National Survey of Small Business Finances

METHODOLOGY REPORT

Sponsored by the Board of Governors of the Federal Reserve System and the Small Business Administration

July 24,1996

Price Waterhouse LLP

METHODOLOGY REPORT
TABLE OF CONTENTS

METHODOLOGY REPORT
APPENDICES

METHODOLOGY REPORT
LIST OF EXHIBITS

METHODOLOGY REPORT
LIST OF EXHIBITS IN APPENDICES

I. INTRODUCTION AND PROJECT OVERVIEW

The 1993 National Survey of Small Business Finances (NSSBF), co-sponsored by the Federal Reserve Board and the U.S. Small Business Administration, was conducted during 1994-1995 to collect information on the availability of credit to small and minority-owned businesses. The 1993 NSSBF was motivated in part by Section 477 of the FDIC Improvement Act of 1991, which requires that the Federal Reserve Board collect and publish information on small business credit availability. A second, but related motivation, was the Federal Reserve Board's responsibility for evaluating bank mergers and acquisitions applications. The information collected by the survey constitutes one of the most comprehensive general-purpose database on small businesses that is available in the public domain. This information should greatly enhance understanding of the financial and credit relationships that small businesses maintain with banks and other financial institutions, and improve the ability of researchers to empirically delineate meaningful economic markets.

The 1993 NSSBF provides information from 5,356 completed interviews of a random sample of small businesses, with stratification by firm size, geographic region of the country, and urban/rural location. In addition, the sample design was structured to yield sufficient numbers of minority-owned firms to conduct separate analyses of minority- and non-minority-owned small businesses. Included among the 5,356 firm sample are 354 Asian-owned firms, 508 Black-owned firms, and 371 Hispanic-owned firms.

Prior to completion of the 1993 NSSBF, the most recent source of comprehensive data on small business borrowing was the 1987 National Survey of Small Business Finances. The data from that survey made important contributions to understanding of small business financing, but those data have become somewhat dated. Moreover, the 1987 NSSBF was much smaller (3,254 firms) and did not include enough minority-owned firms for separate analyses of such firms. Since the 1987 NSSBF, numerous

changes have occurred in the financial services industry. The banking industry has continued to consolidate, especially at money-center and large super-regional banks. Banks are offering new services, such as mutual fund investments and banking by personal computer. Hand in hand with consolidation has been the spread of interstate banking across the country. Small business has also undergone change. A "credit crunch" caused many firms to retrench in order to survive. Both the credit crunch and industry consolidation are rumored to have led banks to become more cautious about financing and lending, thereby limiting the available sources of financing for small businesses. These credit constraints may have led small businesses to explore credit sources other than commercial banks, their traditional lenders.

The 1993 National Survey of Small Business Finance targets all enterprises operating under the current ownership during 1992 and with fewer than 500 full-time equivalent employees, but excluding agricultural enterprises, financial institutions, not-for-profit institutions, government entities, and subsidiaries controlled by other corporations. The survey provides the Federal Reserve Board, the Small Business Administration, and other government agencies with information on the cost and availability of financing for these small businesses. These agencies are concerned with the effects of innovation by financial institutions and regulatory reforms mandated by Congress, and with the effect of the economy in recent years on the range of financial options available to small businesses. The information from the survey will be used to evaluate the impact of public policies on small businesses of different sizes, locations, and ownership characteristics. The 1993 National Survey of Small Business Finances was conducted to answer questions such as the following:

- What experiences have small businesses had with recent credit applications?
- What factors may have increased prices or reduced the availability of credit to small businesses?

- Are credit problems especially severe for small businesses of particular sizes, in particular regions of the country, or with specific characteristics?
- Have government regulations reduced the availability of credit to small businesses?
- Are there sources other than commercial banks to which small businesses can turn to meet a significant portion of their financing needs?
- Do financial institutions actively seek to provide credit to small businesses?

The ensuing report summarizes the methodology used by Price Waterhouse LLP ("Price Waterhouse") to conduct the 1993 National Survey of Small Business Finances. The report describes questionnaire development in Section II, sample selection in Section III, survey disposition in Section IV, pre-interview procedures in Section V, and data collection and data processing in Sections VI and VII.

The survey filtered out-of-scope businesses (e.g., branch offices, subsidiary companies, not-for-profit, financial services, and agricultural industries) from Dun & Bradstreet's November, 1993 DMI file to create the population frame of about 7.3 million potentially in-scope business enterprises. This frame was divided into four partitions—Main, Asian, Black, and Hispanic—to facilitate the sampling of minority-owned firms. We used the filtered population to draw a stratified random sample with the following

independent sample strata:

- Main Partition
  - 90 sampling strata defined by:
    - 9 Census Regions
    - Urban/Rural Businesses (2 groups)
    - 5 Size Groups (1-19, 20-49, 50-99, 100-499 employees, and Unknown)
  - One stratum of businesses with more than 500 employees[1]
- Asian Partition
  - Urban/Rural Businesses (2 strata)
- Black Partition
  - Urban/Rural Businesses (2 strata)
- Hispanic Partition
  - Urban/Rural Businesses (2 strata)

We conducted a two-stage interview consisting of eligibility screening and a main interview. Both stages were conducted by telephone using computer-assisted telephone interviewing (CATI) methodology. After the initial screening, qualifying businesses were mailed an information package to prepare for the main interview and to familiarize themselves with the study.

The survey collected detailed statistics on owner and organizational demographics, sources of financial services, income and expenses, and a complete balance sheet of

[1] While businesses with more than 500 full-time equivalent employees are out-of-scope for the survey, differences in the treatment of part-time employees (Dun & Bradstreet do not differentiate between full-time and part-time employees whereas the NSSBF counts part-time employees as one-half of a full-time employee) and inaccuracies in the DMI file meant that we could not be sure these businesses were out-of-scope. We therefore sampled a small proportion of these businesses to ensure a representative sample of all small businesses.

financial data for each firm. Information ranged from the principal owner's age to a firm's accounts receivables, from number and types of loans to a firm's interest expense, and from length of relationships with financial institutions to the trade credit discount for early payment.

Systematic procedures ensured the collection of quality data. Two formal pretests of the questionnaire and data collection procedures highlighted areas for improvement. Computer assisted telephone interviewing strategies with recurrent range-edits and consistency checks also enhanced the data quality level. Paralleling the previous survey, interviewers requested that respondents send to Price Waterhouse via mail their firm's tax records, worksheets, and financial statements in a business reply envelope. The Federal Reserve used these documents to conduct further data edits and to verify data integrity.

The project's final product is an edited and coded data base. This forms the basis for further editing, analytical imputations, and processing by research staff at the Federal Reserve which will produce a final 1993 NSSBF public-use data base.

Several key facts about the survey are listed below.

- The sample frame chosen is the Dun & Bradstreet November, 1993 DMI file.
- The main effort in designing the questionnaire, from determining content to CATI programming, lasted from October, 1993 to March, 1994.

- Survey data reference several different temporal periods:
  - For screener information and for firm and owner demographic data, the current date was referenced (1994).
  - For questions about sources of financial services such as checking accounts, savings accounts, lines of credit , etc., the most recent complete calendar year was referenced (as of December, 1993).
  - For questions about the firm's balance sheet assets and liabilities and the firm's income and expenses, the firm's 1992 accounting period (whether calendar or fiscal) was referenced. This period was selected in order to allow firms to use complete accounting or tax records in order to enhance data collection.
  - Questions about the firm's most recent credit application referenced the "last three years."

- The first pretest of 24 hard copy interviews was conducted from November, 1993 to January, 1994.

- The CATI survey instrument was programmed from November, 1993 to March, 1994.

- We conducted the second pretest in CATI for both screening and main interviewing stages. The second pre-test screening lasted from January 25, 1994 to January 28, 1994. The main pre-test interviewing lasted from February 9, 1994 to February 22, 1994.

- We fielded the survey screener from March 9, 1994 to November 3, 1994. We fielded the main questionnaire from March 23, 1994 to January 31, 1995. Follow-up calls to individual respondents for crucial missing data were made through February, 1995.

- Answering machines, answering services, receptionists, and pagers severely restricted our ability to reach many appropriate respondents.

- Once the appropriate respondent was contacted, the two major obstacles to conducting a complete interview were confidentiality concerns and time concerns.

- The survey collected data for 5,356 firms, including 354 Asian-owned firms, 508 Black-owned firms, and 371 Hispanic-owned firms.

- The best estimate of the overall survey response rate is 59%, with a lower bound of 50%. (The best estimate assumes that the eligibility rate of non-respondents was the same 68% as that of respondents, while the lower bound assumes that 100% of non-respondents were eligible).

- The average number of calls to complete a screener was 5.1.
  The average number of calls to complete a main interview was 18.7.

- The average length in minutes to complete a screener was 2.2.
  The average length in minutes to complete a main interview 49.7.

The 1993 National Survey of Small Business Finances fulfills the following research objectives:

- Collects data from a representative probability sample of small businesses and of minority-owned small businesses[2]:

  - The target population is all business enterprises operating under the current ownership during 1992 and with fewer than 500 full-time equivalent employees, but excluding agricultural enterprises, financial institutions, not-for-profit institutions, government entities, and subsidiaries controlled by other corporations.

  - The sample is designed to achieve a precision of five percent at the 95 percent level of confidence for a test of differences in proportions between non-minority-owned firms and Black-, Hispanic-, or Asian-owned firms. The sample is large enough to conduct separate analyses of small businesses in sub-groups defined by region, urban versus rural, and size based on number of employees.

- Provides data needed by the Federal Reserve and the Small Business Administration on issues associated with credit availability to small and minority-owned businesses.

[2] The sample frame date is November, 1993.

- Provides data needed by the Federal Reserve to assist in evaluating bank mergers and acquisitions applications.

- Provides data on financial assets and liabilities of small businesses, the use of financial services by small businesses, the location and types of institutions that supply financial services to small businesses, and firm demographics (e.g., organizational structure, industry, age).

- Provides an analysis data base for use by the Federal Reserve, the Small Business Administration, and ultimately the general public that has been thoroughly edited and coded, with analysis weights appended.

- Facilitates a potential follow-up or longitudinal survey.

The research design model for this study is the 1987 National Survey of Small Business Finances, sponsored by the Board of Governors of the Federal Reserve System and the Small Business Administration. Useful reference documents from the current survey (bound separately) include the Survey Questionnaire and the Codebook.

II. QUESTIONNAIRE DEVELOPMENT

A. Introduction

This project was designed to provide information on credit availability to small and minority-owned businesses, to generate an up-to-date general-purpose database on small businesses, and to update information collected by the 1987 NSSBF. To achieve these goals, we created a comprehensive questionnaire that collected large amounts of financial data and information regarding the owners and the firm. The concepts and ideas embedded in the 1987 questionnaire provided a starting point for constructing the 1993 questionnaire. However, the focus of that survey was the definition of banking markets used by small businesses; therefore, it was necessary to substantially revise that questionnaire in order to fulfill the above stated goals of the 1993 survey.

B. Structure and Topics of the 1993 NSSBF Questionnaire

This section describes the rationale behind the ordering, phrasing, and content of the questions that comprise the 1993 NSSBF questionnaire. The structure and topics of the questionnaire are presented in Exhibit II-1.

Section I of the questionnaire collects information on the firm's characteristics. Such questions include the survey's eligibility requirements, personal traits of the owner/top executive, and the firm's demographics. The next section identifies all the financial services that the firm was using for business purposes as of year end 1993 and creates a roster of the different financial sources providing these services. For example, questions explore the firm's use of checking and savings accounts, capital leases, lines of credit, transaction services, and brokerage services. Section II also obtains information regarding the characteristics of the financial sources, the firm's

EXHIBIT II-1: STRUCTURE OF THE 1993 NSSBF QUESTIONNAIRE*

SECTION I - CHARACTERISTICS OF THE FIRM

A. Screening Information
B. Organization Demographics
C. Personal Characteristics of Owners
D. Firm Demographics

SECTION II - SOURCES OF FINANCIAL SERVICES

E. Use of Deposit Services
    1. Checking Accounts
    2. Savings Accounts

F. Use of Credit and Financing
    1. Credit Cards
    2. Lines of Credit
    3. Leases
    4. Mortgages
    5. Motor Vehicle Loans
    6. Equipment Loans
    7. Loans from Partners/Stockholders
    8. Other Loans

G. Use of Other Financial Services
H. Relationships with Financial Institutions
J. Most Recent Credit Application
K. Solicitation
L. Use of Trade Credit
M. New Equity Investments in the Firm

SECTION III - INCOME AND EXPENSES

N. Reference Period
P. Income and Expenses

SECTION IV - BALANCE SHEET

R. Assets
S. Liabilities and Equity
U. Credit Worthiness

*Note that subsections I, O, Q, and T are intentionally omitted.

most recent credit application, solicitation by creditors, the use of trade credit, and new equity investments in the firm.

Section III of the questionnaire collects a detailed accounting of the firm's financial data. Specifically, questions ask for income and expense items ranging from cost of goods or services sold to interest expense. The final section of the questionnaire has two main purposes. The first is to collect key financial information from the firm's balance sheet regarding the firm's assets, liabilities, and equity (e.g., inventory and accrued expenses). The second purpose is to collect information regarding the firm's credit worthiness and conclude the interview.

The order of questions generally followed the rule of asking easier questions early to obtain cooperation and to help establish respondent-interviewer rapport. In order to accommodate anticipated break-off, it was important to place the more difficult questions (e.g., confidentiality concerns, etc.) together at the end.

This specific questionnaire structure unfolded as a result of the dynamic questionnaire development process undertaken by the project's researchers from Price Waterhouse and the Federal Reserve Board, and through the conduct of two live pretests with eligible small businesses. This questionnaire design period lasted approximately six months.[3] The following sections present some motivations and decisions that had significant impact on the development of the final questionnaire.

[3] Of course, other tasks were performed simultaneously, including sample preparation and selection, identification of minority lists, etc.

C. Development of NSSBF Questionnaire

Questionnaire development involved revising the 1987 NSSBF questionnaire to accommodate Board and Small Business Administration requested changes and to cover required topics of interest for the 1993 survey. After finalizing the content, we streamlined the data collection instrument and conducted two pretests with eligible small business owners.

The questionnaire development phase of the project was accomplished through numerous working sessions among the Federal Reserve, SBA, and Price Waterhouse staff. One of the earliest decisions of the questionnaire team was to first build an exhaustive, comprehensive questionnaire that collected every relevant piece of information. Then it was necessary to delete specific items to make the questionnaire an effective and workable data collection survey instrument.[4] This elimination focused on unnecessary duplication and less essential information. The pretests revealed that additional streamlining was required to meet the target interview length of 45 minutes. At this point, items with a relatively lower priority were eliminated. Through this process, we ended up with a compromise questionnaire that attempted to balance the amount of detailed information required with anticipated analysis needs, but which stretched the limits of data collection on the survey topics from the difficult target population of owners of small businesses.

Because of the complexity of the information required from respondents, we retained the general format for asking each question as developed in the previous NSSBF survey. That is, a respondent is first asked a short, concise question. If further explanation is necessary, a prompt containing additional clarification and definitions could be read. This arrangement accommodates almost every respondent's level of

[4] Questions were deleted with their question number. Therefore, in the final version of the questionnaire, several question numbers are intentionally left missing.

financial sophistication. It also results in an optimal amount of explanation for maximizing the efficiency of questionnaire administration.

Board researchers took the lead in determining the content of information to be collected and gave approval for the final questionnaire version summarized in Exhibit II-1. While the question content does mirror the content aggregated in the 1987 NSSBF questionnaire, several major differences exist. The 1993 version seeks additional data regarding credit cards and new equity investments in the firm. A detailed section on the most recent credit application is included. Also, the question on minority ownership was redesigned to reflect the Census definitions. This helped both to track our identification of minority-owned firms and to screen out non-minority firms from those sampled from the specially created black, Asian, and Hispanic minority-owned firm sample partitions. On the other hand, the 1993 NSSBF questionnaire obtains less information in several aspects regarding a firm's relationships with financial institutions than the 1987 questionnaire.

A parallel process to determine the interview content was establishing an appropriate temporal basis for the individual questions. Some questions referred to a period of time (e.g., during FY 1993) while others had to be answered as of a specific point in time (e.g., as of December 31, 1993). Items that are generated over time, such as income and expenses, associate with a period of time, while items that exist for a point in time, such as account balances and amounts owed, relate to a particular date.

The questionnaire provided clear guidelines for establishing several time references. The actual time frames used were 1990, the 1992 fiscal or calendar year, and 1993. We asked a few questions referencing 1990 for comparison purposes. These questions (e.g., the number of employees, total sales) followed their up-to-date counterparts so respondents could provide answers while maintaining focus on the relevant topics. Section II of the questionnaire, Sources of Financial Services, referred

to "as of year end 1993." We used this time period to collect the most recent data for a complete calendar year that were available to all respondents. Section III and IV, Income and Expenses and Balance Sheet, respectively, cited the 1992 fiscal or calendar year. We referenced the 1992 year because many firms surveyed during 1994 would not have completed financial and tax records for their 1993 fiscal or calendar year, and we wanted firms to be able to refer to records in order to improve the accuracy of their responses. Section J on Most Recent Credit Application and Section M on New Equity Investments in the Firm used "during the last three years" as the time period reference. This time period appeared to be an appropriate window to collect accurate and pertinent information—recent enough to satisfy recall concerns, but long enough to provide enough sample for analysis.

We have specifically retained from the 1987 survey Section II's format of probing each financial service for all institutions or other sources that provide that service to the firm. Pre-test experiments conducted for the earlier survey demonstrated the superior effectiveness of this approach over alternatives. (See Cox, Elliehausen, and Wolken, September 1989). The primary alternative structure is to identify an institution or other source of financial services used by a firm and then obtain all services provided by that source. The major advantage of this alternative approach is to take advantage of the fact that many firms use only one financial institution. However, it would be necessary to use detailed follow-up probes for each financial service to avoid missing any sources of financial services. Section II developed a roster of institutions or individuals that are sources of financial services used by the business to support this format. As the respondent identified a new institution or individual as the source, the interviewer entered the source's full name on the next line of the roster beside the corresponding number. The CATI system stored all sources named from previous questions (e.g., checking account) and when the respondent indicated using a particular source for a named service for later questions, the interviewer flagged the source. When a respondent did not know or refused to provide a source name, the interviewer recorded

institution 1, institution 2, . . . , or person 1, person 2, . . . , etc., to facilitate building the roster of sources of financial services used by each small business.

Some valuable information ascertained from the 1987 study was used to inform the ordering of the 1993 questionnaire. For instance, we asked about for-profit status and verifying headquarter location initially in the screening process since these were the two reasons firms still in business were most often screened out. As stated above, the more difficult, sensitive questions were placed near the end of the survey and the questions which would perhaps be more interesting and pique respondent interest were placed toward the beginning of the screener and the main questionnaire.

After deciding an effective ordering of the questions, it was necessary to fine-tune question wording. To the extent possible, exact question wording was maintained for trend questions repeated from the 1987 survey. For new question topics, the primary determining factor in developing question wording was clear, concise language to capture the essential content of data collection needs expressed by the Board and confirmed by analysis of the pretests. The pretests noted several individual questions that generated needless confusion or irritation either to the interviewer and/or the respondent. (The specific questions are detailed in sections D.1 First Pretest and D.2 Second Pretest.) The project staff examined each question, making revisions as needed so as to produce the best question wording. One major objective was to have questions phrased consistently across the questionnaire with the goal of obtaining the most accurate possible data.

We further enhanced the questionnaire by tailoring questions to firms based on their status as a proprietorship, partnership, or type of corporation. Each organizational type of firm was asked different questions to get at the same basic information. This allowed us to use the most appropriate language with each respondent. To accomplish this, skip patterns were embedded in the questionnaire to jump to the appropriate question

based on the firm type. The survey's computer-based (CATI) system effectively allowed this tailoring, without the need for interviewer intervention.

Price Waterhouse accountants provided assistance in the development of the financial sections of the questionnaire. Additional advice was obtained from external accountants who specialize in providing financial services to small businesses. These accountants specifically helped in the financial sections' organization and content, providing advice on the collection of appropriate data for each type of firm sampled (i.e., proprietorships, partnerships, S-corporations, corporations). The Price Waterhouse accountants and tax advisors also aided in the construction of detailed reference prompts to link survey questions with the specific line(s) of appropriate IRS tax forms.

An additional problem to be addressed involved the fact that although IRS tax forms for partnerships and corporations provide appropriate lines for a firm's income, expenses, and balance sheet, proprietorship tax returns do not reference a complete balance sheet. In order to address this inadequacy and to improve general respondent readiness for the interview, all firms were sent a mail package familiarizing the owner with some of the information to be collected (see Appendix C). Also, all respondents were asked to have financial records available at the time of the interview, not just tax returns.

The accountant advisors also provided guidance in establishing the CATI range edits. They pointed out particular financial items that could only be (or are typically) negative values, positive values, or non-negative values. These were used as internal edit checks to help safeguard the collection of data during CATI interviewing. These edit checks are detailed in the project's separately bound "CATI Interviewer Training Manual."

D. Pretest Results

We conducted two formal pretests of the questionnaire and data collection procedures using a convenience sample of businesses randomly selected from the DMI frame. A mix of industries, sizes, and forms of organization was included among those firms selected. We directed the pretests towards examining the extent to which the draft questionnaire was meeting the characteristics of good questionnaire design presented in Exhibit II-2. Each had significant impacts that resulted in refinements to the survey instrument.

EXHIBIT II-2: CHARACTERISTICS OF GOOD QUESTIONNAIRE DESIGN

- High item response
- Low break-off rates
- Respondents easily understand the questions
- Respondents do not misinterpret the questions
- Interviewers do not have difficulty with question administration
- Questions with categorical responses list the appropriate responses

D.1 First Pretest

We conducted the first pretest in December, 1993 and designed it to determine if:

- respondents easily understood the questions
- respondents did not misinterpret the questions
- respondents were familiar with terms/vocabulary used
- respondents were able to obtain the requested information
- interviewers did not have difficulty with question administration
- questions with categorical responses listed the appropriate responses
- the flow of the questionnaire/the order of the questions was effective

We did not conduct this pretest to measure potential response and item-response rates. A pretest conducted to adequately measure such rates would require a

significantly longer administration period and a representative sample. Although we did not formally track item non-response, questions that were consistently misunderstood by respondents or elicited a "don't know" response were noted and revised as described below.

Because of the need to get timely feedback from the first pretest, it was conducted as a paper-and-pencil interview. The convenience sample of 300 small businesses initially consisted of an equal number of firms categorized as small, medium, or large. These firms respectively had 0-49, 50-99, and 100-499 employees as recorded on the initial DMI data base. It was important to know that the questionnaire worked equally well among firms of different sizes. Therefore a random sample of 50 firms was selected from each size category. As the pretest was coming to a conclusion, it was determined that an insufficient number of partnerships was sampled. To rectify that situation, the pretest sample was augmented with a sample of 40 additional firms from the DMI database specially targeted as likely partnerships (e.g., firms with the title of partner as owner in the sample data base). Of the 340 small businesses in this convenience sample, 190 firms were contacted for the first pretest and 150 firms were held as reserve sample.

Exhibit II-3 displays the final sample disposition of the first pretest. Many firms were not finalized because the pretest fielding period ended before the sample could be fully worked with a rigorous call-back protocol.

EXHIBIT II-3: FIRST PRETEST SAMPLE DISPOSITION

Screening Results	Count
Total Firms Contacted	190
■ Ineligible Firms (complete)	65
■ Eligible Firms (complete)	125
Interview Results	**Count**

Total Mail-Outs	125
Completed Interviews	24
■ Corporations	9
■ S-Corporations	9
■ Sole Proprietorships	3
■ Partnerships	3
Scheduled Appointment at Conclusion of Pretest	5
Unresolved at Conclusion of Pretest	70
Refused	26

The basic survey administration procedure planned for the main study (i.e., the same as used in the 1987 NSSBF) was tried and investigated in this first pretest. This involved an initial screening to determine if the firm was eligible for the survey and to confirm name and address information. After screening, eligible firms were sent the mailout package containing a letter from the Chairman of the Federal Reserve Board, Alan Greenspan, a letter from Price Waterhouse, a brochure of answers to commonly asked questions, and worksheets for organizing the firm's financial data. This mailout package is described in greater detail in Section IV: Pre-Interview of this report, and it is displayed as Appendix C: Advance Mailout Package.

It was apparent that the initial contact procedure was very important. Gatekeepers wanted to know why we wished to speak with the owner. Gatekeepers also frequently expressed a concern about whether they should be giving information required by the screening questionnaire over the telephone. Another observation was that conducting the screener with the owner did not necessarily seem to improve the probability of concluding the screener. These factors combined led us to establish a new general procedure for the screening survey. Interviewers would do the following:

- read the introductory script
- ask for selected respondent (owner or top executive)
- if selected respondent not available, ask for their assistant
- if no assistant or they are not available, continue with gatekeeper

By following this sequence, verification of information was obtained nine out of ten times with a gatekeeper. Other lessons learned from the first pretest regarding the screening interview follow:

- Not conveying to the respondent the number of upcoming questions was sometimes a problem. It was found that adding a statement such as "and then I have about 5 or 6 more questions to verify you are a small business" after the introduction worked well.

- The phrase "not-for-profit" often led to confusion. Some firms assumed that an operating loss put them in this category. Saying "non-profit" with some probes worked much better. The final questionnaire adopted this version of obtaining not-for-profit status.

- The question "Does another company own more than 50% of your firm?" put some respondents on guard, and occasionally contributed to respondent unwillingness to continue with the screening interview. An explanation such as "we are looking for the headquarters of the firm" seemed to help in this situation.

- The original question "For tax purposes, is the firm considered to be a sole proprietorship, partnership, S-corporation, or corporation?" led some respondents to answer with C-corporation. We trained interviewers that this is the same thing as a corporation. To enhance the accuracy of the classification, we concurrently trained interviewers to be aware of signals in the firm's name, such as “Inc.”, that might classify the firm as either an S-corporation or corporation as opposed to proprietorship or partnership.

- Another issue was that the phrase "for tax purposes" put respondents on guard. For that reason, the phrase "for tax purposes" was eliminated from the final version of the questionnaire.

- The screener's original closing remarks seemed insincere and prolonged. Specifically, the phrase we will be sending you a "very important package" evoked unnecessary disdain. We revised the script to ensure a more concise and constructive ending to the screener survey.

After experimenting with several approaches during the first pretest, the following interviewer procedures were found to be most effective when making the first contact after the screener:

- Ask respondents "have you received the mail package?"
- Ask respondents "have you had a chance to review it?"
- Ask respondents "do you have any questions?"
- Then conduct the interview or set an appointment.

Being able to conduct the interview at the respondent's convenience was a very important aspect to getting the interview. Examples of special requests for interview times during the first pretest include a 5:00am, a 6:30am, a 7:00am, and a 10:00pm Saturday call. Our suspicion was confirmed that accommodating respondent needs should be an integral part of any interviewing strategy.

The first pretest raised several issues that related to interviewer training, the interviewing process, and the survey in general. One issue noted that worksheet references should be provided as well as tax form references. Another was that having complete financial definitions on the CATI screens would be beneficial to the interviewer and the respondent. A third issue resulted in the observation that interviewer discretion was the best guide in deciding when to read prompts. Another partially supported the 1987 assertion that corporations and partnerships tended to have more sophisticated financial records than proprietorships. The pretest indicated that partnerships and proprietorships were more alike in their financial data organization. This was the main difference across types of firms that surfaced during the first pretesting. In these cases, it was determined that the prompts would help those less sophisticated firms.

There were also three further meaningful points learned from the first pretest. The first was that several respondents refused to address any question that was not in the worksheets. This led to the inclusion of a statement in the worksheets explaining that it does not contain all the questions in the study. The second also involved the worksheets. At the time of the first pretest, the worksheets were rather comprehensive. They were foreboding as each page contained very little white space and was full of

words and boxes. This unnecessary complication contributed to avoidance behavior. It was decided to retain the content in the worksheets, but to edit the language to simple, concise and essential instruction.

The third issue investigated during the first pretest was the formation of interviewer teams. We concluded that organizing the sample for conducting both screener and the questionnaire in terms of interviewer teams would add value to the survey administration process for the following reasons:

- Focusing on teams reduces the number of individuals contacting a firm and reduces the respondent's tendency to become frustrated because they feel as if they are being "bounced around."
- When working with businesses that have to be contacted several times before actually conducting the interview, respondents seem to like having one person, (or a minimal number of people) with whom they must speak to complete the interview.

Cutting the questionnaire was not a paramount issue after the first pretest since it was intended that questionnaire be designed as comprehensive as possible for content during the early stages of development. Since the first pretest was conducted paper and pencil, and not via CATI, this version could not result in confident conclusions as to the length of the impending CATI interview. Still it gave some indication as to where the length stood. With 24 completes, the average total time for an interview was 68.6 minutes. The minimum length was 46 minutes and the maximum length was 106 minutes. This gave a rough indication that the interview needed to be shortened, and cut perhaps by as much as one third.

There were two subsections which had relatively long pretest interview times: Use of Credit and Financing (Section F), and Relationships with Financial Institutions (Section H). These questionnaire sections averaged 8.7 and 8.3 minutes in length respectively. These pretested lengths suggested that these two areas of the questionnaire might have a disproportionate share of time allotment compared to their importance to the

survey. In the final version of the study the questionnaire section on Relationships with Financial Institutions, (Section H), was significantly reduced.

The first pretest raised only a few, minor questionnaire concerns. Specific questions needing attention were J4 and R14. They respectively asked for the timing of the most recent request for a loan or a line of credit (J4) and for other assets not yet described (R14). Adding the phrase "either approved or denied" to the end of J4 helped clarify the question. In addition, we determined that R14 needed a prompt to include examples of what would be considered other assets. We incorporated both of these additions into the survey.

The first pretest also indicated that respondents had difficulty selecting a response to the question about where the firm primarily sold or delivered its products (D4). Respondents had trouble differentiating between the response categories: regionally and locally. A change was made to the response categories to eliminate this confusion. Finally, the first pretest showed that responses to the question about the most important issue affecting your firm (D14), was heavily influenced by the series of questions on problems facing some businesses (D13). In order to eliminate this influence the order of these questions was reversed.

Lastly, Exhibit II-4 displays the use of reference materials in the first pretest. For the questions on Sources of Financial Services (Section II), three-fourths (18) of the 24 respondents interviewed in the first pretest reported using records to answer questions. When asked which types of records were referenced, thirteen (13) respondents indicated that they used tax records and ten (10) indicated that they used financial statements. For the financial questions in the questionnaire (Sections III and IV), four out of five (20) of the respondents reported using records. When asked which types of records they used, the results were the same as for Section II. Any conclusions

regarding record use should keep in mind that the pretest sample did not have a proportionate amount of “very” small firms.

EXHIBIT II-4: FIRST PRETEST USE OF REFERENCE MATERIALS

Question	Response	Count
Think about the questions concerning the firm's use of checking, savings and investment accounts, and credit lines, leases, and loans. Were records used to answer these questions?	Yes No	18 6
What records were used?	Tax records Bank records Financial Statements Others*	13 1 10 6
Now think about the questions concerning the firm's assets, liabilities, income and expenses during (YEAR). Were records used to answer these questions?	Yes No	20 4
What records were used?	Tax records Bank records Financial Statements Others*	13 1 10 6

*Other responses were auditor, balance sheet/Comptroller, trial balance/balance sheets, general ledger, and account statements.

D.2 Second Pretest

We conducted the second pretest in January and February, 1994. Both the screener and the main questionnaire were conducted via CATI. This pretest tested for characteristics of good questionnaire design as well as the CATI program and the questionnaire length.

As in the first pretest, we did not conduct the second pretest to formally measure potential response and item-response rates. A pretest to adequately measure these rates would require a significantly longer administration period. However, questions that respondents consistently misunderstood or responded with a "don't know" were revised.

Because we conducted this pretest in CATI utilizing available range edits, we were able to test the system of problem sheets and exception codes. The CATI program used range checks in order to minimize key entry errors and highlight unusual responses. When an interviewer entered a response that was out of range, they received an error message and a prompt to re-ask the question. If, after probing, the respondent insisted that the response was accurate, the interviewer entered the exception code "EX." The interviewer then recorded the given response along with an explanation of the unusual circumstances on a problem sheet. At the conclusion of the interview the interviewer entered the information recorded on the problem sheet into a machine readable text field at the end of the questionnaire for future coding. Based on the results of this pretest, some ranges were revised for the final questionnaire.

A total of 190 small businesses were randomly sampled from the DMI file for the second pretest. Exhibit II-5 presents the final sample disposition of the second pretest. As in the first, the final outcome of interview attempts for many firms were not resolved

EXHIBIT II-5: SECOND PRETEST SAMPLE DISPOSITION

Screening Results	Count
Total Firms Contacted	190
■ Unresolved	42
■ Passed Screener	111
■ Refusal	4
■ Language Problem	1
■ Non-Working Number	7
■ Confirmed Out of Business	3
■ Call Histories Filled	2
■ Failed Screener	20
Not in Business in 1992	9
Non-Profit Organization	6
Not Headquarters Location	3
Government Owned	1
More than 50% Owned by Another Firm	1

Interview Results	Count
Total Mailouts	111
Completed Interviews	28
■ Proprietorships	15
■ Partnerships	1
■ S-Corporations	6
■ Corporations	6
Unresolved at Conclusion of Pretest	67
Refusals	10
Language Problem	1
Terminate	2
Other Reason for Non-Interview*	3

*These interviews were not able to be conducted for the following reasons: the respondent was on travel for the duration of the pretest, there was a death in the respondent's family, or the business had closed.

because the deadline for completing the pretest arrived before the sample could be fully worked with a rigorous call-back protocol.

The screener interview was conducted over four days. The majority of screeners was resolved in three or fewer calls. Only 22 screeners took more than three calls to resolve. The interviewing sequence of screen, mail, interview was once again tested. However, the procedure this time incorporated many of the concepts developed from the first pretest. The second pretest revealed no major flaw or necessary addition. Overall, the screener performed very well. No new issues were raised during the testing that needed to be addressed.

We conducted the main questionnaire over 14 days. Most interviews were conducted within four calls; however, some required up to eight call-backs. Cases that were not resolved at the time this pretest was concluded had been called between one and twelve times. The interviewers used the approach developed from the first pretest described above to contact respondents after the screening. This worked very well for those respondents who completed an interview. When evaluating the number of calls to complete the interview and the effectiveness of the approach, one should keep in mind that the sample was not fully worked to determine the effect on extremely reluctant respondents. As observed in the first pretest, being able to conduct the interview at the respondent's convenience was a very important aspect in obtaining the interview.

This pretest identified a potential problem regarding the workbook. Several respondents indicated that they would not participate in the study because, in order to complete the workbook, they would have to pay their accountant. During training, the interviewers were made aware of this issue. They were told to explain to these respondents that completing the workbook was not a requirement, and that alternatively, the respondent could use their tax records as a reference.

There were three other main observations from the second pretest:

- First, the interviewer team concept seemed to work very well. A team of 4 to 5 interviewers worked on the test sample. The interviewers became familiar with each firm and the problems and concerns each firm had regarding participation in the study. Also, the gatekeepers became familiar with the interviewers. This established a rapport and added credibility to the survey.

- Second, we recognized the usefulness of PROMPTS. The prompts gave the interviewers additional credibility when, at their discretion, they provided a definition or further clarified a question.

- Third, respondents frequently asked for worksheet references. This reinforced the point already originated in the first pretest. These references were built into the final version of the CATI questionnaire. At the time of the second pretest, however, they were not available.

The phrase "did the firm" in the part of the questionnaire on liabilities and equity (Section S) generated confusion among some sole proprietors as differentiating personal from business liabilities proved difficult. With sole proprietors, we were interested in any personal liabilities resulting from their business activity. However, when we said "the firm," proprietors only referenced liabilities in the firm's name. This resulted in the addition of the prompt, "if the firm is responsible for 50 percent or more of a liability, then please consider it a business liability and include it in your answers." A comparable prompt was added to the questions on assets (Section R) to resolve similar confusion.

Several questionnaire issues surfaced as a result of the second pretesting. The question seeking the firm's total credit card limit (F6), did not include cards without a specific credit limit (e.g., American Express or Diner's Club). As respondents frequently used cards without fixed credit limits, we eventually allowed a zero response to be recorded through the exception procedure (reference Section VI on Data Collection in this report). Note that zero is not a correct response and does not indicate a zero credit limit, but is solely a placeholder indicating no credit limit. The question seeking the total

amount of principal owed on loans from partners/stockholders (F39), frequently elicited a response of zero. Since this was not a correct response, this information was essential to assist training interviewers for the main survey.

Determining the type of institution or source providing a financial service (H2) was also a minor problem. Respondents commonly did not know how to classify the type of financial institution or source they used. When an interviewer prompted, "is it a...", the respondent generally took the first item, whether it was right or wrong. This resulted in a reordering of the response categories to minimize any potential bias in the answer.

A typical answer for the question, "What was the interest rate on this loan?" (J34) had two decimal places. Initially, the response category only allowed for one decimal place. This question was altered to allow two decimal places due to the high incidence of this response.

A final concern relating to a particular question which asks for the cost of treasury stock (S18). It was decided to accept only non-negative answers in the CATI range check. On the tax forms, treasury stock is enclosed in parentheses, signifying it as a negative number. Occasionally, a respondent would insist upon a negative answer. This problem was handled through interviewer training and contributed to the development of a procedure to handle atypical responses.

Since interviews were conducted via CATI, the second pretest's interview length was a better indication of the main interview's duration than the first pretest. However, both resulted in similar conclusions. The interview length was greater than the planned duration. The average interview length for the second pretest was 51.5 minutes. These results do not include the times required to complete the screener, the times of any suspended interviews, and the times of the more difficult interviews that were unresolved.

There was one specific example where the total time required to complete the interview had a significant impact on the data collection process. One respondent spent an hour filling out the worksheets in an effort to reduce the time necessary for the interview. This particular firm dealt with several financial institutions and sources. The second time through the loop in the part of the questionnaire on relationships with financial institutions, Section H, already 60 minutes into the interview, the respondent wondered if the same questions were to be asked of every institution or source. After ascertaining the loop would be completed six times, the respondent wanted to terminate, saying that they had not even started the worksheet questions and they are already over his allotted 45 minutes for the survey. The interview continued up to the part of the questionnaire just before the most recent loan or credit application (Section J), at which point the respondent refused to continue.

As suggested from the first pretest, Sections F and H still had relatively high average completion times of 6.1 and 5.7 respectively. It was also determined that Section T: Comparison Data, had a disproportionate benefit to cost ratio as its average length was 4.2 minutes. These conclusions led to the elimination of an entire section on comparison year data (Section T[5]) and a major revision to H.

The questionnaire was revised prior to the main study. These improvements were not limited as responses to the pretest observations alone, as the dynamic nature of the questionnaire development process raised several additional issues. The end result was the creation of a comprehensive data collection tool that accomplished most of the goals for the survey.

[5] A few questions were retained for comparison purposes from T.

III. SAMPLE SELECTION

A. Introduction

The 1993 National Survey of Small Business Finances was designed to collect data from a representative sample of small businesses that could be extrapolated, using sampling weights, to represent the population of small businesses operating in 1992. The sample design for the 1993 NSSBF, which provides a representative probability sample of small businesses, was driven in large part by contractual requirements regarding statistical comparisons between Asian, Black, or Hispanic-owned firms and non-minority owned firms and between firms with 50-99 employees or 100-499 employees and firms with less than 50 employees. These requirements led to a target sample of 6,000 completed interviews, including at least 400 Asian-owned firms, 400 black-owned firms, and 400 Hispanic-owned firms, and including at least 500 firms with 20-49 employees, 500 firms with 50-99 employees, and 500 firms with 100-499 employees.

The requirement regarding comparisons of minority-owned firms with non-minority firms necessitated the oversampling of minority-owned firms, which, in turn, required some method for identifying minority-owned firms prior to conducting the interview. (The information necessary to identify minority-owned firms in the sample frame was not available.) One way to accomplish this is to screen a sample of sufficient size to contain the required numbers of minority-owned firms. Based upon the incidence of minority-owned firms identified by the 1987 survey, we estimated that we would have to screen approximately 50,000 firms to identify 400 Asian-, 400 Black-, and 400 Hispanic-owned firms. The cost of screening made this approach impractical, and we therefore chose an alternative approach which made use of external information to increase the likelihood of identifying minority firms (see Section D). While this method allowed us to minimize screening costs, it had the disadvantage of increasing the design effects from

stratification, which reduce our ability to make statistical comparisons between groups of firms.

The remainder of this section describes the steps involved in designing and selecting the sample.

B. Definition of the Target Population

The target population is small business enterprises operating under the current ownership during 1992 and with fewer than 500 full-time equivalent employees, but excluding agricultural enterprises, financial institutions, not-for-profit institutions, government entities, and subsidiaries controlled by other corporations. Except for the year, this parallels the population of interest defined in the 1987 NSSBF. An enterprise was defined as a business organization consisting of all establishments under common ownership or control. Business enterprise was the unit of analysis because financial and operating decisions are made at the firm (enterprise) level rather than at the location (establishment) level. For the purposes of this study, we defined a small business as a firm which employs less than 500 full-time equivalent employees. The survey defines the number of full-time equivalent employees as the number of full-time employees plus one-half the number of part-time employees. This differs from Dun & Bradstreet's employment definition as it assigns full-time and part-time employees equal weights of one.

This study excludes all firms in the financial services or agricultural industries because the nature of these industries is such that their use of financial services varies greatly from other small businesses. Subsidiary firms (firms more than fifty percent owned by another company) were also excluded because they do not have absolute control in their financial decision making. In addition, the "parent" company can usually utilize their own resources to ensure the availability of credit to their subsidiary.

C. Choosing a Frame for NSSBF Sampling

For this study, Dun's Market Identifiers (DMI) file was used to construct the sampling frame. While other potential sources such as IRS records were considered, no more comprehensive source was available in a timely fashion. Since the 1987 NSSBF also used the DMI file to construct its frame, there is some consistency across the two surveys, although the DMI file itself changed significantly from 1987 to 1993 (see, for example, Exhibit III-1).

The 1993 DMI file has approximately 9.5 million business establishments. The records are derived from the traditional Dun & Bradstreet credit rating program and business telephone listings. The file is updated monthly based on new publications and new editions of business telephone listings.

The DMI file includes information such as business address, telephone number, and name of owner or top executive, industry (SIC Code), region (state and ZIP code), and size (revenues and employment). In many cases, information about a firm may be missing or out-of-date. This is particularly true for employment and revenues. Additionally, there are gaps in the coverage, particularly among new firms and very small firms.

D. Constructing the Sample Frame

The first step in preparing the sample frame was to drop out-of-scope businesses from the DMI file[6] (e.g., branch offices, subsidiary companies, not-for-profit, financial services, and agricultural industries). As with the previous survey, the following SIC codes are considered out-of-scope and were removed from the frame:

Industry	SIC
Agriculture, Forestry and Fishing	0000-0999
U.S. Postal Service	4311
Finance and Insurance Carriers	6000-6399
Bank Holding Companies	6712
Unit Investment Trusts	6726 [7]
Educational, Religious, & Charitable Trusts	6732
Trusts, Except Educational, Religious, & Charitable	6733
Real Estate Investment Trusts	6798
Membership Organizations	8600-8699
Public Administration	9000-9721

Filtering the 9.5 million establishments on the headquarters and SIC criteria resulted in the current frame of about 7.3 million potentially in-scope business enterprises. After screening sample firms, an additional 2.3 million firms were deemed ineligible, so that the population to which the survey results should be generalized is the remaining 5.0 million small businesses.

[6] We used the November, 1993 DMI file.

[7] Changed from 6724 in previous survey due to change in SIC code.

EXHIBIT III-1: CHANGES IN DMI FRAME FROM PREVIOUS SURVEY					
DMI Frame Categories		**Current Survey***	**Previous Survey***	**Current Minus Previous**	**Percent Difference**
Number of Employees	1 to 4	4,807,653	2,641,509	2,166,144	82.0%
	5 to 19	1,690,988	1,455,073	235,915	16.2%
	20 to 49	345,727	303,827	41,900	13.8%
	50 to 99	116,632	99,107	17,525	17.7%
	100 to 499	73,820	70,736	3,084	4.4%
	500 +	15,189	14,273	916	6.4%
	Unknown	274,841	603,547	(328,706)	-54.5%
	TOTAL	7,324,850	5,188,072	2,136,778	41.2%
MSA Status	Urban (in MSA)	5,698,314	3,952,172	1,746,142	44.2%
	Rural (not in MSA)	1,626,536	1,235,900	390,636	31.6%
	TOTAL	7,324,850	5,188,072	2,136,778	41.2%
Census Region	Northeast	1,589,745	1,140,437	449,308	39.4%
	North Central	1,715,550	1,279,366	436,184	34.1%
	South	2,290,085	1,602,385	687,700	42.9%
	West	1,729,470	1,165,884	563,586	48.3%
	TOTAL	7,324,850	5,188,072	2,136,778	41.2%

*Distribution of numbers after filtering DMI file on headquarters and SIC criteria.

Exhibit III-1 presents a comparison of the current DMI frame to the previous survey frame. The DMI file has grown in all size categories, but primarily in the smaller size categories and particularly in the smallest size category (one to four employees grew by 82 percent). The unknown category is now about half the size of the previous survey. These changes reflect enhanced coverage of the DMI file (particularly for smaller businesses) and improved size information.

In order to enhance survey coverage for minority-owned businesses, we identified businesses on the DMI frame likely to be Black, Asian, and Hispanic owned and partitioned the DMI into four distinct groups -- the three minority groups and a non-minority group.[8] We partitioned each of the three minority groups from the DMI frame into distinct samples while the remaining "non-minority" businesses formed a main partition.

We partitioned the lists because the alternative method of identifying minority-owned businesses -- screening a sample of sufficient size to contain the specified number of minority firms -- was deemed too inefficient and costly because of the low incidence rate of each minority group on the DMI. The minority partitions were developed as follows.

Asian and Hispanic Partitions

The Census Bureau conducts a Survey of Minority-Owned Businesses from which it compiles lists of common surnames of Asian and Hispanic business owners. We acquired these Asian and Hispanic surname lists and provided them to D&B to match to the surname of the "CEO" on the DMI file. This resulted in 254,867 matches for Asian surnames and 245,641 matches for Hispanic surnames. These businesses were then partitioned from the DMI file to create the Asian and Hispanic strata.

Black Partition

The Black business list was compiled from three sources. First, Black Pages directories were collected from around the country. Black Pages directories are

[8] This partition of the DMI file will be called the “main partition” since it includes minority owned businesses not identified prior to survey screening.

independently published business telephone directories (just like Yellow Page directories) that primarily list Black owned businesses.

Exhibit III-2 contains a summary of the Black Pages directories collected and the number of business listings on each. This is a comprehensive list of all actively published Black Pages directories that are members of the National Association of Black Pages.

Second, we used Black Enterprise magazine's "B.E. 100" list, which includes the nation's 100 largest Black Businesses. Many of these businesses have fewer than 500 employees and are complementary to the Black Pages directories, which are believed to contain mostly very small sole proprietorships and partnerships.

Third, we used the Small Business Administration's Procurement Assistant Sourcing System (SBA's PASS) file of applicants for minority status for Federal contracts.

The three sources for the Black lists were then processed for the Black partition. First, Dun & Bradstreet matched the Black Pages directories to the DMI population of 9.5 million establishments using a proprietary field-to-field matching process using the business' name, street address, city, state, zip code, and phone variables. Approximately half (9,692) were successfully matched to the DMI. The SBA's PASS file is a regular D&B product and there are 11,142 PASS businesses on the DMI (i.e., matching was not required). Eighty of the "B.E. 100" were successfully matched to the DMI. The three files were then combined, subjected to the SIC screen to eliminate out-of-scope businesses, and screened for duplicate listings resulting in a total list of 14,704 potential Black owned businesses. Exhibit III-3 illustrates the three sources' disposition and overlap.

EXHIBIT III-2: BLACK PAGES LISTED BUSINESSES BY CENSUS REGIONS			
CENSUS REGION	**CITY**	**COUNT**	**TOTAL**
New England	Brockton, MA	280	
	Dorchester, MA	206	486
Middle Atlantic	New York, NY	647	
	Philadelphia, PA	298	
	Pittsburgh, PA	1,186	2,131
East North Central	Chicago, IL	1,334	
	Detroit, MI	546	
	Cleveland, OH	315	
	Columbus, OH	108	
	Milwaukee, WI	168	2,471
West North Central	Kansas City, KS	781	
	Minneapolis, MN	122	
	St. Louis, MO	683	1,586
South Atlantic	Washington, DC	213	
	Miami, FL	2,712	
	Tallahassee, FL	523	
	Tampa, FL	202	
	Atlanta, GA	791	
	Charlotte, NC	175	
	Charleston, SC	70	
	Columbia, SC	129	
	Lynchburg, VA	24	
	Richmond, VA	351	
	Roanoke, VA	33	5,223
East South Central	Louisville, KY	247	
	Memphis, TN	908	
	Nashville, TN	506	1,661
West South Central	Little Rock, AR	42	
	Baton Rouge, LA	236	
	Dallas, TX	802	
	Houston, TX	805	1,885
Mountain Division	Phoenix, AZ	155	155
Pacific Division	Corona & LA, CA	4,894	4,894
TOTAL			20,492

EXHIBIT III-3: BLACK-OWNED BUSINESS DISPOSITION AND OVERLAP	
MATCHED TO DMI	
Black Pages	9,692
PASS	11,142
B.E. 100	80
	20,914
ELIGIBLE AFTER SIC FILTER	
Black Pages	8,987
PASS	10,331
B.E. 100	74
	19,392
UNIQUE RECORDS	
Black Pages Only	6,643
PASS Only	7,987
B.E. 100	74
	14,704
DUPLICATES	
Black Pages & PASS	4,688

While we attempted to identify other lists of Black-owned businesses, we had little success in doing so. For example, we called Chambers of Commerce and Black Chambers of Commerce in large cities, and numerous other Black organizations that would be potential sources of Black-owned business lists (e.g., Congressional Black Caucus, National Black Chamber of Commerce, etc.). We found very few lists and those that were available were very limited in their coverage (e.g., certain types of businesses in limited geographic areas). We also considered using the subscription list for Black Enterprise magazine. However, the subscription list is large (231,801 active subscribers), the ownership incidence rate is low (about 17 percent), and business owners are not identified as such on the list.

The sample provides relatively inefficient estimates for Black-owned businesses because only a small portion of Black-owned businesses in the frame could be identified as such and segmented into the Black partition. The design effect for Black-owned businesses is 5.4, indicating that the variance of estimates for Black-owned

businesses is 5.4 times as large as the variance of an equal sized simple random sample of Black-owned businesses. In contrast, the design effect for Asian-owned and Hispanic-owned businesses are 1.7 and 1.8, respectively.

A profile of the partitions by Census region, number of employees, MSA status (urban/rural) and industry is found in Appendix A. The next section describes how the sample was selected from the partitions.

E. Sample Selection

The sample was a stratified random sample with the following independent sample strata:

- Main Partition [9]
  - 90 sampling strata defined by:
    - 9 Census Regions
    - Urban/Rural Businesses (2 groups)
    - 5 Size Groups (1-19, 20-49, 50-99, 100-499 employees, and Unknown)
  - One stratum of businesses with more than 500 employees [10]
- Asian Partition
  - Urban/Rural Businesses (2 strata)
- Hispanic Partition
  - Urban/Rural Businesses (2 strata)

[9] Note that there are minority-owned firms in the main as well as the minority partitions.

[10] While businesses with more than 500 full-time equivalent employees are out-of-scope for the survey, differences in the definition of total number of employees (Dun & Bradstreet do not differentiate between full-time and part-time employees whereas the NSSBF counts part-time employees as one-half of a full-time employee) and inaccuracies in the DMI file meant that we could not be sure these businesses were out-of-scope. We therefore sampled a small proportion of these businesses to ensure a representative sample of all small businesses.

- Black Partition
  - Urban/Rural Businesses (2 strata)

Our target number of completed interviews was 6,000 allocated as follows:

- 400 Asian businesses
  400 Black businesses
  400 Hispanic businesses

- 3,250 Non-minority businesses with 1 to 19 employees or an unknown number of employees

- 500 Non-minority businesses with 20 to 49 employees
  500 Non-minority businesses with 50 to 99 employees
  500 Non-minority businesses with 100 to 499 employees

- 50 Non-minority businesses with 500+ employees

Within the Asian, Black, and Hispanic sample strata, the sample was allocated proportionally to urban and rural strata. For the four non-minority size strata that are under 500 employees, the sample was allocated proportionally across Census regions and Urban/Rural strata. There was no further stratification within the "500+ employees" non-minority stratum. Exhibit III-4 shows the distribution of both the DMI frame and the target sample.

EXHIBIT III-4: DMI FRAME SIZE AND TARGET COMPLETED INTERVIEWS

Sample Stratum	DMI Frame Size	Target Completed Interviews	Sampling Rate (1 in)
1-19 employees N. England urban	280,824	145	1937
1-19 employees N. England rural	94,589	49	1930
1-19 employees Mid Atlantic urban	835,491	431	1938
1-19 employees Mid Atlantic rural	81,802	42	1948
1-19 employees E. N. Central urban	735,764	380	1936
1-19 employees E. N. Central rural	226,532	117	1936
1-19 employees W. N. Central urban	258,102	133	1941
1-19 employees W. N. Central rural	237,444	122	1946
1-19 employees S. Atlantic urban	697,797	360	1938
1-19 employees S. Atlantic rural	224,801	116	1938
1-19 employees E. S. Central urban	166,213	86	1933
1-19 employees E. S. Central rural	125,986	65	1938
1-19 employees W. S. Central urban	520,380	269	1934
1-19 employees W. S. Central rural	180,000	93	1935
1-19 employees Mountain urban	226,736	117	1938
1-19 employees Mountain rural	146,905	76	1933
1-19 employees Pacific urban	899,074	464	1938
1-19 employees Pacific rural	109,933	57	1929
20-49 employees N. England urban	16,227	25	649
20-49 employees N. England rural	4,337	7	620
20-49 employees Mid Atlantic urban	47,965	75	640
20-49 employees Mid Atlantic rural	3,699	6	617
20-49 employees E. N. Central urban	47,491	74	642
20-49 employees E. N. Central rural	11,533	17	678
20-49 employees W. N. Central urban	15,418	24	642
20-49 employees W. N. Central rural	9,945	16	622
20-49 employees S. Atlantic urban	39,442	62	636
20-49 employees S. Atlantic rural	11,192	16	700
20-49 employees E. S. Central urban	10,934	17	643
20-49 employees E. S. Central rural	6,328	10	633
20-49 employees W. S. Central urban	23,389	37	632
20-49 employees W. S. Central rural	6,798	11	618
20-49 employees Mountain urban	11,786	18	655
20-49 employees Mountain rural	6,168	10	617
20-49 employees Pacific urban	43,260	68	636
20-49 employees Pacific rural	4,294	7	613
50-99 employees N. England urban	5,250	24	219
50-99 employees N. England rural	1,209	6	202
50-99 employees Mid Atlantic urban	16,037	74	217
50-99 employees Mid Atlantic rural	1,110	5	222
50-99 employees E. N. Central urban	17,135	79	217
50-99 employees E. N. Central rural	3,580	16	224
50-99 employees W. N. Central urban	5,731	26	220
50-99 employees W. N. Central rural	3,212	15	214
50-99 employees S. Atlantic urban	13,333	61	219
50-99 employees S. Atlantic rural	3,522	16	220
50-99 employees E. S. Central urban	3,776	17	222
50-99 employees E. S. Central rural	2,041	9	227
50-99 employees W. S. Central urban	8,118	37	219
50-99 employees W. S. Central rural	2,322	11	211
50-99 employees Mountain urban	4,101	19	216
50-99 employees Mountain rural	1,809	8	226

50-99 employees Pacific urban	15,387	71	217
50-99 employees Pacific rural	1,220	6	203

EXHIBIT III-4: DMI FRAME SIZE AND TARGET COMPLETED INTERVIEWS (cont.)

Sample Stratum	DMI Frame Size	Target Completed Interviews	Sampling Rate (1 in)
100-499 employees N. England urban	3,608	26	139
100-499 employees N. England rural	849	6	142
100-499 employees Mid Atlantic urban	11,020	81	136
100-499 employees Mid Atlantic rural	902	7	129
100-499 employees E. N. Central urban	10,807	78	139
100-499 employees E. N. Central rural	2,639	18	147
100-499 employees W. N. Central urban	3,787	27	140
100-499 employees W. N. Central rural	1,886	14	135
100-499 employees S. Atlantic urban	8,324	60	139
100-499 employees S. Atlantic rural	2,202	15	147
100-499 employees E. S. Central urban	2,403	17	141
100-499 employees E. S. Central rural	1,456	11	132
100-499 employees W. S. Central urban	5,182	37	140
100-499 employees W. S. Central rural	1,320	10	132
100-499 employees Mountain urban	2,371	17	139
100-499 employees Mountain rural	929	7	133
100-499 employees Pacific urban	8,853	64	138
100-499 employees Pacific rural	685	5	137
Unknown employees N. England urban	14,446	7	2064
Unknown employees N. England rural	3,487	2	1744
Unknown employees Mid Atlantic urban	48,648	26	1871
Unknown employees Mid Atlantic rural	3,325	2	1663
Unknown employees E. N. Central urban	36,591	19	1926
Unknown employees E. N. Central rural	7,093	3	2364
Unknown employees W. N. Central urban	8,229	4	2057
Unknown employees W. N. Central rural	3,962	2	1981
Unknown employees S. Atlantic urban	32,753	17	1927
Unknown employees S. Atlantic rural	7,295	3	2432
Unknown employees E. S. Central urban	5,764	3	1921
Unknown employees E. S. Central rural	3,164	2	1582
Unknown employees W. S. Central urban	16,711	9	1857
Unknown employees W. S. Central rural	3,525	2	1763
Unknown employees Mountain urban	10,242	5	2048
Unknown employees Mountain rural	4,551	2	2276
Unknown employees Pacific urban	35,621	18	1979
Unknown employees Pacific rural	3,118	2	1559
500 or more employees	14,418	50	288
Asian urban	225,995	355	637
Asian rural	28,872	45	642
Black urban	14,215	387	37
Black rural	489	13	38
Hispanic urban	215,012	350	614
Hispanic rural	30,629	50	613
Total	7,324,850	6,000	--

We sorted the sampling frame for each stratum by SIC and selected a systematic (every nth) random sample within each stratum. The SIC sort provides an implicit stratification by SIC and thus ensures that the sample includes a mix of firms across the full range of SIC codes.

We considered response and eligibility rates to determine the actual number of records selected in the sample. We assumed a 75 percent response rate and a 65 percent eligibility rate. Combining these two rates, we "expected" to need a sample of approximately 12,000 records to achieve the 6,000 interviews across all four partitions. We selected the sample from the DMI frame using systematic random sampling (taking the n^{th} record from each strata). We set the sample selection interval to the number of records on the frame divided by the number of records needed and rounded the selection interval downward for each strata. The actual number randomly selected was 24,516. (See Exhibit III-12 for the actual amount of sample selected by strata.)

Realizing that we would not need to use all 24,516 records to achieve our target number of interviews, we separated the sample into the four partitions and then, for each partition, selected alternating records and placed them into two distinct groups. The resulting sets of two groups are systematic random samples. In order to facilitate sample control as the survey was administered we divided the sample into replicates. Separately for each of the four partitions, we formed the replicates of 25 records each by randomly sorting the first group and placing the first 25 in replicate 1, second 25 in replicate 2, etc.

Since the eligibility rate was initially unknown, we released the drawn sample in waves for each of the four partitions. Results from the first wave of 12,375 random cases provided the rate used to calculate the additional sample to yield the desired number of

eligible firms. The first wave consisted of 396 replicates from the main partition and 33 replicates from each of the minority partitions. The number of replicates to include in the first wave was based on the target number of completes from each partition and an initial estimate of 75 percent for response rate and an initial estimate of 65 percent for eligibility rate. For the main partition, the estimated required sample was calculated to be 9,846 businesses (i.e., 4,800 / (.65 * .75)). This estimate was rounded up to 9,900 businesses. For each minority partition, the estimated required sample was calculated to be 820 businesses (i.e., 400 / (.65 * .75)). This estimate was rounded up to 825 businesses. The sample that was not used for during the first wave was reserved for later use if it appeared we would not meet the target number of completed interviews as discussed above.

Ultimately, the survey's development required the release of second and third waves. The second wave added 769 Asian and 350 Hispanic random records. The third wave included 2,220 random firms from the main sample partition. The aggregate screening sample of 15,714 resulted in a sufficient number of eligible firms.

Approximately halfway through the survey, we found that we were falling short of sample in the Asian, Hispanic, and main sample. That is, given the response and eligibility rates to that point, our projected number of completed interviews with eligible businesses would fall short of our targets. Our interim response and eligibility rate projections, presented in Exhibit III-5, took into account the status of businesses still in the interview process. Interim response rates and eligibility rates were calculated as of September 1, 1994 for the main sample partition and August 4, 1994 for the minority partitions. The interim response rate was estimated using the number of actual completed interviews to date plus the number of completed interviews expected from the remaining unresolved sample. The expected number of interviews from businesses still in the interview process was calculated by applying estimated response rates to the unresolved sample based on the number of calls made (i.e., for each business still in

the interview process, the greater the number of calls already attempted, the less the chance to obtain a completed interview). The interim eligibility rates are the eligibility rates obtained from the screener questionnaire.

EXHIBIT III-5: INTERIM RESPONSE, ELIGIBILITY, AND HIT RATES

Sample Partition	Interim Response Rate*	Interim Eligibility Rate**	Hit Rate***
Main	56%	69%	N/A
Black	70%	78%	93%
Asian	48%	50%	73%
Hispanic	65%	53%	75%

*The interim response rate was estimated using the number of actual completed interviews to date plus the number of completed interviews expected from the remaining unresolved sample.
**The interim eligibility rates are the eligibility rates obtained from the screener questionnaire.
***The hit rate for each sample partition is the percent of sample in that partition verified through the screener interview to be correctly assigned to that partition.

For the Asian and Hispanic businesses, we were falling short due to lower than expected "hit" rates. The hit rate for each sample partition is the percent of sample in that partition verified through the screener interview to be correctly assigned to that partition (e.g., the percent of sample in the Black partition that are found to be Black owned businesses when screened). The "hit" rate for the Black partition was 93%. However, for the Asian and Hispanic partitions, it was 73% and 75% respectively.

For the main sample partition of primarily non-minority businesses, the problem was attributed to weaknesses in the "number of employees" field available from the DMI file. The DMI file counts both full-time and part-time employees as one for purposes of developing the number of employees. For the NSSBF, part-time employees are counted as one half and full-time employees as one. The frame strata is the initial size strata assignment given to each sampled business based on DMI frame data. The analysis group is the size strata assigned to each business after we determined size

information from the screener survey. Many of the businesses in the larger size frame strata were actually in smaller analysis group size strata. Exhibit III-6 presents the crossover rates for completed interviews as of 9/1/1994. The crossover rate describes the percent of sample that is in one sampling strata and another analysis group. For example, of the sample originally allocated to the 20 to 49 size stratum, 42.5% ended up in the 1 to 19 stratum and only 50.0% in the 20 to 49 stratum.

EXHIBIT III-6: MAIN SAMPLE CROSSOVER RATES* AS OF 9/1/94

Frame Size Strata	Analysis Group**					Total
	Unknown	1 to 19	20 to 49	50 to 99	100 to 499	
Unknown	0.0%	87.8%	6.1%	4.1%	2.0%	100%
1 to 19	0.2%	97.4%	1.6%	0.6%	0.3%	100%
20 to 49	0.8%	42.5%	50.0%	5.0%	1.7%	100%
50 to 99	0.4%	15.8%	21.5%	53.4%	9.0%	100%
100 to 499	0.7%	14.9%	4.4%	12.7%	67.3%	100%
500+	0.0%	50.0%	50.0%	0.0%	0.0%	100%

*The crossover rate describes the percent of sample that are in one sampling strata and another analysis group.
**Analysis Group is the number of employees determined after survey response and using the NSSBF definition of counting part-time employees as one half.

Exhibit III-7 presents interim response and eligibility rates as of 9/1/94. To improve sampling efficiency, we added random sample strategically to those strata where the estimate indicated a projected shortfall of completes, rather than simply adding additional random replicates of 25. The amount of additional main sample necessary to achieve the target completes was estimated in the following manner. First, an iterative proportional fitting was used to determine the number of firms needed in each frame stratum to achieve the target sample size for analysis group. Then, the additional number of completes needed per size strata was divided by the interim response rates and eligibility rates to determine the amount of sample required.

EXHIBIT III-7: MAIN SAMPLE
INTERIM RESPONSE RATES AND ELIGIBILITY RATES
9/1/94

Frame Size Strata	Interim Response Rate	Interim Eligibility Rate
Unknown	65.3%	66.7%
1 to 19	62.0%	70.9%
20 to 49	59.7%	73.6%
50 to 99	57.6%	65.5%
100 to 499	60.8%	58.1%
500+	N/A	N/A

Exhibits III-8 through III-11 summarize the sample's final allocation. Exhibit III-8 shows the sample allocation by frame partition, Exhibit III-9 by employment size, Exhibit III-10 by Census region, and Exhibit III-11 by urban/rural location. Exhibit III-12 shows sample allocation by each of the 97 sample strata.

EXHIBIT III-8: SAMPLE ALLOCATION BY FRAME PARTITION

Sample Partition	DMI Frame	Sample Selected From DMI		Actual Sample Used
		Expected Sample Required	Sample Held in Reserve	
Main	6,809,638	9,900	9,589	12,120
Black	14,704	825	1,013	825
Asian	254,867	825	769	1,594
Hispanic	245,641	825	770	1,175
Total	7,324,850	12,375	12,141	15,714

EXHIBIT III-9: SAMPLE ALLOCATION BY FRAME SIZE

Sample Partition	DMI Frame	Sample Selected From DMI		Actual Sample Used
		Expected Sample Required	Sample Held in Reserve	

Unknown	274,841	377	366	480
1 to 19	6,498,641	8,444	8,412	10,254
20 to 49	345,727	1,190	1,098	1,341
50 to 99	116,632	1,087	1,028	1,830
100 to 499	73,820	1,073	1,031	1,601
500+	15,189	204	206	208
Total	7,324,850	12,375	12,141	15,714

EXHIBIT III-10: SAMPLE ALLOCATION BY FRAME GEOGRAPHY

Sample Partition	DMI Frame	Sample Selected From DMI		Actual Sample Used
		Expected Sample Required	Sample Held in Reserve	
East North Central	1,151,837	1,982	1,934	2,468
East South Central	340,245	595	609	734
Middle Atlantic	1,138,819	1,929	1,881	2,488
Mountain	445,765	663	699	852
New England	450,926	753	672	920
Pacific	1,283,705	2,179	2,189	2,905
South Atlantic	1,118,380	2,008	1,943	2,502
West North Central	563,713	894	871	1,096
West South Central	831,460	1,372	1,343	1,749
Total	7,324,850	12,375	12,141	15,714

EXHIBIT III-11: SAMPLE ALLOCATION BY URBAN/RURAL CLASSIFICATION

Sample Partition	DMI Frame	Sample Selected From DMI		Actual Sample Used
		Expected Sample Required	Sample Held in Reserve	

Urban	5,698,314	10,024	9,812	12,795
Rural	1,626,536	2,351	2,329	2,919
Total	7,324,850	12,375	12,141	15,714

EXHIBIT III-12: SAMPLE ALLOCATION BY STRATUM

Sample Stratum	Expected Sample Required	Sample Held in Reserve	Reserve Sample Used	Total Sample Used
1-19 employees N. England urban	306	275	35	341
1-19 employees N. England rural	94	102	14	108
1-19 employees Mid Atlantic urban	847	880	127	974
1-19 employees Mid Atlantic rural	83	86	8	91
1-19 employees E. N. Central urban	760	761	103	863
1-19 employees E. N. Central rural	235	233	26	261
1-19 employees W. N. Central urban	288	246	33	321
1-19 employees W. N. Central rural	240	251	36	276
1-19 employees S. Atlantic urban	743	699	80	823
1-19 employees S. Atlantic rural	257	208	19	276
1-19 employees E. S. Central urban	163	181	18	181
1-19 employees E. S. Central rural	135	126	19	154
1-19 employees W. S. Central urban	550	526	72	622
1-19 employees W. S. Central rural	181	191	34	215
1-19 employees Mountain urban	234	235	40	274
1-19 employees Mountain rural	149	155	23	172
1-19 employees Pacific urban	936	922	133	1069
1-19 employees Pacific rural	98	129	20	118
20-49 employees N. England urban	59	43	1	60
20-49 employees N. England rural	13	14	0	13
20-49 employees Mid Atlantic urban	166	134	16	182
20-49 employees Mid Atlantic rural	15	8	1	16
20-49 employees E. N. Central urban	157	140	11	168
20-49 employees E. N. Central rural	31	41	4	35
20-49 employees W. N. Central urban	43	54	6	49
20-49 employees W. N. Central rural	24	38	4	28
20-49 employees S. Atlantic urban	127	120	11	138
20-49 employees S. Atlantic rural	43	27	3	46
20-49 employees E. S. Central urban	33	35	3	36
20-49 employees E. S. Central rural	24	16	2	26
20-49 employees W. S. Central urban	80	66	9	89
20-49 employees W. S. Central rural	19	24	4	23
20-49 employees Mountain urban	35	39	6	41
20-49 employees Mountain rural	15	24	3	18
20-49 employees Pacific urban	140	130	15	155
20-49 employees Pacific rural	18	9	1	19
50-99 employees N. England urban	52	45	35	87
50-99 employees N. England rural	15	8	7	22
50-99 employees Mid Atlantic urban	161	136	107	268
50-99 employees Mid Atlantic rural	9	12	10	19
50-99 employees E. N. Central urban	170	148	104	274
50-99 employees E. N. Central rural	40	27	22	62
50-99 employees W. N. Central urban	54	52	35	89
50-99 employees W. N. Central rural	38	22	15	53
50-99 employees S. Atlantic urban	111	136	109	220
50-99 employees S. Atlantic rural	20	46	31	51
50-99 employees E. S. Central urban	32	38	28	60
50-99 employees E. S. Central rural	23	15	11	34
50-99 employees W. S. Central urban	75	76	55	130
50-99 employees W. S. Central rural	28	15	12	40
50-99 employees Mountain urban	39	37	23	62
50-99 employees Mountain rural	18	16	12	30
50-99 employees Pacific urban	149	136	97	246
50-99 employees Pacific rural	14	9	7	21

EXHIBIT III-12: SAMPLE ALLOCATION BY STRATUM (cont.)

Sample Stratum	Expected Sample Required	Sample Held in Reserve	Reserve Sample Used	Total Sample Used
100-499 employees N. England urban	57	49	24	81
100-499 employees N. England rural	15	10	5	20
100-499 employees Mid Atlantic urban	158	166	91	249
100-499 employees Mid Atlantic rural	13	14	4	17
100-499 employees E. N. Central urban	166	152	77	243
100-499 employees E. N. Central rural	34	44	23	57
100-499 employees W. N. Central urban	61	50	28	89
100-499 employees W. N. Central rural	32	24	11	43
100-499 employees S. Atlantic urban	135	110	64	199
100-499 employees S. Atlantic rural	33	32	18	51
100-499 employees E. S. Central urban	33	38	18	51
100-499 employees E. S. Central rural	23	20	10	33
100-499 employees W. S. Central urban	81	72	43	124
100-499 employees W. S. Central rural	18	21	10	28
100-499 employees Mountain urban	27	43	20	47
100-499 employees Mountain rural	10	18	5	15
100-499 employees Pacific urban	130	131	64	194
100-499 employees Pacific rural	10	10	5	15
Unknown employees N. England urban	14	16	2	16
Unknown employees N. England rural	3	4	1	4
Unknown employees Mid Atlantic urban	56	45	6	62
Unknown employees Mid Atlantic rural	5	2	1	6
Unknown employees E. N. Central urban	37	39	9	46
Unknown employees E. N. Central rural	10	5	0	10
Unknown employees W. N. Central urban	12	5	0	12
Unknown employees W. N. Central rural	6	2	0	6
Unknown employees S. Atlantic urban	37	31	5	42
Unknown employees S. Atlantic rural	8	8	4	12
Unknown employees E. S. Central urban	4	8	1	5
Unknown employees E. S. Central rural	4	3	0	4
Unknown employees W. S. Central urban	16	19	4	20
Unknown employees W. S. Central rural	4	3	1	5
Unknown employees Mountain urban	8	13	3	11
Unknown employees Mountain rural	4	5	0	4
Unknown employees Pacific urban	41	33	3	44
Unknown employees Pacific rural	4	3	0	4
500 or more employees	202	199	0	202
Asian urban	730	683	683	1,413
Asian rural	95	86	86	181
Black urban	798	979	0	798
Black rural	27	34	0	27
Hispanic urban	725	671	314	1039
Hispanic rural	100	99	36	136
Total	12,375	12,141	3,339	15,714

F. Sample Selection Probabilities

We used stratified random sampling to select an initial sample to be screened for eligibility. During the screening process, if we found a non-minority business had been included on a minority frame, we had the computer randomly decide whether or not the business would be passed on to the interview process. This random subsampling was done to ensure that sampling rates for non-minority businesses would be similar regardless of which frame they were sampled from (sampling rates for the minority partition frames were higher than those for the main frame). The subsampling rates were 1 in 30 for non-minorities on the Black frame and 1 in 3 for non-minorities on the Asian and Hispanic frames. Specifically, when a non-minority firm was identified in the Black sample partition, it was selected at a rate of 1 in 30 since firms were sampled from the Black frame at a rate 30 times higher than firms from the non-minority frame. When a non-minority firm was identified in the Asian or Hispanic sample partition, it was selected at a rate of 1 in 3 since firms were sampled from the Asian and Hispanic frames at a rate 3 times higher than were firms from the non-minority frame. In summary, the probability that a business was selected in our sample to be interviewed was:

$$P(\text{Sampled}) = (n_h/N_h) \times (\text{subsampling rate})$$

where

n_h = stratum h sample size (total sample used)

N_h = stratum h population count from D&B frame

subsampling rate =	1/30	for non-minorities on the Black frame
	1/3	for non-minorities on the Asian and Hispanic frames
	1	for all other businesses

We refer to the inverse of the sample selection probability as the "design weight" because it is the weight attributable to sample design selection probabilities. Exhibit III-13 summarizes the median and ranges of the survey design weights. We collapsed the "Unknown" main strata with the "1 to 19" main strata to calculate the design weights because the two strata were sampled at the same rate.

EXHIBIT III-13: DESIGN WEIGHTS

Frame/Group		Design Weight Range	Design Weight Median
Non-minority Frame	1 to 19 and Unknown Strata	855 - 1066	910
	20 to 49 Strata	226 - 382	283
	50 to 99 Strata	55 - 70	63
	100 to 499 Strata	42 - 53	45
	500+Strata	72 - 72	72
Black Frame	Minorities	18 - 18	18
	Non-minorities (subsampled)	275 - 275	275
Asian Frame	Minorities	166 - 168	166
	Non-minorities (subsampled)	543 - 622	543
Hispanic Frame	Minorities	216 - 247	216
	Non-minorities (subsampled)	628 - 823	628

IV. SURVEY DISPOSITION AND WEIGHTING

A. Survey Disposition

Businesses sampled to participate in the NSSBF were first administered a screening questionnaire to determine their eligibility and then administered the interview or "main" questionnaire. Exhibit IV-1 summarizes the disposition of the sample.

The top part of Exhibit IV-1 shows that 15,714 businesses were administered the screening questionnaire to assess eligibility and verify contact information. Of these businesses:

- 3,925 businesses were found to be ineligible for the survey (box “A”). The most common reasons for ineligibility were: out of business (34 percent), non-profit (32 percent), branch office (12 percent), or began business after 1992 (12 percent). [11]
- 10,533 businesses were either confirmed eligible or were hard refusals to the screener (box “B”). Most of these businesses (10,141 of the 10,533) were passed on to the main questionnaire stage. The 392 businesses not passed (box “D”) were removed from the sample. They were randomly subsampled from 554 eligible, non-minority businesses found on the minority business frames.
- 1,256 businesses (box “C”) could not be reached after an exhaustive search that included contacting Directory Assistance for the relevant area code and area codes in the near vicinity, the local Chamber of Commerce, the area library reference desk, and the Secretary of State.

[11] Because firms were eliminated sequentially by reason, these numbers are biased downward for all reasons except for branch office, which was asked first.

EXHIBIT IV-1: MICROGRAPHICS FLOWCHART SAMPLE DISPOSITION

The bottom part of Exhibit IV-1 presents the disposition of the 10,141 sampled businesses that we passed on to the main questionnaire:

- 650 businesses were determined to be ineligible for the survey (box “E”)
- 5,356 businesses completed the main questionnaire interview (box “F”)
- 4,135 businesses did not respond to the main survey (box “G”)

B. Cooperation Patterns

We examined patterns in the percentage of businesses that cooperated in the main questionnaire interview process to determine factors that affected response rates. The numerator in the percentage included businesses that completed the main questionnaire (box “F” in Exhibit IV-1) and businesses that cooperated in the main interview process to the extent necessary to determine that the business was ineligible for the survey (box “E” in Exhibit IV-1). The denominator in the percentage included these businesses and survey non-respondents (boxes “C” and “G” in Exhibit IV-1). We refer to this percentage as the “cooperation rate” for the main questionnaire:

$$Cooperation\ Rate = \frac{Cooperated\ With\ Main\ Survey\ (Boxes\ E+F)}{Cooperated + Nonrespondents\ (Boxes\ C+E+F+G)}$$

We use the cooperation rates with the design weights to calculate sampling weights. We calculate sampling weights using this method rather than the more generally used non-response adjustment for two reasons. First, the eligibility of non-respondents, which is needed to calculate a response rate, was unknown (the response rate is the percentage of eligible businesses that complete the interview). Therefore, we would be

forced to make some assumptions in order to estimate a response rate. Second, we had more detailed frame data on size, location, and industry that could be used to adjust for nonresponse by applying a proportional fitting procedure known as raking (discussed in more detail in Section C).

The cooperation rate differs from a response rate and was computed because the eligibility of non-respondents, which is needed to calculate a response rate, was unknown (the response rate is the percentage of eligible businesses that complete the interview). Later in this section we use sampling weights to estimate eligibility and response rates for the survey.

Exhibits IV-2 through IV-6 summarize cooperation rates broken down by the five variables available on the survey frame:

- Sample Partition (Exhibit IV-2). Cooperation rates were lowest for the Asian partition and highest for the Black partition.
- Frame Total Employment (Exhibit IV-3). Larger businesses were more likely to cooperate.
- Geography (Exhibit IV-4). Cooperation was lowest in New York and New Jersey and highest in Maine, New Hampshire, and Vermont.
- Urban/Rural Indicator (Exhibit IV-5). Urban businesses were less likely to cooperate than rural businesses.
- Industry (Exhibit IV-6). Cooperation was lowest for Retail Trade and highest for Insurance/Real Estate.

EXHIBIT IV-2: COOPERATION RATES BY SAMPLE PARTITION		
Sample Partition	**Count**	**Cooperation Rate**
Main	8,854	52.9%
Black	681	65.2%
Asian	1,058	42.1%
Hispanic	804	54.2%
Subtotal	11,397	N/A
Screener Confirmed Ineligible	3,925	N/A
Subsample Removed	392	N/A
Total	15,714	N/A

EXHIBIT IV-3: COOPERATION RATES BY FRAME TOTAL EMPLOYMENT		
Total Employees per Frame	**Count**	**Cooperation Rate**
Unknown and 0	**351**	**41.0%**
Unknown	351	41.0%
0	0	N./A.
1 to 5	**6,008**	**49.1%**
1	1,208	51.3%
2	1,331	52.4%
3	2,034	44.2%
4	902	49.7%
5	533	53.8%
6 to 249	**4,746**	**57.5%**
6 to 9	950	57.7%
10 to 19	730	56.6%
20 to 49	1,003	58.7%
50 to 99	1,238	57.8%
100 to 249	825	53.8%
250 to 499	**245**	**59.2%**
500 or more	**47**	**74.5%**
Subtotal	11,397	N/A
Confirmed Ineligible	3,925	N/A
Subsample Removed	392	N/A
Total	15,714	N/A

EXHIBIT IV-4: COOPERATION RATES BY GEOGRAPHY		
Level/States	**Count**	**Cooperation Rate**
Low	**1,373**	**42.5%**
New Jersey	424	42.5%
New York	949	42.6%
Medium Low	**2,222**	**47.2%**
California	1,712	46.6%
MA,RI,CT	510	49.2%
Medium	**4,526**	**54.6%**
Illinois	546	51.3%
Pennsylvania	502	53.0%
Michigan	395	54.7%
SC, NC, GA, FL	1,225	55.2%
Pacific excluding CA	439	56.0%
OH,IN,WI	826	55.8%
DE,MD,DC,VA,WV	593	55.1%
High	**3,276**	**58.1%**
Texas	866	54.2%
Mountain	589	56.5%
AR,OK,LA	380	57.6%
East South Central	528	60.6%
ME, NH, VT	149	62.4%
West North Central	764	61.3%
Subtotal	11,397	N/A
Confirmed Ineligible	3,925	N/A
Subsample Removed	392	N/A
Total	15,714	N/A

EXHIBIT IV-5: COOPERATION RATES BY URBAN/RURAL CLASSIFICATION		
Classification	**Count**	**Cooperation Rate**
Urban	9,485	51.1%
Rural	1,912	60.5%
Subtotal	11,397	N/A
Confirmed Ineligible	3,925	N/A
Subsample Removed	392	N/A
Total	15,714	N/A

EXHIBIT IV-6: COOPERATION RATES BY INDUSTRY		
Level/SIC	**Count**	**Cooperation Rate**
Low	**8,070**	**50.7%**
Mining/Construction (1000-1999)	1,389	51.7%
Retail Trade (5200-5999)	2,689	49.6%
Services (7000-7999)	2,232	50.7%
Services (8000-8999)	1,760	51.4%
Medium	**2,675**	**56.8%**
Manufacturing (2000-2999)	568	57.9%
Manufacturing (3000-3999)	685	55.8%
Trans/Comm/Util (4000-4999)	442	57.0%
Wholesale Trade (5000-5199)	980	56.8%
High	**652**	**61.0%**
Insurance/Real Estate (6000-6999)	652	61.0%
Subtotal	11,397	N/A
Confirmed Ineligible	3,925	N/A
Subsample Removed	392	N/A
Total	15,714	N/A

We tested whether differences in cooperation rates were statistically significant using Analysis of Variance (ANOVA) both before and after grouping the detailed categories (shown in regular type) in Exhibits IV-2 through IV-6 into the more aggregate categories (shown in **boldface** type). Due at least in part to the large sample size, cooperation rate differences were significant at the 0.0001 level for each of the variables individually, and at the 0.005 level for each variable when all of the variables were considered simultaneously in one ANOVA. We compared the ANOVA which included all of the aggregate groupings to one which included all of the detailed groupings and found that the sum of squared errors explained by the two models were within five percent of each other. This indicated that the aggregate groupings explained nearly all of the differences in cooperation rates that could be explained by the more detailed groupings. We therefore used the more aggregate groupings to develop sampling weights.

C. Final Weights

The final sampling weights reflect the sample design, the cooperation rates detailed in the previous section, and the implied eligibility patterns. The formula used to compute the final weights is:

Final Sampling Weight = Design Weight × Cooperation Weight.

The design weights are detailed in Section III. The cooperation weights are our version of a response/eligibility adjustment and are applied to the 6,006 “cooperators”—firms in boxes E (confirmed ineligibles) and F (confirmed eligibles) in Exhibit IV-1—in order to represent the 11,397 firms in boxes E, F, G (nonrespondents to main questionnaire), and C (nonrespondents to screener, confirmed unreachable). The 3,925 firms in Box A (screener confirmed ineligibles) are not included since they were confirmed ineligible during the screening interview.

From the analyses of variance described in the previous section, we determined that cooperation rates are affected by minority partition, urban/rural location, geographic region, number of employees, and industry. Thus, it is desirable to account for all of these categories in the computation of the final weights. In order to do this, we used an iterative proportional fitting technique (sometimes referred to in the literature as table standardization or raking-ratio estimation) to calculate the cooperation weights.[12] With this technique, the weights are iteratively computed so that they sum to the totals in the adjustment categories. We used the aggregated categories that are shown in bold in Exhibits IV-2 to IV-6.[13] After three iterations of the program, the following held true for totals of sampled businesses in each category of each of the five variables:

$$\sum_{\substack{i\,Cooperated \\ (Boxes\,E,F)}} (Cooperation\ Weight_i \times Design\ Weight_i) = \sum_{\substack{i\,(Cooperated+Nonrespondents) \\ (Boxes\,C,E,F,G)}} Design\ Weight_i$$

The cooperation weights ranged from 1.3 to 2.7. The average design weights ranged from 18 to 927 and the average final weights ranged from 24 to 2,391 as shown by stratum in Exhibit IV-7. The product of the cooperation weight and the design weight is the sampling weight used to extrapolate the data from eligible survey respondents (box "F" in Exhibit IV-1) to the target population of small businesses.

EXHIBIT IV-7: AVERAGE DESIGN, COOPERATION, AND FINAL WEIGHTS BY STRATUM

Sample Stratum	Average Design Weight	Average Cooperation Weight	Average Final Sampling Weight
1-19 employees N. England urban	827	2.21	1,824
1-19 employees N. England rural	876	1.77	1,551
1-19 employees Mid Atlantic urban	853	2.34	2,000
1-19 employees Mid Atlantic rural	878	1.83	1,609
1-19 employees E. N. Central urban	850	1.96	1,662

[12] Agresti, Alan. *Categorical Data Analysis*, John Wiley & Sons Inc., 1990, Section 6.7, p. 196. (describes Table Standardization).

[13] The categories used in the computation are: partition, employees per frame (unknown and 0, 1-5, 6-249, 250-499, 500 or more), geography (low, medium low, medium, high), urban/rural, and standard industrial classification (low, medium, high)

1-19 employees E. N. Central rural	862	1.68	1,451
1-19 employees W. N. Central urban	800	1.92	1,537
1-19 employees W. N. Central rural	856	1.64	1,408
1-19 employees S. Atlantic urban	845	1.96	1,653
1-19 employees S. Atlantic rural	806	1.67	1,344
1-19 employees E. S. Central urban	925	1.92	1,778
1-19 employees E. S. Central rural	817	1.64	1,337
1-19 employees W. S. Central urban	837	1.94	1,623
1-19 employees W. S. Central rural	834	1.67	1,395
1-19 employees Mountain urban	832	1.95	1,621
1-19 employees Mountain rural	861	1.69	1,451
1-19 employees Pacific urban	840	2.21	1,857
1-19 employees Pacific rural	927	1.78	1,653
20-49 employees N. England urban	270	1.92	518
20-49 employees N. England rural	334	1.58	528
20-49 employees Mid Atlantic urban	264	2.05	541
20-49 employees Mid Atlantic rural	231	1.74	403
20-49 employees E. N. Central urban	283	1.73	489
20-49 employees E. N. Central rural	330	1.50	496
20-49 employees W. N. Central urban	315	1.70	534
20-49 employees W. N. Central rural	355	1.42	506
20-49 employees S. Atlantic urban	286	1.77	505
20-49 employees S. Atlantic rural	243	1.50	365
20-49 employees E. S. Central urban	304	1.72	524
20-49 employees E. S. Central rural	243	1.45	353
20-49 employees W. S. Central urban	263	1.71	449
20-49 employees W. S. Central rural	296	1.48	438
20-49 employees Mountain urban	287	1.69	486
20-49 employees Mountain rural	343	1.44	494
20-49 employees Pacific urban	279	1.94	540
20-49 employees Pacific rural	226	1.48	334
50-99 employees N. England urban	60	1.96	118
50-99 employees N. England rural	55	1.62	89
50-99 employees Mid Atlantic urban	60	1.97	118
50-99 employees Mid Atlantic rural	58	1.57	92
50-99 employees E. N. Central urban	63	1.74	109
50-99 employees E. N. Central rural	58	1.50	87
50-99 employees W. N. Central urban	64	1.69	109
50-99 employees W. N. Central rural	61	1.47	89
50-99 employees S. Atlantic urban	61	1.75	106
50-99 employees S. Atlantic rural	69	1.51	104
50-99 employees E. S. Central urban	63	1.73	109
50-99 employees E. S. Central rural	60	1.46	88
50-99 employees W. S. Central urban	62	1.71	106
50-99 employees W. S. Central rural	58	1.50	87
50-99 employees Mountain urban	66	1.75	116
50-99 employees Mountain rural	60	1.52	91
50-99 employees Pacific urban	63	1.96	122
50-99 employees Pacific rural	58	1.64	95

EXHIBIT IV-7: AVERAGE DESIGN, COOPERATION, AND FINAL WEIGHTS BY STRATUM (cont.)

	Average Design	Average Cooperation	Average Final

Sample Stratum	Weight	Weight	Sampling Weight
100-499 employees N. England urban	45	1.92	86
100-499 employees N. England rural	42	1.55	66
100-499 employees Mid Atlantic urban	44	2.01	89
100-499 employees Mid Atlantic rural	53	1.49	79
100-499 employees E. N. Central urban	44	1.73	77
100-499 employees E. N. Central rural	46	1.45	67
100-499 employees W. N. Central urban	43	1.68	71
100-499 employees W. N. Central rural	44	1.44	63
100-499 employees S. Atlantic urban	42	1.74	73
100-499 employees S. Atlantic rural	43	1.48	64
100-499 employees E. S. Central urban	47	1.69	80
100-499 employees E. S. Central rural	44	1.40	62
100-499 employees W. S. Central urban	42	1.70	71
100-499 employees W. S. Central rural	47	1.46	69
100-499 employees Mountain urban	50	1.70	86
100-499 employees Mountain rural	54	1.48	80
100-499 employees Pacific urban	46	1.91	87
100-499 employees Pacific rural	54	1.81	97
Unknown employees N. England urban	827	2.42	2,002
Unknown employees N. England rural	*	*	*
Unknown employees Mid Atlantic urban	853	2.69	2,297
Unknown employees Mid Atlantic rural	878	2.72	2,391
Unknown employees E. N. Central urban	850	2.37	2,013
Unknown employees E. N. Central rural	862	2.16	1,863
Unknown employees W. N. Central urban	800	2.42	1,933
Unknown employees W. N. Central rural	856	2.12	1,815
Unknown employees S. Atlantic urban	845	2.32	1,964
Unknown employees S. Atlantic rural	806	2.12	1,709
Unknown employees E. S. Central urban	925	2.24	2,071
Unknown employees E. S. Central rural	817	2.12	1,733
Unknown employees W. S. Central urban	837	2.29	1,913
Unknown employees W. S. Central rural	834	1.89	1,574
Unknown employees Mountain urban	832	2.09	1,734
Unknown employees Mountain rural	861	1.89	1,624
Unknown employees Pacific urban	840	2.59	2,173
Unknown employees Pacific rural	927	2.22	2,061
500 or more employees	71	1.30	93
Asian urban	200	2.30	453
Asian rural	282	1.88	528
Black urban	20	1.50	29
Black rural	18	1.31	24
Hispanic urban	248	1.81	450
Hispanic rural	335	1.47	498

*No respondents in this sampling strata.

D. Eligibility and Response Rates

After calculating final sampling weights, we applied the weights to an eligibility indicator (eligible=1, ineligible=0). The weighted total of the eligibility indicator is the estimate of the number of eligible businesses on the D&B frame.

Exhibit IV-8 presents lower bound, upper bound, and best estimate survey eligibility rates by major sample partition (i.e., main, black, Asian, and Hispanic sample partitions). The 5,748 eligible businesses includes 5,356 completed interviews and 392 businesses in the non-minority subsample not selected for the main questionnaire (boxes "D" and "F" in Exhibit IV-1). The 4,575 ineligible businesses includes 3,925 businesses determined to be ineligible at the screener stage and 650 businesses determined to be ineligible at the main questionnaire stage (boxes "A" and "E" in Exhibit IV-1). The remainder of the sample, 5,391 businesses with unknown eligibility, includes the 1,256 non-respondents to the screener and the 4,135 non-respondents to the main questionnaire (boxes "C" and "G" in Exhibit IV-1).

EXHIBIT IV-8: SURVEY ELIGIBILITY RATE

Sample Partition	Confirmed Eligible	Confirmed Ineligible	Eligibility Unknown	Total	**Eligibility Rate**		
					Lower Bound*	Upper Bound*	Best Estimate*
Main	4,180	3,767	4,173	12,120	34.49%	68.92%	68.05%
Black	457	131	237	825	55.39%	84.12%	82.21%
Asian	594	387	613	1,594	37.26%	75.72%	68.91%
Hispanic	517	290	368	1,175	44.00%	75.32%	70.42%
Total	5,748	4,575	5,391	15,714	36.58%	70.89%	68.19%

*The **lower bound** rates assume all unknowns are ineligible (e.g., Total = 5,748/15,714).
The **upper bound** rates assume all unknowns are eligible (e.g., Total = (5,748 + 5,391)/15,714).
The **best estimate** rates are calculated by summing the final sampling weights for eligible firms and dividing by the sum of the final sampling weights across all cooperating firms.

The lower bound eligibility rates assume all businesses with unknown eligibility (survey non-respondents) are ineligible. The upper bound rates assume all unknowns are eligible. We calculate the best estimate of eligibility rates within each stratum by summing the final sampling weights for eligible businesses (box “F” in Exhibit IV-1), then dividing by the sum of the final sampling weights across all cooperating firms (boxes “A”, “E”, and “F” in Exhibit IV-1).

Exhibit IV-9 details the reasons why businesses were ineligible. It shows that the four most common reasons for ineligibility were confirmed out-of-business (1,614 cases), non-profit firms (1,321 cases), new businesses since 12/31/92 (620 cases), and branch offices (535 cases). These four reasons account for over 89% of the ineligible businesses (4,090/4,575). The corresponding fraction is 88%, 93%, 89%, and 96% within the main, black, Asian, and Hispanic partitions respectively.

Exhibit IV-10 displays lower bound and best estimate survey response rates by major sample partition. The response rate is the number of respondents who completed interviews divided by the number who completed interviews plus eligible non-respondents. The lower bound response rates assume all non-respondents are eligible. The best estimate rates assume a portion of the non-respondents are eligible based on the best estimate of eligibility described earlier. The exhibit divides the non-respondents into estimated eligible and ineligible counts. The response rate calculations exclude the 4,575 ineligible and the 392 subsampled, not selected businesses (boxes "A + E" and "D" in Exhibit IV-1). The calculations account for the 5,356 completed interviews (box "F" in Exhibit IV-1) and the 5,391 non-respondents (1,256 from the screener and 4,135 from the main questionnaire (boxes "C" and "G" in Exhibit IV-1).

EXHIBIT IV-9: SURVEY INELIGIBILITY STATISTICS*

Reasons for Ineligibility	Sample Partition				Total
	Main	Black	Asian	Hispanic	
Confirmed Out of Business	1,265	60	148	141	1,614
Non-Profit	1,211	29	36	45	1,321
New Business since 12/31/92	425	25	118	52	620
Branch Office	446	9	41	39	535
Subsidiary	174	1	25	2	202
500 or More Employees	153	4	3	5	165
Government Owned	54	2	7	5	68
Not a Business	29	1	9	1	40
Ineligible SIC Code	10	0	0	0	10
Total Ineligibles	3,767	131	387	290	4,575
Total Firms	12,120	825	1,594	1,175	15,714

*Because firms were eliminated sequentially by reason, these numbers are biased downward for all reasons except for branch office, which was asked first.

Exhibit IV-11 presents survey non-response statistics by major sample partition. It shows that the three most common reasons for non-response were refusals (2,900 cases), confirmed unreachables (1,351 cases), and other non-responses (676 cases). These three reasons account for over 91% of the non-respondents (4,927/5,391). The corresponding fraction is 94%, 90%, 74%, and 88% within the main, black, Asian, and Hispanic partitions respectively. When we add the 109 language problem cases to the three common reasons within the Asian partition, that ratio increases to 92%.

EXHIBIT IV-10: ESTIMATED SURVEY RESPONSE RATES

Sample Partition	Completed Interview Respondents	Non-Respondents			Total	Estimated Response Rates	
		Estimated Eligible	Estimated Ineligible	Total Non-Respondents		Lower Bound*	Best Estimate*
Main	4,180	2,795	1,378	4,173	8,353	50.04%	60.12%
Black	414	195	42	237	651	63.59%	68.02%
Asian	382	423	190	613	995	38.39%	47.44%
Hispanic	380	259	109	368	748	50.80%	59.45%
TOTAL	5,356	3,672	1,719	5,391	10,747	49.84%	59.48%

*The **lower bound** rates assume all non-respondents are eligible (e.g., Total = 5,356/10,747).
The **best estimate** rates apply the best estimate of eligibility rate to the non-respondent counts (e.g., Total = 5,356/(5,356 + 3,672)).

EXHIBIT IV-11: SURVEY NON-RESPONSE STATISTICS

Reasons for Non-Response	Main	Black	Asian	Hispanic	Total
Refusal	2,380	118	246	156	2,900
Confirmed Unreachable	996	73	165	117	1,351
Other Non-Response	557	23	44	52	676
Terminated	196	22	48	27	293
Language Problem	34	1	109	16	160
Extended Travel	5	0	0	0	5
Illness/Injured	4	0	1	0	5
Other Reason	1	0	0	0	1
Total Non-Response	4,173	237	613	368	5,391
Total Screened	12,120	825	1,594	1,175	15,714

Exhibits IV-12, IV-13, and IV-14 provide further detail on best estimate eligibility and response rates by sampling strata.

EXHIBIT IV-12: SURVEY ELIGIBILITY RATES* BY SAMPLING STRATA

SIZE OF FIRM	Unknown	1 to 19	20 to 49	50 to 99	100 to 499
	%	%	%	%	%
NON-MINORITY URBAN					
East North Central	84.2	70.7	70.8	65.1	65.1
East South Central	52.8	70.2	71.4	75.7	70.6
Middle Atlantic	34.9	70.2	72.1	59.4	54.5
Mountain	34.4	64.0	81.6	67.2	63.2
New England	33.0	78.1	74.6	60.9	53.7
Pacific	71.7	70.3	75.8	62.2	60.7
South Atlantic	68.2	66.2	71.9	64.7	56.8
West North Central	63.8	74.8	75.1	65.6	62.2
West South Central	56.8	65.6	75.7	56.1	62.4
NON-MINORITY RURAL					
East North Central	51.9	66.4	54.6	48.7	50.3
East South Central	68.0	67.6	71.7	61.1	48.2
Middle Atlantic	57.7	78.5	55.5	57.9	59.9
Mountain	100.0	68.8	47.4	36.4	57.5
New England	**	65.0	59.5	79.5	50.9
Pacific	69.0	74.2	59.6	47.2	15.3
South Atlantic	72.8	62.0	76.5	56.7	46.0
West North Central	86.4	61.7	55.0	48.0	36.8
West South Central	79.1	52.7	48.1	51.2	60.5
OTHER					
Non-Minority 500+	5.9				
Black Urban	82.1				
Black Rural	86.2				
Asian Urban	69.3				
Asian Rural	65.7				
Hispanic Urban	70.5				
Hispanic Rural	70.2				

*The best estimate rates are calculated by summing the final sampling weights for eligible firms and dividing by the sum of the final sampling weights across all cooperating firms.
**No respondents in this sampling strata.

EXHIBIT IV-13: SURVEY RESPONSE RATES* BY SAMPLING STRATA

SIZE OF FIRM	Unknown	1 to 19	20 to 49	50 to 99	100 to 499
	%	%	%	%	%
NON-MINORITY URBAN					
East North Central	51.7	57.5	60.5	66.4	66.0
East South Central	48.6	51.2	56.5	73.9	93.2
Middle Atlantic	37.8	47.7	51.3	57.6	58.5
Mountain	59.2	60.2	68.2	74.1	73.2
New England	43.1	54.7	67.2	69.1	57.4
Pacific	48.8	50.5	71.7	62.5	57.4
South Atlantic	50.9	57.7	66.5	66.3	63.4
West North Central	61.0	60.5	63.3	61.1	71.9
West South Central	71.2	55.3	67.3	63.5	72.4
NON-MINORITY RURAL					
East North Central	21.6	67.4	73.3	69.8	79.9
East South Central	42.4	67.1	83.0	86.8	79.2
Middle Atlantic	36.6	64.1	69.3	75.2	79.6
Mountain	50.0	66.6	67.9	79.4	63.5
New England	**	68.1	73.7	75.1	70.2
Pacific	100.0	62.3	81.7	76.1	62.1
South Atlantic	63.2	65.7	69.8	75.3	74.8
West North Central	63.4	72.3	91.6	79.7	75.1
West South Central	55.8	64.8	84.7	79.6	87.8
OTHER					
Non-Minority 500+	92.7				
Black Urban	67.4				
Black Rural	84.6				
Asian Urban	46.9				
Asian Rural	54.2				
Hispanic Urban	58.4				
Hispanic Rural	69.4				

*This exhibit reflects best estimate of response rates calculated by applying the best estimate of eligibility rates to the non-respondent counts.
**No respondents in this sampling strata.

EXHIBIT IV-14: AVERAGE DESIGN WEIGHT, ELIGIBILITY RATE*, AND RESPONSE RATE* BY STRATUM

Sample Stratum	Average Design Weight	Eligibility Rate *	Response Rate*

1-19 employees N. England urban	827	78.1	54.7
1-19 employees N. England rural	876	65.0	68.1
1-19 employees Mid Atlantic urban	853	70.2	47.7
1-19 employees Mid Atlantic rural	878	78.5	64.1
1-19 employees E. N. Central urban	850	70.7	57.5
1-19 employees E. N. Central rural	862	66.4	67.4
1-19 employees W. N. Central urban	800	74.8	60.5
1-19 employees W. N. Central rural	856	61.7	72.3
1-19 employees S. Atlantic urban	845	66.2	57.7
1-19 employees S. Atlantic rural	806	62.0	65.7
1-19 employees E. S. Central urban	925	70.2	51.2
1-19 employees E. S. Central rural	817	67.6	67.1
1-19 employees W. S. Central urban	837	65.6	55.3
1-19 employees W. S. Central rural	834	52.7	64.8
1-19 employees Mountain urban	832	64.0	60.2
1-19 employees Mountain rural	861	68.8	66.6
1-19 employees Pacific urban	840	70.3	50.5
1-19 employees Pacific rural	927	74.2	62.3
20-49 employees N. England urban	270	74.6	67.2
20-49 employees N. England rural	334	59.5	73.7
20-49 employees Mid Atlantic urban	264	72.1	51.3
20-49 employees Mid Atlantic rural	231	55.5	69.3
20-49 employees E. N. Central urban	283	70.8	60.5
20-49 employees E. N. Central rural	330	54.6	73.3
20-49 employees W. N. Central urban	315	75.1	63.3
20-49 employees W. N. Central rural	355	55.0	91.6
20-49 employees S. Atlantic urban	286	71.9	66.5
20-49 employees S. Atlantic rural	243	76.5	69.8
20-49 employees E. S. Central urban	304	71.4	56.5
20-49 employees E. S. Central rural	243	71.7	83.0
20-49 employees W. S. Central urban	263	75.7	67.3
20-49 employees W. S. Central rural	296	48.1	84.7
20-49 employees Mountain urban	287	81.6	68.2
20-49 employees Mountain rural	343	47.4	67.9
20-49 employees Pacific urban	279	75.8	71.7
20-49 employees Pacific rural	226	59.6	81.7
50-99 employees N. England urban	60	60.9	69.1
50-99 employees N. England rural	55	79.5	75.1
50-99 employees Mid Atlantic urban	60	59.4	57.6
50-99 employees Mid Atlantic rural	58	57.9	75.2
50-99 employees E. N. Central urban	63	65.1	66.4
50-99 employees E. N. Central rural	58	48.7	69.8
50-99 employees W. N. Central urban	64	65.6	61.1
50-99 employees W. N. Central rural	61	48.0	79.7
50-99 employees S. Atlantic urban	61	64.7	66.3
50-99 employees S. Atlantic rural	69	56.7	75.3
50-99 employees E. S. Central urban	63	75.7	73.9
50-99 employees E. S. Central rural	60	61.1	86.8
50-99 employees W. S. Central urban	62	56.1	63.5
50-99 employees W. S. Central rural	58	51.2	79.6
50-99 employees Mountain urban	66	67.2	74.1
50-99 employees Mountain rural	60	36.4	79.4
50-99 employees Pacific urban	63	62.2	62.5
50-99 employees Pacific rural	58	47.2	76.1

EXHIBIT IV-14: AVERAGE DESIGN WEIGHT, ELIGIBILITY RATE*, AND RESPONSE RATE* BY STRATUM (cont.)

Sample Stratum	Average Design Weight	Eligibility Rate*	Response Rate*
100-499 employees N. England urban	45	53.7	57.4
100-499 employees N. England rural	42	50.9	70.2
100-499 employees Mid Atlantic urban	44	54.5	58.5
100-499 employees Mid Atlantic rural	53	59.9	79.6
100-499 employees E. N. Central urban	44	65.1	66.0
100-499 employees E. N. Central rural	46	50.3	79.9
100-499 employees W. N. Central urban	43	62.2	71.9
100-499 employees W. N. Central rural	44	60.5	87.8
100-499 employees S. Atlantic urban	42	56.8	63.4
100-499 employees S. Atlantic rural	43	46.0	74.8
100-499 employees E. S. Central urban	47	70.6	93.2
100-499 employees E. S. Central rural	44	48.2	79.2
100-499 employees W. S. Central urban	42	62.4	72.4
100-499 employees W. S. Central rural	47	36.8	75.1
100-499 employees Mountain urban	50	63.2	73.2
100-499 employees Mountain rural	54	57.5	63.5
100-499 employees Pacific urban	46	60.7	57.4
100-499 employees Pacific rural	54	15.3	62.1
Unknown employees N. England urban	827	33.0	43.1
Unknown employees N. England rural	**	**	**
Unknown employees Mid Atlantic urban	853	34.9	37.8
Unknown employees Mid Atlantic rural	878	57.7	36.6
Unknown employees E. N. Central urban	850	84.2	51.7
Unknown employees E. N. Central rural	862	51.9	21.6
Unknown employees W. N. Central urban	800	63.8	61.0
Unknown employees W. N. Central rural	856	86.4	63.4
Unknown employees S. Atlantic urban	845	68.2	50.9
Unknown employees S. Atlantic rural	806	72.8	63.2
Unknown employees E. S. Central urban	925	52.8	48.6
Unknown employees E. S. Central rural	817	68.0	42.4
Unknown employees W. S. Central urban	837	56.8	71.2
Unknown employees W. S. Central rural	834	79.1	55.8
Unknown employees Mountain urban	832	34.4	59.2
Unknown employees Mountain rural	861	100.0	50.0
Unknown employees Pacific urban	840	71.7	48.8
Unknown employees Pacific rural	927	69.0	100.0
500 or more employees	71	5.9	92.7
Asian urban	200	69.3	46.9
Asian rural	282	65.7	54.2
Black urban	20	82.1	67.4
Black rural	18	86.2	84.6
Hispanic urban	248	70.5	58.4
Hispanic rural	335	70.2	69.4

*The contents of this exhibit reflect the best estimate eligibility rates and best estimate response rates.
**No respondents in this sampling strata.

Exhibit IV-15 shows by stratum the DMI frame size and the estimated number of eligible firms on the frame. The estimated number of eligible firms is the population to which the survey results are representative. This number is calculated as the sum by stratum of each sample firm's final sampling weight. Hence, the survey is representative of 5.0 million firms.

EXHIBIT IV-15: DMI FRAME SIZE AND ESTIMATED NUMBER OF ELIGIBLE FIRMS*

Sample Stratum	DMI Frame Size	Number of Eligible Firms*
1-19 employees N. England urban	280,824	219,324
1-19 employees N. England rural	94,589	61,483
1-19 employees Mid Atlantic urban	835,491	586,515
1-19 employees Mid Atlantic rural	81,802	64,215
1-19 employees E. N. Central urban	735,764	520,185
1-19 employees E. N. Central rural	226,532	150,417
1-19 employees W. N. Central urban	258,102	193,060
1-19 employees W. N. Central rural	237,444	146,503
1-19 employees S. Atlantic urban	697,797	461,942
1-19 employees S. Atlantic rural	224,801	139,377
1-19 employees E. S. Central urban	166,213	116,682
1-19 employees E. S. Central rural	125,986	85,167
1-19 employees W. S. Central urban	520,380	341,369
1-19 employees W. S. Central rural	180,000	94,860
1-19 employees Mountain urban	226,736	145,111
1-19 employees Mountain rural	146,905	101,071
1-19 employees Pacific urban	899,074	632,049
1-19 employees Pacific rural	109,933	81,570
20-49 employees N. England urban	16,227	12,105
20-49 employees N. England rural	4,337	2,581
20-49 employees Mid Atlantic urban	47,965	34,583
20-49 employees Mid Atlantic rural	3,699	2,053
20-49 employees E. N. Central urban	47,491	33,624
20-49 employees E. N. Central rural	11,533	6,297
20-49 employees W. N. Central urban	15,418	11,579
20-49 employees W. N. Central rural	9,945	5,470
20-49 employees S. Atlantic urban	39,442	28,359
20-49 employees S. Atlantic rural	11,192	8,562
20-49 employees E. S. Central urban	10,934	7,807
20-49 employees E. S. Central rural	6,328	4,537
20-49 employees W. S. Central urban	23,389	17,705
20-49 employees W. S. Central rural	6,798	3,270
20-49 employees Mountain urban	11,786	9,617
20-49 employees Mountain rural	6,168	2,924
20-49 employees Pacific urban	43,260	32,791
20-49 employees Pacific rural	4,294	2,559
50-99 employees N. England urban	5,250	3,197
50-99 employees N. England rural	1,209	961
50-99 employees Mid Atlantic urban	16,037	9,526
50-99 employees Mid Atlantic rural	1,110	643
50-99 employees E. N. Central urban	17,135	11,155
50-99 employees E. N. Central rural	3,580	1,743
50-99 employees W. N. Central urban	5,731	3,760
50-99 employees W. N. Central rural	3,212	1,542
50-99 employees S. Atlantic urban	13,333	8,626
50-99 employees S. Atlantic rural	3,522	1,997
50-99 employees E. S. Central urban	3,776	2,858
50-99 employees E. S. Central rural	2,041	1,247
50-99 employees W. S. Central urban	8,118	4,554
50-99 employees W. S. Central rural	2,322	1,189
50-99 employees Mountain urban	4,101	2,756
50-99 employees Mountain rural	1,809	658

50-99 employees Pacific urban	15,387	9,571
50-99 employees Pacific rural	1,220	576

EXHIBIT IV-15: DMI FRAME SIZE AND ESTIMATED NUMBER OF ELIGIBLE FIRMS* (cont.)

Sample Stratum	DMI Frame Size	Number of Eligible Firms*
100-499 employees N. England urban	3,608	1,937
100-499 employees N. England rural	849	432
100-499 employees Mid Atlantic urban	11,020	6,006
100-499 employees Mid Atlantic rural	902	540
100-499 employees E. N. Central urban	10,807	7,035
100-499 employees E. N. Central rural	2,639	1,327
100-499 employees W. N. Central urban	3,787	2,356
100-499 employees W. N. Central rural	1,886	1,141
100-499 employees S. Atlantic urban	8,324	4,728
100-499 employees S. Atlantic rural	2,202	1,013
100-499 employees E. S. Central urban	2,403	1,697
100-499 employees E. S. Central rural	1,456	702
100-499 employees W. S. Central urban	5,182	3,234
100-499 employees W. S. Central rural	1,320	486
100-499 employees Mountain urban	2,371	1,498
100-499 employees Mountain rural	929	534
100-499 employees Pacific urban	8,853	5,374
100-499 employees Pacific rural	685	105
Unknown employees N. England urban	14,446	4,767
Unknown employees N. England rural	3,487	**
Unknown employees Mid Atlantic urban	48,648	46,118
Unknown employees Mid Atlantic rural	3,325	246
Unknown employees E. N. Central urban	36,591	34,908
Unknown employees E. N. Central rural	7,093	1,525
Unknown employees W. N. Central urban	8,229	6,731
Unknown employees W. N. Central rural	3,962	2,195
Unknown employees S. Atlantic urban	32,753	31,607
Unknown employees S. Atlantic rural	7,295	3,195
Unknown employees E. S. Central urban	5,764	2,271
Unknown employees E. S. Central rural	3,164	1,531
Unknown employees W. S. Central urban	16,711	15,742
Unknown employees W. S. Central rural	3,525	832
Unknown employees Mountain urban	10,242	8,255
Unknown employees Mountain rural	4,551	1,497
Unknown employees Pacific urban	35,621	25,540
Unknown employees Pacific rural	3,118	2,151
500 or more employees	14,418	851
Asian urban***	225,995	156,615
Asian rural***	28,872	18,969
Black urban***	14,215	11,671
Black rural***	489	422
Hispanic urban***	215,012	151,583
Hispanic rural***	30,629	21,502
Total	7,324,850	5,024,749

*Estimated number of eligible firms is calculated using the best estimate of eligibility rates.

**No respondents in this sampling strata.
***The table values for these strata are not the total number of Asian-, Black-, or Hispanic-owned firms since some minority-owned firms are identified in the non-minority strata during screening and interviewing.

V. PRE-INTERVIEW PROCEDURES

A. Introduction

Prior to the main interview, two additional areas required preparation—the mailout package and interviewer training. Each had a crucial role in ensuring the collection of quality data.

B. Mailout Package

Mirroring the 1987 NSSBF, a mailout package was created that contained background information on the study and information needed by the respondent to prepare for the main interview. The package included a letter from the Federal Reserve Board chairman, Alan Greenspan, a letter from Price Waterhouse, a brochure of answers to commonly asked questions, worksheets for organizing the firm's financial data, and a postage-paid return envelope (See Appendix B).

The mailout package added credibility and significance to the survey process at minimal cost. While the mailout package did not increase a respondent's aversion to the survey, it did provide an opportunity for avoidance behavior - many respondents would put off dealing with the interviewer by asking that the package be remailed. Another avoidance behavior that we encountered was respondents indicating that they had not yet received or yet reviewed the package. During the beginning of the survey, interviewers generated a great number of requests to remail the package. Remailing resulted in scheduling the interview at least one week later. Several firms exhibited classic avoidance by requesting the package several times. This led to our downplay of the importance of the mailout packet for those firms. We instructed interviewers to suggest alternatives such as faxing the cover letters, the worksheets, or the brochure to

address any concerns. We also instructed interviewers to urge respondents to use tax records or financial statements in order to complete the survey. This shift resulted in a substantial increase in the number of fax requests and a significant decrease in the amount of mail package requests.

The letters and brochure were similar in design to those used in the 1987 NSSBF survey. They were updated to accurately reflect the current purpose and background of the study. The Price Waterhouse letter included a phone number for an incoming line to resolve any additional questions or concerns. As in the previous survey, many respondents called to confirm the project's legitimacy. Also, a number of firms who solely use answering services used the line to contact us.

The worksheets assisted in improving the completeness and accuracy of the data collected. Many respondents would not have known the answers to the quantitative questions without reference to records. The interviews for those respondents who completed the worksheets went very smoothly. In addition, for those respondents who did not complete the worksheet, but at least reviewed it, conducting the interview was significantly more efficient.

A few anxious respondents would complete the worksheets and mail them back before the main interview commenced. Some thought that the entire survey was then completed. The interviewers would explain the intended study purpose and the respondent would either allow the interviewer to conduct the entire interview, refuse to answer any questions, or only respond to sections not covered in the worksheets.

Another peculiar issue regarding worksheets was some respondents' resistance to complete them. In accordance with downplaying the mailout package and accommodating the respondent, interviewers urged the use of alternate forms, especially 1992 tax forms. The survey provided interviewers with tax line references to

specific questions to assist in the collection of accurate and pertinent data. Interviewers asked for these and any other forms to be sent to us to assist in the data cleaning phase.

The postage-paid return envelope was added to the mail package after the pretests. In order to facilitate the resolution of any discrepancies or missing information and provide validation to the data collection process, we decided to request that the firms return the worksheets or any financial records that they used to complete the interview.

C. Interviewer Training

Interviewer training was designed to provide a common foundation for all interviewers to build upon. One of the goals of the training session was to achieve consistency across interviewers at the data collection stage.

Significant resources were devoted in order to provide accurate and relevant instruction. A thorough training manual was created that embodied the concepts, processes, and relationships in the survey. This text covered the survey's personnel, interviewer conduct, the rules for interviewing, contacting respondents, using the CATI software, and understanding the questionnaire content. Exhibit V-1 details the training manual's table of contents.

The interviewers and CATI supervisors participated in a comprehensive training session that lasted one week. Several members of the questionnaire design team were actively involved in the training. Particularly, the Federal Reserve staff and a Price Waterhouse accountant contributed their experience and knowledge that significantly added to the training. The session paralleled the training manual's content.

EXHIBIT V-1: TRAINING MANUAL TABLE OF CONTENTS

EXHIBIT V-1: TRAINING MANUAL TABLE OF CONTENTS (cont.)

After the training session, interviewers practiced conducting the interview by "role playing" with one another. Before any interviewer placed a call to a respondent, they first had to qualify by passing a test interview with a CATI supervisor.

As the data collection phase of the project lasted longer than anticipated, it was necessary to conduct several refresher training sessions with interviewers, and to conduct additional sessions to train new interviewers hired after the initial training session was concluded. Each subsequent session augmented the original training content with useful knowledge gained through conducting interviews. Training for new interviewers typically lasted three days with an audience of fifteen people. The first day focused on understanding the survey's background and purpose as well as CATI procedures. The second day consisted of a review of the entire questionnaire. The third day consisted of role playing with another interviewer and a qualifying interview with a supervisor. New interviewers would sometimes require a few additional days to qualify.

There is no significant evidence that supports definitive conclusions regarding the addition of new interviewers during the collection process. We believe new interviewers have a neutral, if not somewhat positive, effect on the survey's response rate for the following reasons:

- We maintained very high standards for hiring interviewers, and those standards did not decrease. For example, interviewers had to have at least two years of college course work. Interviewers' education ranged from some college to advanced degrees.
- The training did not significantly change. It did, however, have the benefit of notes from previous training sessions and group meetings.
- An analysis of the data showed no clear pattern of interviewer effect on response rate or item non-response.

- The new interviewers provided a positive force. They added a degree of enthusiasm that affected the interviewers who had been there longer.

We began the screening process with four trained full-time equivalent interviewers. We started the main interviewing process with over six trained full-time equivalents. Over 36 weeks of the data collection process, we averaged over 15 interviewers and peaked at 26 full-time equivalents. The average for the entire data collection phase was 12. This included eight closing weeks where fewer interviewer hours were necessary to conclude data collection.

VI. DATA COLLECTION

A. Introduction

The collection of accurate, pertinent, and comprehensive data proved to be a very challenging aspect of the study. The evolving demands of the survey required the system to be dynamic and persistent.

Analogous to its predecessor, this survey's data collection process involved a screening stage, a subsequent mailout, and a concluding main interview. Screening interviews established the appropriate respondent, determined a firm's eligibility, and verified a firm's mailing address. The resulting mailout to eligible firms provided each respondent with the previously discussed mailout package before the main interview commenced. Lastly, the main interview collected the extensive statistics that comprise the survey's content.

B. Sample Administration

Since the eligibility rate was initially unknown, survey managers released the drawn sample in waves. Results from the first wave of 12,375 random cases provided the rate used to calculate the additional sample to yield the desired number of eligible firms. Ultimately, the survey's development required the release of second and third waves. The second wave added 769 Asian and 350 Hispanic random records. The third included 2,220 random firms from the main sample partition. The aggregate screening sample resulted in a sufficient number of eligible firms.

This study demanded an unplanned supplementary stage in addition to screening in order to identify and locate sampled firms. Each wave had a proportionate number of cases that were missing crucial identifying information. Records for a total of 1,588

firms were missing at least a phone number, and 641 of these were also missing other important information, such as contact name or address. Exhibit VI-1 shows how many firms were missing what kind of information.

EXHIBIT VI-1: BREAKDOWN OF TYPES OF INFORMATION MISSING

Information Missing	Incidence
Phone Only	947
Phone & CEO Name	546
Phone & Address	12
Phone, CEO Name & Address	83
TOTAL	1,588

These deficiencies meant that these firms could not be screened. This added another dimension to the survey. We created a separate tracking system to locate these cases and obtain resolution. This process resulted in either the acquisition of pertinent information or the confirmation that no information exists. The former enabled the records to be resolved appropriately or to be added to the screening stage. The latter resulted in a confirmed unreachable resolve code.

Exhibits VI-2 and VI-3 show by sample partition and by employment size, respectively, the number of firms missing at least the phone number.

EXHIBIT VI-2: BREAKDOWN OF SAMPLE MISSING CRUCIAL INFORMATION BY SAMPLE PARTITION		
Sample Partition	**Missing Crucial Information Sample**	**Total Sample**
Main	1,333	12,120
Black	31	825
Asian	115	1,594
Hispanic	109	1,175
TOTAL	1,588	15,714

EXHIBIT VI-3: BREAKDOWN OF SAMPLE MISSING CRUCIAL INFORMATION BY EMPLOYMENT SIZE OF FIRM		
Size (Number of Employees)	**Missing Crucial Information Sample**	**Total Sample**
Unknown	47	480
1 to 19	1,400	10,254
20 to 49	58	1,341
50 to 99	57	1,830
100 to 499	26	1,601
500+	0	208
TOTAL	1,588	15,714

C. Screening Technique and Results

The protocol for conducting initial screening used the following technique:

- two attempts to contact the owner or top executive
- two attempts to screen with the owner's assistant
- “unlimited” number of attempts to screen with a gatekeeper

Each attempt was a distinct call to conduct the screening. The actual steps were an introduction that disclosed who we are and why we are calling, along with an explanation that we would like to verify the firm's mailing address and to determine the firm's eligibility for participation in the study.

Results ranged from auspicious to disappointing as some owners would go out of their way to assist the screening while others would go out of their way to hinder it. In the end, the average number of calls to complete a screening interview was 5.1 and ranged from 1 to 51 attempts. Exhibits VI-4 and VI-5 present details on the number of calls to complete a screener.

For some businesses, we were never able to complete a screener interview since some firms were never able to be reached and others refused. If, for any reason, interviewers could not make contact or appropriately screen a firm, a "dummy screener" was completed. Dummy screeners were completed for each firm where we could verify an address with a reliable third party (e.g., Chamber of Commerce). They were completed by recording “DK” to all questions in the screener. This continued the process of proceeding with the mail package and concluding with the main interview. The mail package sometimes created interest and addressed the concerns that prevented the screening.

EXHIBIT VI-4: NUMBER OF CALLS TO COMPLETE A SCREENER BY SAMPLE PARTITION

Sample Partition	Count	Median	Mean	Standard Deviation	Minimum	Maximum
Main	11,010	3.0	5.1	5.9	1.0	51.0
Black	760	3.0	5.4	6.7	1.0	45.0
Asian	1,423	3.0	5.0	5.6	1.0	44.0
Hispanic	1,050	3.0	5.1	6.3	1.0	44.0
TOTAL	14,243*	3.0	5.1	5.9	1.0	51.0

*Includes 3,925 confirmed ineligibles (Exhibit IV-1 Box 'A') plus 10,318 confirmed eligibles (Exhibit IV-1 Box 'B', 10,533 less 215 hard refusals).

EXHIBIT VI-5: NUMBER OF CALLS TO COMPLETE A SCREENER BY SIZE OF FIRM

Size	Count	Median	Mean	Standard Deviation	Minimum	Maximum
Unknown	416	4.0	6.3	6.8	1.0	43.0
1 to 19	8,974	3.0	5.3	6.4	1.0	51.0
20 to 49	1,290	3.0	4.6	4.9	1.0	48.0
50 to 99	1,782	3.0	4.4	4.8	1.0	40.0
100 to 499	1,574	3.0	4.8	5.0	1.0	48.0
500+	207	3.0	4.5	5.3	1.0	45.0
TOTAL	14,243*	3.0	5.1	5.9	1.0	51.0

*Includes 3,925 confirmed ineligibles (Exhibit IV-1 Box 'A') plus 10,318 confirmed eligibles (Exhibit IV-1 Box 'B', 10,533 less 215 hard refusals).

The screener processed a total of 15,714 firms from the sampling frame, of which 10,141 either were determined to be eligible or were passed on to the mailout phase of the survey through use of a "dummy screener." While the initial screener statistics

allowed information gathered from assistants and gatekeepers to be recorded, this information was regarded as indicative and temporary, but not final. For any firm about whom eligibility information (i.e., fewer than 500 employees, branch office, non-profit, etc.) was not obtained from the owner or senior management , each of the questions that determined eligibility for the survey was repeated at the beginning of the main interview.

D. Interviewing Technique and Results

The screeners, (both eligible and "dummy") were used to generate a mailing. Each firm on the list was sent the information package. Interviewers began telephoning respondents to schedule or obtain an interview no sooner than seven successive days after the mail package was sent.

Exhibits VI-6 and VI-7 present details on the number of calls to complete the main interview. The average number of calls was 18.7 and ranged from 1 to 99 attempts.

EXHIBIT VI-6: NUMBER OF CALLS TO COMPLETE A MAIN INTERVIEW BY SAMPLE PARTITION

Sample Partition	Count	Median	Mean	Standard Deviation	Minimum	Maximum
Main	4,180	12.0	18.2	17.4	1.0	99.0
Black	414	17.0	23.3	20.2	1.0	99.0
Asian	382	14.0	18.9	16.7	1.0	99.0
Hispanic	380	13.0	18.6	18.8	1.0	99.0
TOTAL	5,356*	13.0	18.7	17.7	1.0	99.0

*Includes 5,356 eligible, completed interviews (Exhibit IV-1 Box 'F').

EXHIBIT VI-7: NUMBER OF CALLS TO COMPLETE A MAIN INTERVIEW BY SIZE OF FIRM

Size	Count	Median	Mean	Standard Deviation	Minimum	Maximum
Unknown	122	16.0	23.3	19.7	1.0	89.0
1 to 19	3,495	12.0	17.8	17.4	1.0	99.0
20 to 49	550	13.0	19.2	17.7	1.0	99.0
50 to 99	643	14.0	19.7	17.6	1.0	99.0
100 to 499	535	15.0	21.5	18.6	1.0	99.0
500+	11	24.0	30.3	23.6	8.0	82.0
TOTAL	5,356*	13.0	18.7	17.7	1.0	99.0

*Includes 5,356 eligible, confirmed interviews (Exhibit IV-1 Box 'F').

The pretests refined the main interviewing approach to a four-step sequence. Exhibit VI-8 displays this technique. The pretests indicated that the most effective strategy to contact the respondent was one that is non-confrontational, puts the respondent at ease, and allows them to control their time while resolving any concerns they may have regarding the study.

EXHIBIT VI-8: MAIN INTERVIEW APPROACH

Step 1.	"Did you receive the package in the mail from Price Waterhouse with a letter from Alan Greenspan?"
Step 2.	"Have you had a chance to review it?"
Step 3.	"Do you have any questions or concerns regarding the study?"
Step 4.	Set the appointment or conduct the interview.

There were a number of cases where this approach did not work effectively. These cases led to the development of a second approach.

This second technique was more direct and focused. The issues and ideas embedded in the previous approach were used to address any questions or concerns the respondent initiated. The actual approach involves an introduction of who we are and what we are doing followed by the statement, "if you don't have any questions, we'd like to begin the interview now." This approach resolved many of the evolving demands of the study and avoided giving the respondent the opportunity to engage in avoidance behavior. It minimized the interviewer's introduction of concerns and problems to respondents. It also forced the respondent to either continue with the survey or raise concerns that would be immediately addressed.

Regardless of the adopted approach, there still existed great difficulty in reaching the appropriate contact. Over-protective gatekeepers, answering machines and services, and well-traveled respondents accounted for this inconvenience. The study dealt with this issue by persistence, an incoming line, and tailored letters. Perseverance was one way of combating this impediment. Consistently calling at different times of the day and different days of the week helped reach the appropriate person. The incoming line was also of assistance. It was a toll-free private line for respondents that served as a help, complaint, and contact line. It was used for starting as well as concluding interviews. Interviewers left the number with gatekeepers, answering machines, or answering services. The number was also listed on the mailout package's brochure and the tailored letters. Many respondents did call the private incoming telephone line. They verified the survey's authenticity, voiced concerns, sought information to help their business, declined to participate, and established appointments to begin the survey. However, the majority of respondents who received the number never used it.

The core group of respondents that we rarely were able to reach led us to generate tailored letters informing them of their inclusion in the survey. These letters, tailored to the firm's expressed concern or their type of avoidance behavior, conveyed the study's importance and purpose.

Once an interviewer established a rapport with a respondent, avoidance behavior almost always surfaced. The three most frequent concerns were regarding legitimacy, confidentiality, and time to complete. Respondents used the incoming line to verify the project's legitimacy. Some also used it to seek a credible reference at the Federal Reserve or Small Business Administration. A few firms went further to contact their local Price Waterhouse or Federal Reserve office. This frequently led to cynicism as those offices would convey little substantial support despite our attempts to keep all parties informed about the survey.

Every hedging firm expressed confidentiality and time concerns. The study's guidelines effectively resolved the former concern by immediately addressing the issue and providing explanations on the strict rules being followed. In rare cases, respondents requested signed agreements assuring their data's privacy. The study, however, did not proficiently rectify every respondent's time concerns. After offering to be available at the respondent's convenience and to conduct the interview during several short segments, respondents still raised time considerations.

Another important data collection protocol feature was the use of special appointments. If respondents could not begin or continue the survey when the interviewers called, they would be asked for a more convenient time for an appointment to conduct the survey. These appointments ranged from very definite to very doubtful. However, they did add significant benefits as they captured a portion of the respondents who remained steadfast to the appointment, conveyed our commitment, and imparted our consideration of the respondents' time. The other side of the coin is that some respondents consistently used this as the main mechanism to evade data collection. However, interviewers became familiar with who was using this tactic and adapted accordingly.

Once an interviewer effectively resolved a respondent's aversion, many other challenges inherent to the interview itself had to be addressed. A notable proportion of interviews did not go from start to finish in one sitting. If a respondent was not willing to continue at some point, an interviewer would suspend the interview for later retrieval. Lack of time, refusals, referrals, and desire to obtain better information accounted for the majority of suspends. All of these reasons led some respondents to stop the interview after a given point. A common break point was at the beginning of Section P: Income and Expenses. Our CATI software allowed us to resume the interview at any point at which the breakoff occurred.

At any point in the survey, respondents could refer us to someone else, whether he/she was within the firm or an outside accountant. This situation occurred when respondents either did not want to do the survey or did not have enough knowledge regarding the financial questions. This usually resulted in the cessation of the interview and a subsequent call to the referral. Outside accountants could not be counted upon to answer the questions regarding the owner or the firm, so interviewers attempted to obtain owner and firm demographic information from the respondent for those questionnaire sections before seeking the accountant. Interviewers also used a hard copy of Section U: Creditworthiness, to capture the owner's responses to prevent a call-back after the referral provided answers to the financial questions. These recorded answers would then be entered by the interviewer after the telephone interview with the referral or during the data processing stage.

Some accountants provided essential help while others created major opposition. Upon referral, accountants would periodically counsel their clients to discontinue their participation. Subsequent calls and faxes attempted to convert these respondents. However, these conversion attempts did not optimally work as, at best, we would continue with the less knowledgeable owner/respondent, thereby gaining less accurate estimates rather than actual figures.

Some respondents did not know any of their 1992 financial figures and were not willing or able to obtain the associated tax forms, but they did have 1993 tax records. For these instances where the respondent was adamant about not retrieving any other information, we collected the 1993 data while noting it.

Special circumstances surfaced during some interviews where an interviewer was unable to deal with a particular response. In these cases, the interviewer had to note the problem so that it could be corrected after the interview had been completed. Interviewers indicated these problems on problem/comment sheets (see Exhibit VI-9).

These problem sheets allowed the data collection system to adapt to and handle any impediment. Respondents' and interviewers' concerns were recorded and stored along with comments, explanations, and other responses. Initially, interviewers generated an inordinate number of sheets that reflected issues ranging from the respondent being out of the office to the company no longer being in business. Often this duplicated effort as these issues were already accounted for using the final resolve codes for call outcomes. We then trained the interviewers to minimize this duplicated data. Problem/comment sheets then ranged from such information as:

- indicating that the reported motor vehicle loans should be at a different bank than the motor vehicle loans recorded by CATI, or

EXHIBIT VI-9: PROBLEM/COMMENT SHEET

National Survey of Small Business Finances

INTERVIEWER	
DATE/TIME	

PWCODE:	**Page ____ of ____**
TELEPHONE #:	( ) -

NOTES*

CASE STATUS ASSIGNED:

*Be sure to clearly indicate which survey question number you are referencing if any.

- that these motor vehicle loans were recorded incorrectly as capital leases, or
- that question F6 should be 0 (F6 asks for the firm's total credit limit on the credit cards), or
- question P10's response was included in P9.1's answer (P10 asks for the amount of the officers' compensation/guaranteed payments to partners. P9.1 asks for the total amount the firm paid in salaries and wages, excluding jobs credit and owners' compensation).

In addition to the problem/comment sheet, a key aspect to this exception procedure was the use of an exception code. Coupled with an explanation on a problem/comment sheet, an interviewer would record an exception as a question's response (e.g., a response considered out of range and therefore not allowed by CATI) to continue with the survey.

An integral part of the interviewing strategy was to work in teams. A group of interviewers worked on a specific subsample of firms. This sample segment resulted in the team working the same group of small businesses for the main interview. The specific segments were Asian, Black, Hispanic, and main (non-minority) sample partition. The interviewer had the opportunity to offer the specific names of one of their team members for any call requiring a call-back. This assisted the team in establishing a rapport so that once contact was established, there was minimal "cold calling" done by an interviewer not recognized by the respondent. Also, once a difficult or reluctant respondent had been contacted and successfully converted, the respondent felt more comfortable working with a few particular interviewers. Furthermore, teams enhanced the respondent's confidence in the study and established the importance of their participation.

This organization of interviewer teams revealed unique issues common to one or two groups. The Asian and Hispanic sample teams had to deal with considerable language

barriers. The former sample involved several dialects, while the latter entailed mainly one. Therefore, it was logistically necessary to hire and train Spanish speaking interviewers. These interviewers first attempted to conduct the interview in English. When this was not possible, they administered the survey in English while providing prompts and definitions in Spanish when necessary. As a last resort, the interview was conducted solely in Spanish. Spanish interviews were conducted by our bi-lingual interviewers using the English version of the survey instrument.

Systematic information regarding surveys conducted in Spanish is minimal because an indicator that the interview was done partially or completely in Spanish did not exist. Interviewers conjectured that Spanish-speaking respondents were more receptive to Spanish-speaking interviewers and that there were not any unique problems in administering the survey beyond the language barrier.

In an effort to achieve the sampling target number of completed interviews and adequately represent Asian-owned small businesses, a Korean-speaking interviewer was hired during the final data collection stage to help overcome the language barrier. However, success in converting initial refusals was limited.

In order to resolve cases that persisted for several months, we adopted a variation on the team concept. A team leader, who was more adept at countering avoidance behavior, persuaded the respondent to begin or continue with the survey. He/She would then pass the interview to a team member to conclude the interview.

Eliciting response was a mounting problem throughout the study. As in the previous survey, respondents were averse to reporting sensitive and confidential data. Also, establishing contact with the correct individual at an appropriate time proved difficult. Both directly resulted in the need to make more than twenty call-back attempts on average to complete a case (screener and main). The overall average time to resolve

a case was about three hours. The time required to complete an interview was dominated by the time spent establishing contact, gaining cooperation, reestablishing contact after breakoff, and scheduling appointments often not kept. At the end of the data collection period, the number of call-backs had exceeded 35 for the sample still not resolved and the average interviewer productivity had dropped to 0.2 completes per hour (i.e., five hours to obtain a completed interview).

In an effort to overcome the severe opposition in providing what some respondents felt to be private and sensitive information, interviewers sought several types of responses. They first attempted to collect actual figures. When opposition to providing actual figures occurred, interviewers sought to address and resolve specific respondent concerns which prevented the collection of these figures. If the interviewers felt that further probing for the actual figures jeopardized completion of the interview, they would prompt the respondent for estimates. If a respondent was still unwilling to provide a response, the interviewer asked for ranges. If a respondent was reluctant to provide ranges, interviewers attempted to determine if the firm had the particular item or financial service or not.

E. Use of Records and Worksheets

Nearly one-half of the respondents (47%) said they had records available to help them answer questions about income, expenses, and the balance sheet. These respondents reported the following types of records available during the interview:

- Tax records (53%)
- Survey Worksheets (19%)
- Financial statements (29%)
- Bank statements (7%)
- Other records (13%)

Exhibits VI-10 and VI-11 present the disposition, by sample partitions and by size of firm, of the 483 respondents who indicated during the survey that they used the worksheets.

EXHIBIT VI-10: DISPOSITION OF RESPONDENTS WHO INDICATED THEY USED WORKSHEETS BY SAMPLE PARTITION

Sample Partition	Count	Percent
Main	416	86.1
Black	30	6.2
Asian	16	3.3
Hispanic	21	4.3
TOTAL	**483**	**100.0**

EXHIBIT VI-11: DISPOSITION OF RESPONDENTS WHO INDICATED THEY USED WORKSHEETS BY SIZE OF FIRM

Size	Count	Percent
Unknown	9	1.9
1 to 19	273	56.5
20 to 49	69	14.3
50 to 99	67	13.9
100 to 499	65	13.5
500+	0	0.0
TOTAL	**483**	**100.0**

Despite our efforts to encourage the return of the worksheets, only 673 (12.5%) returned worksheets or tax records. We occasionally received worksheets for unresolved cases. For these cases, interviewers would first attempt to conduct the survey without referencing the worksheets. If a respondent demanded the use of the

worksheet, we would then reference the worksheets while conducting the interview. As previously noted, some respondents would refuse to review any questions that the worksheets covered. After concluding with the respondent, the interviewer would then transcribe the worksheet information.

VII. DATA PROCESSING

A. Introduction

The project's final product is an edited and coded data base. To create this, the collected data underwent several processing steps. These systematic procedures ensured the construction of a data base that the Federal Reserve will use to perform additional editing and analytical imputations.

B. Computer-Assisted Telephone Interviewing

Computer-assisted telephone interviewing (CATI) facilitated conducting the NSSBF. Interviewers read questions displayed on a computer monitor, recording the respondent's answer. The computer played a vital up-front role enforcing a minimum quality level in the collected data. The computerized interview script allowed for automatic branching and on-line feedback for ineligible responses. This aided in resolving discrepancies on the spot, reducing transcription errors, and reducing the incidence of missing or inconsistent data.

Designed range-edits were the main checks safeguarding data integrity. The computer would only accept responses that fell in a certain predefined range. The design team established these confines to reflect the most probable answers. If a response fell outside the preordained bounds, the computer would alert the interviewer. The interviewer would then probe to guarantee the response's accuracy and pertinence. Once the out-of-scope answer was determined to be valid, the interviewer would follow the exception procedure and continue with the survey.

Diverging from the 1987 NSSBF, this study did not include any elaborate cross-item CATI edit checks akin to the one ensuring that the balance sheet balanced. However,

it did contain a web of comprehensive edits that reduced errors. Exhibit VII-1 illustrates several specific range edits. The training manual is the complete reference guide to every individual check.

C. Problem Resolution

Since the survey is very long and complicated, special circumstances occasionally surfaced where an interviewer was unable to deal with a certain response while the interview was live. In these cases, the interviewer had to note the problem so that the data could be corrected after the interview was completed. We designed a post-interview recoding system to automate this process.

This system involved hardware and software platforms in addition to an aligned editing process. Dedicated computers and a stand-alone recoding data base formed the backbone of the technical support. The recode data base was merged onto the original data to reflect the appropriate modifications while maintaining originally recorded responses. The process detailed the means to consistently modify data accurately and to produce the final edited and coded deliverable.

Three distinct phases comprised recoding: problem/comment sheets, exception codes, and respondent changes to SIC codes. Since the recoding was executed in batches, all three stages were performed simultaneously.

EXHIBIT VII-1: EXAMPLES OF SPECIFIC RANGE EDITS

- The principal partner's ownership percentage should be less than 100 percent.
- An S-corporation should have less than or equal to 35 stockholders.
- The number of sites located in the same area as the main office should be less than or equal to the total number of sites.
- The total amount of principal owed as of year end on any loan or lease should be greater than 0 dollars.
- If the lender required a real estate appraisal or environmental survey for the most recent credit application, the respective cost should be greater than 0 dollars.
- The most recent credit application's original interest rate should be between 4 and 24 percent.
- A trade credit supplier's discount period for early payment should be between 0 and 30 days.
- A trade credit supplier's monthly penalty for late payment should be between 0 and 5 percent.
- The firm's projected annual sales growth for the five-year period following the most recent equity financing should be between -100 and 100 percent.
- The firm's total sales should be greater than or equal to 0 dollars.
- The total value of the firm's inventories, if any, should be greater than 0 dollars.
- An S-corporation's or corporation's combined amount of capital stock and additional paid-in-capital should be greater than or equal to 0 dollars.

NOTE: A complete reference to range edits by question is found in Section VII of the "CATI Interviewer Training Manual" (separately bound).

The first stage was recoding the problem/comment sheets. As noted earlier, the problem/comment sheet and exception procedure circumvented the CATI guidelines. Whenever a response was made that the system would not accept, that response appeared on this problem report. Generally, these notes would be key entered following an interview directly into the text box at the end of the CATI questionnaire. Occasionally, interviewers would fail to enter their paper notes into the computer. The most common reason for this was that the problem survey was frequently completed in steps. Therefore, an interviewer may have been unaware of problem/comment sheets from an earlier part of the interview that needed to be entered.

Exhibit VII-2 ranks the top 25 of 167 survey variables having exception responses prior to any recoding. These variables account for nearly half (47.3%) of the total number of exception responses.

EXHIBIT VII-2: VARIABLES AND FREQUENCIES OF EXCEPTION RESPONSES BEFORE RECODING		
Variable Name	**Variable Description**	**Count**
F6	Total credit card limit	377
J34	Interest rate on most recent loan	103
L12	Monthly penalty on unpaid balance	85
R17C	Book value of asset	84
R17B	Book value of asset	72
L9	Early payment discount period length	61
P2	1990 total sales	60
F101	Total credit line limit	51
H6C2	Zip code for source	51
H42	Years conducted business with source	50
H6B2	State for source	50
C28	Age of principal owner	45
E31	Monthly checking account balance	45
F301	Motor vehicle loan principal	45
C33	Ownership share of principal owner	43
S11C	Book value of liability	42
H82	Number of miles from main office	41
P10	Amount of officers' compensation	40
F351	Equipment loan principal	39
J28	Length of most recent line of credit	37
R18	Total dollar amount of all assets	37
D8	1993 sales percentage outside U.S.	36
J35	Percentage points paid to close loan	36
H6B3	State for source	34
H6C3	Zip code for source	34
Subtotal	**Exception Responses in Top 25**	**1,598**
TOTAL	**Exception Responses**	**3,378**

We explain the reason for an "EX" response for the six variables that were most coded "EX."

F6-total credit card limit	Firms exclusively used credit cards such as American Express that have no fixed credit limit. Therefore, there is no maximum amount the firm could charge on these accounts.
J34-interest rate on most recent loan	Firms did not know the exact interest rate, often they knew only the amount over or under prime.
L12-monthly penalty on unpaid balance	Respondents indicated that the trade credit penalty for paying after the due date is a loss of trade credit or a fixed dollar amount, not a percentage on the unpaid balance.
R17C and R17B-book value of asset	Firms generally had only one "other asset," however, the CATI program allowed for three. When the interviewer got to questions two and three he/she mistakenly entered "EX" instead of skipping those questions.
L9-early payment discount period length	Firms indicated that the discount period was over 30 days (i.e., the maximum entry allowed).

Since some problem/comment notes may exclusively exist on paper, the computer text was not comprehensive. The data editors had to translate the remaining paper notes into a suitable machine readable form. This procedure involved a verbatim transcription into the computer.

To accurately recode all problem/comment sheets, it was necessary to go through both the written and computer printed notes. If a discrepancy existed, it was corrected in the recode data base. The correction involved an entry into the recode data base of all

problem/comment sheet information for each individual case. Once this updated material was merged onto the old, all hard copy data were permanently stored in machine form with the appropriate case.

The second recoding stage involved modifying individual case data where exception codes were indicated. Creating a comprehensive list of text notes enabled the confident recoding of these data. To begin this next phase, it was necessary to produce lists of the relevant text notes and the recorded data and text responses. These were used as reference materials to illustrate what changes needed to be made, explain why the modifications were made, and verify the problem's legitimacy.

The actual process was very consistent and conservative, modifying a total of 3,488 surveys (65% of the 5,356 eligible, completed interviews). The modifications noted in the text note limited the scope. That is, the text note information was the sole vehicle for creating recodes. Discovering unrelated inconsistencies did not result in recoding unless the discovery warranted immediate action. For example, ascertaining that the financial institution roster contained two identical banks led to the appropriate recoding of the duplicate and its associated questions. However, discovering that the fundamental asset values did not sum to the total did not lead to applicable recoding.

Sometimes, an issue highlighted by the text note or discovered through inspecting the data could not be sufficiently rectified through a conservative recode. Examples range from too little to too much information provided. Occasionally, a variable would have an exception code, but no corresponding explanation. Other times, the note provided either an ambiguous summary or a detailed account with inadvisable facts. For these cases, the data editor entered a concise explanation into a special variable that specifies what could not have been recoded. This allows the Federal Reserve processors to utilize this stored information to finalize recoding or to accept the original values as deemed appropriate.

Creating a recode data base revealed several recurring edit and recoding issues that needed to be resolved. Exhibit VII-3 ranks the top 25 of 550 variables having recodes. These variables account for one-third (33.7%) of all recodes. In all, 3,488 of the 5,356 completed interviews had at least one recode. These are attributed to the CATI range edit, question wording, or the respondent's or interviewer's misunderstanding. Simple to difficult data adjustments corrected these issues.

The last recoding phase involved the editing of the firm's standard industrial classification code. The script would query the respondent if the firm's activities fell under a specific description. CATI obtained this characterization based on a two-digit SIC code on file. At that level, the narrative was very vague and general. Therefore, some respondents would say no to the general characterization of their industry, but indicate a focused portrait as a separate text response. The recoding stage processed these text responses while simultaneously assigning a correct SIC code.

A significant portion of this data cleaning step could have been eliminated if the interviewer was better versed on the SIC description. For 881 cases where the respondent provided a description of the firm's business activities, the original SIC code on file turned out to be the appropriate code. We did change the SIC code for 540 cases. Of these, only 13 proved to be out-of-scope (i.e., firms in the financial services or agricultural industries). We provide the complete listing of out of scope SIC codes above (see Section III.D). These cases, in addition to a few which reported ambiguous descriptions, were contacted an additional time to verify their SIC assignment.

EXHIBIT VII-3: VARIABLES AND FREQUENCIES OF RECODES		
Variable Name	**Variable Description**	**Count**
B6	Yes/No of recorded SIC description	881
NEWSIC	New SIC code	540
RECODPB	Recode note	450
F6	Total credit card limit	402
P1	Total sales	273
S1	Yes/No of any loans, mortgages, notes and bonds	252
TOTBANKS	Total number of banks	239
NBANKS	Number of banks excluding most recent	235
S2	Amount of loans, mortgages, notes, and bonds	234
TOTEMP	Number of full-time equivalents	231
P1V	Verification of P1	212
B11	Number of full-time employees	202
SHORT02	Short name of bank	168
BANK02	Long name of bank	167
U3	Number of personal obligations delinquent	153
U4	Number of business obligations delinquent	151
B13	Number of part-time employees	135
U6	Yes/No for rendered judgments	131
S1_1	Yes/No of liabilities earlier discussed	128
ELIG	True/False eligible	126
SHORT03	Short name of bank	126
BANK03	Long name of bank	120
U1	Yes/No for declared bankruptcy	115
SHORT01	Short name of bank	113
BANK01	Long name of bank	112
Subtotal	**Recoded Data Elements in Top 25**	**5,896**
TOTAL	**Recoded Data Elements**	**17,502**

D. Completeness Checks

The Federal Reserve project staff produced reports on item non-response in an attempt to define when a case would be counted as a completed interview. When possible, we conducted a final data retrieval strategic call-back for critical variables specified by the Federal Reserve (e.g., total assets, total liabilities, etc.).

E. Final Data Base Delivery

The final data files consist of the actual questionnaire responses, recoded questionnaire corrections, and analysis weights. The first two are cumulative data bases that, when ultimately merged, create an edited and coded file. This consolidated file forms the basis for further editing, analytical imputations, and processing by the Federal Reserve which will produce a final NSSBF user data base.

VIII. BIBLIOGRAPHY

Agresti, Alan. *Categorical Data Analysis*, John Wiley & Sons, Inc., New York. 1990.

American Statistical Association, *Proceedings of the International Conference on Establishment Surveys*, Alexandria, VA., 1993.

Cochran, William G. *Sampling Techniques* (Third Edition). John Wiley & Sons, New York, NY. 1977.

Cox, Brenda G., Binder, D., Chinnappa, B.N., Christianson, A., Colledge, M.J., Kotts, P.J., eds. *Business Survey Methods*, John Wiley & Sons, Inc., New York. 1995.

Cox, Brenda G., Elliehausen, Gregory E., and Wolken, John D. "Surveying Small Businesses About Their Finances." ASA: Proceedings of the Section on Survey Research Methods. 1989.

Cox, Brenda G., Elliehausen, Gregory E., and Wolken, John D. "The National Survey of Small Business Finances: Description and Preliminary Evaluation." Finance and Economic Discussion Series, Federal Reserve Board. November 1989.

Cox, Brenda G., Elliehausen, Gregory E., Wolken, John D. *The National Survey of Small Business Finances: Final Methodology Report*, RTI/4131-00F, Research Triangle Institutes, Research Triangle Park, North Carolina. September, 1989.

Deming, W. Edwards. *Sample Design in Business Research*. John Wiley & Sons, New York, NY. 1960.

Dilman, Don A., et. al. "The 1992 Simplified Questionnaire Test: Effect of Questionnaire Length, Respondent-Friendly Design, and Request for Social Security Numbers on Completion Rates." Prepared for the US Bureau of the Census, Annual Research Conference, March 1993.

Elliehausen, Gregory E., and Wolken, John D. "Banking Markets and the Use of Financial Services by Small and Medium-sized Businesses." Board of Governors of the Federal Reserve System. September 1990.

Federal Committee on Statistical Methodology. *Federal Longitudinal Surveys* (Statistical Policy Working Paper 13). Statistical Policy Office, Office of Management and Budget. 1986.

Federal Committee on Statistical Methodology. *Quality in Establishment Surveys* (Statistical Policy Working Paper 15). Statistical Policy Office, Office of Management and Budget. 1988.

Forsgren, R.. "Increasing Survey Response Rates: Methods for Small Business Researchers." *Journal of Small Business Management*, October 1989.

General Accounting Office. "Agencies Use of Consistent Race and Ethnic Definitions." December 1992.

Groves, Robert M. *Survey Errors and Survey Costs*. John Wiley & Sons, Inc., 1989.

Groves, Robert M. *Telephone Survey Methodology*. John Wiley & Sons, Inc., New York, NY, 1988.

Huberlein, Thomas A., et. al. "Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Review of the Published Literature." *American Sociological Review*, Vol. 43, No. 4, 1987.

Kish, Leslie. *Survey Sampling*. John Wiley & Sons, New York, NY. 1965.

Office of Management and Budget. *Standard Industrial Classification Manual*. Executive Office of the President, Office of Management and Budget. 1987.

Rovine, Michael J. and von Eye, Alexander. *Applied Computational Statistics in Longitudinal Research*. Academic Press, Inc. Boston, MA. 1991.

Singer, Eleanor. *Survey Research Methods*. The University of Chicago Press, Il. 1987.

Small Business Administration. *The State of Small Business: A Report of the President*. 1992.

U.S. Bureau of the Census, 1987 Economic Census. *Characteristics of Business Owners*.

U.S. Bureau of the Census, 1987 Economic Census. *Minority-Owned Business Enterprises*.

U.S. Bureau of the Census, 1987 Economic Census. *Women-Owned Businesses*.

U.S. Bureau of the Census, 1987 Economic Census. Censuses of retail trade, wholesale trade, service industries, transportation, manufactures, mineral industries, and construction.

Von Eye, Alexander. *Principles and Structuring Change* (Statistical Methods in Longitudinal Research Vol. I). Academic Press, Inc. Boston, MA. 1990.

Von Eye, Alexander. *Time Series and Categorical Longitudinal Data* (Statistical Methods in Longitudinal Research Vol. II). Academic Press, Inc. Boston, Ma. 1990.

Wolken, John D. "Geographic Market Delineation: A Review of the Literature." Board of Governors of the Federal Reserve System. October 1984.

APPENDIX A: DEMOGRAPHIC PROFILES OF DMI MINORITY PARTITIONS

EXHIBIT A-1: PROFILES OF THE MAIN BUSINESS PARTITION

Profile Category		Frequency	Percent
Census Region	East North Central	1,101,846	15.7%
	East South Central	328,867	4.6%
	Mid Atlantic	1,052,765	15.5%
	Mountain	416,179	6.1%
	New England	425,755	6.2%
	Pacific	1,123,347	17.5%
	South Atlantic	1,042,935	15.3%
	West North Central	548,783	7.7%
	West South Central	769,161	11.4%
	TOTAL	6,809,638	100.0%
Number of Employees	1 to 4	4,500,614	64.6%
	5 to 19	1,547,759	23.1%
	20 to 49	320,206	4.7%
	50 to 99	108,893	1.6%
	100 to 499	69,223	1.0%
	500+	14,418	0.2%
	Unknown	248,525	3.8%
	TOTAL	6,809,638	99.0%
MSA Status	Rural	1,566,546	22.2%
	Urban	5,243,092	77.8%
	TOTAL	6,809,638	100.0%
Industry	Construction	888,671	12.8%
	Manufacturing	445,018	6.5%
	Trn/Cm/Util	230,769	3.4%
	Wholesale	512,406	7.7%
	Retail	1,596,149	23.9%
	Insurance/Re	484,164	6.9%
	Services	2,619,277	38.3%
	Mining	33,184	0.5%
	TOTAL	6,809,638	100.0%

EXHIBIT A-2: PROFILES OF THE BLACK BUSINESS PARTITION			
Profile Category		**Frequency**	**Percent**
Census Region	East North Central	2,399	16.3%
	East South Central	1,101	7.5%
	Mid Atlantic	1,801	12.2%
	Mountain	292	2.0%
	New England	452	3.1%
	Pacific	2,304	15.7%
	South Atlantic	4,104	27.9%
	West North Central	954	6.5%
	West South Central	1,297	8.8%
	TOTAL	14,704	100.0%
Number of Employees	1 to 4	7,826	53.2%
	5 to 19	4,724	32.1%
	20 to 49	999	6.8%
	50 to 99	320	2.2%
	100 to 499	349	2.4%
	500+	42	0.3%
	Unknown	444	3.0%
	TOTAL	14,704	100.0%
MSA Status	Rural	489	3.3%
	Urban	14,215	96.7%
	TOTAL	14,704	100.0%
Industry	Construction	2,160	14.7%
	Manufacturing	1,100	7.5%
	Trn/Cm/Util	580	3.9%
	Wholesale	1,421	9.7%
	Retail	1,920	13.1%
	Insurance/Re	403	2.7%
	Services	7,109	48.3%
	Mining	11	0.1%
	TOTAL	14,704	100.0%

EXHIBIT A-3: PROFILES OF THE ASIAN BUSINESS PARTITION			
Profile Category		**Frequency**	**Percent**
Census Region	East North Central	27,810	10.9%
	East South Central	6,896	2.7%
	Mid Atlantic	44,857	17.6%
	Mountain	9,605	3.8%
	New England	9,940	3.9%
	Pacific	94,912	37.2%
	South Atlantic	32,159	12.6%
	West North Central	8,528	3.3%
	West South Central	20,160	7.9%
	TOTAL	254,867	99.9%
Number of Employees	1 to 4	149,199	58.5%
	5 to 19	72,405	28.4%
	20 to 49	12,643	5.0%
	50 to 99	3,522	1.4%
	100 to 499	2,062	0.8%
	500+	362	0.1%
	Unknown	14,674	5.8%
	TOTAL	254,867	100.0%
MSA Status	Rural	28,872	11.3%
	Urban	225,995	88.7%
	TOTAL	254,867	100.0%
Industry	Construction	15,006	5.9%
	Manufacturing	14,757	5.8%
	Trn/Cm/Util	6,205	2.4%
	Wholesale	26,606	10.4%
	Retail	88,497	34.7%
	Insurance/Re	11,817	4.6%
	Services	91,480	35.9%
	Mining	499	0.2%
	TOTAL	254,867	99.9%

EXHIBIT A-4: PROFILES OF THE HISPANIC BUSINESS PARTITION			
Profile Category		**Frequency**	**Percent**
Census Region	East North Central	19,782	28.1%
	East South Central	3,381	1.4%
	Mid Atlantic	39,396	16.0%
	Mountain	19,689	8.0%
	New England	14,779	6.0%
	Pacific	63,142	25.7%
	South Atlantic	39,182	16.0%
	West North Central	5,448	2.2%
	West South Central	40,842	16.6%
	TOTAL	245,641	100.0%
Number of Employees	1 to 4	150,014	61.1%
	5 to 19	66,100	26.9%
	20 to 49	11,879	4.8%
	50 to 99	3,897	1.6%
	100 to 499	2,186	0.9%
	500+	367	0.1%
	Unknown	11,198	4.6%
	TOTAL	245,641	100.0%
MSA Status	Rural	30,629	12.5%
	Urban	215,012	87.5%
	TOTAL	245,641	100.0%
Industry	Construction	30,748	12.5%
	Manufacturing	16,517	6.7%
	Trn/Cm/Util	9,987	4.1%
	Wholesale	21,083	8.6%
	Retail	64,610	26.3%
	Insurance/Re	12,468	5.1%
	Services	89,620	36.5%
	Mining	608	0.2%
	TOTAL	245,641	100.0%

APPENDIX A: DEMOGRAPHIC PROFILES OF DMI MINORITY PARTITIONS

APPENDIX B: ADVANCE MAIL PACKAGE

APPENDIX C: INCOME TAX FORMS