

## RECONCILIATION OF THE ECONOMIC CENSUSES RESULTS AND CURRENT SURVEYS PROGRAM

Malcolm E. Bernhardt and Sol D. Helfand  
U.S. Bureau of the Census

### Introduction

The Bureau of the Census collects and publishes basic economic data about businesses, industries and the professions. These reports are the primary source of facts about the structure and functioning of the Nation's economy. They provide information essential for government, business, industry, and the general public. The programs include monthly, quarterly, and annual surveys, and quinquennial economic censuses. These statistics are widely used in evaluating current economic conditions and in forecasting future developments. The data also provide key inputs into other significant economic series, particularly the estimates of the gross national product.

Historically, comparisons of census and survey data have been made after each of the economic censuses. The last comparison covered the 1972 Economic Censuses and is described in (1). This paper describes the reconciliation of the 1977 Economic Censuses with comparable current survey estimates for some major economic series, including manufacturing, retail trade, wholesale trade, and service industries.

The reconciliation had a number of objectives:

- To measure the extent of differences between the censuses and current surveys in coverage, classification, and data;
- To determine why the differences occurred;
- To identify systematic errors, make corrections during census and survey processing, and take steps to minimize similar future errors;
- To identify random errors, including reporting errors, and make corrections, thus improving the annual and census estimates;
- To improve coverage in the manufacturing survey through the addition of births;
- To improve the quality of the annual and census levels used for benchmarking current survey estimates;
- To serve as a guide in planning for future surveys and censuses.

In addition to meeting the objectives of the reconciliation, the Census Bureau has had a series of other programs designed to evaluate the quality of the data and to improve the surveys and censuses. Some of the programs are described in the final section of this paper. The most significant of these programs overcame most of the deficiencies of the current surveys sampling

frame identified in previous reconciliations. This was accomplished by developing a common frame for the Bureau's economic surveys and censuses.

The primary component of the universe for economic surveys and censuses is the Census Bureau's Standard Statistical Establishment List (SSEL)--which effectively covers all employers who made social security payments for their employees under the Federal Insurance Contributions Act (FICA). For the 1977 Censuses of Retail Trade and Service Industries, data for all nonemployers, i.e., firms with no paid employees, were obtained from 1977 Federal income tax records. In the current surveys of retail trade and service industries, the primary sample component is the list sample--a probability sample of employer firms selected from the SSEL. The secondary component is the area sample, a probability sample of land segments which represents businesses without employees, and employers--mainly recent "births"--not represented in the list sources.

The balance of this paper describes the reconciliation for retail trade, wholesale trade, and service industries; the reconciliation for manufacturing; and conclusions and summary.

### Retail Trade, Wholesale Trade, and Service Industries

"Reconciliation of two independently conducted surveys such as the Census of Retail Trade and the current retail trade sample survey inevitably involves a costly, time consuming, and often frustrating attempt to match individual business units identified by descriptions obtained from different sources at different times by different people under different procedures. Moreover, the matching must be done with extreme care if the results of the reconciliation are to be meaningful".

This quote is as pertinent today as it was about 20 years ago when it appeared in Bureau of the Census Technical Paper No. 9, (8).

The fact that differences exist between the Economic Censuses and the current sample surveys is not unexpected. While the Censuses of Retail Trade, Wholesale Trade, and Service Industries and the corresponding current sample surveys attempt to measure the same universe of distributive trade and service industries, the methods are different. The results are determined in part by the sources of information, rules of classification, and procedures available to each. In addition, some differences can be expected between current estimates, based on a sample, and a full census, because of inherent sampling variability in estimates derived from the selected sample and because of conceptual differences in coverage.

There were significant differences between the census and survey data which were reflected in an understatement of total sales and receipts in the monthly series for 1972. Based on classifications from the 1967 Standard Industrial Classification (SIC) Manual (4), data for 1972 from the Economic Censuses were more than current survey estimates (the sum of the 12 months of 1972) by 4.8 percent for retail trade, by 7.8 percent for wholesale trade, and by 8.0 percent for service industries. In each case these differences exceeded the amounts expected due to sampling variation in the monthly surveys. Part of the differences were attributed to improved coverage in the 1972 Economic Censuses through more effective use of administrative records. Other differences resulted from deficiencies in the monthly series, such as an incomplete and outdated sampling frame, misclassification of firms, and shortcomings of the sampling units.

A new sample design was launched in 1975 and introduced in November 1977, with separate historical reports for retail trade, wholesale trade, and service industries. (5), (7), and (6). This program represented the most ambitious sample revision ever undertaken by the Bureau's Business Division, and involved more than two and a half years of intensive effort. The revision is described in (3) and (9). The new design is believed to have remedied nearly all of the deficiencies noted above. Most of the shortcomings of the sampling frame were eliminated. Increased sample efficiency under the new design permitted a reduction in the size of the sample and, at the same time, reduced the levels of sampling variability for most kinds of business.

Plans for a reconciliation study covering the 1977 Economic Census were begun in 1977. A decision was made to start the work in 1978 so that census errors discovered during the match of census and current data could be corrected during census processing. Earlier reconciliation studies were conducted after the census results were published. When a census error was identified, a correction was published if there was a significant impact on the published estimates.

The reconciliation plan which called for coverage checks and data matches was not fully implemented. The plan for coverage checks consisted of computer matches of census and sample records to determine potential undercoverage. Except for some spot checks which showed no coverage problems in the census or sample, the coverage operation was dropped. Since the primary components of both the census and sample universe came from the same source, the SSEL, it was assumed that differences caused by coverage would be negligible.

Since the decision was made to bypass the coverage checks, the reconciliation was limited to a data match whereby sales (receipts) and wholesale inventories reported by firms in the current monthly and annual sample surveys were matched to comparable data collected in the economic censuses. The data match was restricted to the large firm or "certainty" component of the current samples representing about 3,500 companies. For sales and receipts, there was a three-way match:

(a) the sum of 1977 monthly survey data; (b) data collected in the 1977 annual surveys; and (c) of survey and census end-of-year inventories for wholesale trade.

With only about 3,500 companies subject to the data match, comparisons were made manually rather than by computer. Differences between census, annual, and the sum of monthly data were measured. A few companies reported the same figures in the surveys and censuses or had differences which were negligible. For many companies, however, differences amounted to millions of dollars. Experience from prior reconciliation studies had shown that data resolution requires an intensive effort by the analysts and clerical staff. A decision was made to concentrate our investigation on companies with the largest differences. Cutoffs, based on annual sales or receipts were set as follows: for wholesale, the approximately 225 companies with differences of \$100 million or more; for retail, about 70 companies with differences of at least \$50 million; and for service, about 35 companies with differences of \$25 million or more.

Targeted companies were multiunit firms (those operating establishments at more than one location). Therefore, the first step consisted of determining if the company was reporting for the same establishments in the census and in the current surveys. In the census, large multiunit firms report separate data for each of their establishments; in the annual surveys, a consolidated company report is requested. In the monthly surveys, consolidated company information is requested in addition to a geographic breakdown for some wholesale companies. In the monthly retail survey, large firms with many retail establishments are asked to provide sales for a sample of their establishments in order to develop sales estimates by geographic area. Firms with few retail establishments are asked to report sales for each of their establishments.

It was discovered that many of the firms reporting large differences were not covering the same establishments in the census and current surveys, or were not reporting as instructed. Reporting differences also resulted from different people completing the questionnaires, from dissimilar instructions, and from timing differences. Requirements for early reporting in the monthly surveys often result in estimates, whereas the data for the annual surveys and censuses are usually based on accounting records. Similar problems are also found in the reconciliation of the manufactures surveys. Some additional coverage and reporting errors were found:

1. In wholesale trade there were numerous inconsistencies between the census and current surveys in the assigned "type-of-operation" classification. The current survey is primarily limited to merchant wholesalers who take title to the goods they sell, such as wholesale merchants or jobbers, industrial distributors, voluntary group wholesalers, exporters, importers, cash-and-carry wholesalers, retail cooperative

warehouses, terminal elevators, farm products assemblers, and wholesale cooperative associations. The Census of Wholesale Trade also covers other types of operations, including: (a) sales branches and sales offices maintained by domestic manufacturing or mining enterprises apart from their plants or mines for the purpose of marketing their products; and (b) agents, merchandise or commodity brokers, and commission merchants.

2. Kind-of-business classifications differed in the census and current surveys for some establishments. Differences between retail trade and wholesale trade were particularly common for lumber and other building materials dealers and for paint, glass, and wallpaper stores. Establishments are classified as retail if they sell to the general public even if sales to contractors account for a larger proportion of total sales; these establishments are known as "retail" in the trade. Establishments primarily selling these products but not selling to the general public are classified in wholesale trade.

3. Central administrative offices and auxiliary locations (captive warehouses, and other units which service or are not classified in the industry of the operating units, are auxiliary to the other establishments within the same organization). Some retail, wholesale, and service reports, particularly in the current surveys, included data for their auxiliary locations.

4. Retail firms are asked to include sales for departments and concessions operated by other firms in their retail stores, but to exclude sales for departments and concessions operated by their firm in retail stores of other firms. The census and annual retail questionnaires have specific inquiries about departments and concessions, but the monthly forms only refer to them in the general instructions for reporting sales. As a result, reporting of sales for departments and concessions was not consistent for some companies.

5. A number of firms have franchised operations as well as company-owned stores. Where franchised stores are independently owned and operated, data are requested only for the company-owned stores. In some cases, duplicate sales were reported for franchised stores by the franchisor.

6. Businesses are asked to report sales and receipts excluding sales and excise taxes collected directly from customers and paid directly to a local, state, or federal tax agency. Sales and receipts should include excise taxes, such as those on gasoline, liquor, and tobacco, which are levied on the manufacturer or wholesaler and included in the cost of goods purchased by the company. A few errors in tax coverage were identified. However, such differences were far fewer than in previous reconciliations, when census and current survey instructions provided for the inclusion of all such taxes in sales and receipts.

Estimates by the type of error are not available. Differences were most significant in wholesale trade. A match of the "certainty" wholesale companies showed gross differences between the sales from the annual survey and census sales of \$86 billion. Most of the errors were caused by coverage differences or by inconsistencies in the assigned type-of-operation classification.

If the errors were the result of incorrect reporting on the census or annual questionnaires, corrections were made during the reconciliation. As a result of these corrections, differences between the published census and survey data were minimized. For example, the 1977 Census of Retail Trade published sales of \$723 billion. A coverage adjustment of +\$2 billion was made to reflect estimated sales of nonemployer direct sellers which are included in the surveys but excluded from the census. The adjusted census sales of \$725.2 billion were less than 0.2 percent greater than the sales estimate of \$724.0 billion published in the 1977 Annual Retail Trade Report. This is the closest the two estimates have ever been.

For the 12 months of 1977, the sum of the estimated monthly retail sales was \$708.3 billion, 2 percent below the census and annual survey sales estimates. In April 1979, a revised series which made use of "benchmark" data derived from the 1977 Annual Retail Trade Survey was issued. Revised estimates of monthly retail sales and inventories covering the period January 1967 - December 1978 were published. The series was revised using a method referred to as the "link relative" revision procedure. For a detailed description of the revision methodology, see (2).

The benchmark process is being continued and resulted in the April 1980 publication of revised monthly retail sales estimates by kind of business for the United States for the period January 1973 through December 1979. The basic series was processed in a manner designed to constrain and link the series to specified benchmark levels and a designated link point. Data for the 1977 Census of Retail Trade and the 1978 Annual Retail Trade Survey were used as the benchmark levels, and the December 1972 sales estimate from the current monthly sample was established as the "link" point.

#### Manufacturing

In the manufacturing area, the reconciliation involved the resolution of differences between the results of the 1977 Census of Manufactures and the Current Industrial Reports (CIR) series. The CIR program consists of surveys designed to collect detailed product data for select industries or categories e.g., apparel, chemicals, steel and electronics. Survey frequency may be monthly, quarterly, or annual. While the objectives of the manufactures reconciliation were basically the same as for retail and wholesale trade and service industries, there are major differences between the current surveys in the two areas which affected the procedures used in problems encountered.

1. For trade and services, the data being compared represent total sales (or receipts from operations) of all establishments of the company in the applicable industry classification; for manufactures, data represent shipments of specified products, with separate reports usually submitted by each plant.

2. For trade and services, there are three comprehensive surveys (covering retail trade, wholesale trade, and selected service industries). The manufactures program includes 70 surveys involving over 450 product group comparisons.

3. The trade and services surveys are based on complex probability samples. The manufactures surveys generally cover all known producers of the products covered, although, for many surveys, small firms for which census data were derived from administrative records are excluded.

4. Small companies are handled differently. In the Censuses of Retail Trade and Service Industries, nonemployers and most small employers were not required to file a census report. Data for these firms were obtained from administrative records. In the Census of Wholesale Trade, all employee firms were mailed a census questionnaire, regardless of size. In the Census of Manufactures, small employee companies were covered by the use of administrative records. Nonemployers are excluded from Censuses of Manufactures and Wholesale Trade.

5. The number of reporting units reconciled (approximately 30,000) far exceeded the 3,500 companies subject to matching in the trade and service programs. Because of this, the initial comparisons were done by computer through the use of a data base.

Historically, comparisons of census and survey data have been done as part of the review of the census of manufactures so that errors discovered during the match of the two programs may be corrected during census processing. Census data comparisons are also made to other government agency data, and to trade association data. During the intercensal years, the CIR surveys are also reconciled to the Annual Survey of Manufactures (ASM). However, since the ASM is based on a sample survey, the intercensal reconciliations are limited to comparisons of data for establishments included in both programs.

Each of the 70 CIR surveys involved in the reconciliation with the 1977 Census of Manufactures covers specific products representing an average of six product groups per survey. Rather than requesting the same data in both programs, detailed product data were collected only in the CIR series; in the census of manufactures, aggregate product data were collected. These aggregates are typically for product classes, which are groups of similar products (e.g., data for cathode ray tubes (TV tubes), were collected as a one line item in the census of manufactures, whereas, in the CIR series, data were collected by type and/or size of tube.

Since a reporting unit may manufacture items included in a number of product groupings, the number of matches multiply. Because of basic

differences in the census and CIR programs, which will be discussed below, matching each product group 100 percent would be a virtually impossible task.

Consequently, the following "tolerance" ratios for reconciling value of shipments data were applied:

Size of Product Group	Tolerance
Under \$250 million	.90 - 1.10
\$250 - \$499 million	.92 - 1.08
\$500 - \$999 million	.94 - 1.06
\$1 - \$2.5 billion	.96 - 1.04
Over \$2.5 billion	.98 - 1.02

The reconciliation was done by survey and, within survey, by product groups with differences resolved and corrected wherever possible. The majority of matches (over 275 groups) reconciled only dollar value. However, for some product groups quantity data were also available from each source. Since quantity data may be collected in a variety of measures, reconciliation criteria were determined on a survey-by-survey basis. However, to the extent possible, tolerance measures were the same as those used for value of shipments.

Shown below is an illustration of a Current Industrial Report (MA-36A, Switchgear and Switchboard Apparatus) that had been reconciled.

Comparison of data reported in the 1977 Census of Manufactures (Census) and the Current Industrial Reports Series (CIR) for switchgear and switch board apparatus was as follows:

#### SHIPMENTS OF SWITCHGEAR AND SWITCHBOARD APPARATUS, 1977

(Millions of Dollars)

Product Group Code	Product Group Title	Originally Reported			Revised		
		Census	CIR	Census + CIR	Census	CIR	Census + CIR
3613100	Switchgear, Except Ducts & Relays	834.7	814.2	1.025	841.4	845.1	1.004
3613200	Power Circuit Breakers, All Voltages	280.4	248.0	1.131	261.2	260.5	.997
3613300	Low Voltage Panelboards, 750 Volts & Under	841.6	856.7	.982	878.2	873.6	.995
3613400	Fuses and Fuse Equipment, Under 2300 Volts	303.5	169.9	1.786	201.6	189.0	.938
3613500	Molded Case, Circuit Breakers, Under 750 Volts	375.5	481.5	.780	492.0	511.1	1.039
3613600	Duct, Plug-in Units, Accessories, Under 750 Volts	96.8	84.1	1.151	88.5	84.0	.949
3613700	Relays, Control Circuit	320.6	350.5	.915	349.3	344.0	.985
3622000	Industrial Controls	2384.1	2119.1	1.125	2429.1	2381.0	1.020

This survey concerned shipments of products relating to the electrical switchgear industries, such as common household circuit breakers, industrial switching mechanisms, and electrical controls used in a variety of manufacturing processes. Because of the wide range of applications, units of quantity were meaningless; thus only dollar figures were collected on this survey, and on the census of manufactures.

The types of problems which were the source of discrepancies were of three types: (1) misclassification of data, (2) reports being completed by different individuals, and (3) small companies not being mailed census forms. These three items are discussed on the following pages.

In addition to the general problems of coverage, timing, and dissimilar instructions which were described earlier, a number of problems were identified which are unique to the manufactures programs. Among the most significant were the following:

1. Administrative Records--As mentioned earlier, many small companies were not mailed report forms for the census of manufactures, but were included through the use of administrative records. These administrative record sources (Internal Revenue Service and the Social Security Administration) provide data only at the industry (4-digit SIC) level. Consequently, there was no way of determining in which of the product groups the companies shipments should be included. Many of these small firms are included in the CIR surveys. In the entire census of manufactures, as well as for those industries involved in the reconciliation, administrative records accounted for 1.7 percent of shipments. The largest contribution by industry was 6.6 percent. A few industries, typified by large establishments, had no administrative records.

2. The general problem of reports being completed by different individuals was a major cause of differences between the census and CIR reports. In cases of this type, the census reports were typically completed at the company headquarters, where the requested general statistics (employment, payroll, workhours, etc.) often are available centrally for all establishments of the company. In contrast, the CIR report may request only detailed product information, which may be more accessible from plant records. In such cases, the broad product groups reported by the headquarters often were inconsistent with the product detail reported in the CIR by the plant.

3. Even when reports were completed by the same individual, proper classification of products was a major problem. Some respondents did not relate the product detail in the CIR to the aggregate product groups for which data were reported in the census of manufactures.

4. The census of manufactures permits the reporting of shipments as "not specified by kind" (n.s.k.). This pseudo-classification is used both for reporting shipments of products which the respondent is unable to classify and by the Census Bureau to account for all shipments, including shipments on improperly completed or delinquent reports. Because of the extensive followups of large company reports, most of the n.s.k. data represented shipments of small firms. Product information reported in the CIR was used during the reconciliation to sharply reduce the n.s.k. classification. However, the residual in this "unclassified" category remained a significant source of difference between the two programs.

#### Conclusions and Summary

The reconciliation studies identified a number of problems which can be resolved or avoided in planning future surveys and censuses. The problems identified in the 1972 study led to improvements of the current business surveys. There will continue to be improvements and sample updates, including the following:

1. Extensive planning has begun for a revision of the current survey samples for retail trade, wholesale trade, and service industries. The new samples are expected to be in operation by 1982. A major goal of the program is to redistribute the reporting burden of smaller firms by reselecting new reporting panels.

2. An attempt will be made to minimize nonsampling errors, although there will continue to be some nonsampling and response error biases.

3. To improve response rates and reporting accuracy, questionnaires are being redesigned to make them smaller, simpler, and easier to complete.

4. Procedures will be developed to provide better identification and handling of concessions, leased departments, and franchise operations to avoid omissions or duplication.

5. A strong effort will be devoted to having frequent updates of the content and coverage of multiestablishment sampling units.

6. Benchmarking the monthly series to the results of the Annual Retail Trade Survey will be continued. The annual sales estimates developed from the larger annual sample are more reliable than the sum of the monthly estimates due to the availability of "book figures."

In the manufactures area, the reconciliation provides the basis for a broadly-based program of improvements to the Current Industrial Reports surveys. Following is a summary of the main aspects of this program, which is now underway:

1. A systematic program of panel maintenance. Approximately 13,000 potential "births" were identified as a result of the reconciliation. With a panel maintenance program, most of these cases would have been identified and added to the surveys on a flow basis, or at least annually. A panel maintenance program based on the SSEL is currently being developed. The program, which will provide a systematic annual updating of the survey lists, should minimize future reconciliation differences resulting from census "births." The program will provide for plants to be added to the survey panel based on the same size criterion on which the existing panel is based. Plants which have increased in size since the last sample-update, as well as qualifying new firms will be added to the survey panel.

2. Benchmarking of CIR surveys to the census of manufactures. Because of the n.s.k. problem cited above, benchmarking will be limited to broad levels of comparison, e.g., 4-digit (SIC) industry. Since some surveys do not cover all products in the 4-digit classification, it will not be possible to benchmark every survey. This limited benchmarking operation is currently underway, with results reflected in the 1979 annual (or summary) reports for the affected series.

3. Match and Update Identification Numbers in the CIR and SSEL Lists. These differences, which usually result from a failure to report change of ownership in both files, were resolved during the reconciliation process. Future differences will be corrected annually during the panel maintenance operation. The elimination of most of these differences should facilitate future census-CIR reconciliations.

4. Development of General Purpose Reconciliation Program and Techniques. Within the CIR program are several annual CIR surveys which collect data that are also included in less or different detail in monthly or quarterly surveys. The sum of the monthly and quarterly data must be reconciled to the annual data. With minor modifications, the programs and techniques used for the census-CIR reconciliation can be applied to these inter-CIR reconciliations.

5. Development of a system by which census schedules as they are received could be matched on a reporting unit basis to similar data received in the current program. If differences in product data between the two reports appear, they can be referred and resolved during this early stage of processing. This would result in fewer corrections being made after the data are tabulated. More important, the data could be reviewed and processed on a more timely basis.

The reconciliation of the 1977 Economic Censuses and the related current surveys is part of a broader program within the Census Bureau to evaluate and improve the quality of its statistical products, to expedite the release of data, and to reduce and redistribute the reporting burden of smaller firms.

Among the other programs to evaluate and improve the 1977 Economic Censuses, was the evaluation of the census content, coverage, and data. Another program, designed to update the industrial classification of business and industrial organizations, was initiated in 1979. It involved canvassing approximately 700,000 firms to get more up-to-date information on their principal activity and the number of establishments operated by the firm. Papers describing these programs are being presented in this session.

#### REFERENCES

- ( 1 ) Eisen M., et al. "The Quality of Census Bureau Economic Time Series Indicators-- Past Performance and Planned Improvements." Proceedings of the Business and Economics Statistics Section, American Statistical Association, Atlanta, Georgia, 1975.
- ( 2 ) Helfand, S.D., Monsour, N.J., and Trager, M.L., "Historical Revision of Current Business Survey Estimates." Proceedings of the Business and Economics Statistics Section, American Statistical Association, 1977.
- ( 3 ) Isaki, C.T., et al. "Sample Redesign of the Census Bureau's Monthly Business Surveys." Proceedings of the Business and Economic Statistics Section, American Statistical Association, 1976.
- ( 4 ) Standard Industrial Classification Manual: 1967. Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.
- ( 5 ) U.S. Department of Commerce, Bureau of the Census. Monthly Retail Sales: January 1967 - August 1977 (Revised); Washington, D.C.
- ( 6 ) U.S. Department of Commerce, Bureau of the Census. Monthly Selected Services Receipts: January 1972 - August 1977 (Revised); Washington, D.C.
- ( 7 ) U.S. Department of Commerce, Bureau of the Census. Monthly Wholesale Trade - Sales and Inventories: January 1967 - August 1977 (Revised); Washington, D.C.
- ( 8 ) U.S. Department of Commerce, Bureau of the Census, Technical Paper No. 9; Reconciliation of the 1958 Census of Retail Trade with the Monthly Retail Trade Report.
- ( 9 ) Wolter, K.M., et al. "Sample Selection and Estimation Aspects of the Census Bureau's Monthly Business Surveys." Proceedings of the Business and Economics Statistics Section, American Statistical Association, 1976