DOCUMENTATION OF NONSAMPLING ERROR ISSUES IN THE SURVEY OF INDUSTRIAL RESEARCH AND DEVELOPMENT

Douglas Bond

ABSTRACT

This paper summarizes past, current, and future work to control and measure nonsampling errors in the Survey of Industrial Research and Development, an annual survey of the U.S. Bureau of the Census. Errors are grouped into five categories: specification, coverage, response, nonresponse, and processing. Census and the survey's sponsor, the National Science Foundation, have recognized many of the sources of nonsampling errors in this survey, and have taken some appropriate steps to control and measure these errors. They are continuing this process by conducting an improvement project and other evaluation studies. The paper recommends a number of ways to further study, control, and measure nonsampling errors in the survey.

KEY WORDS: Nonsampling errors, research and development, industry surveys

TABLE OF CONTENTS

SUMMARY

The Survey of Industrial Research and Development (the "R&D Survey") provides information that is needed by industry, government, and academia to evaluate and make decisions about R&D spending by U.S. industry. This paper documents what is currently known about nonsampling errors in this survey, conducted annually by the U.S. Bureau of the Census (Census) under the sponsorship of the National Science Foundation (NSF). The NSF supported the research to prepare this paper as part of a project to maintain and improve the quality of the R&D Survey.

The paper groups nonsampling errors into five categories, using the structure of Statistical Policy Working Paper 15 of the Office of Management and Budget (OMB 1988): specification, coverage, response, nonresponse, and processing. For each category, there is a discussion of possible sources of the error and methods to control and measure the error, followed by a summary of past, current, and future efforts to control and measure the error in the R&D Survey.

The paper concludes that the NSF and Census have recognized many of the sources of nonsampling errors in the R&D Survey, and have taken some appropriate steps to control and measure these errors. Other work currently underway or planned will continue this process for improving the quality of the survey. R&D Survey documentation is generally up-to-date and clearly written. Major recommendations for action by Census and/or the NSF are:

·   Prepare more of the documentation, especially specifications, in a consistent format, in one location.

·   Find more ways to study, control, and measure nonsampling errors as part of the survey process, rather than with special studies outside the survey.

·   Continue to seek and accept input on uses and needs for R&D Survey data, and consider implementing several additional methods for receiving this input.

·   Do more to measure coverage error.

·   Continue conducting response analysis surveys and do a record check study to measure response error.

·   Compute response and nonresponse rates using standardized methods for a variety of purposes and cross-classifications.

·   Study ways to reduce nonresponse, such as with incentives or alternative follow-up procedures.

·   Investigate whether processing errors are causing data in the database to differ from data reported on forms by respondents.

DOCUMENTATION OF NONSAMPLING ERROR ISSUES IN
THE SURVEY OF INDUSTRIAL RESEARCH AND DEVELOPMENT

Douglas Bond

1. BACKGROUND

The U.S. Bureau of the Census (Census) has conducted the Survey of Industrial Research and Development (the R&D Survey) annually since 1957. Industry, government, and academia need the results of this survey to evaluate the state of science and technology in the U.S. and to develop government and corporate policy. Industry performs or funds the majority of R&D in the U.S., and the R&D Survey is the only complete source of national data on R&D spending by industry ($104 billion in 1990) and the number of scientists and engineers in industry R&D (710,000 full-time equivalent in January 1991). The survey provides detailed information on R&D that is not available anywhere else, such as R&D spending by source of funds, character of work (basic, applied, development), industry, size of company, field of research, and geographic location.

The National Science Foundation (NSF), sponsor of the R&D Survey, has historically supported efforts to maintain and improve the quality of this survey. The NSF continued this support by funding an improvement project that Census is performing during FY 93-95. There are five tasks in this project:

(1)   A study of methods that are used to follow up R&D Survey nonrespondents and measure and compensate for nonresponse, including imputation.

(2)   An evaluation of the quality of the survey by studying the sources, control, and measurement of nonsampling errors, and follow-up studies where more work is needed on nonsampling errors.

(3)   An investigation of the scope, frame construction, and sample design.

(4)   An evaluation of survey forms and letters, including interviews of respondents to find out how they interpret survey questions and concepts.

(5)   An evaluation of recommendations from tasks 1 through 4.

This paper presents work on task 2. It documents the current state of our knowledge about nonsampling errors in the R&D Survey. Sources for this paper include conversations with R&D Survey staff and staff in the Research and Methodology Branch of the Manufacturing and Construction Division (MCD), memoranda, specifications, technical reports, and the statistical literature.

2. ORGANIZATION OF THIS PAPER

Section 3 describes the scope, frame, sample design, data collection forms and data, and timing of the survey. Section 4 summarizes what has already been done to control and measure nonsampling errors and discusses how tasks 1, 3, and 4 will contribute to the study of nonsampling errors. Section 5 contains conclusions and recommendations for action and possible follow-up studies.

3. THE R&D SURVEY

This section describes the R&D Survey as it was conducted for the 1992 reference year. Some features, such as frame construction and sample design, changed dramatically for 1992. The survey design and procedures may change in the future as recommendations from this improvement project and other studies are implemented.

3.1 Scope

The target population of the R&D survey is all private, nonfarm businesses (manufacturing and nonmanufacturing, excluding government, university, and other nonprofit R&D performers) in the U.S. that perform R&D within the company in the U.S., or that fund R&D anywhere in the world, including R&D contracted out to other organizations or individuals in the U.S. 1992 R&D Survey materials defined R&D as follows: basic and applied research in the sciences and engineering, and the design and development of prototypes and processes. According to the survey materials, R&D includes activities by people trained, formally or by experience, in the physical sciences (including related engineering) and the biological sciences (including medicine but excluding psychology), if the purpose of these activities includes one or more of the following:

·   To pursue a planned search for new knowledge, whether or not the search has reference to a specific application.

·   To apply existing knowledge to problems in creating a new product or process, including work to evaluate possible uses.

·   To apply existing knowledge to problems in the improvement of a present product or process.

R&D excludes the following:

·   Capital expenditures.

·   Routine product testing.

·   Research in social sciences or psychology.

· Geological and geophysical exploration activities.

· Technical services such as quality control, plant sanitation, troubleshooting, advertising to promote or demonstrate new products or processes, and assistance in preparing speeches and publications for people not doing R&D.

Wages and salaries, materials and supplies (including computer software), depreciation and overhead, property and other taxes, maintenance and repairs, and overhead are all examples of R&D costs as defined in the survey materials.

3.2 Frame

Census constructed a new frame for the 1992 survey, and will likely build new frames annually. Before 1992, Census constructed frames about every five years; the previous frame was built for the 1987 survey. For 1992, company names were selected from Census's Standard Statistical Establishment List (SSEL) in the following way. For multi-unit companies, Census consolidated establishments to the company level, and classified each company as manufacturing or nonmanufacturing, depending on the predominant activity of the company (measured by value of payroll). Then each multi-unit company was assigned a 2-digit Standard Industrial Classification (SIC) code and a 3-digit SIC code to subdivide the 2-digit code, again using allocation of payroll as the classification criterion. Names of in-scope multi-unit companies were added to a list of in-scope single-unit companies from the SSEL (already classified by SIC code) with five or more paid employees. The resulting list of company names was the frame for the R&D Survey.

The frame was stratified by 3-digit SIC code for manufacturing industries and 2 and 3-digit SIC codes for certain nonmanufacturing industries. Companies in the nonmanufacturing sector which were not likely to conduct R&D formed a single stratum. The frame contained nearly 1.8 million companies for the 1992 survey, up from about 155,000 companies for the 1987 survey. Census matched the frame against a list of R&D companies assembled by R&D Survey staff (Havard 1993). For the 1992 survey, this list was built from the following sources:

· Companies from the previous year's survey that were known to perform or fund R&D.

· "Business Week" R&D Scoreboard of the top 100 R&D performers, based on annual reports and Form 10-K from the Securities and Exchange Commission (SEC). Publicly-held companies with "material" (as interpreted by their accountants) R&D expenditures are required by law to report these expenditures to the SEC on Form 10-K. Most of the large R&D companies therefore report on this form. The definition of "R&D" is a little different from what is used in the R&D Survey, and it includes R&D performed in foreign countries.

· "Inside R&D" weekly, which listed the largest companies and emerging companies.

- CompuStat database, a list of about 2,500 known R&D companies.

- Moody's, which had information on about 10,000 companies (about 1,600 had R&D expenditures). Census has stored this information on CD-ROM.

- A file of over 1,100 companies that reported over $1,000 of R&D spending in 1987 on Economic Census Form ES-9200. This form was mailed to "administrative offices" (headquarters, separate R&D facilities, and so forth) of multi-establishment companies.

- Electric Power Research Institute annual report.

- "R&D Ratios & Budgets," published by Schonfeld & Associates, who gathered information with their own questionnaire. Over 2,900 companies were listed with R&D spending.

- "Defense News," which listed top R&D companies.

- Information from the Department of Defense on R&D contracts, including the top 500 R&D contractors.

- Department of Commerce R&D consortia register.

- Industrial Research Institute, which had about 275 members.

- Corporate Technology Directory.

- Directory of American Research and Technology.

- "Government Executive," which listed the top 100 R&D contracts.

If a match occurred, the company was made a "predetermined certainty," i.e., it was in the sample, with a sampling weight of 1. Companies were also designated predetermined certainties if they met one of the following criteria:

- The unit had 1,000 or more employees (based on information in the SSEL).

- The unit received a Form RD-1 or RD-1-S (see explanation in section 3.4) in the previous year's survey.

There were 9,764 predetermined certainties for the 1992 survey. In addition, there were 1,394 "analytical certainties" - companies with R&D expenditures that exceeded cutoff levels set by Census for each stratum. For more details on frame construction, see Greenberg (1993b). Champion (1993) discussed the evolution of R&D frame size, stratification, and sample size and allocation from 1981 to 1992.

3.3 Sample Design

Census computed stratum-level sample sizes to achieve desired levels of precision (2% or 5% relative standard error) for estimating total R&D costs for 1992. Strata with significant R&D spending and/or separate publication cells were considered more important and were targeted for higher precision (2% relative standard error) (Champion 1993). As discussed above, some companies were already in the sample with certainty. Census chose additional units (non-certainties) to reach the desired sample sizes, using probability proportional to size (PPS) Poisson sampling. The measure of size for PPS sampling was previous-year total R&D expenditures. If there were no reported R&D expenditures for the previous year, Census imputed the measure of size by multiplying total employment (which was on the SSEL for every company) by "R-factors" - ratios of R&D expenditures to total employment for companies that had R&D expenditure data for the 1991 survey (Greenberg 1993a). R-factors were not computed for individual strata; instead, they were computed for groups of strata, called R&D "recodes."

If the measure of size was very small for a unit, the probability of selection would be very small, and the weight would be very large. Then the expanded totals for survey items, though unbiased, would be inflated, as would the variances. Census controlled this problem by forcing probabilities of selection to be no less than designated minimums (thus putting a cap on weights). This raised sample sizes somewhat, but should have reduced variances. Most non-certainty units in the frame for 1992 were assigned the minimum probability of selection. Kusch (1993) documented the theory and implications of minimum probability PPS sampling in the R&D Survey.

3.4 Forms and Data Items

Census collects data with two different forms each year, Form RD-1A and Form RD-1 or RD-1-S. Forms RD-1 and RD-1-S, the "long" and "short" forms, respectively, are used in alternating years to reduce respondent burden. Census mailed Form RD-1-S for the 1992 survey, to companies that reported $1 million or more in R&D expenditures in the previous year's survey, and to additional companies so that at least 95% of total R&D expenditures were covered in each stratum; a total of 1,599 RD-1-S forms were mailed out. All companies that were mailed an RD-1-S were certainties. Census mailed Form RD-1A to the remaining companies in the sample. The burden was reduced for these companies by asking fewer questions than on Form RD-1-S.

In summary, the 1992 sample was made up of the following components:

        Predetermined certainties   9,764
        Analytical certainties   1,394

Total certainties	11,158	(1,599 RD-1-S, 9,559 RD-1A)
Non-certainties	12,218	(All RD-1A)
Total sample	23,376	(1,599 RD-1-S, 21,777 RD-1A)

Table 1 shows the items that were collected by the 1992 R&D Survey and that will be collected by the 1993 survey. Only four items were mandatory as of the 1992 survey.

3.5 Timing of the Survey

The Data Preparation Division (DPD) prints and mails out questionnaires early in the year, usually in March. Mailout was delayed until June 1993 for the 1992 survey. The DPD checked in, manually reviewed, and keypunched questionnaires, and the data went through a computer edit. Nonrespondents were mailed follow-up packages about 30, 60, 90, and 120 days after mailout. R&D Survey staff phoned the 300 largest R&D companies that did not respond after the 60-day follow-up, to collect key data items and encourage reporters to mail in their questionnaires.

Census normally prepares preliminary estimates for the NSF by July, based largely on reports from the top 300 R&D companies. These estimates were delayed in 1993 until early November because of the late start. Most final 1992 estimates were delivered to the NSF by March, 1994. The NSF then reviews and publishes the data. Recently, the timing was as follows: the NSF published selected tables for the 1990 survey in June 1992, and a final report on the 1989 survey (National Science Foundation 1992) at about the same time.

4. NONSAMPLING ERRORS

The Office of Management and Budget (OMB) (1988) grouped nonsampling errors in establishment surveys into five categories:

·   Specification.

·   Coverage.

·   Response.

·   Nonresponse.

·   Processing.

Table 1: Data collected on the 1992-93 R&D questionnaires

Item	RD-1	RD-1-S	RD-1A	Mandatory
Performed in the U.S.:				
Sales	X	X	X	X
Employment	X	X	X	X
R&D Scientists and engineers	X	X	X	
- By source (Federal, company & other)	X			
Total R&D costs within company	X	X	X	X
- By source (Federal, company & other)	X	X	X	X
- By character (basic, applied, development)	X	X	X	
- By type of expense (wages, supplies, etc.)	X			
- By state	X			
-By source (Federal, total)	X			
- By category (energy, pollution abatement)	X	X		
- By subcategory (fossil, geothermal, etc.)	X	X[1]		
- By source (Federal, total)	X	X		
- Projected	X	X		
Projected R&D costs within company, excl. Federal	X	X	X	
Federal R&D costs within company				
- By agency (DOD, NASA, etc.)	X			
Basic research costs within company				
- By field (chemistry, physics, etc.)	X			
Applied R&D costs within company				
- By product group (food, textiles, etc.)	X			
- By source (Federal, total)	X			
Company R&D funds, performed outside company	X	[2]	X	
Performed outside the U.S.:				
Company R&D funds	X	X	X	
- By country	X			

[1]   Not as many detailed subcategories on RD-1-S.

[2]   Not on 1992 RD-1-S, but will be on 1994 RD-1-S.

In this section, I use this grouping as a framework for discussing what has already been done, or is planned, to control and measure nonsampling errors in the R&D Survey, and to talk about other tasks in the NSF-sponsored improvement project. I reference the OMB report so many times that I refer to it as simply "the OMB." For each error category, the OMB provided a definition, listed sources of the error, and described ways to control and measure the error. This section has the same structure.

4.1 Specification Error

Specification error occurs when the specification of data to be collected by the survey is inadequate or inconsistent with respect to the objectives of the survey. For example, if the definition of "R&D" in the instructions was different from the definition that the NSF intended, or if it was so poorly written that respondents misunderstood it, a specification error would have occurred.

4.1.1 Sources of Specification Error

The OMB listed three possible sources of specification error in a survey. They follow, with examples of how these sources could exist in the R&D Survey:

(1)   <u>Inadequate specification of uses and needs</u> The NSF specifies what data should be collected by Census based on the NSF's needs and the needs of other data users. If the specified data were not the same as data needed by the NSF and others, this would be a specification error. This could result from a poor understanding of users' needs by the NSF or failure to keep current with changing needs, for example. It could also be caused by the NSF writing unclear specifications for Census.

(2)   <u>Inadequate specification of concepts</u> If the NSF wrote a definition of "R&D" in the specifications for Census that did not conform to the expectations of data users, this would be a specification error. The concept of "R&D" in the survey would not be the same as the concept understood by users.

(3)   <u>Inadequately specified data elements</u> This source refers to items on the questionnaire that are not consistent with the purposes of the survey. This problem could result from: ambiguous definitions in survey materials; poorly worded questions and instructions; and the inability to collect data that directly measure a concept, such as "applied research."

4.1.2 Control of Specification Error

The OMB suggested the following techniques for controlling specification error in an establishment survey:

(1)   <u>Identify uses and needs</u> One way to identify uses and needs, according to the OMB, is a requirements review; in the case of the R&D Survey, Census or the NSF would contact R&D data users to see if the data currently being collected met their needs, and if they wanted new data. Appropriate changes would be made to the questionnaire and instructions if needs were not being met.

(2)   <u>Evaluate questionnaire and instructions</u> The OMB cited several methods:

    ·   Consult with respondents (actual or potential); ask them about definitions, instructions, availability of data, and so forth, and about proposed changes to the questionnaire.

    ·   Pretest the questionnaire.

    ·   Perform cognitive studies of respondents, which can identify troublesome questions and instructions.

    ·   Have a panel of subject matter experts review the survey materials to see how well the specifications conform to the needs of users.

Most of the above control techniques have been applied to the R&D Survey. The following is a list of actions and studies that have contributed to the control of specification error (and other nonsampling errors) in the R&D Survey:

(1)   The NSF conducted a workshop that was similar to a requirements review, although it was not specific to the R&D Survey. Twenty experts on industrial science and technology (S&T) were brought together in June 1989, to "identify and discuss emerging industrial S&T policy issues and data needs of the nineties" (Applied Management Sciences 1989). These people identified many needs for additional R&D data, including: R&D of cooperative arrangements of companies (consortia, joint ventures, and so forth); the impact of mergers, divestitures, and so forth on R&D spending and employment; tracking emerging technologies; and R&D performed in the U.S. by foreign companies. Census made changes to the questionnaire based on some of the needs raised at this workshop.

(2)   "An NSF senior-level committee of outside experts met in 1990-91 to provide advice and guidance concerning current and emerging science and technology policy issues and the kinds of data collection methods and analyses that would help illuminate them," according to the OMB clearance package for the 1993-95 R&D Surveys.

(3)   The NSF established an outside advisory committee to give expert advice on data collection and other activities, in October 1992.

(4)   NSF staff regularly consult with industry R&D directors and government leaders. Several names are listed in the OMB clearance package for the 1993-95 R&D Surveys.

(5)   The NSF identifies uses and needs for R&D Survey data in the course of responding to numerous requests for data.

(6)   Adams and Champion (1992) identified changes they wanted to see in the R&D Survey: increases in coverage in nonmanufacturing industries, mandatory reporting for all items, and more use of the phone for follow-ups. They also recommended collecting more data, as follows:

·   All R&D by field of science, not just basic R&D.

·   More detail for the engineering field.

·   Add computer hardware to the list of product groups.

·   Information on patents, joint ventures, and company-university connections.

·   R&D lost and gained through mergers and divestitures.

·   Data to estimate R&D capital stock, such as the rate of depreciation.

The recommendation to increase coverage in nonmanufacturing industries has already been implemented, through improvements in the frame and sample, as noted elsewhere in this paper. The other recommendations require further study; item (7) below is an example of the kind of study that is needed.

(7)   Census staff interviewed 84 companies (49 of them were top 200 R&D performers) in the R&D Survey sample face-to-face or by phone, from May through November 1992, to evaluate the questionnaire and instructions (U.S. Bureau of the Census 1993a). Most industry and company size categories were represented. The objectives were to see how well respondents understood proposed survey changes, how clear definitions were, the ability to report data, and to get recommendations for survey improvements. These interviews led to changes in the format and wording of the questionnaire, additional detail for some questions, and clearer definitions and wording in survey materials. The changes are listed in the 1993-95 OMB clearance package (section A.13.b.) for the R&D Survey. The interviews also suggested that Census should not increase the scope of the survey to include social science research and should not add a separate question about independent R&D, and recommended further research before asking for more detailed information on contracted-out R&D (R&D performed outside the company with company R&D funds).

(8)   Staff from Census's Center for Survey Methods Research are conducting a cognitive study of respondents and potential respondents during 1993-94, under task 4 of the R&D Survey improvement project. They are evaluating cover letters, follow-up materials, the questionnaire, and instructions, to identify problems in content and format and make suggestions for improvements.

4.1.3 Measurement of Specification Error

The OMB mentioned the following measures of specification error. These are measures of other types of nonsampling errors, too; it is often difficult or impossible to separate the effects of different types of errors.

(1)   <u>Direct measures</u> Record check studies directly measure nonsampling errors, including specification error. In these studies, data obtained by other means on individual sample units are compared with survey data.

(2)   <u>Indirect measures</u> A validation study can indirectly measure specification error. In this type of study, each stage of data collection is examined to see where errors are caused by inadequate specifications. It usually involves visits to respondents to see how data are measured and whether intended concepts are being measured.

A cheaper indirect measure offered by the OMB is a comparison of survey estimates with estimates derived from independent sources. If there were differences, individual records that caused these differences would be sought, with the goal of finding specification problems in the survey materials.

Several record check studies have been conducted for the R&D Survey; they are described in section 4.3.3. As far as I know, no validation study has been conducted; this is not surprising, because validation studies are expensive. However, survey estimates were compared with independent estimates for the early 1980s (NSF 1985). The following is a list of the sources of the estimates, and results (if any) of comparisons with R&D Survey estimates of totals. Table 2 shows comparisons of estimated percent year-to-year changes.

·   "Inside R&D" newsletter aggregated company-funded R&D for the top 50 (in 1981-82) or 100 (in 1983) R&D spenders, based on SEC Form 10-K data.

·   R&D Scoreboard, assembled by "Business Week," was based on 10-K data for about 800 companies that had $1 million or more R&D spending, or that spent 1% or more of sales on R&D. The total for these companies was about 10% below the comparable R&D Survey level in the early 1980's.

·   McGraw-Hill Publications surveyed companies to obtain R&D expenditures, obtaining responses from "more than 300." They did not provide respondents with definitions or instructions for survey items, and their response rate was only 25%, even after a

nonresponse follow-up. Their computed totals for 1981-83 were 94% to 101% of
comparable R&D Survey totals.

·   "Research and Development" magazine used data from the 10-K, Federal budget data,
        company annual reports, and an "informal" phone survey of about 45 companies.

·   Battelle Memorial Institute projected R&D company funds for each year based on the
        preceding year's R&D Survey total, 10-K data, and "other" sources (no survey),
        including data on economic conditions. Their totals were 6% to 10% below the
        comparable R&D Survey totals during 1981-83.

Table 2: Percent change in company R&D funds from previous year

Year	R&D Survey	Inside R&D	R&D Scoreboard	McGraw-Hill	R&D Magazine	Battelle
1981	+16	+13	+15	+17	+14	+13
1982	+13	+13	+12	+17	+12	+11
1983	+9[1]	+10	+10	+8	+10	+8

[1] Later revised to +11%.

The R&D Survey results were similar to the independent estimates of annual changes;
however, there are several reasons why these comparisons, and the comparisons of totals, do
not have much value for improving the quality of the R&D Survey:

·   The level of the R&D Survey could be grossly in error, but year-to-year changes could
        be perfectly reasonable. For example, if respondents erroneously excluded Federally
        funded R&D spending every year because of poor instructions, levels would be
        reduced but annual percentage changes might be affected very little. Then a
        comparison such as in Table 2 would not shed any light on the problem.

·   None of the independent sources had the coverage or statistical rigor of the R&D
        Survey. Therefore, their estimates were probably less reliable.

·   Estimates were not exactly comparable, because of differences in definitions and
        inclusion or exclusion of R&D performed in foreign countries.

·   Suppose estimates from different sources disagreed. How would you know which
        were correct? The independent sources in this case were subject to nonsampling
        errors just like the R&D Survey was. How would you know why the estimates
        disagreed? Comparisons of aggregated levels or changes do not identify which parts
        of the survey process were right or wrong.

4.2 Coverage Error

Groves (1989) provided a more thorough definition of coverage error than the OMB. He defined coverage error as the situation where there is not a one-to-one correspondence between units in the frame and units in the target population. He listed four cases where there are coverage errors:

·   Some units in the target population are not in the frame (undercoverage). R&D Survey example: a company with R&D spending is left out of the frame.

·   Some units in the frame are not in the target population (overcoverage). R&D Survey example: a nonprofit company is included in the frame.

·   Multiple units in the frame correspond to one unit in the target population (duplication, another type of overcoverage). R&D Survey example: a company is listed twice in the frame, under two different names.

·   Multiple units in the target population correspond to one unit in the frame. R&D Survey example: a company has divided into two companies, but there is just one listing in the frame because no one at Census knew about the division.

There is little doubt that all of these situations exist in the R&D Survey.

4.2.1 Sources of Coverage Error

The OMB divided the sources of coverage error for establishment surveys into two categories:

(1)   <u>Defective frame</u> A frame could be defective because of poor definitions of the sampling units, improper procedures for creating and maintaining the frame, failure to keep the frame up to date, miscoding (or no coding) of sample units (such as SIC, measure of size, or company structure), duplication, inclusion of out-of-scope units, outdated mailing addresses, and so forth. The frame might be good for national estimates, but if SIC coding were defective, estimates by SIC would be erroneous.

(2)   <u>Defective processing of the sample</u> If selected units were erroneously dropped from the sample, nonselected units were added, or mistakes were made in selecting the sample, coverage errors would have occurred. These errors could have been caused by errors in the computer program that drew the sample, for instance.

4.2.2 Control of Coverage Error

The OMB reported many ways to control coverage error. They can be grouped into frame

maintenance and improvement, and control of processing.

(1)   <u>Frame maintenance and improvement</u> The following methods were suggested by the OMB under this heading:

· Integrate multiple lists for frame development.

· Conduct special frame improvement surveys.

· Use two-phase sampling. In the R&D survey, this could be done by drawing an initial sample that included unclassified units with no measure of size. The purpose would be to obtain coding information and a measure of size, as a basis for drawing a second-phase sample.

· Update the frame for births and deaths.

· Update the frame for structural changes, such as mergers, divestitures, and so forth.

· Include units in the frame that are <u>possibly</u> in-scope, to increase the chances of capturing all units that are in-scope.

· Use independent control counts. For example, compare an independent estimate of company R&D spending with the frame total.

· Do internal consistency checks on frame records.

· Identify and remove duplicate records from the frame.

· Include seasonally out-of-scope units in the frame.

(2)   <u>Control of processing</u> The following are the OMB's suggested control methods:

· Validate the sample, i.e., compare sample expansions with frame totals.

· Have well-documented, correct, clear procedures for frame maintenance and sample selection.

· Monitor processing of the frame and sample.

Several of the above control methods have been implemented in the R&D Survey, and additional steps are planned for controlling coverage error. Details are given below. Some control methods have never been implemented, such as special R&D Survey frame improvement surveys and two-phase sampling. Much of the control of coverage error comes

from the maintenance of the SSEL, the source of the R&D Survey frame. Documentation of procedures for frame maintenance and sample selection is generally complete, well written, and up-to-date.

(1)   As mentioned previously, the frame was constructed from the SSEL. Census continuously maintains the SSEL, using information from economic and agricultural censuses, periodic surveys, the annual Company Organization Survey, and administrative data from the Internal Revenue Service, the Social Security Administration, and the Bureau of Labor Statistics. Greenberg (1993c) documented the structure of the SSEL and the methods that ensure it is updated for births, deaths, changes in ownership, new business arrangements, and so forth. Census recently took several steps to improve the quality of the SSEL and reduce processing costs: automation of SIC coding; automation of edit resolutions; and transfer of the SSEL to a microcomputer to allow on-line access and updating (Hanczaryk and Trager 1990).

(2)   A question on Forms RD-1A and RD-1-S asks, "Was this company owned or controlled by another company on December 31, 1992?" The name and address of the new owner are requested. This obtains information on acquisitions by other companies.

(3)   Census updates information on the frame by comparing the frame with lists of R&D performers from other sources; I described this procedure in section 3.2.

(4)   When there is a change in a unit's name, address, phone number, and so forth, which is detected on a form returned by the postal service (a postmaster return, or "PMR"), DPD staff can immediately update the R&D Survey database using an on-line system (U.S. Bureau of the Census 1993b, Ch. 15, Subch. C, Doc. 1). This system provides menus and formatted screens to help the user. For changes made by respondents on their forms, DPD data entry staff key the changes as part of the data entry system (see section 4.5 for details of data entry procedures). R&D Survey staff further investigate name and address changes when they think organizational changes have occurred.

(5)   Census improved coverage for 1992 because the frame was designed to cover many more small R&D performers and nonmanufacturing industries than the previous frame, created for the 1987 survey (Champion 1993). The new design also gave better coverage at the stratum level for certain industries. The 1992 frame, over 10 times larger than the 1987 frame, included many more units that are marginally in-scope than before; this is one of the control methods suggested by the OMB for establishment surveys.

(6)   Coverage should be improved because the frame will probably be constructed annually from now on, rather than every five or six years. It will not "age" as much as in the past.

(7)   Under task 3 of the R&D Survey improvement project, staff from MCD's Research and Methodology Branch are doing work to improve the frame. Areas of study include stratification and the design of an efficient sampling plan.

4.2.3 Measurement of Coverage Error

The OMB listed several techniques to derive measures of coverage error:

(1)   Direct measures These include:

·   Post-enumeration survey. After the survey, use more extensive procedures to identify the sources and effects of nonsampling errors, including coverage errors.

·   Match a sample of units known to be in the population against the frame.

·   Match the frame against other lists.

·   Compare survey estimates with independent estimates, either for the whole population or subpopulations.

·   Study components of the frame, such as classifications of units.

(2)   Indirect measures There are several rates which are indirect measures of coverage error. Most of them can be computed for the frame and/or the sample: birth rate; out-of-business rate; unclassified rate; misclassified rate; duplication rate; and attrition rate. Tracking these rates over time can show trends in coverage error.

Two of the direct measurement techniques have been used for the R&D Survey, and are cited elsewhere in this paper: matching the frame against other lists, and comparing survey estimates with independent estimates. As far as I know, none of the other measurement techniques have been applied to the R&D Survey.

4.3 Response Error

Response error is defined by the OMB as the difference between the value collected and the actual value. There may be a relatively constant response error over all units in the sample, or the error may vary from unit to unit; the magnitude of the error may be related to the actual value of the unit. The error may even vary over repeated measurements on the same unit.

4.3.1 Sources of Response Error

The OMB classified sources of response error into three groups: task, respondent, and

interviewer. Goldenberg et al. (1993) mentioned another: information system. Virtually all of the R&D data are collected by a self-administered mail questionnaire; R&D Survey staff collect a small amount of data by phone only from nonrespondents among the 300 largest companies, to meet the deadline for preliminary estimates. Therefore, interviewers are not an important source of nonsampling errors in this survey. The other sources are described in more detail below.

(1)   <u>Task</u> Goldenberg et al. (1993) defined the task in establishment surveys as "obtain information from a sample of responding establishments about some aspect of the organization." In the R&D Survey, the respondent's role in the task is to locate and provide the information as prescribed by the questionnaire and instructions, usually without help from an interviewer. Task errors could occur if a questionnaire or instructions appeared to be long and complicated, or unprofessional, for example; then respondents would not want to complete the questionnaire, or to spend much time on it. Ambiguous questions and poor wording and layout of questions could cause task errors. Task errors could also occur if a company refused to give data thought to be sensitive or confidential, or if they felt burdened by the survey. In the latter case, the respondent might only provide estimates rather than look in company records for data.

(2)   <u>Respondent</u> A respondent might fail to report correct values because of: memory problems - for example, forgetting that certain data were available, or, in the case of the R&D Survey, forgetting that there were R&D workers during the reference period; poor survey timing - such as a survey that was conducted too long after the end of the reference period, so that recall was more difficult; failure to consult the information system or to accurately extract data from it; dishonesty; the particular respondent being inappropriate; and too much detail on the questionnaire, so that the respondent relied on estimating data rather than consulting records.

(3)   <u>Information system</u> There could be errors if the information system could not provide the data specified by the survey, either because the data were not available or because the system had not been updated with the latest information. The information system might be able to provide aggregated or disaggregated data requested on the questionnaire only with great difficulty, or not at all. If the information system was not automated, errors could arise when the respondent had to do special computations or retrieve data to satisfy needs of the questionnaire that were not part of the respondent's data needs.

4.3.2 Control of Response Error

Some of the following methods given by the OMB for controlling response error are mentioned elsewhere in this paper for controlling other types of nonsampling errors:

(1)   Cognitive studies.

(2)   Questionnaire pretesting.

(3)   Improvement of questionnaire wording and format.

(4)   Cues and examples on the questionnaire and instructions to aid the respondent's
       memory.

(5)   Consultations with respondents, to see what data and information system are available.

(6)   Post-survey evaluation.

(7)   Record check studies.

(8)   Computer editing.

(9)   Accepting data in a form that is produced by the respondent's information system, to
       avoid the errors that might occur if the respondent tried to adapt the data to the
       questionnaire.

Goldenberg et al. (1993) used a response analysis survey (described below in (2)) to control
response error.

Many of these control methods have already been applied to the R&D Survey, or they are
under other tasks in the R&D improvement project. Some are mentioned in other parts of
this paper, including the following: cognitive studies, improvement of questionnaire and
wording, and consultations with respondents. Section 4.3.3, on measurement of response
error, discusses record check studies that have been done. The following is a description of
methods to control response error that have been or will be used in the R&D Survey but have
not been described anywhere else in this paper:

(1)   In an ongoing effort to control response error, R&D Survey staff manually review all
       questionnaires after they are keyed by the DPD, and make necessary corrections.
       Then the keyed data (including data corrected by R&D Survey staff) are computer
       edited. There is a "pre-edit" and a "main edit." The pre-edit checks for such things as
       current-to-previous ratios, totals not equal to sums of parts, and check boxes not
       completed. The main edit includes checks of current-to-previous ratios and inter-item
       ratios (such as R&D costs/total sales), consistency of check boxes with reported data,
       and R&D costs over $1 million (form RD-1A only). The main edit also performs
       imputations. Census documented the error checks performed by the edit (U.S. Bureau
       of the Census 1993b, Ch. 16, Subch. E, Doc. 1). R&D Survey staff review a listing of
       errors and correct errors using an interactive correction system.

(2)   Census conducted a response analysis survey in the early 1980's for the R&D Survey, and another on the 1987 R&D Survey. For the latter study, R&D Survey staff surveyed 76 respondents, mostly from large R&D companies, in face-to-face interviews with a 14-page questionnaire (U.S. Bureau of the Census 1989). The purposes were to get information on company organization, find out about difficulties in reporting due to company information systems, identify unclear instructions and definitions, check for understanding of survey concepts, and get general comments about how companies reported on the survey.

This study led to improved definitions and wording in the questionnaire and instructions, and changes in the questionnaire format. Census obtained a better understanding of how companies organize their R&D facilities and collect and report financial data for them. Also, the study gave a measure of the time required to complete the form: 1 to 400 hours for the long form (RD-1) (average of 63 hours), and 1 to 225 hours for the short form (RD-1-S) (average of 30 hours). These times were probably upper bounds, because most of the companies in the study were large.

(3)   The study described in (7) of section 4.1.2 (U.S. Bureau of the Census 1993a) included some of the features of a response analysis survey.

4.3.3 Measurement of Response Error

The OMB reported the following methods for measuring response error in establishment surveys:

(1)   Direct methods Examples are record check studies and reinterview studies.

(2)   Indirect methods Methods include computation of edit failure rates, debriefing respondents, questionnaire pretesting, and cognitive studies.

Goldenberg et al. (1993) computed other indirect measures of response error, in their response analysis surveys at the Bureau of Labor Statistics: the percent of respondents who followed instructions correctly, did not interpret a concept correctly, and so forth.

All of the measurement techniques listed above, except reinterview studies and the measurements of Goldenberg et al., have been used in the R&D survey. They are described elsewhere in this paper or below:

(1)   The response analysis survey for the 1987 survey, discussed in section 4.3.2 (U.S. Bureau of the Census 1989), included a record check study. R&D Survey staff compared R&D expenditures reported on the survey with those in the "Business Week" R&D Scoreboard, which come from data reported on SEC Form 10-K. Figures agreed for about half the companies in the response analysis survey. Reasons for differences included:

· Data were reported for the calendar year on the survey but for the fiscal year for some 10-K reports.

· Some 10-K reporting for expenditures was out of the scope of the R&D survey.

· Some 10-K reports omitted benefits, overhead costs, and so forth that were reported in the R&D Survey.

· Federal funds, foreign expenditures, and independent R&D were handled differently on the survey and 10-K forms.

As a result of this record check study (and the 1987 response analysis survey and the evaluation study described in (7) of section 4.1.2), the NSF and Census changed the format of the questionnaires to a matrix; this should cut down on errors in reporting the distribution of expenditures into domestic, foreign, contracted-out, Federal, and so forth.

(2)   In an earlier study, referenced in section 4.1.3 (NSF 1985), R&D survey results were compared with SEC 10-K figures (from "Business Week") for the top 200 R&D performers, after adjusting as much as possible for differences in the scope of the two sources (10-K figures include contracted-out and foreign R&D). On a company basis, the percent differences (sign ignored) between the sources were as follows:

Number of companies	Percent difference
62	0 - 3
32	3 - 10
44	10 - 25
26	25+
36	No 10-K
TOTAL: 200	

The aggregated difference rose from 3% in 1976 and 1981 to 8% in 1983 (apparently the aggregated 10-K figures were always lower). The major reason for differences, according to the NSF, appeared to be the inclusion of engineering and routine technical services on the 10-K form. (This does not seem to be a good explanation, because aggregated 10-K figures were apparently lower than the R&D Survey figures.)

(3)   Long and Ravenscraft (1993) did a record check study in an evaluation of the effect of corporate restructuring on R&D. They compared annual sales figures derived from

quarterly financial reports (QFRs, collected by Census) with sales figures reported on the R&D Survey from 1977-88. The R&D Survey figures were in a longitudinal file that was constructed from data sets of varying completeness and quality. For example, data received late were not always in these data sets. Also, the original data sets for some years had not been saved; they had to be reconstructed by rekeying questionnaires (with little of the normal quality control) or by taking reported prior year data from current year data sets.

Of the 7,136 R&D Survey figures checked, 75% fell within $\pm$10% of the QFR-derived figures. The authors gave three major reasons for the larger differences:

·   Some companies reported by divisions on the R&D Survey. In some cases, they only reported for the divisions that performed R&D, so total sales for the entire company were underreported.

·   Some companies may have included foreign sales on the R&D Survey in their domestic sales figure.

·   Mergers and divestitures could have affected reporting on the QFRs and R&D Survey differently. For example, if a company made an acquisition in the fourth quarter, its aggregated annual sales on a QFR would only include the acquisition's fourth quarter sales. But its annual sales for the R&D Survey would probably include the annual figure for the acquisition.

Long and Ravenscraft also found "a number of cases" where there were data on the R&D Survey questionnaires that had never gotten into their longitudinal file; in the file, the data were missing or imputed. This could have occurred from using some incomplete and reconstructed data sets to construct their file, and not from any errors in survey processing. Based on all their findings, they cautioned researchers about using R&D data, and suggested verifying data in their file by referring to the paper questionnaires, if an analysis was to be done on company-level data.

(4)   Siegel and Andrews (1989) examined R&D data that are maintained by Census's Center for Economic Studies (CES) in a company-level R&D database. (Adams and Peck (1993) wrote a guide to this database.) Siegel and Andrews did the following edit checks on data for 1972-85:

·   Longitudinal ratio check: inter-item ratios of selected items, such as total employment/total company sales, and the growth rates of these ratios.

·   Longitudinal current-to-previous check: current-to-previous ratios of all survey items.

·   Presence check: mandatory items present or missing.

·   Balance check: check that detailed items summed to total items.

These types of edits are already done in the operational edit, except for the growth rate of inter-item ratios. Siegel and Andrews computed the number of outliers (for example, the number of current-to-previous ratios that fell outside the interval 0.5 - 2.0) and reported counts such as the number of mandatory items missing and the number of reports by SIC and item. They determined that some outliers were errors, due to keypunching, rounding, and incorrect reporting, but that some outliers were legitimate, due to mergers, acquisitions, divestitures, and leveraged buyouts. True errors are being corrected in the CES database as time and resources are available. The primary purpose of the analysis by Siegel and Andrews was to provide a guide for economists who planned to use R&D Survey data at the company or industry level. They mentioned two other studies that found "accurate and comprehensive information provided by respondents to the NSF R&D survey," and they concluded that "the quality of these data are excellent."

4.4 Nonresponse Error

The OMB defined nonresponse error as the error that comes from the failure to collect complete data on all units in the sample.

4.4.1 Sources of Nonresponse Error

The OMB listed three major categories of nonresponse error sources in establishment surveys:

(1)   Noncontacts Examples that could occur in the R&D Survey are a questionnaire getting lost during processing in the DPD or in the mail, an incorrect address for mailing, an incorrect phone number for nonresponse follow-up, or a seasonal closing of a company.

(2)   Unit nonresponse This could occur when a company is contacted, but does not respond. There could be unit nonresponse when the company did not have time or enough staff, it had too big a survey reporting burden (from this survey or from multiple surveys), it did not have the data at the appropriate level of detail, or when it was company policy to refuse. Unit nonresponse could also result when a company received a questionnaire that was addressed to the wrong division or contact, or because of frame problems; for example, when a company was selected twice, or a selected company was out-of-scope (such as a pension plan), it would be unlikely to respond.

(3)   Item nonresponse Item nonresponse would occur when a responding unit did not answer all the items on the questionnaire. This could be due to lack of staff, an information system that could not provide all the data, company policy to answer only mandatory items, poor wording of some of the questions, a lengthy or complicated

questionnaire, or ambiguous or sensitive questions.

4.4.2 Control of Nonresponse Error

The OMB's suggestions for controlling nonresponse error include the following:

(1)   Noncontact Maintain correct mailing addresses (and phone numbers) and adopt process and quality control procedures for the mailout and check-in operations.

(2)   Unit nonresponse The OMB stressed the importance of getting responses from the large units in establishment surveys, since a small number of large units often account for most of the economic activity in the target population. Techniques for reducing nonresponse for any sample unit included:

·   Do a presurvey notification, with letters, phone calls, personal visits, and so forth.

·   Offer incentives to respond, such as free copies of the survey release.

·   Follow up nonrespondents with reminder cards, additional mailings of the questionnaire, and phone calls.

·   Offer special reporting arrangements, including different reporting formats such as computer printouts or electronic media (for example, diskettes).

·   Keep the questionnaire as simple as possible.

·   Reduce respondent burden; for example, rotate the sample and avoid surveying one respondent in too many surveys.

Timing of survey mailouts can affect unit nonresponse; Hidiroglou et al. (1993) obtained the best response rates for Canadian manufacturers by contacting them three months after the end of their fiscal years, when data were most available.

(3)   Item nonresponse The OMB listed several ways to reduce item nonresponse; these techniques have been discussed before in this paper:

·   Ensure that the requested data conform to what companies can provide from their information systems, such as by pretesting the questionnaire and by conducting response analysis surveys.

·   Follow up nonrespondents.

·   Keep the questionnaire simple, clearly worded, and organized.

Census and the NSF have already used some of these methods to control nonresponse error in the R&D Survey. Some are described elsewhere in this paper: procedures for maintaining and updating names and addresses in the frame are in section 4.2.2; and procedures for controlling errors in mailout and check-in are in section 4.5.2. Other methods include the following:

(1)   Census sends letters, written by Census and by the NSF, to each sample unit along with the questionnaire and instructions, in the initial mailout. These letters explain the importance of the survey, assure the respondent that data will be kept confidential, and provide phone numbers for asking questions or requesting more questionnaires or additional time to report.

(2)   The DPD mails follow-up packages to nonrespondents 30, 60, 90, and 120 days after the initial mailing. All packages contain letters from Census and from the NSF. The 60-day package also contains a questionnaire, instructions, and return envelope, in case the respondent misplaced them. All packages are routed through the Statistical Methods Quality Control Branch in the DPD, for verification that packages were assembled correctly (this is also done for the initial mailing). R&D Survey staff also follow up nonrespondents, focusing on the large companies; recall that they phone the 300 largest R&D companies that do not respond after the 60-day follow-up, to collect important data items and encourage completion of the questionnaire. Census thoroughly documented follow-up procedures (U.S. Bureau of the Census 1993b, Ch. 8, Subch. C, Doc. 1).

(3)   In late 1993, Census began a study of nonrespondents to the 1992 survey. The purposes were to identify reasons for nonresponse, both to mandatory and voluntary items, and to get direction for future efforts to improve response.

(4)   Under task 1 of the R&D Survey improvement project, R&D Survey staff are doing the following things to control nonresponse error:

·   Document current procedures for dealing with nonresponse, including nonresponse follow-up.

·   Describe and examine procedures and assumptions for imputation.

·   Compute cross-classified nonresponse rates, such as by SIC and size of company.

(5)   Census imputes for data that may be missing due to unit or item nonresponse. Only selected items are imputed. Missing data are imputed for all RD-1 or RD-1-S forms, and the same procedures were used for the 1992 survey as for recent previous surveys. But Census adopted new procedures for imputing missing data for RD-1A forms for the 1992 survey. Missing data are imputed only for R&D performers: that is, for

reporters who had nonzero total within-company R&D expenditures in the current or prior year (item 3A.3, column 3 or 6, on Form RD-1A). All items except the distribution of costs by basic, applied, and development research can be imputed for RD-1A data. Most RD-1A imputation is done by computing <u>recode-level</u> year-to-year ratios or inter-item (within-year) ratios, and applying these ratios to reported or imputed company-level values. (Recall that a recode is a group of strata.) For contracted-out and foreign R&D costs, the Bureau imputes by computing the <u>company-level</u> ratio of current-to-prior company & other funds and applying this ratio to current or prior reported or imputed company-level values. Imputation of RD-1A data was thoroughly documented (U.S. Bureau of the Census 1993b, Ch. 16, Subch. E, Doc. 2).

4.4.3 Measurement of Nonresponse Error

The OMB suggested the following measures of nonresponse error in establishment surveys:

(1)   <u>Direct measures</u> One can measure the difference between data from respondents and data from nonrespondents; data can be obtained from nonrespondents by sampling them and intensively following them up or from a census or administrative records.

(2)   <u>Indirect measures</u> One indirect method is to measure response and nonresponse rates, by unit and by item, overall and cross-classified by size of company, SIC, and so forth, unweighted and weighted by size of company. It is important to obtain reasons for nonresponse, to see where future efforts to reduce nonresponse should be targeted.

The following is a description of measures of nonresponse error that have already been made for the R&D Survey. They are only indirect measures; as far as I know, data have never been obtained from R&D Survey nonrespondents to measure nonresponse error.

(1)   Census computes simple unit response rates (percent of forms mailed out that were returned with data) every week during data collection, based on check-in codes assigned by clerks in the DPD. (See section 4.5.2 for more details.) Figure 1, at the end of this paper, shows these response (return) rates for the RD-1 (long) and RD-1-S forms, for the 1991 and 1992 surveys (1992 data collection was not complete). Figure 2 shows response rates for the 1992 RD-1A form. Before the 1992 survey, Form RD-1A was only mailed when a new frame was constructed, every five or six years (the last mailout was for the 1987 survey). It will probably be mailed annually from now on. Note that return rates have exceeded 80% for all the forms in Figures 1 and 2.

(2)   Adams and Peck (1993) at the CES computed item response rates for the questions that break applied R&D costs into product group, basic R&D costs into field of research, and total R&D costs into geographic area (states), for 1974-89. They noted steadily declining item response rates for all these items through 1987, with some recovery in 1989 (these data were on the long form, so they were not collected every

year). For example, they reported that the response rate for applied costs by product group fell from 82.2% (878 out of 1068) in 1974 to 27.0% (465 out of 1724) in 1987, with an improvement to 46.8% (758 out of 1619) in 1989.

Unfortunately, Adams and Peck did not clearly define their response rates; there were no formulas, and no definitions of components of response and nonresponse. The denominator was apparently the number of forms mailed out (but it varied slightly between two sets of tables!), not the number of eligible reporting units, so the response rates could have been a function of more than response behavior. Also, the numerator was the number of positive, non-imputed values; valid zero responses were not included.

(3)   The NSF (1992) stated that response to voluntary questions (such as reported by Adams and Peck) is a problem; they reported that response rates to certain voluntary items have been declining recently.

(4)   Collins (1990) reported that for the 1988 survey, 84% of the companies in the sample reported their total R&D costs, but only 62% of the companies in the sample provided information on costs by character of work (basic, applied, or development). So, Census did not attempt to distribute 30% of total estimated R&D costs by character of work; nearly $30 billion in costs were not distributed.

(5)   Census computes imputation rates (as percents of estimated total values), and the NSF publishes them for selected items and industries in its annual publications (for example, see NSF (1992), Table A-2).

4.5 Processing Error

According to the OMB, processing error comes from the failure to correctly carry out properly planned methods.

4.5.1 Sources of Processing Error

The OMB broke the sources of processing error in establishment surveys into several major areas. They assumed that the sample design was correct and that the questions asked of respondents, and their responses, were correct:

(1)   Questionnaires Examples are typos or failure to put in arrows or boxes, because of inadequate proofreading.

(2)   Data collection Examples are sending the wrong version of the questionnaire or sending the questionnaire to the wrong contact or division in the company.

(3)   Clerical handling of questionnaires Errors could occur during stuffing and sorting of envelopes, or during check-in (such as losing a questionnaire).

(4)   Data processing by clerks and analysts Manual editing and review of error printouts could cause errors. Keypunch errors could occur. Analysts could make mistakes when manually imputing data, or overedit. Clerical staff could miscode questionnaires (SIC code, for example).

(5)   Data processing by computer Errors in computer programs, or programs that do not match specifications, are possible sources of processing error.

4.5.2 Control of Processing Error

The OMB listed several methods for controlling processing error:

(1)   Use acceptance sampling and quality control procedures. According to the OMB, these are appropriate for such things as stuffing envelopes, coding, and keypunching.

(2)   Run data through the computer edit a second time, after analysts have corrected the data that were flagged by the first run of the edit.

(3)   Test and review computer programs.

Census controls processing error by documenting survey procedures, using quality control methods, and by testing programs. Details of these controls are listed below:

(1)   To control errors during preparation and mailout of survey packages (forms, cover letters, instructions, and return envelope), Berry (1993) wrote specifications for an address file and Census wrote a detailed description of what the packages should contain (U.S. Bureau of the Census 1993b, Ch. 7, Subch. C, Doc. 1). Also, the Statistical Methods and Quality Control Branch in the DPD verifies that survey packages contain the specified contents, by inspecting informally chosen samples of packages.

(2)   When the DPD receives forms and correspondence from respondents, they sort them into batches according to detailed specifications (U.S. Bureau of the Census 1993b, Ch. 11, Subch. C, Doc. 1). Batches are processed with the Automated Control System (ACS), which assigns status codes on the basis of check-in codes entered by clerks (U.S. Bureau of the Census 1993b, Ch. 11, Subch. C, Doc. 2). Census uses the check-in codes to compute response and nonresponse rates, and uses the status codes to determine where forms are in the processing flow (for example, in keypunching or with an analyst). The ACS provides a series of computer menus and messages to help the user through the check-in procedure. Errors are further controlled by "wanding" a bar code for the 20-digit ID, rather than keypunching it. The ACS flags invalid IDs as

they are entered, so that they can be corrected immediately. Census also controls errors by providing thorough instructions for the check-in procedure (references were already cited in this paragraph).

(3)   The DPD keys data; procedures are well documented (U.S. Bureau of the Census 1993b, Ch. 13, Subch. E, Doc. 1). Census uses quality control (QC) procedures to ensure the quality of the keypunching. Richardson (1990) described QC for keying SSEL data; these procedures are also used for the R&D Survey. Data entry staff key (verify) a random sample of questionnaires from a previously keyed batch. This sample may be the entire batch if, for example, the batch is small, the data entry person is new, or it is early in the survey processing cycle. For the 1992 R&D Survey, the specifications said to verify 20% of the forms (except for situations, noted above, that warranted 100% verification). The next steps are determined by whether a sample or the entire batch is verified:

·   If a sample, and the error rate is no more than a specified level (1.4 to 2.8% for most batches, depending on batch size), the errors are corrected and the batch is accepted. Otherwise, the entire batch is verified and all errors are corrected.

·   If the entire batch is verified, the batch is accepted only after all errors have been corrected.

The goal of these QC procedures is to minimize the keypunch error rate, not to achieve perfect keypunching.

(4)   Census does most of the computer processing for the R&D Survey with its Current Industrial Reports (CIR) system or customized applications of this system. The CIR system is a parameter-driven package, using item code data rather than fixed field data, that Census uses to process about 70 monthly, quarterly, and annual industry surveys. Thus, the R&D Survey is a specific application of a generalized computer system that is regularly maintained.

(5)   R&D Survey staff verified the accuracy of the new computer program to impute for missing data on Form RD-1A, by examining output and comparing it with expected results.

(6)   MCD staff verified the accuracy of the R&D tab (summary) program, by writing a SAS program to produce the same output and comparing results.

4.5.3 Measurement of Processing Error

Measures given by the OMB are listed below:

(1)   <u>Direct measures</u> Direct measures of processing error are rare, because these errors are mixed in with errors from the other sources discussed in this paper.

(2)   <u>Indirect measures</u> Examples are keypunch error rates, edit failure rates, SIC reclassification rates, coding error rates, and imputation rates.

No direct measures of processing error have been computed for the R&D Survey. The following are two indirect measurements that have been used:

(1)   As discussed in section 4.5.2, the DPD computes keypunch error rates and bases decisions for correcting errors and reverifying keying on these rates.

(2)   The NSF publishes imputation rates in its annual publication (see, for example, Table A-2 in NSF (1992)).

5. CONCLUSIONS AND RECOMMENDATIONS

This section summarizes the documentation of nonsampling error issues. It begins with some general observations about the sources, control, and measurement of nonsampling errors in the R&D Survey. This is followed by some conclusions about each category of nonsampling error, suggestions for action, and possible follow-up studies.

5.1 General Observations

(1)   The NSF and Census have recognized many of the sources of nonsampling errors in the R&D Survey, and have taken some appropriate steps to control and measure these errors. The other tasks under the improvement project and the study of nonrespondents currently underway continue this process for improving the quality of the survey.

(2)   The documentation is generally up-to-date and clearly written; however, I had to go to many sources to get this documentation, and the format of the sources was inconsistent: memoranda (for the record or to an individual); reports (internal and external); the OMB clearance package; and a specifications and procedures manual. More of the documentation, especially specifications, should be in a consistent format, in one location.

(3)   Some of the studies that have been completed or are planned are special studies, outside the regular survey, that typically require additional funding. Examples are the response analysis surveys and the study of nonrespondents which has already begun. In this era of dwindling funds, we should try to find more ways to study, control, and measure nonsampling errors as part of the survey process.

5.2 Specification Error

The NSF has received input from many sources who identified their uses and needs for R&D Survey data. This input has led to improvements in the survey, such as in questionnaire design, so that data users are better served. The NSF should continue to seek and accept input. Some of the input has not been specific to the R&D Survey, so the NSF should seek input that is directly related to this survey in its future consultations with industry and government leaders and at workshops.

There are several additional, inexpensive ways to identify uses and needs for R&D Survey data which could be easily implemented. They are:

·   Ask members of appropriate advisory committees about their uses and needs for R&D data.

·   Whenever the NSF fills a user's request for R&D data, follow up by asking whether the user will complete a brief mailed questionnaire on their uses and needs. Maintain a file of users and their uses and needs.

·   Include a page in the front of all NSF publications on R&D that requests information on uses and needs from users of those publications.

I do not recommend evaluating the questionnaire and instructions under task 2, because so much has already been done or is planned in this area, and we are awaiting final results from task 4.

Census or the NSF should compare R&D Survey results with estimates from independent sources, because this has not been done for about ten years. As I noted earlier, this comparison has limitations, but it does give the user a better understanding of the quality and concepts of the different estimates, and an appreciation of the quality of R&D Survey estimates relative to other estimates. If good independent estimates could be found, the comparison could also identify and quantify specification and other nonsampling errors.

5.3 Coverage Error

Census seems to do a good job of controlling for coverage errors by constructing the frame from the SSEL, which is continuously maintained, consulting many sources for additional information on frame units, and updating names and addresses of frame units during survey processing. The 1992 frame, which is over 10 times larger than its predecessor, the 1987 frame, should have provided better coverage of small and nonmanufacturing R&D companies. With the move to annual frame construction, Census will be able to keep the frame more up-to-date than previously. Work under task 3 of the improvement project should lead to frame improvements and a more efficient sample design.

Only two techniques for measuring coverage error have been used. More should be done to

measure coverage error; for example, indirect measures (out-of-business rate, attrition rate, and so forth) could be computed.

5.4 Response Error

Response analysis surveys have been very useful for improving the questionnaire and instructions and for learning about the organization of companies and their information systems. These studies should be repeated periodically, say, every five years. Results should include indirect measures of response error, such as percents of respondents who followed instructions correctly or did not interpret a concept correctly.

Several record check studies have been done to measure response error, at irregular intervals. These studies should be done regularly; perhaps they could be built into the edit system. Other sources of data might be used, including: the list assembled by R&D staff for matching with the frame; Form BE-11, the Annual Survey of U.S. Direct Investment Abroad, used by BEA (a measure of foreign R&D); and Economic Census Forms ES-9100 (Enterprise Summary Report, mailed to all companies with 400 or more employees, that includes a question on total R&D costs) and ES-9200.

Census may want to measure edit failure rates. Census could also do a study of who actually responds on the questionnaires, to find out their level of responsibility and authority in the firm, their access to data, and so forth; this could lead to recommendations for strategies to get better respondents assigned to completing the questionnaire. More work is probably needed to see how well the respondents' information systems conform with the data requested by the survey.

5.5 Nonresponse Error

Response and nonresponse rates should be computed in a standard way, for comparisons across years and across the various Census surveys. In section 4.4.3, I mentioned some problems that can arise when rates are not clearly defined. Hidiroglou et al. (1993) provided a framework of response and nonresponse components that has been adopted as the standard for Statistics Canada's surveys. This framework could be used, with few changes, for the R&D Survey and other Census surveys; see Figure 3 at the end of this paper. Possibly boxes 9 (Temporarily Out-of-Scope Units) and 11 (Refusal Conversions) in Figure 3 would be removed for the R&D Survey; otherwise I think this represents the response and nonresponse components of the R&D Survey.

Unit and item rates should be computed, weighted and unweighted, overall and by industry, geographic area, and size of company. There are many rates that could be computed, including the resolved rate, the in-scope rate, the refusal rate, the no-contact rate, and the response rate (varying the denominator for the inclusion or exclusion of out-of-scope units). These rates are measures of the quality of the frame, the effectiveness of nonresponse follow-up, the quality of data collection procedures, and so forth. Unfortunately, there is not

sufficient information in the database to compute all of these response and nonresponse rates; it is impossible, for example, to compute a refusal rate. Additional codes are needed in the database to show which response or nonresponse category each unit falls into. Possibly fields that are already available in the CIR database, but are not currently used for the R&D Survey, could be used. Some changes might also be needed in the clerical check-in procedure at the DPD described in section 4.5.2.

Census should obtain reasons for refusals. The information should be coded in the database.

Nonresponse follow-up procedures may need to be studied. Are they statistically sound? Could the follow-up procedures be more aggressive?

More study is probably needed on incentives to respond to the R&D Survey. What additional incentives could be offered? Are respondents offered free publications? If not, perhaps they should be.

Jabine (1990, pg. 35) mentioned a cheap method for studying nonrespondents in his quality profile of SIPP. Data were examined for nonrespondents for earlier waves of the survey when they had been respondents. Then it was possible to compare the characteristics of nonrespondents with respondents; for example, renters, ages 15-24, tended to be nonrespondents in later waves of the SIPP. This method could be used in the R&D Survey, because the larger R&D companies are in the survey year after year.

5.6 Processing Error

Procedures for assembling survey packages, mailout, and check-in were well documented by Census. But there are no documented measures of the quality of these procedures. If there are measures, they should be documented; if not, measures should be computed, as they already are for keypunching.

Long and Ravenscraft (1993) found cases where data were on the forms but not in their computer file. While this could have been due to using some incomplete and reconstructed data sets to build their file, it would still be wise to compare a sample of forms with what is in the computer, to see if this is still a problem, and if so, to identify reasons. One might begin by investigating the data for companies that have undergone mergers, divestitures, and so forth, since these cases can be especially troublesome.

Census should verify that estimators and their variances are consistent with the survey design and definitions, if this has not already been done.

REFERENCES

Adams, J.D. and Champion, E.J. (1992), "Restructuring Research and Development Statistics at Census: a Blueprint for Change," presented at the meetings of the Census Advisory Committee of the American Economic Association, October 22-23.

Adams, J.D. and Peck, S. (1993), "A Guide to R&D Data at the Center for Economic Studies, U.S. Bureau of the Census," Center for Economic Studies report (draft).

Applied Management Sciences, Inc. (1989), "NSF Workshop on Industrial S&T Data Needs for the 1990s, Final Report," prepared for the National Science Foundation.

Berry, J. (1993), "Transmittal of Specifications," internal Census Bureau memoranda (Spec. nos. 2 and 3) to D. Tasky, May 28.

Champion, E. (1993), "Description of 1992 Survey of Industrial Research and Development Sample Design," internal Census Bureau memorandum to J. Berry and others, April 20.

Collins, E.L. (1990), "Estimating Basic and Applied Research and Development in Industry," NSF 90-322, Washington, D.C.

Goldenberg, K.L., Butani, S., and Phipps, P.A. (1993), "Response Analysis Surveys for Assessing Response Errors in Establishment Surveys," Proceedings of the International Conference on Establishment Surveys, Alexandria, VA: American Statistical Association, pp. 290-299.

Greenberg, B.V. (1993a), "Development of R-factors for the 1992 Research and Development Survey," internal Census Bureau memorandum to J. Berry, March 25.

Greenberg, B.V. (1993b), "Creation of the Sampling Frame for the Survey of Industrial Research and Development," internal Census Bureau memorandum to D. Tasky, April 2.

Greenberg, B.V. (1993c), "Documentation of Coverage Activities," internal Census Bureau memorandum to J. Govoni, June 1.

Groves, R.M. (1989), Survey Errors and Survey Costs, New York: John Wiley.

Hanczaryk, P.S. and Trager, M.L. (1990), "Improving the Quality of Standard Statistical Establishment List Data," Proceedings of Statistics Canada Symposium 90, Ottawa: Statistics Canada, pp. 211-222.

Havard, J. (1993), "Sources for Selecting the R&D Predetermined Certainties," internal Census Bureau memorandum for the record, April 6.

Hidiroglou, M.A., Drew, J.D., and Gray, G.B. (1993), "A Framework for Measuring and Reducing Nonresponse in Surveys," Survey Methodology, 19, 81-94.

Jabine, T.B. (1990), SIPP Quality Profile, Washington, D.C.: Department of Commerce.

Kusch, G. (1993), "Remarks on Minimum Probability," internal Census Bureau documentation, April 9.

Long, W.F. and Ravenscraft, D.J. (1993), "The Impact of Corporate Restructuring on Research and Development," prepared for the National Science Foundation.

National Science Foundation (1985), A Comparative Analysis of Information on National Industrial R&D Expenditures, NSF 85-311, Washington, D.C.

National Science Foundation (1992), Research and Development in Industry: 1989, NSF 92-307, Detailed Statistical Tables, Washington, D.C.

Office of Management and Budget (1988), Quality in Establishment Surveys, Statistical Policy Working Paper 15, Washington, D.C.: OMB.

Richardson, W.J. (1990), "Data Entry Quality Control," internal Census Bureau memorandum to J.S. Harris, June 15.

Siegel, D. and Andrews, S.H. (1989), "Final Report: Examining the Company-Level R&D Database," Center for Economic Studies report.

U.S. Bureau of the Census (1989), "Evaluation of Response to the 1987 Survey of Industrial Research and Development," Industry Division report prepared for the National Science Foundation.

U.S. Bureau of the Census (1993a), "Evaluation of Proposed Changes to the Survey of Industrial Research and Development," Industry Division report prepared for the National Science Foundation.

U.S. Bureau of the Census (1993b), Survey of Industrial Research and Development Specifications and Procedures, internal Industry Division documentation.