview.

[25] The field test experimental design consisted of three randomly assigned early-response incentive groups who were offered $0, $20, or $30 to complete the self-administered questionnaire over the Internet within 3 weeks of the initial mailing and two nonresponse incentive groups of $0 and $30 for those who had not completed the survey by a certain date during data collection. The early-response incentive yielded 31 and 34 percent response rates for the $20 and $30 incentives, respectively, compared with a 16 percent response rate for the control group. The nonresponse incentive yielded a 47 percent response rate for those offered $30 and a 34 percent response rate for the control group. The differences between the treatment and the control groups were statistically significant for both phases of the experiment; however, the apparent difference in amounts ($20 versus $30) for the early-response incentive period, while in the expected direction, was not statistically significant.

Nonresponse incentives were introduced after 8 weeks of CATI prompting of all nonrespondents who had not already been offered the refusal incentive. Letters and e-mail prompts were sent periodically to nonrespondents throughout the data collection period. All correspondence mentioned the incentive when it was available to sample members. Table 21 provides a breakdown of the types of incentives offered and the results of each incentive period.

Table 21. Faculty response rates, by incentive period: 2004

	Number[1]		Percent[2]		
Incentive offered	Eligible	Responded	Eligible[3]	Responded[4]	Response rate[5]
Total	34,330	26,110	76.1	100.0	76.1
Early-response incentive	34,330	15,010	43.7	57.5	43.7
Period of no incentive	19,320	5,250	27.2	20.1	15.3
Refusal incentive	2,410	570	23.6	2.2	1.7
Nonresponse incentive	11,660	5,280	45.3	20.2	15.4

[1] Numbers rounded to the nearest 10.
[2] Percentages are based on original unrounded numbers.
[3] Percentages are based on the number of eligible sample members within the row under consideration.
[4] Percentages are based on the total number of respondents.
[5] Percentages are based on the total number of eligible sample members.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

These results indicate that the combination of early-response incentive and other later incentives was required to reach the targeted response rate within the data collection schedule. While the early-response incentive was effective in getting 44 percent of the eligible sample members to complete the survey within the initial 4 weeks of the study, and an additional 15 percent of the sample members completed within 8 weeks after the initial incentive period, the cumulative response rate after 12 weeks was only 59 percent. The refusal and nonresponse incentives were undoubtedly helpful in attaining the additional 17 percent needed to reach the 76 percent response rate.

3.3 Burden and Effort

3.3.1 Faculty Survey Completion Timing

Like the institution timing analysis, the faculty timing analysis was conducted by embedding time stamps in the programming code for each form (screen) in the survey. From these time stamps, the number of seconds spent on each screen (on-screen time) and the transit time between screens (i.e., the time required to transmit data to the server, the time for the server to store the data and assemble the next page, and the time for the page to be transmitted and loaded on the computer) were calculated. A cumulative on-screen time and a cumulative transit time for the faculty survey also were calculated from the time stamps. The sum of the cumulative on-screen and transit times was the total instrument time (i.e., the number of minutes it took to administer the faculty questionnaire).

Following the 1999 cycle of NSOPF—which averaged over 50 minutes—the faculty questionnaire was shortened substantially, with a goal of achieving a 30-minute survey. The NSOPF:04 field test averaged 42 minutes. Based on the time stamps for each form in the full-scale instrument, the time to complete the entire survey ranged from 8 minutes to 3 hours and

6 minutes,[26] with an average time of 30 minutes. Of these 30 minutes, approximately 26 minutes, on average, were spent answering questions (on-screen time) and 3 minutes, on average, were spent saving data and loading forms (transit time).

Table 22 presents the overall timing data by mode for completed surveys (excluding partial and abbreviated interviews). Average on-screen time was significantly longer for CATI respondents than for web respondents (27 and 26 minutes, respectively; $t = –4.46$, $p < .0001$), while the average transit time was significantly shorter for CATI respondents than for web respondents (1 and 4 minutes, respectively; $t = 34.94$, $p < .0001$). Presumably, the longer on-screen time for CATI respondents is due to the time it takes to read text out loud, and to the fact that the respondent may ask questions. The shorter transit time for CATI is likely due to the use of a high-speed internet connection by interviewers. Some web respondents may have used slower dial-up connections, which increase transit time. Overall, the interview took less time for CATI respondents than for web respondents (29 minutes and 30 minutes, respectively; $t = 7.80$, $p < .0001$).

Table 22. Average on-screen, transit, and total survey completion time, in minutes, for the faculty questionnaire, by mode: 2004

	All respondents		Web respondents		CATI respondents	
Portion of interview	Average time	Number of cases[1]	Average time	Number of cases[1]	Average time	Number of cases[1]
Total	29.7	24,360	30.1	18,630	28.5	5,730
Onscreen	26.5	24,360	26.3	18,630	27.1	5,730
Transit	3.2	24,360	3.8	18,630	1.4	5,730

[1] Numbers rounded to the nearest 10. Abbreviated and partial interviews excluded.
NOTE: Detail may not sum to totals because of rounding. CATI = computer assisted telephone interview.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

The onscreen, transit, and total times were significantly shorter for surveys that were completed during business hours (Monday through Friday, 9:00 a.m. to 6:00 p.m.) compared to those completed during evening and weekend hours (onscreen: 26 and 27 minutes, respectively; $t = 4.79$, $p < .0001$; transit: 3 and 4 minutes, respectively; $t = 17.71$, $p < .0001$; total: 29 and 31 minutes, respectively; $t = 10.29$, $p < .0001$), as shown in table 23. This may be due to faster internet connections for web respondents at their offices compared to their homes or time pressures during the workday.

[26] Excludes the two highest outliers, both with transit times greater than 4 hours.

Table 23. Average on-screen, transit, and total completion time, in minutes, by time of day and mode: 2004

Portion of interview	Web respondents				CATI respondents			
	Weekdays 9am–6pm		Evenings/weekends		Weekdays 9am–6pm		Evenings/weekends	
	Average time	Number of cases[1]	Average time	Number of cases[1]	Average time	Number of cases[1]	Average time	Number of cases[1]
Total	29.2	11,620	31.4	7,010	28.5	3,710	28.6	2,020
Onscreen	26.0	11,620	26.9	7,010	27.0	3,710	27.2	2,020
Transit	3.3	11,620	4.6	7,010	1.5	3,710	1.4	2,020

[1] Numbers rounded to the nearest 10.
NOTE: Abbreviated and partial interviews excluded, as well as two outliers. Detail may not sum to totals because of rounding. CATI = computer assisted telephone interview.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Table 24 provides the average and maximum times (in seconds) to complete each form in the faculty instrument. Seven forms (screens) in the faculty survey took more than 1 minute to administer, on average. These tended to be the more complicated forms and those that collected multiple pieces of information on a single screen. These forms are described in greater detail below.

Table 24. Average and maximum completion time, in seconds, for forms in the faculty instrument: 2004

Questionnaire form	Description	Time in seconds		Number of cases[1]
		Average	Maximum	
Q1	Instructional duties, any	19	102	26,110
Q2	Instructional duties related to credit courses/activities	12	64	24,310
Q3	Faculty status	10	97	26,110
Q4	Principal activity	16	87	26,110
Q5	Employed full or part time at this institution	7	56	26,110
Q6	Part-time employment is primary employment	8	46	8,370
Q8	Part-time but preferred full-time position	8	63	8,340
Q9	Year began current job	21	131	26,110
Q10	Rank	14	85	26,110
Q11	Rank, year attained professor or associate professor	27	194	9,500
Q12	Tenure status	12	100	26,110
Q13	Tenure, year attained at any postsecondary institution	19	160	8,440
Q14	Union status	9	74	26,110
Q15	Union status, reason not a member	12	82	20,850
Q16VS	Principal field of teaching-verbatim	24	134	26,110
Q16AC	Principal field of teaching-autocode	30	143	23,590
Q16CD	Principal field of teaching-manual code	54	230	7,480
Q17A1	Highest degree	18	131	26,110
Q17A1B	Hold PhD and professional degree	5	27	2,120
Q17A2	Highest degree date awarded	12	94	25,870
Q17A3VS	Highest degree field-verbatim	16	96	25,860
Q17A3AC	Highest degree field-autocode	16	104	24,710
Q17A3CD	Highest degree field-manual code	30	189	6,570
Q17A4	Highest degree institution-code	51	245	25,850

See notes at end of table.

Table 24. Average and maximum completion time, in seconds, for forms in the faculty instrument: 2004—Continued

Questionnaire form	Description	Time in seconds		Number of cases[1]
		Average	Maximum	
Q17A4A	Highest degree institution-info for later coding	27	87	1,270
Q17D	Bachelor's degree date awarded	13	96	23,460
Q18	Other current jobs, number of jobs	15	82	26,110
Q19A	Other current jobs, full-time employment	7	66	8,290
Q19B	Other current jobs, number in postsecondary instruction	10	75	8,130
Q21	First postsecondary job, current job is first	17	86	26,110
Q23	First postsecondary job, year began	18	99	14,310
Q24	First postsecondary job, part or full time	8	55	26,110
Q26	First postsecondary job, tenure status	11	68	14,780
Q27	Other jobs, any outside postsecondary since degree	11	75	26,110
Q28	Other jobs, sector of previous job	24	123	26,110
Q31	Hours worked per week	94	338	24,580
Q32	Percent distribution of work activities	78	311	24,330
Q35A	Number of classes taught, credit and noncredit	44	205	23,600
Q35B	Number of classes taught, remedial and distance education	22	127	21,240
Q36	Teaching assistant in any credit class	9	66	20,230
Q37	Number and types of classes taught (up to five classes)	99	402	20,220
Q38	Tools instructor used to evaluate undergraduate students	68	234	16,430
Q39	Website for any instructional duties	14	140	23,020
Q41	Hours per week, e-mailing students	16	87	23,020
Q46	Individual instruction, any	15	94	24,550
Q47	Individual instruction, number of students	20	124	8,230
Q47B	Individual instruction, number of hours	22	131	7,880
Q48	Hours per week, committees/advisees/office hours	61	251	24,530
Q52A	Career publications/presentations	100	427	24,490
Q52B	Recent publications/presentations	47	275	21,190
Q53	Scholarly activity, any	13	93	24,470
Q54VS	Scholarly activity, principal field-verbatim	25	203	540
Q54AC	Scholarly activity, principal field-autocode	14	53	410
Q54CD	Principal research field-manual code	29	139	270
Q56	Scholarly activity, description	17	86	14,000
Q55	Scholarly activity, any funded	12	76	13,930
Q61	Satisfaction, authority/resources/salary/benefits	60	212	24,450
Q65	Retirement plans/history	18	91	24,440
Q66	Income, from institution/other sources	111	403	24,420
Q66B	Amount of total individual income (range)	25	194	2,730
Q67	Type of contract, length of unit	18	190	24,410
Q68	Income paid per course/credit unit or term	11	75	6,260
Q69	Amount of income paid per course/credit unit or term	20	184	5,170
Q70A	Amount of total household income	33	171	24,400
Q70B	Amount of total household income (range)	13	80	3,400
Q71	Gender	9	152	25,990
Q72	Age, year of birth	6	47	25,990
Q73	Race/ethnicity, Hispanic/Latino	4	45	25,980
Q74	Race	9	65	25,980

See notes at end of table.

Table 24. Average and maximum completion time, in seconds, for forms in the faculty instrument: 2004—Continued

Questionnaire form	Description	Time in seconds		Number of cases[1]
		Average	Maximum	
Q75	Disability, any	10	70	25,980
Q77	Marital status, fall 2003	7	49	25,980
Q79	Dependent children, number	8	48	25,980
Q80	United States birth/citizenship status	7	55	25,970
Q82	Opinion, institution fairness	32	125	24,360
Q83	Opinion about choosing an academic career again	8	57	24,360

[1] Numbers rounded to the nearest 10.
NOTE: The number of cases per form varies due to the interview skip logic. Outliers for each form were topcoded (mean + 3 standard deviations).
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Q31 and Q32. The questions that asked for the number of hours per week spent on work activities, Q31 (by paid and unpaid activities at the target institution and outside that institution), and the percent distribution of work activities, Q32, took 94 and 78 seconds, respectively, to administer. Each of these forms took longer when administered by telephone interviewers than when self-administered via the web instrument. Q31 averaged 91 seconds for web respondents compared with 103 seconds for CATI respondents ($t = –13.64$, $p < .0001$). Web respondents averaged 76 seconds on Q32 compared with an average time of 83 seconds for CATI respondents ($t = –7.59$, $p < .0001$). The complexity of these questions may have led to the longer times for CATI administration, as respondents often asked interviewers to repeat the question and examples, and asked questions about the appropriate category for certain types of activities.

Q37 and Q38. Two consecutive forms, Q37 and Q38, asked for a great deal of information on a single screen. Q37 was a matrix-style question that asked six questions about each of the credit classes (up to five) the respondent taught. This form took 99 seconds, on average, to administer, with CATI respondents taking significantly less time than web respondents (94 and 100 seconds, respectively, $t = 4.26$, $p < .0001$). The matrix of items on Q37, visually different from the rest of the forms in the questionnaire, likely took web respondents extra time to make sense of and answer.

Q38 asked respondents to identify which of 10 different types of student evaluation tools were used in their classes and whether they were used in all, some, or none of the classes. This form took an average of 68 seconds to administer, with CATI respondents taking significantly longer than web respondents (93 and 60 seconds, respectively, $t = –49.69$, $p < .0001$).

Q48. This form asked for the number of hours per week the respondent spent on four activities (thesis/dissertation committees, administrative committees, with advisees, and office hours). On average, respondents took 61 seconds to complete this form, with CATI respondents taking significantly longer than web respondents (72 and 57 seconds, respectively, $t = –24.37$, $p < .0001$).

Q52A. Q52A, which asked for the number of career publications or presentations in seven categories, took an average of 100 seconds to complete. This may have required respondents to locate their curricula vitae and count the number of publications. CATI respondents spent significantly more time on this item than web respondents (106 and 98 seconds, respectively, $t = –5.88$, $p < .0001$).

Q66. The form asking about respondents' compensation from the target institution and from other sources, Q66, took 111 seconds to complete, on average. This form consisted of six income questions, which were considered to be among the most sensitive items in the questionnaire. Average time to complete this form was shorter for web respondents (109 seconds) than for CATI respondents (118 seconds; $t = -7.60$, $p < .0001$).

3.3.2 Help Desk

To gain a better understanding of the problems encountered by faculty members attempting to complete the web self-administered questionnaire, software was developed to record each help desk incident that occurred during data collection. For each occurrence, help desk staff confirmed contact information for the sample member, recorded the type of problem, described the problem and resolution, noted its status (pending or resolved), and recorded the approximate time it took to assist the faculty member. Help desk staff were trained not only to answer any calls received from the help desk hotline, but also to conduct telephone interviews when needed. Help desk staff members assisted sample members with questions about the web instrument and provided technical assistance to sample members who experienced problems while completing the self-administered web survey. Help desk agents also responded to voice mail messages left by respondents when the call center was closed.

Help desk staff assisted 3,860 faculty members (11 percent of the sample). Eighty-one percent of these cases called the help desk only once, 12 percent called twice, 4 percent called three times, and 3 percent called four or more times. Of the 3,860 faculty members who called the help desk, 2,940 (76 percent) eventually completed the survey.

Twenty-nine percent of the problems reported by faculty members who called the help desk were for miscellaneous issues. The miscellaneous issues were first coded into specific issues and then these issues were coded into five broader categories as shown in table 25. First time calls included setting an appointment for the CATI interview, providing a new phone number or e-mail address, promising to complete by phone at a later date, or promising to complete the survey on the Web. Nearly 7 percent of help desk contacts were faculty members calling in to refuse. Follow-up calls to the help desk (6 percent) included faculty members checking on the incentive, or verifying that they had completed the survey. Other miscellaneous issues were less than 2 percent of all contacts. Slightly more than 1 percent of help desk calls reported that the faculty member was not at the phone number, e-mail address, or college that was contacted.

Other specific issues handled by the help desk included requests to complete the survey by telephone (21 percent), questions about the study (19 percent), browser setting and computer problems (14 percent), requests for study ID and/or password (12 percent), errors in questionnaire programming (3 percent), questions about questionnaire content (2 percent), website being down or unavailable (1 percent), and routing/skip problems (less than 1 percent).

Table 25. Response pattern, by help desk problem type: 2004

Type of problem	Number	Percent
Total	5,151	100.0
Miscellaneous	1,491	29.0
First time calls (set call back date/time, etc.)	698	13.6
SM called in to refuse	352	6.8
Follow-up calls (checking on incentive, verifying complete)	284	5.5
Other	94	1.8
SM not at this number/college	63	1.2
Called in to complete by phone	1,078	20.9
Question about study	964	18.7
Browser settings/computer problems	694	13.5
Study identification (ID) code/password	626	12.2
Program error	130	2.5
Questionnaire content	114	2.2
Website unavailable	47	.9
Routing/skip problems	7	.1

NOTE: Detail may not sum to totals because of rounding. SM = sample member.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

3.3.3 Interviewer Hours

A total of 17,639 telephone interviewing staff hours (including help desk staffing, telephone follow-up calls, and CATI interview hours) were expended during faculty data collection. These hours do not include supervision, monitoring, administration, and Quality Circle meetings. The average time spent per completed CATI interview was 2.7 hours and per completed interview overall (including web completes) was 0.7 hours. The average time to administer the CATI was 29 minutes, which shows that a majority of interviewer time was spent on other activities. These other activities focused on contacting and locating the sample member, with a small portion of time devoted to bringing up a case, reviewing its history, and closing the case (with the appropriate reschedule, comment, and disposition). A significant proportion of the web completes occurred after the period of telephone follow-up began and were completed only after several CATI follow-up calls had been made to the respondent.

3.3.4 Number of Calls

Telephone interviewers made 226,777 call attempts to faculty members during the NSOPF:04 data collection period (see table 26). The number of calls per case ranged from 0 to 152. On average, six calls[27] were made to each sample member. Those who were not interviewed received the highest average number of calls. An average of four call attempts were required for respondents compared to an average of 13 call attempts for nonrespondents ($t = 60.9$, $p < .0001$). Faculty members who completed the web self-administered questionnaire were called significantly fewer times, with an average of three call attempts per completed survey, compared to an average of eight calls to CATI respondents ($t = 41.5$, $p < .0001$).

[27] This figure includes cases where no call attempts were made, either because the respondent completed the questionnaire via the Web before CATI calling began, or the individual could not be located.

Table 26. Total and average number of calls, by completion status and mode of completion: 2004

Completion status/mode	Number of cases[1]	Number of calls	Average calls per case
Total	35,630	226,777	6.4
Interviewed	26,110	102,946	3.9
Not interviewed	9,520	123,831	13.0
By mode	26,110	102,946	3.9
Web complete	19,780	53,621	2.7
CATI complete	6,330	49,325	7.8

[1] Number of respondents rounded to nearest 10.
NOTE: Detail may not sum to totals because of rounding. CATI = computer assisted telephone interview.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Call screening is a growing problem for studies that rely on the telephone as a mode of contact. Devices such as telephone answering machines can be used to screen unwanted calls. Of the 19,394 cases called by telephone interviewers, 15,183 cases (78 percent) reached an answering machine at least once (see table 27). Interviewers made significantly more calls to cases where an answering machine had been reached at least once (mean attempts = 13) than they did to cases where no answering machine was reached (mean attempts = 5; $t = –46.81$, $p < .0001$). Likewise, cases where an answering machine had been reached at least once were less likely to have completed the interview (54 percent) than cases where no answering machine was reached (63 percent; $\chi^2 = 92.4$, $p < .0001$).

Table 27. Average call attempts, by reached answering machine: 2004

Result of call attempt	Cases called in CATI		Completed cases	
	Number of cases	Average number of calls	Number of cases[1]	Average number of calls
Reached answering machine at least once	15,183	13.4	8,230	11.2
Never reached an answering machine	4,211	5.4	2,630	4.0

[1] Numbers rounded to the nearest 10.
NOTE: Excludes 16,240 completed cases that were never called by telephone interviewers because they completed the self-administered questionnaire during or soon after the early-response period of data collection. Some of the cases called by telephone interviewers actually completed the web self-administered questionnaire.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Looking only at completed cases, significantly fewer calls were required to obtain a completed interview when no answering machine was reached (mean attempts = 4) compared to cases in which an answering machine was reached at least once (mean attempts = 11, $t = -40.69$, $p < .0001$). Those who possessed answering machines were included in the survey definition of "accessible"; however, it took considerable persistence and resources (in the form of repeated call attempts) to reach these faculty members. This finding demonstrates that answering machines and other call screening devices are increasing the effort that must be expended to reach these cases, thereby driving up interviewing costs.

In addition, cases where an answering machine was reached on more than one-half of the call attempts required significantly more effort to contact and interview. The mean number of attempts for cases that reached an answering machine less than one-half of the time was 10 compared to 13 ($t = -15.3$, $p < .0001$) for cases that reached an answering machine more than one-half of the time. Similarly, among completed cases, significantly fewer calls were needed to

complete an interview with cases where an answering machine was reached less than one-half of the time (mean attempts = 8) compared with those where an answering machine was reached more than one-half of the time (mean attempts = 11; $t = -12.9$, $p < .0001$).

3.4 Conclusions

Of the 1,070 eligible institutions, 980 (91 percent, unweighted and weighted) provided faculty lists and 920 (86 percent) completed the institution questionnaire. A total of 26,110 faculty and instructional staff completed the faculty survey for a 76 percent response rate. Approximately three-quarters (76 percent) of respondents completed the web self-administered questionnaire rather than the CATI (24 percent). Strategies that helped attain this response rate included tracing, e-mail contacting, and refusal conversion efforts, along with targeted incentives.

Chapter 4
Evaluation of Data Quality

Evaluations of data quality serve to identify problems with the data collection processes and instruments in order to remedy them for the next cycle of the study. Project staff evaluated faculty list quality, item nonresponse, item mode effects, breakoffs, coding, quality control monitoring of interviewers, and interviewer feedback. The results of these evaluations are presented in this chapter.

4.1 List Quality

4.1.1 List Types

Faculty lists may be characterized both by type of media—whether they are electronic or hardcopy—and method of transmission (e.g., fax or mail, e-mail, electronic upload). For the 2004 National Study of Postsecondary Faculty (NSOPF:04), institutions were asked to provide a single, unduplicated (i.e., duplicate entries of names removed) electronic list of faculty in any commonly-used and easily processed format (e.g., ASCII fixed field, comma delimited, spreadsheet format). These preferred electronic file formats are far less labor intensive to process than paper lists and more easily unduplicated by ID number. However, as in previous cycles, paper lists were accepted, as were multiple files (e.g., separate files of full- and part-time faculty) and lists in electronic formats that did not lend themselves to electronic processing (such as word processing formats).

For the first time, institutions were given the option to transmit their electronic faculty lists via a secure upload to the National Study of Faculty and Students (NSoFaS:04) website and were encouraged to do so. (In previous cycles, direct upload was available only by file-transfer protocols, an option that few institutions utilized). Institutions were also given the option of sending a CD-ROM, diskette, or paper list containing the list data or sending the list via e-mail (as an encrypted file, if necessary).

As shown in table 28, the vast majority of lists received were in electronic formats. Of 980 participating institutions, 830 (85 percent) supplied an electronic list by upload, e-mail, CD-ROM, or diskette. Institutions showed a clear preference for uploading their list by direct upload; 590 institutions (60 percent of lists overall and 71 percent of electronic lists) delivered their data in this manner.

NSOPF:04 clearly benefited from the increased capability and willingness of institutions to supply lists in electronic formats, compared to previous cycles. As table 29 shows, 65 percent of institutions supplied an electronic list for NSOPF:99, with a majority of them in CD-ROM or diskette formats sent by mail.

Table 28. Number of submitted faculty lists, by type of institution and transmittal mode: 2004

Institution type	Number of sample institutions[1]	Number of institutions providing lists via six transmittal modes						
		Total	Upload	Electronic & paper	Diskette	Paper	Abstracted from web directory	E-mail
Total	1,080	980	590	#	40	#	140	200
Public doctoral	190	180	120	#	#	#	30	30
Public master's	120	100	60	#	10	#	10	30
Public bachelor's	30	30	20	#	#	#	10	#
Public associate's	340	290	170	#	30	#	30	60
Public other	10	10	10	#	#	#	#	#
Private not-for-profit doctoral	110	100	70	#	#	#	20	20
Private not-for-profit master's	80	80	50	#	#	#	10	10
Private not-for-profit bachelor's	130	120	70	#	#	#	20	30
Private not-for-profit associate's	10	10	10	#	#	#	#	#
Private not-for-profit other	60	60	30	#	#	#	20	10

Rounds to zero.
[1] Number of institutions rounded to nearest 10. Detail may not sum to totals because of rounding and because of duplicative forms of list transmittal modes.
NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Table 29. Faculty list types, by NSOPF cycle: 1999 and 2004

Type of list	NSOPF:99		NSOPF:04	
	Number of institutions	Unweighted percent[1]	Number of institutions	Unweighted percent[1]
Total	820	100	980	100.0
Paper	290	35.0	10	0.6
Electronic (ftp or upload) [2]	10	1.1	590	60.4
Electronic (E-mail)	220	26.6	200	20.7
CD-ROM or diskette	310	37.2	40	4.0
Abstracted from web resource[3]	—	—	140	14.2

— Not available.
[1] Percentages are based on original unrounded numbers.
[2] FTP was utilized only in 1999; upload was utilized only in 2004.
[3] In 1999, lists abstracted from web resources were processed as, and included with paper lists.
NOTE: Numbers rounded to the nearest 10. Detail may not sum to totals because of rounding. FTP = file transfer protocol.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

For institutions that indicated they lacked the staff or resources to compile a list of faculty on their own within schedule constraints, it was sometimes possible to abstract a list from employee directories, course schedules, or course catalog listings available on the institution's website or through other web resources. As in past cycles (where course catalogs or directories were used as lists of last resort), all such lists were reviewed to ensure they were sufficiently complete for sampling (i.e., included both full- and part-time faculty, did not systematically exclude any subset of faculty and instructional staff). It should be noted that in past cycles, course catalogs and directories comprised a large percentage of lists supplied on (or processed

as) paper. While the web listings utilized for NSOPF:04 required more processing than electronic lists (including reformatting into a spreadsheet or re-keying), they proved, overall, to be far less problematic for processing and sampling than an equivalent paper list. Only 15 percent of institutions submitted paper lists or had lists abstracted from web resources for NSOPF:04; this compares to 35 percent of institutions who submitted paper lists (including lists abstracted from web resources) in NSOPF:99.

4.1.2 List Data Quality

As in prior administrations of this study, secured faculty lists were evaluated for accuracy and completeness of information before they were processed for sampling. To facilitate quality control, faculty list counts were compared against counts obtained from the following supplementary sources:

- the institution questionnaire and/or the file layout form, if a questionnaire was not completed but an overall faculty count was supplied;
- the 2001 Integrated Postsecondary Education Data System (IPEDS) Fall Staff Survey;
- the Contact Information and File Layout (CIFL) form (which included faculty counts, and used when questionnaire data was unavailable); and
- NSOPF:99: frame data from the 1999 survey.

Discrepancies in counts of full- and part-time faculty between the faculty list and other sources that were outside the expected range were investigated. All institutions with submitted lists that failed any checks were recontacted to resolve the observed discrepancies.

Because of time and definitional differences between NSOPF and IPEDS, it was expected that the faculty counts obtained from the institutions and IPEDS would include discrepancies. Consequently, quality control checks against IPEDS were less stringent than those against the institution questionnaire. However, list count comparisons against IPEDS and NSOPF:99 data were useful in identifying systematic errors, particularly those related to miscoding of the employment status of faculty members. Table 30 shows the types of discrepancies encountered by type of institution.

Table 30. Type of discrepancies encountered, by type of institution: 2004

Institution type	Sampled institutions	Discrepant with		Unreadable	Insufficient data	
		IPEDS	QUEX		Needed for sampling	CIFL
Total	1,080	300	280	10	180	190
Public doctoral	190	70	50	#	40	30
Public master's	120	30	30	#	10	20
Public bachelor's	30	10	10	#	10	10
Public associate's	340	80	90	#	30	60
Public other	10	#	#	#	#	#
Private not-for-profit doctoral	110	40	30	#	20	20
Private not-for-profit master's	80	20	20	#	20	10
Private not-for-profit bachelor's	130	30	40	#	30	20
Private not-for-profit associate's	10	#	#	#	#	#
Private not-for-profit other	60	20	20	#	20	10

Rounds to zero.
NOTE: IPEDS is the National Center for Education Statistics (NCES) Integrated Postsecondary Education Data System; QUEX refers to the institution questionnaire; CIFL refers to the contact information and file layout forms. Numbers rounded to the nearest 10. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Table 31 shows the percent differences between the three sources of data for all cycles of NSOPF (1988, 1993, 1999, and 2004). The discrepancies between the faculty lists and institution questionnaire counts have declined over time. Also, table 32 shows mean differences between sources of data across all cycles of NSOPF. More details regarding the quality of faculty lists secured for NSOPF:04 are provided in appendix H.

Table 31. Percentage differences between sources of data across all cycles of NSOPF: 1988 to 2004

Comparison	Year	Number of institutions	Percent difference in faculty counts						
			<-50	-50 to -31	-30 to -11	-10 to 10	11 to 30	31 to 50	>50
LIST-IPEDS	1988	410	8.0	5.6	14.9	35.4	16.6	7.6	12.0
	1993	660	5.0	5.2	11.3	25.4	23.8	13.3	16.0
	1999	770	6.4	6.5	13.6	33.7	23.0	6.8	9.9
	2004	980	2.4	4.1	12.2	32.4	23.1	9.3	16.6
QUEX-LIST	1988	410	1.9	3.9	16.6	51.2	15.1	2.4	8.8
	1993	750	3.7	6.5	13.2	41.7	12.3	6.1	16.5
	1999	770	1.4	2.7	7.0	72.3	5.6	3.2	7.8
	2004	900	1.2	1.3	3.8	82.6	4.9	2.0	4.1
QUEX-IPEDS	1988	410	3.9	6.8	15.9	34.6	20.0	7.8	11.0
	1993	690	2.3	4.5	9.2	26.6	25.4	12.6	19.3
	1999	790	3.3	6.6	11.4	40.7	22.0	6.7	9.5
	2004	900	1.6	2.1	9.2	37.1	24.6	9.5	16.1

NOTE: LIST refers to the faculty list provided by sampled institutions; IPEDS is the National Center for Education Statistics' Integrated Postsecondary Education Data System; QUEX refers to the institution questionnaire. Numbers rounded to the nearest 10.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Table 32. Mean differences between sources of data across all cycles of NSOPF: 1988 to 2004

Comparison	Year	Number of institutions	Mean difference	Standard error[1]	Mean percent difference in faculty counts	Standard error[1]
LIST-IPEDS	1988	410	3.0	(17.3)	14.1*	(3.8)
	1993	660	88.4*	(22.6)	24.8*	(3.1)
	1999	770	24.8	(13.8)	9.8*	(2.1)
	2004	980	57.5*	(13.1)	29.0*	(3.2)
QUEX-LIST	1988	410	8.5	(16.1)	11.4*	(3.2)
	1993	750	23.5	(16.7)	142.4	(106.8)
	1999	770	16.1	(11.2)	14.9*	(2.7)
	2004	900	7.6	(8.8)	5.1*	(1.2)
QUEX-IPEDS	1988	410	11.6	(14.7)	15.8*	(3.6)
	1993	690	96.3*	(21.5)	36.4*	(5.2)
	1999	810	53.5*	(12.8)	18.5*	(2.7)
	2004	900	69.0*	(9.4)	30.2*	(3.3)
LIST-IPEDS[2]	1988	330	-12.3	(10.9)	1.2	(1.1)
	1993	520	34.2*	(9.4)	7.4*	(1.0)
	1999	640	9.8	(9.8)	2.7*	(0.8)
	2004	790	10.0	(9.3)	5.7*	(0.7)
QUEX-LIST[2]	1988	370	-12.1	(8.4)	-1.1*	(0.8)
	1993	600	-22.0	(7.9)	-0.1*	(0.8)
	1999	700	-18.5*	(6.0)	-0.1*	(0.9)
	2004	850	2.0	(2.4)	0.6	(0.3)
QUEX-IPEDS[2]	1988	350	1.5	(9.1)	1.4	(1.1)
	1993	540	35.2*	(8.2)	8.6*	(0.9)
	1999	690	6.7	(8.5)	2.7*	(0.7)
	2004	740	29.9*	(8.9)	7.8*	(0.7)

* Statistically significant at alpha = .05, based on paired t-test.
[1] Standard errors assume simple random sampling.
[2] Observations with percent differences greater than 50 in absolute value were excluded.
NOTE: LIST refers to the faculty list provided by sampled institutions; IPEDS is the National Center for Education Statistics' Integrated Postsecondary Education Data System; QUEX refers to the institution questionnaire. Numbers rounded to the nearest 10.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

4.2 Item Nonresponse

Recent studies (for example, DeRouvray and Couper 2002) using web self-administered questionnaires have shown higher than usual rates of missing data when the "refuse" and "don't know" options are presented on the screen. To limit the rate of nonresponse in the institution and faculty instruments, the refusal option was unavailable to respondents and the "don't know" option was limited to selected screens where the respondent might not know the answer (e.g., expected age at retirement). On the information page at the start of the questionnaire, respondents were instructed to click the "continue" button to proceed to the next question if they wished to decline to answer a question.

For the institution questionnaire, items with a high rate of missing responses are often those that required lookup by an office on campus other than the Institution Coordinator's (e.g., Human Resources or Academic Affairs) and might reflect a lack of cooperation from those other offices. Two of the 90 items in the questionnaire had more than 15 percent of the data missing.

Details of institution item nonresponse, including the nonresponse bias analysis, are presented in appendix I.

Thirty-four of the 162 items in the faculty questionnaire had more than 15 percent of the data missing.[28] With the exception of the income items, which were expected to have higher rates of refusal due to their sensitive nature, the primary reason item nonresponse exceeded 15 percent for these items is that each applies to a relatively small subset of respondents (i.e., small denominator) and these items were not included in the abbreviated instrument. The nonresponse bias analysis and details of faculty item nonresponse are presented in appendix I.

4.3 Faculty Data Quality

4.3.1 Item Mode Effects

The NSOPF:04 faculty instrument was designed to minimize potential mode effects by using a single instrument for both self-administration and CATI. However, whenever multiple modes are used for data collection, the possibility of mode effects is inherent. Because respondents were offered the option of completing the interview by themselves on the Web or with an interviewer, there was the potential for bias due to self-selection or other factors which cannot be accounted for. Therefore, these results should be interpreted as how respondents in different modes of administration answered the survey questions and not as true mode differences.

Due to the large sample size, nearly all test statistics used to measure differences between self-administered and CATI respondents were significant. Reporting all of these statistically significant differences is not substantively meaningful; therefore, only differences of five percentage points or greater are reported.[29,30]

For this analysis, 47 variables were selected, covering the following topic areas: demographic variables, descriptive items, factual items, and opinion-based questions. Criteria for selection of items included importance to the content of this study. Items for which project staff had concerns that there might be mode effects (e.g., complex matrix items) were also selected. Although not presented in tables, the following discussion on item mode effects is based on special tabulations from the 2004 NSOPF faculty data.

Demographics

Compared to their CATI counterparts, web respondents were more likely to be White (Q74E: 87 percent versus 81 percent, $z = 11.65$, $p < .001$). Conversely, CATI respondents were

[28] The items included in this analysis are listed in appendix K (Q1 through Q83). The number of items differs from the number of faculty items reported elsewhere in this document. For example, the difference between the number of analysis variables (162) and the number of items in the faculty questionnaire (183) occurs because some items in the faculty questionnaire were for internal use; similarly, there are fewer stochastically imputed variables (144) than analysis variables (162) because some variables had no missing data after logical imputations were performed.
[29] For questions where means were used, the unit of measurement, range of answers, and standard deviation was evaluated to determine which statistically significant differences to report.
[30] Footnotes are used to report differences in the 3 to 4 percent range since these could be seen as indicative of a substantively important difference.

more likely to be Black or African American (Q74C: 14 percent versus 7 percent, $z = -16.98$, $p < .001$) than their web counterparts. No mode differences were observed for gender or age.[31]

Descriptors

Web respondents were more likely than CATI respondents to report research as their primary activity (Q4: 11 percent versus 6 percent, $z = 11.62$, $p < .001$), be employed full-time (Q5: 73 percent versus 53 percent, $z = 29.72$, $p < .001$), be an assistant professor (Q10: 19 percent versus 12 percent, $z = 17.81$, $p < .001$), be tenured (Q12: 34 percent versus 28 percent, $z = 8.79$, $p < .001$) or be on the tenure track (Q12: 16 percent versus 10 percent, $z = 11.68$, $p < .001$), and not be employed outside the target institution (Q18: 72 percent versus 59 percent, $z = 61.15$, $p < .001$). CATI respondents were more likely than web respondents to report teaching as their principal activity (Q4: 77 percent versus 70 percent, $z = -10.73$, $p < .001$), be an instructor (Q10: 27 percent versus 17 percent, $z = -17.49$, $p < .001$), not be on the tenure track (Q12: 51 percent versus 41 percent, $z = 28.58$, $p < .001$), and be employed outside the target institution (Q18: 31 percent versus 22 percent, $z = -14.54$, $p < .001$).[32]

Factual items

Twenty-four factual items were chosen, based on their importance to the study objectives. These factual items were expected to show few, if any, mode differences. These questions centered on eight main topic areas: number of classes taught, year began first postsecondary job, employment sector of previous job, hours per week spent on various tasks, percent time spent on various tasks, use of various methods in the classroom, other activities, and publications.

Classes taught. There were no significant differences observed in mean number of credit and noncredit classes taught at the target postsecondary institution (Q35A1 and Q35A2).

Year began first postsecondary job. There was no significant difference in the mean year web respondents began their first postsecondary job (Q23) compared to their CATI counterparts.

Employment sector of previous job. Web respondents were more likely to have no other job prior to their current position (Q28: 10 percent versus 5 percent, $z = 12.21$, $p < .001$) than were CATI respondents.[33]

Hours per week spent on various tasks. Web respondents reported spending more time on paid tasks at the target institution (Q31A), on average, than their CATI counterparts (37 hours versus 31 hours, $t = 22.70$, $p < .001$), while CATI respondents reported spending more time on paid tasks outside the institution (Q31C) than web respondents (12 hours versus 7 hours, $t = -20.37$, $p < .001$). No significant differences were found on hours spent on unpaid tasks at the

[31] Two measures showed differences at the 3 and 4 percent level. Web respondents (57 percent) were more likely to be male (Q71) than their CATI counterparts (53 percent, $z = 5.58$, $p < 0.001$), and web respondents (7 percent) were more likely to be Asian (Q74B) than their CATI counterparts (4 percent, $z = 8.44$, $p < 0.001$).

[32] Three measures showed differences at the 3 and 4 percent level. Web respondents were more likely to report administration as their primary activity (Q4: 9 percent versus 6 percent, $z = 7.52$, $p < 0.001$) and be an associate professor (Q10: 17 percent versus 13 percent, $z = 7.54$, $p < 0.001$). CATI respondents were more likely than web respondents to not report an academic title (or use the "other" category) (Q10: 23 percent versus 19 percent, $z = -6.92$, $p < 0.001$).

[33] CATI respondents were more likely to have been employed in an elementary or secondary school prior to their current position (Q28: 20 percent versus 16 percent, $z = -7.37$, $p < 0.001$).

institution (Q31B), unpaid tasks outside the institution (Q31D), or hours spent e-mailing students each week (Q41).

Percentage of time spent on various tasks. Respondents were asked to provide the percentage of time they spent on undergraduate instructional activities (Q32A), graduate instructional activities (Q32B), research activities (Q32C), and other activities (Q32D). CATI respondents reported spending a greater percentage of their time each week on instructional activities with undergraduates than web respondents (61 percent versus 53 percent, $t = -13.71$, $p < .001$).[34]

Use of various methods in the classroom. Of the 11 methods in question, only four showed a significant difference by mode. Compared to web respondents, CATI respondents were more likely to report using multiple choice exams (Q38A: 61 percent versus 55 percent, $z = -6.64$, $p < .001$), using essay midterm or final exams (Q38B: 61 percent versus 51 percent, $z = -10.99$, $p < .001$), and to report using service learning experiences (Q38J: 33 percent versus 26 percent, $z = 25.23$, $p < .001$). Web respondents were more likely to report using a website for instructional duties (Q39) compared to CATI respondents (45 percent versus 35 percent, $z = 13.47$, $p < .01$).[35]

Publications. The average number of articles published in refereed journals in their careers (Q52AA) was no different for web and CATI respondents.

Opinion

Thirteen opinion-based questions were evaluated for mode differences. Eight of these questions asked how satisfied respondents were with various aspects of their job, including: authority to make decisions, technology-based activities, equipment/facilities, institutional support for teaching improvement, workload, salary, benefits, and job overall (Q61 and Q62). As shown in table 33, CATI respondents were significantly more likely to report being either somewhat or very satisfied with five of the eight items—including equipment/facilities, institutional support for teaching improvements, workload, salary, and job overall—compared to web respondents.[36] These differences may be due to the effect of social desirability on responses when an interviewer is involved.

[34] Web respondents reported spending a greater percentage of their time each week on research (Q32C) compared to CATI respondents (15 percent versus 12 percent, $t = 11.63$, $p < 0.001$).

[35] Three items showed differences at the 4 percent level. Compared to web respondents, CATI respondents were more likely to report using multiple drafts of written work (Q38E: 43 percent versus 39 percent, $z = -4.48$, $p < 0.001$), oral presentations by students (Q38F: 65 percent versus 61 percent, $z = -4.52$, $p < 0.001$), and student evaluations of each other's work (Q38H: 41 percent versus 37 percent, $z = -4.52$, $p < 0.001$).

[36] Two additional questions showed significant differences at the 3 and 4 percent level. CATI respondents were more likely than web respondents to report being somewhat or very satisfied with institutional support for technology based instructional activities (Q61B: 89 percent versus 85 percent, $z = -6.96$, $p < 0.001$) and the benefits available to them (Q62C: 74 percent versus 71 percent, $z = -4.15$, $p < 0.001$).

Table 33. Satisfaction with various aspects of job, by mode of administration: 2004

Item	Description	Percent	
		Web	CATI
Q61C	Satisfaction with equipment/facilities	77.5	83.9
Q61D	Satisfaction with institutional support for teaching improvement	68.8	82.7
Q62A	Satisfaction with workload	77.0	83.6
Q62B	Satisfaction with salary	60.7	70.2
Q62D	Satisfaction with job overall	87.3	92.4

NOTE: Percentages are based on those indicating they were somewhat or very satisfied with that aspect of their job.
CATI = computer assisted telephone interview.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

The remaining five opinion-based questions asked respondents to indicate whether they agreed or disagreed that teaching was rewarded, part-time faculty were treated fairly, female faculty were treated fairly, and racial minorities were treated fairly (Q82); and whether they would choose an academic career again (Q83). CATI respondents were more likely than web respondents to somewhat or strongly agree that good teaching was rewarded (82 percent versus 76 percent, $z = -9.35$, $p < .001$) and part-time faculty were treated fairly (75 percent versus 65 percent, $z = -13.62$, $p < .001$).[37]

4.3.2 Breakoffs

A total of 27,350 sample members started the faculty interview. Of these, 800 were deemed ineligible. Of the 26,550 eligible sample members who started the interview, 26,110 completed either a full, abbreviated,[38] or partial interview.[39] An additional 10 cases either refused to be included as respondents or provided insufficient data to be useful. The remaining 430 broke off before completing the workload section (C) and were not considered to be partial completes.

Table 34 lists the forms (screens) that had more than 15 breakoffs. In most cases, the forms with the highest number of breakoffs required detailed recall or requested sensitive information.

[37] One additional question showed a significant difference at the 3 percent level. CATI respondents were more likely than web respondents to either somewhat or strongly agree that female faculty members are treated fairly (Q82: 91 percent versus 88 percent, $z = -6.04$, $p < 0.001$).
[38] The abbreviated interview consisted of sections A (nature of employment), B (academic/professional background) and G (sociodemographic characteristics) of the faculty interview.
[39] Interviews that broke off after completing section C (workload) were considered partial completes. Of the 140 respondents who did so, 48 percent broke off in the scholarly activities section (D), 9 percent in the job satisfaction section (E), 29 percent in the compensation section (F), 11 percent in the characteristics section (G), and 4 percent in the opinions section (H).

Table 34. Faculty instrument forms where more than 15 sample members terminated the interview: 2004

Forms[1]	Description	Number of breakoffs
Q2	Instructional duties related to credit courses/activities	30
Q3	Faculty status	30
Q4	Principal activity	20
Q17A4	Highest degree school coding	30
Q31	Hours per week, paid and unpaid tasks at institution and elsewhere	40
Q32	Percent of time spent on instruction, research, and other activities	40
Q37	Description of each class taught (number of weeks, credits, students, etc.)	30
Q52A	Career publications/presentations	30
Q66	Income, from institution/other sources	20

[1] The faculty/instructional staff questionnaire was divided into forms (screens) and items. Each form was structured to include related items.
NOTE: Numbers rounded to the nearest 10.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

4.3.3 Classification of Instructional Programs (CIP) Coding

The assisted coding system was designed for the NSOPF:04 full-scale study to decrease respondent burden by reducing the time and effort needed to code responses. The assisted coding system was used to code field of teaching, highest degree field, and principal field of scholarly activity. The codes for each of these fields were identical (see appendix J for a list of codes). Respondents were asked to provide a verbatim string. The assisted coding system parsed the string, looking for key words or phrases that matched categories in the database. If a match was located, a list of possible fields was provided for the respondent to choose from. In the event a match was not located or the respondent rejected the fields provided by the system, the respondent could manually code the field. This involved choosing a general category from the 32 categories provided in a drop-down box, and then selecting the specific category within the general category. There were a total of 136 specific categories, but within a general category there were never more than 19 specific categories to choose from.

The anticipated benefit to performing this coding in the interview for web respondents is obvious; the sample member can see the categories and select the appropriate general and specific categories. For telephone-administered interviews, this real-time coding may also improve data quality by capitalizing on the availability of the respondent to clarify coding choices at the time the coding was performed; interviewers were trained to use probing techniques to assist in the coding process.

The assisted coding system coded 75 percent of field of teaching strings, 79 percent of highest degree strings, and 50 percent of field of research strings. The assisted coding matches were accepted more readily by CATI interviewers than by web respondents for field of teaching and highest degree (teaching: 86 percent versus 69 percent, $\chi^2 = 703.7$, $p < .0001$; highest degree: 90 percent versus 74 percent, $\chi^2 = 711.8$, $p < .0001$) but the difference was not significant for field of scholarly activity (57 percent versus 48 percent, $\chi^2 = 2.0$, $p < 0.16$).

As part of the data evaluation activities, a random sample of 10 percent of the results for each of the three Classification of Instructional Program (CIP) codings (teaching, research, and highest degree) was selected. An expert coder evaluated the verbatim strings for completeness

and for the appropriateness of the assigned codes, determining whether a string was too vague to code or whether a different code should be assigned.

Overall, 71 percent of those sampled for recoding were coded correctly, 13 percent were incorrectly coded, and 15 percent of the strings were too vague to determine whether they were correctly coded. Table 35 shows the results of the 10 percent recode, by mode. The expert coder agreed with the coding performed by the web respondent more often than that done by the CATI interviewer ($\chi^2 = 9.69$, $p = 0.002$).

Table 35. Summary of coding results for fields of teaching, research, and highest degree, by mode of administration: 2004

Classification of Instructional Programs (CIP) field item	Web respondents				CATI respondents			
	Coding attempts sampled	Percent coded correctly	Percent recoded	Percent too vague to code	Coding attempts sampled	Percent coded correctly	Percent recoded	Percent too vague to code
Total	3786	72.3	14.0	13.7	1184	67.7	11.6	20.8
Teaching	1949	72.3	14.4	13.2	651	67.4	12.4	20.1
Research	39	53.8	28.2	17.9	5	60.0	20.0	20.0
Highest degree	1798	72.7	13.2	14.1	528	68.0	10.4	21.6

NOTE: Detail may not sum to totals because of rounding. CATI = computer assisted telephone interview.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

In addition to the 10 percent recode, all strings that were not coded, were partially coded (into a general area but not a specific discipline), or were coded "other" were evaluated by the expert coder and upcoded into the appropriate CIP categories, where possible. Of the 52,018 verbatim strings provided, a total of 1,506 strings (3 percent) qualified for this upcoding; 79 percent of these were web respondents and 21 percent were CATI respondents. Of these 1506 strings for which upcoding was attempted, 82 percent were upcoded, 18 percent were too vague to code, and less than 1 percent were correctly coded as "other."

4.3.4 IPEDS Coding

The faculty instrument included a coding system that assisted web respondents and interviewers in collecting postsecondary institution information. This system was designed to improve data quality by allowing respondents to clarify coding choices at the time coding was performed. To assist in the coding process, web respondents were given detailed instructions on screen that enabled them to locate the postsecondary institution. In addition to these on-screen instructions, interviewers were given additional supervised training on how to effectively probe and code respondents' answers.

The institution coding system assigned a six-digit IPEDS identifier for the postsecondary institution that awarded the respondent's highest degree. To facilitate coding, the coding system requested the state and city in which the school was located; from that information, a list of possible schools was displayed, allowing the respondent to select the correct school. The system relied on a look-up table of institutions constructed from the IPEDS institution database.

Of the approximately 25,760 institutions coded over the course of data collection, 1,130 were initially deemed uncodeable. However, based on the information collected (institution name, location, level, and control), 1,025 institutions were positively identified and recoded

during the data file editing stage of the project. Of the remaining 105 uncodeable institutions, 65 provided insufficient data, 20 were identified as closed, 10 were identified as foreign, and 10 were online institutions for which no IPEDS ID was available.

4.3.5 Monitoring

Regular monitoring of telephone data collection serves a number of goals, all aimed at maintaining a high level of data quality. These objectives are to identify problem items; to improve interviewer performance by reinforcing good interviewing behavior and discouraging poor behavior; to detect and prevent deliberate breaches of procedure, such as data falsification; and to assess the quality of the data collected.

Two types of monitoring were performed during the NSOPF:04 data collection. The first type was monitoring by project staff, which involved listening to the interview and simultaneously viewing the progress of the interview on screen, using remote monitoring telephone and computer equipment. Project staff evaluated such things as whether the interviewer sounded professional, probed for complete answers, and handled refusal cases appropriately. Interviewers received feedback on their skills, and additional training was provided, if necessary. When monitoring interviews, project staff also evaluated whether the interview was functioning properly and identified questions in the interview that were difficult to administer so that those items could be revised in future studies.

The second type of monitoring, quality assurance monitoring, was conducted by specially trained monitoring staff within the call center. Similar to project staff monitoring, the monitoring system provided for simultaneous listening and viewing of the interview. Monitors evaluated the interviewer-respondent interchange on whether the interviewer (1) delivered the question correctly and (2) keyed the appropriate response. Each of these measures was quantified and daily, weekly, and cumulative reports were produced. Monitoring took place throughout data collection, although monitoring efforts were scaled back around the 19th week due to lighter caseloads corresponding with the end of the academic year for many schools.

Of the 3,221 items monitored, a total of 28 question delivery errors and 14 data entry errors were observed.[40] This yielded an average error rate of 0.9 percent for question delivery and 0.4 percent for data entry.

4.3.6 Interviewer Feedback

Quality Circle meetings

Quality Circle meetings provided opportunities for interviewers, supervisors, and project staff to discuss data collection issues. These meetings were scheduled regularly throughout the data collection period to ensure that CATI interviews were being conducted in the most effective manner. Interviewer representation was determined by a supervisor so that all staff would have the opportunity to attend these meetings. Project staff updated interviewers and supervisors on the progress of data collection and gathered information to solve problems encountered by interviewers while conducting interviews. The minutes from these meetings were prepared by project staff and were distributed to all interviewers and supervisors. Meeting minutes were

[40] The data for one monitoring session on April 20, 2004 (61 items observed) were removed from the analysis because the monitor did not follow proper procedure.

available in hardcopy and online. Examples of issues raised in Quality Circle meetings included the following.

Progress of data collection. Project staff provided updates regarding the interviews completed to date and goals for the upcoming week. This information benefited both the interviewers and technical staff by recognizing interviewers' efforts and encouraging continued professionalism.

CATI Case Management System (CMS) issues. Interviewers had an opportunity to report CMS issues that required project staff review and discussion. Using the information provided by interviewers, project staff resolved these issues throughout data collection.

Data collection reminders. Several issues were stressed throughout data collection: reminders to verify address information for cases that needed to be remailed and for addresses for incentive checks, how to handle eligibility questions, and tips for locating sample members who are part-time employees. Interviewers were also reminded to complete problem sheets (see later section in this chapter) for any cases that needed attention.

Instrument issues. During the Quality Circle meetings, project staff clarified specific items in the instrument for the interviewers. These items were brought to the attention of project staff in problem sheets, project staff monitoring, or during the Quality Circle meetings themselves. Discussions focused on how to properly code responses (e.g., for Q10, adjunct faculty should be coded as "other," for questions expecting a numeric response, answers between zero and one should be rounded up to one).

Coding. The majority of online coding during data collection was accurate, based on evaluation of verbatim strings and the codes assigned (see earlier section in this chapter on CIP coding), although in some cases the verbatim string was too vague to code. Interviewers were reminded to ask the sample member for the necessary level of detail while entering the verbatim string.

Web issues. A number of web-related issues were raised during Quality Circle meetings. Some sample members reported problems connecting to the website so interviewers were asked to first try to collect the data via CATI or to have someone from the help desk assist the sample member to get connected. Interviewers were reminded to clearly state the study web address (URL) to sample members.

Problem sheets

When interviewers encountered problems during an interview, a description of the issue was documented in the form of an electronic problem sheet. Project and interviewer supervisory staff regularly reviewed these problem sheets and worked on resolving problems, as appropriate. Approximately 1,169 problem sheets were submitted during the data collection period.

Problem sheets were used as follows:

- To address technical CMS issues. Interviewers documented details of the front-end issues so that a programmer could resolve them.
- To report system and web delays or access problems.
- To document sample member contact information as a workaround for front-end issues.

- To alert project staff to questions about sample member eligibility, contact information, and refusals.
- To record incorrect data that were entered (but not corrected) for a case. Interviewers noted cases where project staff needed to take specific action. Project and interviewer supervisory staff ensured that issues pertinent to data collection were resolved as soon as possible.

Interviewer debriefing

A debriefing meeting was held at the end of data collection. The purpose of this meeting was to elicit feedback from the interviewers on various aspects of the data collection process, particularly the administration of the faculty questionnaire. In attendance were telephone interviewers, help desk operators and their supervisors, selected project staff, and the study project officer. The debriefing session was highly informative and gave project staff a wealth of information that will inform instrumentation and data collection activities for future studies.

Project staff asked interviewers which items in the instrument were problematic. Interviewers responded with general comments as well as item-specific ones, based on their interviewing experience.

General comments. Interviewers reported that sample members repeatedly indicated that parts of the questionnaire did not apply to them. Typically these respondents were part-time faculty or those who taught at community colleges, medical, or other specialty schools.

Interviewers felt that that the pop-up boxes used to confirm out-of-range values were intrusive, and slowed the pace of the interview unnecessarily. They recommended that pop-up boxes be used sparingly in future web questionnaires.

Question 1. Interviewers felt that the first question in the interview, which asked whether the respondent had instructional duties, was too long and "wordy." They recommended that the question be shortened or broken into parts.

Question 3. Interviewers reported that adjunct faculty did not know what was meant by faculty status.

Question 9. The second sentence in the wording of this item ("consider promotions in rank as part of the same job") was confusing for respondents. Interviewers suggested restructuring the question to include that information before the respondent attempts to answer the question.

Question 15. Q15 (reason for not being a member of a union) had a high rate of don't know responses. Interviewers said this was because adjunct faculty often did not know whether unions were available.

Questions 16, 17, and 54. Interviewers were quite pleased with the new assisted coding system for field of teaching, highest degree, and scholarly activity. It proved to be less burdensome for them, although they indicated some difficulties finding the exact categories that the respondent wanted.

Question 17. Interviewers reported that IPEDS coding screens (Q17A4) were easy to use. One concern was that some schools were listed in the wrong city.

Question 31. Sample members had difficulty distinguishing between paid and unpaid activities, and their ideas of each often differed from the examples provided in the instrument. Some respondents were upset at having to account for their time. Interviewers reported that respondents found this set of questions (Q31, Q32) difficult to answer as it was a lot of information to account for and difficult to break it down precisely.

Questions 31, 41, 47B, 48. Interviewers pointed out that sample members had a hard time providing answers to the hours per week questions when it is something they only do a couple of weeks out of the term (e.g., advising students). They thought some other unit of time might make it easier to collect this information.

Questions 32, 37, 47. Project staff questioned whether there was any confusion over "first-professional students." Interviewers indicated that some faculty at technical schools did not know what was meant by first-professional students.

Question 35. Interviewers reported that sample members often were unclear what was meant by the term "distance education" in Q35C and suggested including the words "Internet courses" in the question wording.

Question 37. Interviewers indicated that the screen takes a lot of time to complete and those who teach unstructured courses found these items difficult to answer.

Question 50. Advising of students (Q50) was a difficult concept for some sample members in the field test and the wording was changed in the full-scale instrument to clarify the meaning. Interviewers indicated this was still a problem and suggested changing the definition provided on-screen. Respondents also wanted clarification of whether this was designated advisees only or whether it included other advising.

Question 52. Interviewers reported that Q52 (number of scholarly works) was administered fairly smoothly; most respondents had a general idea of the number of publications and presentations although a few consulted their resumes. A small number of respondents had numbers of publications that exceeded the maximum allowed and became upset that their volume of scholarly activity was not properly reflected. Interviewers reported that respondents seemed to get tired around this point in the instrument and felt that combining screens for Q52A and Q52B would improve the flow and reduce burden on the respondents.

Question 53. In the field test, respondents sometimes reported confusion over what was meant by scholarly activity. The question text was revised, and this problem was not reported in the full-scale questionnaire. However, some respondents were unsure whether to report only scholarly activities associated with the target institution or all scholarly activities.

Question 62. In the field test, Q62C (satisfaction with benefits) was not answered by many respondents (mostly part-timers) because they did not receive benefits. The wording was slightly altered for the full-scale questionnaire; however, interviewers reported that many part-time and adjunct faculty still could not answer this question. In particular, some respondents were unsure whether the question was asking about medical benefits or other benefits.

Questions 66 and 70. Sample members complained that Q66 and Q70 (income) items were intrusive. Interviewers suggested that having scripted text for why this question is asked would be helpful. Interviewers felt that income questions were unnecessarily repetitive.

Question 74. Respondents insisted that "Caucasian" be listed among the response options (in parentheses after "White"). Interviewers suggested adding scripted text to explain why race is asked about on this form.

Question 82. Q82D (racial minorities treated fairly) had more than 10 percent missing when administered in CATI. Interviewers explained that some part-time and adjunct faculty did not have an opinion on this set of items. They suggested adding a "no opinion" option for each item on this form.

Interviewers who worked on the field test requested that Q84 (feedback textbox) be put back in the instrument as many sample members wanted to provide feedback.

4.3.7 Instrument Feedback

Two issues with the faculty instrument became apparent in the data editing process. The first issue had to do with Q1, whether the respondent had any instructional duties. Despite question wording intended to get the respondent to think beyond classroom teaching, half of the respondents who said they did not have any instructional duties provided responses indicating they did have instructional duties on other items in the instrument (i.e., taught one or more credit or noncredit classes [Q35A1>0 or Q35A2>0], provided any individual instruction [Q46=1], spent time on thesis or dissertation committees, comprehensive exams or orals committees [Q48>0], indicated that teaching was their principal activity [Q4=1], or spent time on undergraduate or graduate instructional activities [Q32A>0 or Q32B>0]). Items Q32A and Q32B contradicted Q1 most often. Rather than reconciling in the data editing phase, future cycles of NSOPF would benefit from asking follow-up questions immediately after Q1 for those respondents who said they did not have instructional duties.

The other issue concerned items asking about first-professional students (Q32B, Q37E, Q47A, and Q47B). This term, first-professional student, was apparently misunderstood by many faculty and instructional staff at 2-year institutions who indicated they taught first-professional students at that institution. While on-screen examples of first-professional programs were available on some of these forms, in the future it is advised that a check against level of target institution be inserted into the instrument logic for questions concerning first-professional students.

4.4 Comparisons with NSOPF:99

To assess the consistency of survey estimates between the current and prior administrations of NSOPF, weighted estimates were obtained from the 1999 and 2004 survey data for a series of key analytical variables. The results of these assessments are summarized in table 36.

Table 36. Weighted estimates obtained based on the 1999 and 2004 survey data for a series of key analytical variables: 2004

Variable	Weighted estimates	
	1999	2004
Percent of full-time faculty who were tenured	53.1	47.3
Percent of part-time faculty who were not on tenure track	78.3	86.2
Percent of faculty who were part time	42.6	43.7
Percent of part-time faculty who had retired from another position	15.4	19.6
Percent of full-time faculty whose principal activity was teaching	64.5	62.5
Percent of full-time faculty whose principal activity was research	11.3	14.2
Average percent of time that full-time faculty taught undergraduates	41.3	43.2
Percent of full-time faculty with a doctorate	57.7	59.6
Average number of hours full-time instructional faculty taught per week	11.0	11.1
Average number of hours part-time instructional faculty taught per week	7.3	7.7
Average number of recent refereed publications, full-time faculty	3.9	2.2
Average number of recent refereed publications, part-time faculty	1.2	0.5
Average number of career refereed publications, full-time faculty	16.0	16.0
Average number of career refereed publications, part-time faculty	4.4	4.1
Average basic income of full-time faculty	$56,841	$67,239
Average basic income of part-time faculty	11,613	11,010
Average consulting income of full-time faculty who consulted	8,221	7,379
Average consulting income of part-time faculty who consulted	10,579	10,908
Average household income, full-time faculty	163,127	117,702
Average household income, part-time faculty	125,693	92,636
Percent of full-time faculty who were Asian	5.5	9.0
Percent of part-time faculty who were Asian	2.9	3.8
Percent of full-time faculty who were Black	4.9	5.8
Percent of part-time faculty who were Black	4.3	5.7
Percent of full-time faculty who were Hispanic	3.4	3.5
Percent of part-time faculty who were Hispanic	3.9	3.5
Percent of full-time faculty who were White	85.1	81.7
Percent of part-time faculty who were White	87.6	86.9
Percent of full-time faculty who were female	36.3	38.6
Percent of full-time faculty in agriculture and home economics	0.6	2.5
Percent of full-time faculty in business	7.4	6.4
Percent of full-time faculty in education	8.7	7.6
Percent of full-time faculty in engineering	2.4	4.9
Percent of full-time faculty in fine arts	9.3	6.4
Percent of full-time faculty in health sciences	12.2	13.9
Percent of full-time faculty in first-professional health science[1]	4.3	6.6
Percent of full-time faculty in humanities	18.1	13.5
Percent of full-time faculty in natural sciences	16.1	22.4
Percent of full-time faculty in social sciences	9.6	10.5
Percent of full-time faculty in all other programs	15.7	12.1

[1] First-professional health science is a subset of health sciences (previous row).

NOTE: Detail may not sum to totals due to rounding. Differences in estimates between NSOPF:99 and NSOPF:04 may be due to a number of factors, including actual changes over time, differences in how an item was asked between the two years (see table 8), and data editing and imputation procedures (see chapter 5).

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Chapter 5
Data File Development and Imputation

This chapter provides an overview of all procedures used in the development of data files, including descriptions of data editing processes, data swapping, statistical imputations, and derived variable creation.

5.1 Overview of the NSOPF:04 Data Files

Data obtained from the 2004 National Study of Postsecondary Faculty (NSOPF:04) faculty and institution questionnaires are contained in two restricted data files (faculty and institution), which are available on a CD-ROM to researchers who have applied for and received authorization from the National Center for Education Statistics (NCES) to access restricted research files. The restricted data files are documented by an Electronic Codebook (ECB), a Windows-based interface that allows users to view descriptive information and statistics about variables and to select variables for extraction into SAS or SPSS data files. The faculty and institution data files can be merged together for joint analysis. The following files were produced:

Faculty data file. Provides faculty-level questionnaire data collected from 26,110 respondents. These data have been edited, swapped, and imputed. The file contains survey variables (variables that start with Q), derived variables (variables that start with X), and study weights for the faculty file (WTA00) and for the combined faculty and institution files (WTC00—or contextual weight). It also contains replicate weights for variance estimation for the faculty file (WTA01-WTA64) and for the combined faculty and institution files (WTC01-WTC64), the imputation flags (variables that start with F), and INSTID (the IPEDS ID) that will allow faculty file data to be merged with institution file data.

Institution data file. Provides institution-level data collected from 920 institutions. These data have been edited, perturbed, and imputed. The file contains the institution survey variables (variables that start with the letter I), derived variables (variables that start with the letter X), study weight for the institution file (WTB00), replicate weights for variance estimation for the institution file (WTB01-WTB64), imputation flags (variables that start with the letter FI), and INSTID (the IPEDS ID) that will allow data on the institution file to be merged with data on the faculty file.

The faculty and institution files can be merged together for joint analysis by performing a match merge using the variable INSTID. Please note that not all institutions that completed the institution questionnaire have responding faculty, and not all faculty have associated institution questionnaire data. For this reason, when analyzing the faculty and institution data together, responses should be weighted using the contextual weight variables on the faculty file.

The NSOPF:04 institution and faculty analysis variables are presented in appendix K.

5.2 Data Coding and Editing

The NSOPF:04 data were coded and edited using procedures developed and implemented for previous NCES-sponsored studies. These coding and editing procedures were refined during the field test for use in the processing of NSOPF:04 full-scale data.

A large part of the data editing and coding was performed in the data collection instruments, including range edits; across-item consistency edits; and coding of fields of teaching, scholarly activities, and highest degree. During and following data collection, the data were reviewed to confirm that the data collected reflected the intended skip-pattern relationships. At the conclusion of data collection, special codes were inserted in the database to reflect the different types of missing data. There are a number of explanations for missing data; for example, the item may not have been applicable to certain respondents or a respondent may not have known the answer to the question. Table 37 lists the set of consistency codes used to assist analysts in understanding the nature of missing data associated with the NSOPF:04 data elements. With the exception of the not applicable codes, missing data were stochastically imputed (see section 5.4). Moreover, for hierarchical analyses and developing survey estimates for faculty members corresponding to sample institutions that provided faculty lists and responded to the institution survey, contextual weights were produced for such subsets of the responding faculty members. These weights, which aggregate to a number less than the weighted total for all responding faculty and instructional staff, are named *WTC00* and can be found in *weights.dat* on the ECB file.

Table 37. Description of missing data codes: 2004

Missing data code	Description
-1	Don't know; later set to missing and imputed
-3	Not applicable (item was intentionally skipped)
-5	Not applicable (item was asked but respondent indicated it was not applicable)
-7	Item was not administered (abbreviated interview) or reached (partial interview); later imputed
-9	Respondent did not provide an answer; later imputed

SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

The data cleaning and editing process for NSOPF:04 consisted of the following steps:

Step 1. *Review of one-way frequencies for every variable to confirm no missing or blank values and to check for reasonableness of values.* This involved replacing blank or missing data with -9 for all variables in the instrument database and examining frequencies for reasonableness of data values.

Step 2. *Review of two-way cross-tabulations between each gate-nest[41] combination of variables to check data consistency.* Legitimate skips were identified using the interview programming code as specifications to define all gate-nest relationships and replace -9 (missing values that were blank because of legitimate skips) with -3 (legitimate skip code). Additional checks ensured that the legitimate skip code

[41] Gate variables are items that determine subsequent instrument routing. Nest variables are items that are asked or not asked, depending on the response to the gate question. For example, in the faculty questionnaire, Q1 (which asks whether the respondent had instructional duties) determines whether Q2 (which asks whether the respondent's instructional duties were related to credit courses/activities) is asked. Q1 is a gate item and Q2 is a nested item. Q2 is only asked if the response to Q1 was "yes."

was not overwriting valid data and that no skip logic was missed. In addition, if a gate variable was missing (-9), then the -9 was carried through the nested items.

Step 3. *Identify and code items that were not administered due to a partial or abbreviated interview.* This code replaced -9 values with -7 (item not administered) based on the section completion and abbreviated interview indicators.

Step 4. *Recode "don't know" responses to missing.* This code replaced -1 (don't know) values with -9 (missing) for later stochastic imputation. For selected items for which "don't know" seemed like a reasonable response, variables were created both with and without the "don't know" category.

Step 5. *Identify items requiring recoding.* During this stage, previously uncodeable values (e.g., text strings) collected in the various coding systems were upcoded, if possible (see sections 4.3.3 CIP coding and 4.3.4 IPEDS coding).

Step 6. *Identify items requiring range edits, logical imputations, and data corrections.* Descriptive statistics for all continuous variables were examined. Values determined to be out-of-range were either coded to the maximum (or minimum) reasonable value or set to missing for later imputation. Logical imputations were implemented to assign values to legitimately skipped items whose values could be implicitly determined from other information provided. Data corrections were performed where there were inconsistencies between responses given by the sample member.

Concurrent with the data cleaning process, detailed documentation was developed to describe question text, response options, recoding, range edits, logical imputations, data corrections, and the "applies to" text for each delivered variable.

5.3 Data Perturbation

A restricted faculty-level data file was created for release to individuals who apply for and meet standards for such data releases. While this file does not include personally identifying information (i.e., name and Social Security number), other data (i.e., institution, Integrated Postsecondary Education Data System [IPEDS] ID, demographic information, and salary data) may be manipulated in such a way to seem to identify data records corresponding to a particular faculty member. To protect further against such situations, some of the variable values were swapped between faculty respondents. This procedure perturbed and added additional uncertainty to the data. Thus, associations made among variable values to identify a faculty respondent may be based on the original or edited, imputed and/or swapped data. For the same reasons, the data from the institution questionnaire were also swapped to avoid data disclosure.

5.4 Imputation Methodology

The NSOPF:04 data files include institution-level and faculty-level data obtained from the institution and faculty surveys. All non-verbatim and non-text variables on the NSOPF:04 that had missing variables have been imputed. Specifically, a total of 144 variables were stochastically imputed for the faculty data, and 87 variables were stochastically imputed for the institution data. All remaining missing data were deemed not suitable for imputation, such as the postsecondary institution that awarded the highest degree of a faculty respondent. Most of these

variables were imputed using a weighted sequential hot-deck imputation procedure. A number of variables, including gender and race/ethnicity, were imputed using a combination of cold-deck and logical imputation during the data editing process before the data file was considered ready for stochastic imputation. The specific imputation method used for each variable is specified in the imputation flags on the final restricted datasets.

Table 38 shows the number of variables that were imputed based on the percent missing (imputed) for faculty and institution survey data. Accordingly, data for 26 of the 144 faculty variables were imputed for less than 1 percent of all faculty respondents, whereas data for 7 of the faculty variables were imputed for more than 15 percent of the faculty respondents.

Table 38. Prevalence of missing/imputed data for the faculty and institution surveys: 2004

Percent imputed	Faculty variables	Institution variables
Total	144	87
Less than 1 percent	26	0
Between 1 and 5	11	58
Between 5 and 10	93	15
Between 10 and 15	7	11
Over 15	7	3

NOTE: There are fewer stochastically imputed variables for the faculty and institution questionnaires than their corresponding analysis variables, since a subset of such items had no missing values after application of logical imputation.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

5.4.1 Imputation Methods

In broad terms, there are three methods of imputation: logical, cold-deck, and hot-deck imputation. Logical imputation is a process that aims to infer or deduce the missing values from answers to other questions. Cold-deck imputation involves replacing the missing values with data from sources such as data used for sampling frame construction. While resource intensive, these methods often obtain the actual value that is missing. Consequently, attempts were made to fill in the missing values of data using these two methodologies, to the extent possible. In contrast, stochastic imputation methods, such as sequential hot-deck imputation, rely on the observed data to provide replacing values (donors) for records with missing values.

Sequential hot-deck imputation involves defining imputation classes, which generally consist of a cross-classification of covariates, and then replacing missing values sequentially from a single pass through the survey data within the imputation classes. When this form of imputation is performed using the sampling weights, the procedure is called weighted sequential hot-deck imputation. This procedure takes into account the unequal probabilities of selection in the original sample to specify the expected number of times a particular respondent's answer will be used as a donor. These expected selection frequencies are specified so that, over repeated applications of the algorithm, the weighted distribution of all values for that variable—imputed and observed—will resemble that of the target universe. Under this methodology, while each respondent record has a chance to be selected for use as a hot-deck donor, the number of times a respondent record can be used for imputation will be controlled

To implement the weighted sequential hot-deck procedure, imputation classes and sorting variables that are relevant (strong predictor) for each item being imputed were defined. For this study, imputation classes were developed by using a Chi-squared Automatic Interaction

Detection (CHAID) analysis. The CHAID segmentation process divides the data into groups based on the most significant predictor of the item being imputed. Subsequently, this procedure will be repeated using the remaining predictor variables to split each of the emerging groups into smaller subgroups. In this process, a number of subgroups created during a previous iteration might get merged back to form new subgroups. This splitting and merging process continues until no more statistically significant predictors are found, at which point imputation classes are defined from the resulting segments. When dealing with categorical variables, the CHAID process may merge certain categories of such variables that are found not to be significantly different. Similarly, continuous variables are categorized to create the strongest categorical predictors of the item in question.

Using RTI's sequential hot-deck method of imputation, once imputation classes are constructed, items within each class are sorted before the process of donor selection begins. If more than one sorting variable is chosen, a serpentine sort will be performed where the direction of the sort (ascending or descending) changes each time the value of a variable changes. The serpentine sort minimizes the change in the respondent's characteristics every time one of the variables changes its value.

It should be noted that, for this study, distinction was made between legitimate and non-legitimate missing items for imputation. All responses that were left missing as a result of refusal were set to missing and then imputed. Additionally, if the interview was terminated early and some questions were not asked of the respondent, then the value of missing was assigned in those cases as well. However, respondents could legitimately skip questions that did not apply to them. In these cases, the missing responses were coded as legitimate skips (-3) and were not imputed.

5.4.2 Imputation of Faculty Data

Item imputation for the faculty questionnaire was performed in several steps. In the first step, the missing values of gender, race, and ethnicity were filled—using cold-deck imputation—based on the sampling frame information or institution record data. These three key demographic variables were imputed prior to any other variables since they were used as key predictors for all other variables on the data file.

After all logical and cold-deck imputation procedures were performed, the remaining variables were imputed using the weighted sequential hot-deck method. Initially, variables were separated into two groups: unconditional and conditional variables. The first group (unconditional) consisted of variables that applied to all respondents, while the second group (conditional) consisted of variables that applied to only a subset of the respondents. That is, conditional variables were subject to "gate" questions. After this initial grouping, these groups were divided into finer subgroups as detailed next.

The unconditional group was divided into two subgroups based on the percent of missing values: less than 1 percent versus greater than 1 percent missing. The conditional variables were divided into three subgroups based on the level of conditionality where this level was essentially determined by the sequence of the questionnaire. For variables in the conditional group, the questionnaire skip patterns were reviewed and variables were grouped according to which variables determine the values of other variables. After these subgroups were constructed, missing values of the variables were imputed in order from lowest percent missing to highest

percent missing within each subgroup, first for the unconditional variables and then for the conditional variables in an ascending level of their conditionality.

All unconditional variables that had less than one percent missing were imputed using imputation classes defined by a combination of gender, race, and ethnicity. Moreover, institution type,[42] institution size, and faculty type[43] were used as sort variables to place like records in closer proximity to improve the donor selection process. The imputation classes for the remaining unconditional variables (that had more than one percent missing) and all conditional variables were determined by a CHAID analysis based on key demographic variables that were logically imputed and all imputed variables that had less than one percent missing. After all variables were imputed, consistency checks were applied to the entire faculty data file to ensure that the imputed values did not conflict with other questionnaire items, observed or imputed. This process involved reviewing all of the logical imputation and editing rules as well.

5.4.3 Imputation of Institution Data

The imputation process for the missing data from the institution questionnaire involved similar steps to those used for imputation of the faculty data. The missing data for variables were imputed using the weighted sequential hot-deck method. Analogous to the imputation process for the faculty data, the variables were partitioned into conditional and unconditional groups. The unconditional variables were sorted by percent missing and then imputed in the order from the lowest percent missing to the highest. The conditional group was partitioned into three subgroups based on the level of conditionality for each variable, and then imputed in that order. The imputation class for both unconditional and conditional variables consisted of the institution sampling stratum, and the sorting variables included the number of full-time and part-time faculty members.

5.4.4 Evaluation of Imputations

A common measure for determining whether an imputation method produces acceptable results (donors) is based on the similarity of the before and after imputation distributions within imputation classes. For evaluation of the imputation results, distributions were considered to be similar when absolute differences were less than 5 percent, where the absolute difference was calculated by comparing the before and after imputation weighted frequencies. If absolute differences were greater than 5 percent, the unweighted distributions were examined to see if the large differences were due to small imputation cells. When possible, such cases were evaluated and resolved by collapsing neighboring imputation classes. The before and after imputation distributions of several key variables are presented in table 39 for the faculty data and table 40 for the institution data. For more information regarding the bias due to item nonresponse, refer to appendix I.

[42] Institutional type consisted of a cross-classification of control (public verses private not-for-profit) and degree type (doctoral, master's, baccalaureate, associate's, and other).
[43] Faculty type (stratum) is based on faculty demographics, such as gender, race/ethnicity, and employment status.

Table 39. Before and after imputation distributions of key faculty questionnaire variables: 2004

Variable description	Variable category	Before imputation		After imputation	
		Number	Percent	Number	Percent
Faculty status	Total	26,050	100.0	26,110	100.0
	No faculty status	1,670	7.1	1,680	7.2
	Had faculty status	24,390	92.9	24,430	92.8
Principal activity	Total	26,030	100.0	26,110	100.0
	Teaching	18,660	73.2	18,710	73.2
	Research	2,470	8.9	2,470	8.9
	Public service	260	1.0	260	1.0
	Clinical service	1,260	4.6	1,270	4.7
	Administration	2,070	7.2	2,070	7.3
	On sabbatical	380	1.4	380	1.4
	Other activity	940	3.7	950	3.7
Employed full or part time at this institution	Total	26,100	100.0	26,110	100.0
	Full time	17,750	62.4	17,750	62.4
	Part time	8,350	37.6	8,360	37.6
Rank	Total	26,090	100.0	26,110	100.0
	Not applicable	640	2.7	640	2.7
	Professor	5,220	18.9	5,220	18.9
	Associate professor	4,210	14.8	4,210	14.8
	Assistant professor	4,620	16.1	4,620	16.1
	Instructor	5,050	20.5	5,050	20.5
	Lecturer	1,230	5.3	1,230	5.3
	Other title	5,140	21.7	5,140	21.7
Tenure status	Total	25,930	100.0	26,110	100.0
	Tenured	8,390	30.5	8,420	30.5
	On tenure track but not tenured	3,840	13.4	3,860	13.4
	Not on tenure track	11,330	47.5	11,430	47.6
	Not tenured-no tenure system	2,380	8.6	2,390	8.6
Highest degree	Total	26,090	100.0	26,110	100.0
	No degree	250	1.1	250	1.1
	Doctorate degree	12,180	44.5	12,180	44.5
	First-professional degree	2,010	7.4	2,010	7.4
	Master of fine arts/social work	1,190	4.6	1,190	4.6
	Other master's degree	8,080	32.5	8,090	32.5
	Bachelor's degree	1,870	7.8	1,870	7.8
	Associate's degree or equivalent	390	1.6	390	1.6
	Certificate/diploma-undergrad program	140	0.5	140	0.5
Household income (range)	Total	21,500	100.0	26,110	100.0
	$1-24,999	460	2.4	560	2.3
	25,000-49,999	2,520	11.6	2,990	11.4
	50,000-74,999	4,690	22.1	5,630	21.8
	75,000-99,999	4,500	20.6	5,520	20.9
	100,000-149,999	5,460	25.2	6,690	25.5
	150,000-199,999	2,110	9.9	2,590	10.0
	200,000-300,000	1,330	6.3	1,600	6.1
	More than 300,000	430	2.0	530	2.0

See notes at end of table.

Table 39. Before and after imputation distributions of key faculty questionnaire variables: 2004—Continued

Variable description	Variable category	Before imputation		After imputation	
		Number	Percent	Number	Percent
Race/ethnicity, Hispanic/Latino	Total	26,030	100.0	26,110	100.0
	Not Hispanic/Latino	24,320	95.9	24,400	95.9
	Hispanic/Latino	1,700	4.1	1,700	4.1
Race, American Indian or Alaska Native	Total	25,590	100.0	26,110	100.0
	Not American Indian/Alaska Native	25,060	98.2	25,570	98.2
	American Indian/Alaska Native	530	1.8	540	1.9
Race, Asian	Total	25,590	100.0	26,110	100.0
	Not Asian	24,030	94.0	24,480	93.8
	Asian	1,560	6.0	1,630	6.2
Race, Black or African American	Total	25,590	100.0	26,110	100.0
	Not Black/African American	23,450	93.7	23,940	93.8
	Black/African American	2,140	6.3	2,170	6.2
Race, Native Hawaiian or other Pacific Islander	Total	25,590	100.0	26,110	100.0
	Not Native Hawaiian/Pacific Islander	25,500	99.7	26,020	99.7
	Native Hawaiian/Pacific Islander	90	0.3	90	0.3
Race, White	Total	25,590	100.0	26,110	100.0
	Not White	3,670	12.2	3,780	12.3
	White	21,920	87.9	22,330	87.7

NOTE: Numbers rounded to the nearest 10. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

5.5 Derived Variables

For NSOPF:04, a total of 45 institution-level and 130 faculty-level derived variables were constructed to simplify access to standard queries useful to analysts, as well as to enhance substantive analysis. Since research questions often require independent or control variables, this set of derived variables was added to the faculty data files. The 45 institution-level derived variables were also added to the institution data files. Multiple sources of data were used to create institution-derived variables, including selected 2000–01 and 1997–98 IPEDS surveys, the Carnegie classification system, and NSOPF:04 sampling information.

Table 40. Before and after imputation distributions of key institution questionnaire variables: 2004

Variable description	Variable category	Before imputation		After imputation	
		Number	Percent	Number	Percent
Full-time benefit: medical insurance	Total	900	100.0	920	100.0
	Not Applicable	#	0.4	#	0.4
	All	860	95.8	880	95.9
	Some	40	3.7	40	3.7
	None	#	0.1	#	0.1
Full-time benefit: dental insurance	Total	900	100.0	920	100.0
	Not Applicable	#	0.4	#	0.4
	All	820	88.1	830	88.2
	Some	40	4.3	40	4.2
	None	40	7.2	40	7.1
Full-time benefit: disability insurance	Total	890	100.0	920	100.0
	Not Applicable	#	0.4	#	0.4
	All	790	86.8	810	86.7
	Some	60	6.8	60	6.8
	None	40	5.9	40	6.2
Full-time benefit: life insurance	Total	900	100.0	920	100.0
	Not Applicable	#	0.4	#	0.4
	All	810	92.2	840	92.3
	Some	50	4.4	50	4.4
	None	30	2.9	30	2.9
Full-time benefit: child care	Total	880	100.0	920	100.0
	Not Applicable	#	0.4	#	0.4
	All	200	16.2	210	16.2
	Some	50	3.1	60	3.2
	None	620	80.3	650	80.2
Full-time benefit: retiree medical insurance	Total	860	100.0	920	100.0
	Not Applicable	#	0.5	#	0.4
	All	470	50.6	520	50.4
	Some	210	21.6	220	21.0
	None	180	27.4	190	28.2
Full-time benefit: cafeteria-style plan	Total	880	100.0	920	100.0
	Not Applicable	#	0.4	#	0.4
	All	260	29.4	270	29.1
	Some	20	1.8	20	1.8
	None	600	68.4	630	68.7
Full-time benefit: wellness program	Total	880	100.0	920	100.0
	Not Applicable	#	0.4	10	0.4
	All	570	65.2	600	65.2
	Some	30	3.8	30	3.9
	None	270	30.6	290	30.6
Part-time benefit: medical insurance	Total	900	100.0	920	100.0
	Not Applicable	10	0.4	10	0.4
	All	80	8.1	80	8.1
	Some	390	34.4	400	34.4
	None	420	57.1	430	57.2
Part-time benefit: dental insurance	Total	900	100.0	920	100.0
	Not Applicable	10	0.4	10	0.4
	All	70	7.9	70	7.8
	Some	330	30.0	340	29.8
	None	490	61.8	500	62.1

Rounds to zero.
NOTE: Institution counts are rounded to the nearest 10 to protect the confidentiality of faculty and institutions. However, percentages cited are based on the original unrounded numbers. Detail may not sum to totals due to rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

Chapter 6
Weighting and Variance Estimation

Three sets of analysis weights were calculated for this administration of the 2004 National Study of Postsecondary Faculty (NSOPF:04), details of which are provided in this section. First, a set of analysis weights was calculated for institutions responding to the institution survey. Next, analysis weights were constructed for responding faculty, which reflected the selection probabilities of institutions providing faculty lists and selection of faculty members within sample institutions. In addition, a set of contextual weights was calculated to use when linking faculty and institution survey data. These analysis weights were constructed as the product of corresponding sampling weights and adjustment factors for frame multiplicity, nonresponse, and poststratification to known control totals. As detailed in the following sections, each component of the final analysis weights represents either the inverse of a selection probability or a weight adjustment to reduce bias.

The institution analysis weights were computed as the product of the following five weight components and adjustment factors:

(1) institution sampling weight (WT1);

(2) institution multiplicity adjustment factor (WT2);

(3) institution nonresponse adjustment factor (WT3);

(4) institution poststratification adjustment factor (WT4); and

(5) institution ratio adjustment factor (WT5).

In order to compute the analysis weights for faculty, first a set of primary sampling unit (PSU) weights were created for institutions providing faculty lists. These interim weights, which are of no analytical utility, were only used as component weights for construction of the final analysis weights for faculty members. Ultimately, the faculty analysis weights were computed as the product of the following nine weight components and adjustment factors:

(1) institution sampling weight (WT1);

(2) institution multiplicity adjustment factor (WT2);

(3) institution nonresponse adjustment (WT3)[44];

(4) institution poststratification adjustment factor (WT4);

(5) faculty sampling weight (WT5);

(6) faculty multiplicity adjustment factor (WT6);

(7) faculty unknown eligibility adjustment factor (WT7);

(8) faculty nonresponse adjustment factor (WT8); and

(9) faculty poststratification adjustment factor (WT9).

[44] Note that here separate sets of nonresponse and poststratification adjustment factors (WT3 and WT4) were constructed for each institution as compared to those calculated above, since the set of institutions providing faculty lists was not the same as that responding to the institution questionnaire.

Analogous to the calculation of analysis weights for the faculty, a set of contextual weights was constructed for the subset of faculty for whom their corresponding institutions had responded to the institution survey. Table 41 summarizes the distribution of institutions providing faculty lists and responding to the institution questionnaire by sampling strata.

Table 41. Counts of sampled, eligible, and participating institutions, by institution type: 2004

Institution type	Sampled institutions	Eligible institutions	Responded	
			Faculty list	Questionnaire
Total	1,080	1,070	980	920
Public doctoral	190	190	180	170
Public master's	120	120	100	110
Public bachelor's	30	30	30	30
Public associate's	340	330	290	290
Public other	10	10	10	10
Private not-for-profit doctoral	110	110	100	90
Private not-for-profit master's	80	80	80	70
Private not-for-profit bachelor's	130	130	120	110
Private not-for-profit associate's	10	10	10	10
Private not-for-profit other	60	60	60	50

NOTE: Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04)

RTI's weighting software GEM (Generalized Exponential Modeling) (Folsom and Singh 2000) has been used for calculation of all weight adjustment factors. Taking advantage of an iterative proportional fitting algorithm and the logit method, GEM provides a comprehensive weighting program that can utilize a large number of predictor variables for creating a more balanced set of weights while automatically curtailing extreme weights that can reduce the efficiency of weighted estimates. For more details on the GEM adjustment procedure, see appendix L. This section provides details of steps taken to construct the resulting weights.

6.1 Institution Weights

The institution sampling frame for the NSOPF:04 included a total of 3,380 eligible units, detailed composition of which is provided in section 2.1.1. Reflecting the probability proportional to size scheme of sample selection, the probability of selection for institution *i* in stratum *r* was calculated by:

$$\pi_{ri} = \begin{cases} \dfrac{n_r S_{ri}}{S_{r+}}, & \text{for non-certainty selections} \\ 1, & \text{for certainty selections} \end{cases}$$

where:

n_r = sample size for stratum *r*,

S_{ri} = composite measure of size for institution *i* in stratum *r*, and

S_{r+} = composite measure of size for all institutions in stratum *r*.

The initial sample consisted of 1,220 institutions. However, this sample was reduced to a subsample of institutions, since a smaller sample was deemed adequate to secure all precision requirements of NSOPF:04. Therefore, the sampling weight for institution *i* in stratum *r* was calculated as a function of its initial and subsequent selection probabilities. With R_r representing the subsampling rate in stratum *r*, the sampling weight for the *i*-th institution in that stratum was calculated by:

$$WT1_{ri} = \frac{1}{\pi_{ri}} \times \frac{1}{R_r}$$

It should be noted that during the sample refreshing step, institutions were added to the sample of institutions, resulting in total sample of 1,080 institutions for NSOPF:04.

6.1.1 Adjustment for Institution Multiplicity

During the institution recruitment and faculty list sampling stages, a number of institutions were identified that had two or more records listed on the Integrated Postsecondary Education Data System (IPEDS). In some cases this was caused by institutions that had recently merged, while in other cases the sample institution had sent a single faculty list covering multiple campuses. For sampling purposes, combined faculty lists that could not be separated were treated as merged institutions and identified under a single IPEDS ID for purposes of tracking survey results.

For institutions with more than one chance of selection, a multiplicity adjustment factor was calculated by estimating, as if the selections were independent, the probability that each record could be selected. Consequently, when an institution had *n* chances of selection, its probability of selection was calculated by:

$$1 - \prod_{i=1}^{n} \left(1 - \pi_{ri}\right)$$

Next, a multiplicity adjustment factor for the *i*-th sample institution was calculated by:

$$WT2_i = \frac{1}{\left[1 - \prod_{i=1}^{n} \left(1 - \pi_{ri}\right)\right] \times WT1_i}$$

If the given institution did not require such adjustment, its multiplicity adjustment factor was set to unity. This way, the product of *WT1* and *WT2* equals the reciprocal of the resulting multiple chance of selection for the institutions with positive multiplicity, and equals *WT1* for all other institutions.

6.1.2 Nonresponse Adjustment

For calculating the analysis weights for institutions responding to the institution questionnaire, an institution (questionnaire) level nonresponse adjustment factor (WT3) was constructed using the product of the institution sampling weights (adjusted for multiplicity) and

the faculty counts from the frame.[45] For this purpose, the institutional respondent definition provided in section 3.1.1 was used to identify the institution subset belonging to the denominator of this adjustment factor. The resulting adjustment factors, which were calculated using GEM within cells defined by the 10 sampling strata and region, aimed to reduce or eliminate nonresponse bias in survey estimates. Construction of the nonresponse adjustment cells was based on variables that were deemed to be predictive of response status and available for both respondents and nonrespondents.

6.1.3 Poststratification Adjustment

A set of poststratification adjustment factors (*WT4*) was calculated for the 920 institutions responding to the questionnaire using the GEM program. Specifically, nonresponse-adjusted weights for these institutions were ratio-adjusted to the counts of institutions obtained from the sampling frame. Moreover, an additional adjustment factor was calculated to ensure that weighted counts of faculty obtained from the institution survey data would coincide with those obtained from the faculty survey data. As detailed in the next section, the final analysis weights for faculty included ratio adjustments to counts of faculty obtained from the Employees by Assigned Position Survey (EAP) conducted in the Winter 2003–04 IPEDS data collection cycle. In order to achieve the needed concurrence between the weighted estimates obtained from the institution and faculty surveys, the poststratified weights of the 920 institutions were ratio adjusted to the corresponding weighted totals from the faculty data. With this last adjustment factor computed (*WT5*), the final analysis weight for each responding institution (WTB00) was calculated by:

$$WTB00 = WT1 \times WT2 \times WT3 \times WT4 \times WT5$$

6.2 Faculty Weights

The final analysis weights for faculty were constructed as the product of the final institution weights for the 980 institutions that provided faculty lists (PSU weights), inverse of selection probabilities for faculty, and a series of adjustment factors at the faculty level. The needed PSU level weights, which are different from those calculated above for the 920 institutions responding to the institution questionnaire, were calculated by calibrating the product of the institution sampling weights (adjusted for multiplicity) and the faculty frame counts to the institution counts within each of the sampling strata. Note that since a minimum weighted response rate of 85 percent was secured overall and within each of the sampling strata for institutions providing faculty lists, a nonresponse adjustment factor was not calculated for these institution. Operationally, these institutions were assigned a nonresponse adjustment factor of unity, i.e., WT3 = 1.

6.2.1 Selection Probability for Faculty

The overall faculty sampling strata were defined as the institution sampling strata crossed with the faculty strata within institutions. The sample faculty members were systematically

[45] The NSOPF:04 sample of institutions was selected using probabilities proportional to the number of faculty and instructional staff in each institution. Consequently, calculation of the analysis weights included multiplication of sampling weights (adjusted for multiplicity) by the faculty counts within each of the sampling strata. This means that a subset of institutions—particularly those that had small sampling weights, such as certainty institutions, could end up with weights that are less than one.

selected from the faculty lists at institution-specific rates that were inversely proportional to the institution's probability of selection, as dictated by the sample design. That is, the overall stratum sampling rate divided by the institution's probability of selection:

$$f_{s|i} = \frac{f_s}{\pi_{ri}}$$

where f_s represented the overall faculty sampling rate, and π_{ri} represented the institution's probability of selection. The sampling weights (*WT5*) for each of the 35,630 sample faculty members were calculated as the reciprocal of the above institution-specific faculty sampling rates.

6.2.2 Adjustment for Faculty Multiplicity

Faculty members who worked at more than one eligible institution during the 2003–04 academic year had multiple chances of being selected, since they could have been selected from any of the eligible institutions they attended. When this was the case, the resulting multiplicity was adjusted for by dividing the sampling weight of the given faculty by the number of institutions he/she worked at that were eligible for sample selection. Specifically, the faculty multiplicity weight adjustment factor was defined as WT6 = 1/M, where M is the multiplicity or number of institutions attended by sample faculty, based on the interview data.

6.2.3 Adjustment for Unknown Eligibility Status

For nonresponding faculty members whom project staff were unable to contact, the final eligibility status could not be determined. These faculty members were treated as eligible, and their weights were adjusted to compensate for the small portion of faculty members who were actually ineligible. These weight adjustment factors (*WT7*), which were calculated within cells defined by a cross-classification of institution and faculty types, represented the estimated eligibility rates among faculty members with known eligibility status. For faculty members known to be eligible the weight adjustment factor was set to one.

6.2.4 Nonresponse Adjustments

As reported earlier, faculty-level response rates were less than 85 percent, both overall and within a number of sampling strata. Subsequent to a nonresponse bias analysis, details of which are provided in appendix I, adjustment factors were calculated within cells indexed by a cross-classification of the faculty and institution strata and length of time to respond. Again, the weighting program, GEM, was used to create the needed nonresponse adjustment factor (*WT8*) for each of the 26,110 responding faculty members.

6.2.5 Poststratification/Raking Adjustment

To ensure population coverage, nonresponse adjusted weights were further adjusted to match published faculty totals. Specifically, these weights were raked along two dimensions to control totals that were constructed using the Winter 2003-04 Employees by Assigned Position Survey (EAP:03). This source was used to obtain the total number of full- and part-time faculty members by institution type. Moreover, the NSOPF:04 sampling frame was used to generate the distribution of faculty members by race/ethnicity and gender, detailed construction of which is

provided in section 2.1.4 and appendix A. The resulting two raking dimensions are summarized in tables 42 and 43. The raking adjustment factors (*WT9*) were calculated using GEM.

Table 42. Faculty counts obtained from EAP:03, by institution type and employment status

Employment status	Institution type										
	Total	1	2	3	4	5	6	7	8	9	10
Total	1,185,661	306,119	143,540	20,459	355,577	13,473	138,161	99,021	66,803	7,392	35,116
Full-time	662,407	238,168	90,183	11,213	116,491	6,514	94,688	44,198	41,709	3,881	15,362
Part-time	523,254	67,951	53,357	9,246	239,086	6,959	43,473	54,823	25,094	3,511	19,754

NOTE: Institution types are defined as follows: 1 = public doctoral, 2 = public master's, 3 = public bachelor's, 4 = public associate's, 5 = public other, 6 = private not-for-profit doctoral, 7 = private not-for-profit master's, 8 = private not-for-profit bachelor's, 9 = private not-for-profit associate's, 10 = private not-for-profit other.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), 2003 Employee by Assigned Position (EAP).

Table 43. Faculty counts obtained from EAP:03 and NSOPF:04 sampling frame, by institution type, race/ethnicity, and gender

Race/gender	Institution type										
	Total	1	2	3	4	5	6	7	8	9	10
Asian	80,704	35,993	8,804	980	11,968	567	14,203	3,583	2,430	184	1,992
Asian male	53,201	25,463	5,843	627	6,123	371	9,772	2,241	1,373	97	1,291
Asian female	27,503	10,530	2,961	353	5,845	196	4,431	1,342	1,057	87	701
Black	68,790	11,908	10,974	1,920	25,133	882	6,621	4,430	4,614	426	1,883
Black male	33,699	6,126	5,557	990	11,164	396	3,625	2,205	2,371	197	1,068
Black female	35,091	5,782	5,416	930	13,969	486	2,996	2,224	2,243	229	816
Hispanic	41,833	9,564	5,118	557	17,405	255	4,124	2,414	1,397	173	825
Hispanic male	23,177	5,654	2,761	317	9,284	133	2,451	1,243	717	103	513
Hispanic female	18,656	3,910	2,358	239	8,121	122	1,673	1,171	680	70	312
Other	994,330	248,656	118,643	17,002	301,071	11,769	113,212	88,594	58,360	6,608	30,416
Other male	575,155	158,016	66,289	9,751	154,787	6,751	74,456	49,303	33,395	3,450	18,956
Other female	419,175	90,639	52,354	7,251	146,284	5,018	38,756	39,291	24,966	3,158	11,460
Faculty	1,185,661	306,119	143,540	20,459	355,577	13,473	138,161	99,021	66,803	7,392	35,116
Male	685,238	195,259	80,451	11,686	181,358	7,652	90,305	54,994	37,857	3,848	21,828
Female	500,423	110,860	63,089	8,773	174,219	5,821	47,856	44,027	28,946	3,544	13,288

NOTE: Institution types are defined as follows: 1 = public doctoral, 2 = public master's, 3 = public bachelor's, 4 = public associate's, 5 = public other, 6 = private not-for-profit doctoral, 7 = private not-for-profit master's, 8 = private not-for-profit bachelor's, 9 = private not-for-profit associate's, 10 = private not-for-profit other.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS), 2003 Employee by Assigned Position (EAP).

Finally, eligibility definitions for NSOPF:04 include non-faculty who provide instruction but do not have teaching as their principal activity. Since there were no published counts for such faculty members that could be used for weighting, nonresponse-adjusted weights were used to develop an estimate for this small subgroup. For this purpose, all emerging 560 non-faculty members retained their nonresponse adjusted weights—totaling to 26,803—as their final weights and were not included during the above raking process.[46] That is, a value of one was assigned to the raking adjustment factor for these respondents when calculating their final analysis weights. In general, however, the final analysis weights (WTA00) were computed as the product of the institution and the faculty component weights by:

$$WTA00 = WT1 \times WT2 \times WT3 \times WT4 \times WT5 \times WT6 \times WT7 \times WT8 \times WT9$$

[46] This means that the sum of the weights for all 26,110 respondents, which includes the 560 non-faculty members, is 1,211,744 = 1,185,661 + 26,083.

6.3 Variance Estimation

The 2004 National Study of Postsecondary Faculty (NSOPF:04) sampling design was a stratified two-stage design. A stratified sample of postsecondary institutions was selected with probabilities proportional to a composite measure of size at the first stage, and a stratified systematic sample of faculty and instructional staff was selected from sample institutions providing lists at the second stage. Because of this complex sampling design, statistical analyses should be conducted using software packages that properly account for the employed survey design through use of survey weights.

Most commonly used statistical procedures assume that data are obtained from a simple random sample; that is, that the observations are independent and identically distributed. When the data have been collected using a complex sampling design, the simple random sampling assumption usually leads to underestimating the sampling variance, which would lead to artificially narrow confidence intervals and liberal hypothesis test results; that is to say, rejecting the null hypothesis when it is true more often than indicated by the nominal Type I error level. (Carlson et al. 1993).

Statistical strategies that have been developed to address this issue include: first-order Taylor series expansion of the variance equation; balanced repeated replication; and the Jackknife approach (Wolter 1985). Software packages that have been developed for analyzing complex sample survey data include SUDAAN, WesVar, Stata, and SAS. SUDAAN is a commercial product developed by RTI. Further information can be obtained from the website http://www.rti.org/sudaan. WesVar is a product of Westat, Inc., for which additional information can be obtained from the website http://www.westat.com/wesvar. Stata is a product of StataCorp LP; additional information about Stata can be found at the following website: http://www.stata.com. SAS information may be found on the SAS corporate website: http://www.sas.com. Also, the National Center for Education Statistics (NCES) has developed a software tool called the Data Analysis System (DAS) for analysis of complex survey data. Information about DAS is available from the website http://nces.ed.gov/das.

The variance estimation strategy chosen for NSOPF:04 has aimed to satisfy the following requirements and design features:

- variance reduction due to stratification at all stages of sampling;
- unequal weighting effects due to nonresponse adjustment and poststratification;
- variance inflation due to clustering;
- estimation of linear and nonlinear statistics such as quantiles; and
- variance reduction due to finite population corrections at the PSU (institution) stage of sampling and the high sampling rates in certain strata.

Commonly applied bootstrap variance estimation techniques satisfy the first four requirements. To meet the last requirement, however, the methodology developed by Kaufman was applied (Kaufman 2004). This methodology incorporates the finite population correction factors at both stages of sampling. However, for NSOPF:04, application of this method reflected the finite population correction factor at the first stage only where sampling fractions were often high. At the second stage, where the sampling fractions were generally low, the finite population correction factor was set to 1.00.

The Kaufman methodology was used to develop a vector of 64 bootstrap sample weights that are included on the analysis file, along with the full sample analysis weights. Replicate weights were set to zero for units not selected in a particular bootstrap sample while weights for other units were inflated for the bootstrap subsampling. Note that analogous to the full sample weights, these replicate weights were also poststratified to the same set of control totals for calibration.

The number of replicate weights was set at 64 based on an empirical investigation of the behavior of variance estimates as the number of replicates increased. This investigation showed that the stability of variance estimates improved with increasing numbers of replicates and became fairly stable for most estimates when between 50 and 55 replicate weights were used. Also, a similar process of generating replicate weights was used for the institution file except that all procedures relating to the second stage of sampling were omitted.

The vector of *B* replicate weights allows for computing additional estimates for the sole purpose of estimating a variance. With the 64 sets of replicate weights, the variance of any statistic, $\hat{\theta}$, can be estimated by separately calculating the statistic of interest from each replicate and then using the variability among the resulting estimates to calculate the variance of the given statistic by:

$$\text{var}(\hat{\theta}) = \frac{\sum_{b=1}^{64} (\hat{\theta}_b^* - \hat{\theta})^2}{B},$$

where $\hat{\theta}_b^*$ is the estimate based on the *b*-th replicate weights.

Once the replicate weights are provided, this estimate can be produced by survey software packages such as SUDAAN, STATA, and WesVar. Here, the analyst should specify the full study and replicate weights appropriate for the given analysis. In this case, the analyst should specify the full study and replicate weights, which are appropriate for the given analysis. Below is an example of a generic SUDAAN code for producing point estimates and their associated standard errors using replicate weights that reflect the reduction in variance due to finite population correction (*fpc*) at the institution stage of sampling. The symbols /* and */ in the code indicate the beginning and end of a comment. Note that the dataset does not need to be sorted.

```
proc descript data=/* insert filename*/ design=brr;
weight STUDYWEIGHT;
repWgt BRRWT01–BRRWT64;
var /*insert variables*/;
subpopn /* insert domain of interest if analysis domain is a subset of faculty members*/;
print nsum mean semean / style=nchs;
run;
```

Again, it should be noted that there are three sets of study (analysis) weights and their corresponding replicate weights. These weights are:

1. Institution weights

Analysis: *WTB00*
Replicate: *WTB01 – WTB64*

2. Faculty weights
   Analysis: *WTA00*
   Replicate: *WTA01 – WTA64*
3. Contextual weights
   Analysis: *WTC00*
   Replicate: *WTC01 – WTC64*

Should analysts decide to use the Taylor Series Linearization method for approximating standard errors of estimates, the design structure (level of clustering) should be specified by identifying the analysis strata and primary sampling units (PSU). Below is an example of generic SUDAAN code to produce estimates and standard errors using Taylor Series approximation. The symbols /* and */ in the code indicate the beginning and end of a comment. Note that the dataset must be sorted by analysis strata and analysis PSUs.[47]

proc descript data=/ insert filename*/ design=wr;*

nest analysis stratum analysis PSU;

weight STUDYWEIGHT;

*var /*insert variables*/;*

subpopn / insert domain of interest if analysis domain is a subset of faculty members */;*

print nsum mean semean / style=nchs;

run;

For each of the three types of analyses—institution, faculty, and contextual (merged)—specific design variables, which are generically named *analysis stratum* and *analysis PSU* in the above code, need to be identified. These variables are available on corresponding final datasets as described below.

- Institution design variables
  Analysis PSU: *IPSU*
  Analysis Stratum: *ISTRATUM*
- Faculty design variables
  Analysis PSU: *FPSU*
  Analysis Stratum: *FSTRATUM*
- Contextual design variables
  Analysis PSU: *FPSU*
  Analysis Stratum: *FSTRATUM*

[47] Please note that DAS uses Taylor linearization approach for variance estimation and calculation of DEFF and DEFT.

6.4 Design Effects and Standard Errors

Table 44 provides estimates of design effects for selected faculty data. These estimates, which consist of the ratio of variance of estimates under the employed design and simple random sampling (DEFF)[48] and the square root of this ratio (DEFT),[49] are typically used as measures for the efficiency of a sample design. The larger the design effect, the larger the variance of the estimate relative to what would have been obtained under simple random sampling where all units have the same chance of selection.

The standard errors were calculated using SUDAAN with the replicate weights that were calculated for these data, details of which are provided section 6.3. The average design effect for the listed key faculty estimates in table 44 was 1.88. Briefly, this indicates that due to differential sampling and weight adjustments, the resulting sample is 1.88 times less effective as compared to a simple random sample with 100 percent response rate. That is, the original sample size should be divided by 1.88 to obtain the effective sample size under the employed design. More detailed tables are available in appendix M.

Table 44. Design effects (DEFFs) and root design effects (DEFT) for faculty statistics: 2004

Item	Number[1]	Percent estimate	Standard error		DEFF	DEFT
			Design	SRS		
Q1: Percent with instructional duties	26,110	97.0	0.14	0.11	1.69	1.30
Q2: Percent with some credit instruction	26,110	90.2	0.32	0.18	2.94	1.71
Q3: Percent who had faculty status	26,110	92.2	0.26	0.17	2.39	1.54
Q4: Percent whose principal activity was teaching	26,110	73.8	0.34	0.27	1.58	1.26
Q4: Percent whose principal activity was research	26,110	8.8	0.18	0.18	1.06	1.03
Q6: Percent part-time is primary employment	8,360	18.0	0.34	0.24	1.99	1.41
Q8: Percent part-time preferred full-time	8,360	28.0	0.38	0.28	1.83	1.35
Q10: Percent with academic rank of professor	26,110	21.2	0.31	0.25	1.53	1.24
Q12: Percent with tenure	26,110	39.8	0.45	0.30	2.24	1.50
Q15: Percent nonunion union not available	20,880	11.7	0.27	0.20	1.91	1.38
Q19A1: Percent with other job that is full-time	26,110	11.5	0.25	0.20	1.62	1.27
Q35A1: Percent teaching a single credit class	26,110	73.2	0.37	0.27	1.87	1.37
Q37F1: Percent with no TA in first class	21,460	93.7	0.19	0.15	1.66	1.29
Q37C2: Percent meet > 3 hours for second class	15,280	27.2	0.34	0.28	1.56	1.25
Q39: Percent with website for instruction	26,110	84.9	0.30	0.24	1.55	1.24
Q62A: Percent not “very satisfied” workload	26,110	31.5	0.62	0.38	2.69	1.64
Q64: Percent retired from another position	26,110	57.6	0.33	0.31	1.20	1.10
Q68: Percent paid by the course exclude salary	6,740	34.1	0.62	0.52	1.43	1.20
Q77: Percent marital status single	26,110	34.8	0.58	0.52	1.22	1.11
Q77: Percent marital status married	26,110	69.0	0.54	0.32	2.88	1.70
Q81: Percent United States citizen	26,110	36.9	0.97	0.59	2.71	1.65

[1] This number reflects the total number of eligible respondents for each item, rounded to the nearest 10.
SOURCE: U.S. Department of Education, National Center for Education Statistics, 2004 National Study of Postsecondary Faculty (NSOPF:04).

[48] The design effect (DEFF) is the variance estimate of an estimated parameter under the survey design divided by the variance estimate of an estimated parameter for a simple random sample of the same size.
[49] The root design effect (DEFT) is the square root of the design effect (DEFF).

References

Abraham, S.Y., Steiger, D.M., Montgomery, M., Kuhr, B.D., Tourangeau, B., Montgomery, B., and Chattopadhyay, M. (2002). *1999 National Study of Postsecondary Faculty (NSOPF:99) Methodology Report* (NCES 2002–154). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Berger, A., Kirshtein, R., Zhang, Y., and Carter, K. (2002). *Developing the 2004 Faculty Survey: Themes from the Literature on Postsecondary Education.* Washington, DC: American Institutes for Research.

Carlson, B.L., Johnson, A.E., and Cohen, S.B. (1993). An Evaluation of the Use of Personal Computers for Variance Estimation with Complex Survey Data. *Journal of Official Statistics*, 9(4):795-814.

Chen, P., Penne, M.A., & Singh, A.C. (2000). Experience with the Generalized Exponential Model of Weight Calibration for the National Household Survey on Drug Abuse. American Statistical Association (ASA) *Proceedings on Survey Research Methodology Section*, 604-609.

Chromy, J.R. (1979). Sequential Sample Selection Methods. *Proceedings of the American Statistical Association, Section on Survey Research*, pp. 401–406.

Chromy, J.R. (1987). "Design Optimization with Multiple Objectives." *Proceedings of the American Statistical Association Section on Social Statistics*, pp. 194-199.

Cohen, S.B. (1997). An Evaluation of Alternative PC-Based Software Packages Developed for the Analysis of Complex Survey Data. *The American Statistician*, 57(13): 285-292.

Cominole, M., Siegel, P., Dudley, K., Roe, D., and Gilligan, T. (Forthcoming). *National Postsecondary Student Aid Study 2004 (NPSAS:04) Methodology Report (NCES 2006–180)*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

DeRouvray, C., and Couper, M.P. (2002). Designing a Strategy for Reducing "No Opinion" Responses in Web-Based Surveys. *Social Science Computer Review, 20*(1):3–9.

Deville, J.C., & Särndal, C.E. (1992). Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association* (JASA), 87, 376-382.

Education Sciences Reform Act of 2002. (2002). Public Law 107-279, Title I, Section 153. Available: http://www.ed.gov/policy/rschstat/leg/PL107-279.pdf

Folsom, R.E. and Singh, A.C. (2000). The Generalized Exponential Model for Sampling Weight Calibration for Extreme Values, Nonresponse, and Poststratification. *Proceedings of the American Statistical Association, Section on Survey Research*.

Folsom, R.E., & Witt, M.B. (1994). Testing a New Attrition Nonresponse Method for SIPP. *ASA Proceedings on Survey Research Methodology Section*, 428-433.

Heuer, R.E., Cahalan, M., Fahimi, M., Curry-Tucker, J.L., Carley-Baxter, L., Curtin, T.R., Hinsdale, M., Jewell, D.M., Kuhr, B.D., and McLean, L. (2004). *2004 National Study of Postsecondary Faculty (NSOPF:04) Field Test Methodology Report* (NCES 2004–01). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Kaufman, S. (2004). Using the Bootstrap in a Two-Stage Design when Some Second-Stage Strata have only One Unit Allocated. *Proceedings of the American Statistical Association, Section on Survey Research Methods*.

Kojaku, L. and Nuñez, A. (1998). *Descriptive Summary of 1995-96 Beginning Postsecondary Students, With Profiles of Students Entering 2- and 4-Year Institutions* (NCES 1999–030). Washington DC: U.S. Department of Education, National Center for Education Statistics.

Nettles, M.T., Perna, L.W., and Bradburn, E.M. (2000). *Salary, Promotion, and Tenure Status of Minority and Women Faculty in U.S. Colleges and Universities* (NCES 2000–173). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

U.S. Department of Education (2001). *National Postsecondary Student Aid Study: Student Financial Aid Estimates for 1999–2000* (NCES 2001–209). Washington, DC: National Center for Education Statistics.

U.S. Department of Education (2002). *Classification of Instructional Programs: 2000 Edition* (NCES 2002–165R). Washington, DC: National Center for Education Statistics.

Wolter, K. (1985). *Introduction to Variance Estimation*, Springer-Verlag: New York, NY.

Zimbler, L.J. (2001). *Background Characteristics, Work Activities, and Compensation of Faculty and Instructional Staff in Postsecondary Institutions: Fall 1998* (NCES 2001–152). Washington, DC: U.S. Department of Education, National Center for Education Statistics.