

Michael G. Farrell and Carl A. Konschnik, Bureau of the Census

Thomas B. Jabine, Committee on National Statistics

I. Introduction

Under the auspices of the Administrative Records Subcommittee of the Federal Committee on Statistical Methodology, an inter-agency working group is reviewing industry coding systems used by Federal agencies to classify establishments and other economic units for statistical purposes. By industry coding systems are meant the methods and procedures for assigning industry codes, rather than the technical aspects of constructing a classification scheme. The objective of this Industry Coding Working Group is to review and document the existing industry coding systems with a view toward ultimately improving the comparability and quality of the data classified by industry. This paper describes the activity of the Working Group, gives the status of the work at this time, and presents some preliminary findings.

The Working Group's effort is responsive to at least two recommendations made by a predecessor group, the Subcommittee on Statistical Uses of Administrative Records, which also worked under the auspices of the Federal Committee on Statistical Methodology. In its final report (OFSPS, 1980), that Subcommittee recommended:

"Recommendation 2. - The quality of administrative records to be used for statistical purposes should be evaluated systematically to determine the appropriateness of the records for the proposed use."

"Recommendation 3. - Consistent procedures should be used in administrative and statistical data collection efforts for defining reporting units, identifying and coding reporting unit characteristics, and developing standards for data tabulation."

These recommendations apply with particular force to industry classification and coding, because of its widespread use in economic statistics programs.

In order to get some idea of the magnitude of the industry code assignment by the Federal government, consider the following. Annually, the Internal Revenue Service (IRS) assigns industry codes to over 13 million business units as part of its revenue processing of the tax returns. Additionally, more than 200,000 units are coded for the IRS Statistics of Income Program. Similarly, the Social Security Administration (SSA) assigns industry codes to over 900,000 new business units each year, with most of these (an estimated 875,000) coded in the Single-unit Employer Identification (EI) Number File coding operation.

As part of the Employment Security Program, the Bureau of Labor Statistics (BLS) maintains an industry-coded file of about 4.5 million units. Each year about 500,000 new units are coded, and codes are reviewed and updated, where appropri-

ate, for about one-third of the existing units.

At the Census Bureau, as part of the Annual Company Organization Survey, over 900,000 establishments of multi-unit firms have their codes reviewed, and changed if appropriate, while about 75,000 new multi-unit establishments are industry coded. In addition to this, about 50,000 new business births are coded each year. For the quinquennial Economic Censuses, the Bureau mails census forms covering about half of the total universe of 6.7 million establishments in scope to the Censuses. This mailing covers all large employers, based on payroll cutoffs, and a sample of the smaller employer establishments. No establishments coded out-of-scope of the Censuses are included, except through periodic Census evaluation surveys of the out-of-scope universe (e.g., for the Finance, Insurance, and Real Estate Group). (Over time the nonmail segment of the Censuses has grown because of respondent burden and cost considerations.) All returned forms are reviewed for correct industry classification, initially via computer checks on the response provided, followed by analyst review where appropriate. Also, as part of the Censuses, another 200,000 or so unclassified establishments are coded via a classification form mailing.

The figures cited account for a substantial percentage of the volume of industry coding done either by, or under the auspices of, the Federal government. However, this is not the whole picture, as will be apparent when the Working Group's final report is issued.

It is also easy to see that the costs associated with industry code assignment are substantial. No attempt has been made here to quantify them. This would indeed be difficult, since the industry coding is a necessary (and in many instances a small) component of the overall administrative or statistical work which is being conducted concurrently.

Industrial classification of identical or overlapping populations of economic units by different agencies has led to problems of comparability for analysts and other users who try to compare and combine data from different agency sources. One example of this is in the area of productivity measurement. A recent report on this subject (Committee on National Statistics, 1979) said that "A major problem with the comparability of the basic data has been that different agencies assign the same establishments to different industry classifications, as a consequence, aggregated data at the industry level are not in fact comparable from agency to agency" (p. 178). Similar problems occur in connection with the preparation of the National Income and Product Accounts, in manpower studies, in the development of a data base for small businesses, and in other uses of economic statistics.

Several review groups have examined these problems (for example, the Central Statistical Board, 1939; the Hoover Commission, 1949; the President's Commission on Federal Statistics, 1971; the Committee on National Statistics Panel to Review Productivity Statistics, 1979; and the General Accounting Office, 1979). Without exception, they have recommended creation of a central listing of establishments and other economic units, classified by industry, which would be available to other Federal and possibly State agencies for statistical purposes. The Census Bureau's Standard Statistical Establishment List (SSEL) was in fact developed for this purpose (Bureau of the Census, 1979), but existing statutory restrictions on the release of Census Bureau information have so far made it impossible for other agencies to use the SSEL, except in a very limited sense.

At the technical level, several studies of relationships between reporting unit definitions and industry classification practices in different agency systems were undertaken by inter-agency working groups, under the general direction of the Office of Statistical Standards of the Bureau of the Budget, in the early 1950's. Several of these studies, which were begun in an attempt to account for observed discrepancies between manufacturing employment totals from the 1947 Census of Manufactures and the BLS's Current Employment Statistics, involved matching individual reports for selected companies and establishments. These studies identified numerous problems that often impaired uniform reporting, many of which were solved by the working groups or referred to the Office of Statistical Standards SIC Technical Committee for action. The work during this period showed that significant progress toward comparability could result from carefully conducted studies of the coding principles and procedures used by different agencies and their application to particular units (Bureau of the Budget, 1961).

Since that time, however, there does not seem to have been any comprehensive and detailed technical review of the existing industry coding systems: their coverage, the classification principles followed, the coding procedures, and the uses of the industry codes assigned and of aggregate data classified by these codes. The findings suitably updated should greatly assist the implementation process.

II. Description of the Industry Coding Working Group Project

A. Composition of the Industry Coding Working Group

The Working Group is made up of representatives of the Office of Management and Budget (OMB), the Internal Revenue Service (IRS), the Bureau of the Census, the Social Security Administration (SSA), the Bureau of Labor Statistics (BLS), the Bureau of Economic Analysis (BEA), the Department of Treasury, and the Committee on National Statistics (of the National Research Council). For the most part, four of these agencies - IRS, SSA, BLS, and the Census Bureau

do the great majority of the industry coding which is done by the Federal government or under Federal-State programs. These representatives met for the first time in May of 1981 and have conducted monthly meetings since then.

B. Specific Objectives

The Working Group felt that a fundamental first task was to review and document the major industry coding systems. Once this is accomplished, analysis and comparison can follow, leading to the desired improvements in the comparability and quality of the industry codes. As a further application of this work, a user or potential user of data classified by industry can be provided with essential information concerning the usability and relative quality of the data.

C. Development of the Basic Documentation for the Federal Industry Coding Systems

1. Scope of the Review

The following 17 coding systems have been included in the Working Group's review:

a. Bureau of Economic Analysis (BEA):

(1) Direct Investment Statistics

b. Bureau of Labor Statistics (BLS):

(1) Employment and Wages Program (ES-202 Report)

c. Bureau of the Census (CEN):

(1) Agriculture Census; (2) Business Births; (3) Company Organization Survey; (4) County Business Patterns; (5) Economic Censuses; (6) Population Census and Current Population Survey (CPS)

d. Federal Trade Commission (FTC):

(1) Quarterly Financial Report

e. Internal Revenue Service (IRS):

(1) Statistics of Income (SOI) - Sole Proprietorships; (2) SOI-Partnerships; (3) SOI-Corporations; (4) Revenue Processing (RP)-Sole Proprietorships; (5) RP-Partnerships; (6) RP-Corporations

f. Social Security Administration (SSA)

(1) Single-unit Employer Identification (EI) Number File; (2) Multi-unit EI Number File

The systems selected for review include some used only for statistical purposes (e.g., all Census systems) and some that were established primarily for non-statistical uses (e.g., the IRS revenue processing systems). Most of them have broad coverage in terms of Standard Industrial Classification (SIC) divisions; however, there are some exceptions, such as the Agriculture Census system. All are of a more or less permanent character, i.e., the universe or a sample of it is coded periodically, or the coding is continuous in support of accretions or changes to a cumulative file. Most systems have a relatively large volume of coding, and together they are believed to account for a substantial proportion of the industry coding of establishments and other business units that is done by the Federal Government (and by State agencies under Federal-State cooperative programs). One of the systems reviewed - the Population Census and CPS - assigns industry codes to individuals based on information they provide about the establishments in which they are employed. The Working Group's primary interest in this system is in how these codes

compare with those based on information provided by employers.

It was necessary to distinguish between an industry coding system and the principal file in which the codes reside. To illustrate this, most industry codes assigned to establishments by the Census Bureau are placed in the Standard Statistical Establishment List (SSEL). (Industry codes assigned to agriculture establishments during the agriculture census processing are not placed in the SSEL, while agricultural services codes are.) However, the separate industry coding activities done at various times and based upon different source documents are treated as separate industry coding systems.

2. The Industry Coding Questionnaire

a. Content

The Working Group constructed a questionnaire on industry coding which requested basic information needed to assess the systems and make comparisons among them. This questionnaire covered the following main areas: (1) the basic coding unit (the unit to which an industry code is assigned), the source or source document from which the coding is done, and the industry classification system used; (2) the volume, timing, coding procedures, resource material used, and quality measures associated with the coding; (3) general characteristics of the principal file(s) in which the codes reside; (4) updating of the codes and recent or planned changes to the coding system; and (5) the uses and users of the industry codes.

Within each of these areas specific questions were asked. Also, related documentation was requested, principally the forms or source documents from which the coding is done, code lists and instructions concerning classification system variations, and any available data bearing on the quality of the coding.

b. Method of Administration

Members of the Working Group identified industry coding systems within their own agencies which fit into the scope of the review. At the same time, they identified key persons who were most knowledgeable about each coding system. The survey questionnaires were then delivered to these respondents by the Working Group members.

Each completed questionnaire was reviewed by one or more members of the Working Group and a meeting was arranged with the respondent for clarification or further information. As a result of the meeting, responses to the questionnaire were revised, and frequently additional documentation of the system was obtained.

3. System Descriptions

A summary system description is prepared from each questionnaire and the associated materials. These descriptions are designed to put the collected information in a standardized, concise format for easy reference, comparison, and analysis. It is expected that the collection of these summary descriptions will form the basis of the Working Group's final report.

4. Status as of September 1, 1982

As of September 1, 1982, responses to the

questionnaires have been prepared for most of the systems. Most of these have undergone the follow-up review. Based on this review, six summary descriptions are either final or very near to final. These are: (1) BEA - Direct Investment Statistics; (2) BLS - Employment and Wages Program (ES-202 Report); (3) FTC-Quarterly Financial Report; (4) IRS - SOI for Sole Proprietorships; (5) SSA - Single-unit EI Number File; (6) SSA - Multi-unit EI Number File. Drafts have been prepared for several of the other systems. The work on producing final summary descriptions for the remaining systems is proceeding.

D. Analysis

Once the data base consisting of the final systems descriptions is reasonably complete, the analyses by the Working Group will proceed in three stages. These stages are: (1) comparative analysis of characteristics of industry coding systems; (2) descriptions of systems relationships; (3) analysis of comparability and quality of information in the systems studied. Each of these stages is described below.

It is expected that these analyses will provide a stimulus to the agencies maintaining the systems to make changes aimed at increasing comparability with other systems and at improving the accuracy of codes and reducing the cost of coding in their own systems. In addition, the information developed will make possible a technical evaluation of possible new arrangements for inter-agency code sharing, subject to legal restrictions on such exchanges. Finally, the results should help users of data from these systems to understand their structure and limitations and the extent to which data from different systems are comparable.

1. Comparative Analysis

An initial step is to identify the system characteristics or dimensions to be compared. The primary dimensions that have been identified are:

a. Coverage. Systems coverage has three sub-dimensions which can be described by the answers to three questions: (1) What kinds of units are coded? (2) Which of these units are included in the target population? (3) Is coding for all units or for a sample?

Concerning the first aspect of coverage, the kinds of units that are classified by industry vary widely. The Standard Industrial Classification (SIC) was developed for classification of establishments by industry. Its offshoot, the Enterprise Standard Industrial Classification (ESIC) was developed for classification by industry of enterprises or companies, many of which may consist of two or more establishments (Office of Management and Budget, 1972, 1974; and Office of Federal Statistical Policy and Standards, 1977b.)

The second aspect of coverage is to identify which of the specified units are included in the target population for the system. The five principal criteria are: (1) geographic location; (2) legal form of organization; (3) presence of employees; (4) SIC division; and (5) size.

The third aspect of coverage is whether or not sampling is used. If it is, the particular sample design will affect the frequency with which coding will be required and the potential for sharing industry codes with other systems. Examples of sample-based systems are the IRS Statistics of Income systems, the FTC Quarterly Financial Report system, and the Census Bureau's business birth coding system.

b. Frequency and Timing of Initial Coding and Updating. The extremes of this dimension can be represented by the IRS revenue processing coding systems and the SSA Single-unit EI Number system. In the IRS systems, industry codes are assigned annually to businesses reported on tax returns, without reference to prior year codes. In the SSA systems, each covered employer is assigned an industry code at the time of entry into the system, which occurs when the employer applies for an Employer Identification (EI) Number. This code is generally retained in the system unless and until updated by matching against Economic Census codes for the employers in the file. Many systems lie somewhere in between these extremes.

c. Classification System Used. All of the systems studied use a classification scheme based on the SIC. The systems which classify groups of establishments, e.g., the IRS systems for corporations, use systems based on the ESIC, which in turn ties into the SIC.

Nearly all of the systems vary from the 4-digit SIC or ESIC classifications in the level of detail shown. The IRS systems, in particular, combine many categories at the 3 and 4-digit levels. The BLS and SSA systems, with minor exceptions, classify units to the full 4-digit SIC level. For the Economic Censuses and the Census Bureau's annual Company Organization Survey, the Census Bureau Industry and Product Classification (IPC) provides additional detail at the 6-digit level, principally for the wholesale, retail and service divisions.

All of the systems have conformed to SIC revisions; in addition, many of them have introduced other changes from time to time, usually in the direction of showing more detail.

d. Classification Principles. Given the general principle of adherence to the SIC, there remain several conceptual issues to be dealt with in order to develop the procedures to classify establishments or other units by industry (Simmons, 1953). These include: (1) classification of units with multiple activities; (2) the length and definition of time interval on which the classification is based; (3) the relationship between the reference period for code determination and the period or periods to which data collected and assigned the code refer; (4) the effect of previous classifications on the current classification; and (5) classification of certain ancillary or auxiliary activities, such as central administrative offices, manufacturers sales branches, laboratories, and warehouses.

e. Information Used as Input to Coding. There is a wide variation in the kinds of information used in different systems to classify units by

industry and assign codes that are associated with these units in agency data files. These inputs to coding can first be divided into two broad categories: codes and other information. (1) Codes used as input to coding systems come from several sources: codes supplied by respondents, i.e., the units which are being classified; codes supplied by other agencies; prior codes assigned to the units by the same agency; and codes available from commercial directories, such as Dun and Bradstreet, Standard and Poor's, or State industrial directories. (2) Information, other than codes, supplied by the units to be coded. Identification information can be used in two ways. In some cases a unit can be completely or partly coded on the basis of its name, for example, A's Gasoline Station, B's Diner, or C's Beauty Shop. Secondly, the name and other identifiers can be used to locate industry codes in commercial directories, as mentioned earlier.

Other information used for coding may vary from a short description of the unit's principal activity, product or service, all the way to a listing, in absolute or relative terms, of gross receipts from each of the unit's products or services. Information on related subjects, such as type of organization, kinds of raw materials used, type of facility, and type of customers or clients, is often requested. Some of the items on source documents have closed response categories, some are completely open, and others are mixed. An example of the latter is the principal activity item on the form used by the Census Bureau to code previously unclassified units prior to Economic Censuses. Check boxes are provided for the most common activities; other activities are to be described by the respondent.

f. Coding Procedures. This category or dimension covers several different kinds of activities. First, in many systems it is necessary to distinguish between initial coding and updating; these may be done quite differently. In either case, there are several steps to be considered. Not all of the steps listed are relevant for each of the systems studied: (1) manual coding and review; (Sometimes the manual coding process is computer assisted, for example, by having the coder select a key word or phrase from the activity description on the source document and keying it on a terminal, which causes an associated CRT to display the titles and codes associated with that particular key word.) (2) data entry; (3) computer coding; (4) quality control of manual operations; (5) computer consistency checks; (6) treatment of missing data.

In the comparative analysis, all of the systems studied will be systematically compared with respect to each of the dimensions that have been identified. It is hoped that these comparisons may suggest improvements in some systems and help to evaluate proposed code sharing arrangements or code assignment techniques.

2. Descriptions of Systems Relationships

In considering possible new code sharing arrangements it is useful to know something

about the linkages that already exist among the industry coding systems that are being reviewed by the Working Group. These are of two kinds: intra-agency and inter-agency. Intra-agency linkages are not inhibited by legal restrictions; technical and operational factors determine their feasibility and desirability.

Inter-agency linkages are subject to legal restrictions. In general, codes residing in files of statistical agencies cannot be transferred to other agencies for non-statistical purposes. There are also severe restrictions on inter-agency transfers of industry codes from administrative systems for statistical purposes. Nevertheless, some transfers are permitted and do occur.

In this phase of the analysis the Working Group, in addition to identifying existing inter-system linkages, will review some of the technical problems that have been encountered in inter-agency linkages, such as differences in reporting unit definitions and classification systems.

3. Analysis of Comparability and Quality of Industry Codes in Different Systems

a. Comparability

One way to define comparability is to ask: if the input data for each coding system were complete and accurate and the coding procedures were followed without error, to what extent would the codes in any two systems be identical for units covered in both systems? A somewhat weaker requirement is imposed if the question is whether codes in one system can be converted to codes in another system. This can be done: either by simple transformation of numeric codes for units that match, or by combining units in one system to make them correspond to units in another system, assigning codes to the combined units, and then making any necessary transformations of numeric codes. The extent of comparability between systems depends on several of the characteristics or dimensions described earlier. In particular, it will be a function of: definition of the units coded; reference periods on which the current codes are based; the degree of adherence to the SIC categories and their definitions; the bases for classifying units with multiple activities; the resistance rules, if any, used in updating codes; and the kinds of information (e.g., product and service detail) available for each unit to be coded.

There are many other differences that affect the comparability of codes assigned in different systems; these will be systematically reviewed in this section of the Working Group's final report.

b. Quality

In broad terms, quality of data is often viewed as having three main aspects: relevance, timeliness and accuracy. Each of these will be examined by the Working Group, but major emphasis will be placed on accuracy.

In looking at accuracy, it may be useful to think both about the accuracy of individual codes and about the errors in data classified by industry based on the codes in a particular file. Both aspects are important; errors in codes assigned to units that are small in terms of

employment, payroll, gross receipts, etc., clearly have less impact on the quality of aggregate data than errors in coding larger units. However, most of the information available to the Working Group relates to errors in codes rather than the effects of these errors on data classified by industry.

Information about the accuracy of industry codes is fragmentary, at best. Direct measures of accuracy are especially hard to come by. The kinds of information available to the Working Group can perhaps be put into four different groups: (1) Comparisons of aggregate data by industry from different systems, (examples of such comparisons appear in Office of Management and Budget (1961), Office of Federal Statistical Policy and Standards (1977a, 1980), and Small Business Administration (1982)).

(2) Comparisons of industry codes for individual units from different systems. One example is a comparison of Bureau of the Census and Bureau of Employment Security (BES) payroll data for 1963 for the retail SIC division in the State of Delaware (Bureau of the Census, 1965b). A two-way match was conducted; a sample of 1963 Economic Censuses establishments was matched against BES records, and a sample of BES reporting units (groups of establishments) was matched against Census records. All matched establishments and reporting units with classification differences (retail vs. other) were reviewed jointly by Census and BES staff and, where necessary, telephone calls were made to arrive at agreed-on final classifications. Coding into the wrong SIC division was found to explain a substantial proportion of the differences that occurred. Another example is the Employer Record Check conducted by the Bureau of the Census (1965a) following the 1960 Census of Population. For a sample of individuals, the industry codes assigned to them on the basis of information supplied by them (or a member of their household) in the Population Census were compared with industry codes from the SSA single-unit and multi-unit files for the establishments or reporting units where they were employed. Various measures of accuracy were computed. The report suggests that it may be useful to take the SSA code for each person as the standard in interpreting the observed results; however, it recognizes that these codes are also subject to error. The Working Group has identified only a few studies involving inter-agency micro-comparisons; the findings from these studies will be summarized and their implications discussed in the final report.

(3) Information about components of error in individual coding systems. Errors can occur at various stages in arriving at the final codes for units in a particular system. The Working Group's data base will include any available quantitative information about these components of error for each of the industry coding systems reviewed. This information will provide a partial "error profile", along the lines illustrated by the Office of Federal Statistical Policy and Standards (1978), for each system.

(4) Comparison of system characteristics. The system characteristics that are most

relevant to accuracy include the sources and kinds of information used for coding, the treatment of units with incomplete data, the basic coding procedures (manual or computerized), the experience of coding personnel, the quality control of manual operations, the computer checks and correction procedures, and (if applicable) the updating procedures. Information on all of these characteristics has been obtained for each industry coding system and will be reviewed and analyzed in this phase.

III. Future Directions

The development of a data base consisting of detailed descriptions of major Federal and Federal-State industry coding systems has been a substantial task. It has required considerable effort by the Working Group members and could not have been done without the contributions of those other employees of the agencies involved who completed questionnaires, provided documentation, answered follow-up inquiries and reviewed drafts of systems descriptions.

Because of the effort invested, the Working Group feels an obligation to make the additional effort needed to do a careful and thorough analysis of the information obtained and to present its findings and recommendations in a timely fashion and in a form that will be most useful and accessible to the agencies involved and to the users of data classified by the industry codes in these systems. This should have a substantial payoff in making more effective use of administrative records for statistical purposes.

[A fuller version of this paper, including tables and a sample system description, is available from: Carl Konschnik, Business Division, Bureau of the Census, Washington, D.C. 20233]

Acknowledgments

The authors would like to note the many contributions to this work by members of the Working Group. Members of the Industry Coding Working Group are: Carl A. Konschnik, Census Bureau, Chairman; Bruce Levine, Bureau of Economic Analysis; Susan Hostetter, James Millette and Michael Searson, Bureau of Labor Statistics; Thomas B. Jabine, Committee on National Statistics; James R. Hobbs, Beth Kilss and Linda B. Taylor, Internal Revenue Service; Barbara Haas, Department of the Treasury; and Linda M. Dill, Social Security Administration. Maria E. Gonzalez, Chairman of the Federal Committee on Statistical Methodology and Fritz Scheuren, Chairman of the Administrative Records Subcommittee are ex-officio members.

R E F E R E N C E S

Bureau of the Budget:

1961--Brief history of the movement in the Federal Government for a central directory and of related efforts aimed at improving quality and comparability of economic statistics.

Office of Statistical Standards, unpublished report.

Bureau of the Census:

1965a--Evaluation and Research Program of the U.S. Censuses of Population and Housing 1960. The Employer Record Check. Series ER60, No. 6.

1965b--Final results of BES-Census retail payroll reconciliation for the State of Delaware. Memorandum from Peter Ohs and Ralph Woodruff to Harvey Kailin and William Hurwitz, July 22.

1979--The Standard Statistical Establishment List Program. Technical Paper No. 44

Committee on National Statistics:

1979--Measurement and Interpretation of Productivity. Report of the Panel to Review Productivity Statistics. Washington: National Academy of Sciences.

Central Statistical Board:

1939--Annual Report for Fiscal year 1939. Washington, D.C.

Commission on Organization of the Executive Branch of the Government (Hoover Commission):

1949--Task Force Report on Statistical Agencies (Appendix D). January. Washington, D.C.

General Accounting Office:

1979--After six years, legal obstacles continue to restrict government use of the Standard Statistical Establishment List. Report to the Congress by the Comptroller General of the United States.

Office of Federal Statistical Policy and Standards:

1977a--Gross National Product Data Improvement Project Report. Report of the Advisory Committee on Gross National Product Data Improvement.

1977b--Supplement. Standard Industrial Classification Manual.

1978--An Error Profile: Employment as Measured by the Current Population Survey. Statistical Policy Working Vol. 3, by Camilla A. Brooks and Barbara Bailar.

1980--Report on Statistical Uses of Administrative Records. Statistical Policy Working Paper 6.

Office of Management and Budget:

1972--Standard Industrial Classification Manual. Statistical Policy Division.

1974--Enterprise Standard Industrial Classification Manual. Statistical Policy Division. President's Commission on Federal Statistics:

1971--Federal Statistics: Report of the President's Commission. Washington. Government Printing Office.

Simmons, Walt R.:

1953--The Elements of an Industrial Classification Policy. Journal of the American Statistical Association, Volume 48, Number 263; 429 - 439.

Small Business Administration:

1982--The Annual Report on Small Business and Competition of the Small Business Administration (in The State of Small Business: A Report to the President, transmitted to the Congress, March 1982). Appendix B. The small business data and other sources of business information: recent progress.