Session 1

ECONOMIC CLASSIFICATION REVISIONS

Economic Classifications in the New North American Industry Classification System (NAICS)

Jack E. Triplett
Chief Economist, Bureau of Economic Analysis
Chairman, Economic Classification Policy Committee

Prepared for the Seminar on New Directions in Statistical Methodology[1]
Washington, D.C.
May 25, 1994

I. Why Has the United States Mounted this Effort?

The Economic Classification Policy Committee (ECPC), acting on behalf of the Office of Management and Budget (OMB), Statistics Canada, and Mexico's Instituto Nacional de Estadística, Geografía e Informática (INEGI), have agreed to create a new North American Industry Classification System (NAICS). Differences in the internal pressures on the three countries' respective statistical systems create corresponding differences in our motivations, and in the constraints we face. That the three countries have joined together in this effort suggests that the similarities among us may be more significant than the differences. Nevertheless, some of the pressures that have influenced classifications in the United States are not irrelevant to international as well as national discussions of classification systems.

I should first emphasize that the U.S. Standard Industrial Classification (SIC) system is quite old (it dates from the 1930's), it has been well tried in statistical agency collection programs, and it has often been revised (the last time in 1987) in an attempt to keep it up to date. Data users have had abundant experience with the industry data that this classification system provides, and they have had many years to learn of its strengths and shortcomings. Statistical agencies have had many opportunities to react to user experiences, <u>within the parameters</u> that have guided this classification system for the past 50 years.

Yet, increasing public dissatisfaction with the U.S. SIC system has been expressed through its last several revisions. Discussions of the adequacy of the SIC in the United States have occurred not merely or solely in professional exchanges between economists and statisticians, and have not occurred solely within the boundaries of narrowly technical dialogues. In the United States, focus on problems of the SIC has extended widely to the popular press and the business press.

---

[1] This paper was originally presented at a meeting of The Statistical Programme Committee (NACE), Statistical Office of the European Communities (Eurostat), Luxembourg, March 17, 1994.

Perhaps this public attention to industrial classifications is unique to the United States. To the extent that it is unique, we believe that, at least in part, it is a response to a mature industry classification system that has increasingly been viewed as inappropriate for generating the data that are needed for economic analysis.

The United States has experienced increasing private sector and government sector demands for data for the purpose of economic analysis. As one example, a major statistical initiative has in the last several years been launched [8], and a major priority within that initiative is to improve U.S. data for the measurement and analysis of productivity. We must, therefore, ask: Does the existing SIC system produce the industry data that are wanted for economic analysis, including the industry data needed for productivity measurement? These are exactly the questions that were posed for classification systems at the Williamsburg international conference on classifications [1].

II. The ECPC Examination of Classifications Starts from the Use of Economic Data

The ECPC was established by OMB in 1992 to conduct a "fresh slate" examination of economic classification systems. The ECPC began its investigation with an examination of the uses of economic data that are produced using classification systems. What is the purpose of classification systems? When economic data are produced from these systems, what are the uses of the data for which classifications are designed? The review of these issues is contained in ECPC Issues Paper No. 1, "Conceptual Issues" [3].

The ECPC's approach is a departure from the traditional approach to classifications, at least as it has developed in the United States. In the traditional view, there are many uses for data, and because there are many uses, it has been believed that the classification system must produce data for all the uses. This means, effectively, that the uses of data have relatively little ultimate role in the design of the classification system, because the requirements for different users tend to cancel each other. The traditional classification is a compromise between competing ends. The nature of these compromises is not dictated by the use of the data, nor do the designers of the classification system have a framework from which to examine the costs to the data user of the compromises incorporated into the system.

The ECPC, in common with the traditional view, also recognizes that there are multiple uses for data that are produced by classification systems. However, in contrast to the traditional view, the ECPC has concluded that the economic use of data must determine the design of the classification system; if it does so,

this will assure that the data produced by the classification system meet the intended use.

If there are alternative demands for data that must be grouped or classified, and if these alternative demands for data have different implications for the classification system, the ECPC concludes that different classification systems must be constructed. In the ECPC's view, when different requirements for classifications grow out of different data uses, this implies that different classification systems should be set up, each one designed to meet the intended need. In the traditional approach in the United States, a compromise has been sought to meet everyone's needs within a single classification system.

If tailoring classification systems to the needs for data implied a very large number of different classification systems, the ECPC's position might be impractical. We also have concluded, however, that the major analytic needs for classified data can themselves be grouped into two major classes of uses. This matter is discussed at some length in ECPC Issues Paper No. 1 [3].

Briefly, one class of uses requires that a classification system be erected on a production-oriented concept (which may also be called a supply-based concept). A second major class of uses requires that data be grouped according to a market-oriented concept (which may also be called a demand-based concept).

Thus, the major difference between the ECPC's position and the traditional one in the United States condenses to the following questions: Should there be two different classification systems, each designed to produce data for one of the two classes of uses? Or should there be only one compromise system for both classes of uses?

In international discussions of classifications, the situation is a little different from the U.S. tradition, because multiple classifications already exist. The United Nations systems include the International Standard Industrial Classification of all Economic Activities, Third Revision (ISIC), and the Central Product Classification (CPC), and Europe has Nomenclature des Activités économiques des Communautés Européennes (NACE), Classification des Produits associée aux Activités (CPA), and another system called PRODCOM.

Yet, the issues that have been debated are quite similar: What are the data uses for which different classifications are required? Should some of the classification systems (usually, the product classification system) be connected in some manner to the others (the industry classifications)? And whatever distinctions can be drawn in principle among the various systems (by reading their introductions and statements of principles, for

example), their implementations are marked by compromise. The connections between data use, classification concept, and construction of the classification system have been influenced by compromise among competing demands for data, rather than by the determination to tailor each classification system to a major economic use of grouped data.

Though seldom stated explicitly in this language, almost all of the literature discussing classification principles, whether in the United States or in international forums, can be understood as a conflict between designing a production-oriented, or supply-based, classification system and designing a market-oriented, or demand-based, classification system. That is, much discussion of classification problems concerns, at the root of the matter, the conflict between providing data for production-side economic analysis, on the one hand, and for market- or demand-oriented analysis on the other. The ECPC's issues papers, particularly ECPC Issues Paper No. 1, attempt to make more clear and explicit what has unfortunately remained implicit in much of the past discussion of economic classifications.

The distinction made in ECPC Issues Paper No. 1, however, is actually quite old. After this project was well under way, David Wharton of Statistics Canada called my attention to a very enlightening article on economic classifications published by R.H. Coats in 1925. Coats' pertinent observations for our time include the following:

> "...the basic principle in classification is that mutually exclusive concepts may not be united on an equality in the same category.... It is precisely this elementary rule that statisticians too often ignore. Called upon for statistics of aggregates from many and diverse standpoints, they attempt to meet the demand within the limits of a single classification. This leads inevitably to confusion as between principles....
>
> "...To state as was stated in the resolution originally tabled at the Geneva Conference on Labour Statisticians that a combination of principles must be adopted is surely to abandon the issue prematurely...." [2, emphasis in original].

Allowing for changes in economic language over the past 70 years, a subsequent passage of Coats' article can be interpreted as discussing production-oriented and market-oriented principles as alternative concepts for classification systems. Coats proposed also a third principle--distinguishing the stage of process in the hierarchy of the classification system.

III. Why Was the Production-Oriented Concept Chosen for NAICS?

A statement adopted by the ECPC, Statistics Canada, and Mexico's INEGI reads, in part:

> "The uses of industrial statistics which include measuring productivity, unit labor costs, and the capital intensity of production require that information on outputs and inputs be used together. Moreover, statistical agencies in the three countries expect to be called upon to produce information on inputs and outputs, industrial performance, productivity, unit labor costs, employment, and other statistics in order to analyze the effects of the North American Free Trade Agreement. An industry classification system erected on a production-oriented, or supply-based, conceptual framework will assure maximum usefulness of industrial statistics for these and similar purposes. Therefore, the three countries agree that the new North American Industry Classification System should conform to a production-oriented economic concept" [5].

The reasoning behind the three countries' decision may be summarized as follows. An industry is grouping of economic activities. Though it inevitably groups the products of the economic activities that are included in the industry definition, it is not solely a grouping of products.

Put another way, an industry groups producing units. Accordingly, an industry classification system provides a framework for collecting the variables that describe production--inputs and outputs--together on a consistent basis. The industry system thus groups data for analyses for which it is important that inputs and outputs be used together.

What uses of economic data require that inputs and outputs be used together, and be collected on the same basis? Such uses include production analyses, productivity measurement, studying input usage and input intensities, and so forth. For these uses, producing units should be grouped together by similarities in their production processes, which is exactly the production-oriented concept discussed in ECPC Issues Paper No. 1.

Thus, the North American countries have chosen the production-oriented concept as the framework for industry statistics (a) because important production analysis uses of data require groupings of producing units, and (b) because these uses are the ones that require that inputs and outputs be collected together on a comparable basis. The production-oriented concept for classification systems is discussed at greater length in ECPC Issues Paper No. 1

IV. Planning for the Alternative, Product Grouping System

A classification system that groups or aggregates products is a very different system from an industry classification system. An example of a product grouping system is the Central Product Classification (CPC) prepared by the United Nations Statistical Office.

A product grouping system[2] satisfies a different need--a different use of data--from the one served by an industry classification system. A product grouping system is used for analyses from the demand side--to define markets to study market power or to conduct marketing studies, for demand estimation, for determining the extent of substitution among commodities, and so forth. One does not want a product grouping system for studying productivity; an industry classification system produces the data for productivity analysis.

A product grouping system has the following characteristics:

(a) It should incorporate, and facilitate the analysis of, relationships among products--demand relations, substitution relations, marketing relationships, uses by consumers or by other ultimate purchasers.

(b) For demand and market analyses, the inputs to production generally do not matter for the intended data use. As a

---

[2] The term "product system" has been used to encompass at least three different ideas, only one of which is the product grouping system. The first, which might be termed the "product enumeration system," provides a list of all the products (goods and services) that exist. For example, the Harmonized System (HS) of the Customs Cooperation Council provides in principle a listing of all the products that move in international trade. A product enumeration system can also contain a grouping system, and frequently does so for organizational reasons, though the enumeration system's grouping system is not necessarily constructed to facilitate economic analysis, and is usually not suitable for analytic purposes. The listings in the product enumeration system provide building blocks for the product grouping system, which is described in the text. Finally, one often needs to list the outputs in each industry of the industry classification system. Such a listing has sometimes been given its own name; other times it has been referred to simply as the "index items" or "indexes"--for example, in the United States, the products in the "alphabetical index" of the *Standard Industrial Classification Manual, 1987* [9]. This "index system" also uses, with certain exceptions, the listings in the product enumeration system.

consequence, only the outputs matter in a product grouping system, no information on inputs need be collected.

(c) Accordingly, product groupings may cut across the producing relationships in establishments, or other producing units. Establishment outputs may be separated and assigned to different product groupings, as the principles of the product grouping system dictate.

Because it satisfies a different data use, a product grouping system is appropriately constructed on a different economic concept from the one that is used for an industry system. A product grouping system requires a market-oriented, or demand-based, economic concept. The market-oriented, or demand-based, concept for economic classifications is discussed at greater length in ECPC Issues Paper No. 1.

Moreover, there is no reason to integrate a product grouping system with an industry classification system, and there is every reason to avoid linking the two where they are in fact different. A product grouping system is intended to meet its own needs, and should meet those needs independently of the industry classification system, which is properly designed to serve a different purpose.[3]

The three North American countries have agreed that product grouping systems should be established on their own merits, as indicated in the following paragraph from their joint statement:

> "The statistical agencies of the three countries also agree that market-oriented, or demand-based, groupings of economic data are required for many purposes, including studies of market share, demands for goods and services, import competition in domestic markets, and similar studies. Each country will provide product data compiled within the framework of its respective statistical system, to meet the need for such information. Recognizing the increasing international trade in goods and services, each country will work cooperatively to help improve commodity classification systems, including the Harmonized System (HS) of the Customs Co-operation Council and the United Nations provisional Central Commodity Classification (CPC) system for services, by coordinating efforts and keeping each agency informed of proposals for changes" [5].

---

[3] Some controversy exists on this point, which I believe has arisen out of failure to distinguish between the purposes for which a product grouping system is needed and the quite different functions of a system that lists the index items in the industry classification system (see footnote 1).

V. Is a Conceptual Classification System Practical?

The approach followed in constructing NAICS involves: (a) taking the economic uses of industry data as a starting point, and (b) deriving an economic concept for industry classifications, making use of the economic theory that underlies the economic analyses that use industry data. Though it is clearly desirable that data be constructed in accordance with the implications of economic theory, and constructed so that the data meet the requirements for economic analysis, there has been some justifiable concern about the practicality of such an endeavor. Can we analyze pragmatically and empirically our present economic classifications with respect to the theoretical requirements for a conceptually based classification system? Can we design new and improved classification systems making use of the theory?

The ECPC and Statistics Canada have produced a number of studies that suggest that the task we have set ourselves is indeed practical.

A. The matrix papers

In two separate studies [6] [10], U.S. and Canadian 4-digit SIC industries were reviewed. Teams in each country asked whether individual industries embodied a production-oriented economic concept, or a market-oriented economic concept.

As explained in ECPC Report No. 1, "Economic Concepts Incorporated in the Standard Industrial Classification Industries of the United States" [6], these two reviews combined understanding of the economic concepts, as developed in ECPC Issues Paper No. 1, with informed judgments about the technologies and the markets that pertain to each detailed 4-digit SIC industry. The reviews were, first, tests to see whether the economic concepts could be implemented in a pragmatic way, using mainly the type of information about industries that has been used in the past to make decisions about the U.S. and Canadian SIC systems. These reviews use the available information to assess economic concepts.

Secondly, the two reviews provide a preliminary assessment of the concepts embodied in the U.S. and Canadian systems by past decisions. Their results are subject to revision on the basis of industry expertise.

Some present 4-digit SIC industries are already constructed along production-oriented lines, or could be, with relatively small adjustments to definitions. In the United States, the study suggests that a little under a fifth (19 percent) of manufacturing shipments come from 4-digit industries that are fully defined on the production-oriented concept, and another two-fifths (actually 45 percent) originate from industries that

could be made consistent with the concept by combining and/or subdividing existing industries (the details for these estimates are contained in [6]).

On the other hand, one could emphasize the other side of the picture. Fully two-fifths of manufacturing industries have no discernible production-oriented basis and nearly as large a proportion of manufacturing shipments arise in these industries; to this one could add the two-fifths of the manufacturing industries that require some adjustments to be fully consistent with the production-oriented concept (as noted in the previous paragraph). From those numbers, it is evident that the U.S. system as presently developed does not conform to the production-oriented concept.

The situation is about the same for the 150 services industries that were reviewed in the U.S. study. Actually, a slightly higher proportion of services industries has been defined to be consistent with the production-oriented concept, but much additional refinement of service industry definitions will be required to produce adequate industrial data.

A little under a quarter (23 percent) of U.S. manufacturing shipments come from SIC industries that have been defined on a market-oriented basis. Nearly half of those (10 percent of manufacturing shipments) are industries that meet the conditions for both production-oriented and market-oriented conceptual systems: These were designed "Ideal" industries in the review, because statistics for them are appropriate for both of the major classes of economic analysis.

Another 35 percent of shipments arise in industries that have some market-oriented basis in their definitions. Many of those are cases where production-oriented and market-oriented reasoning has been combined into a compromise industry definition that fully satisfies neither.

In the traditional view, the classification problem is to find ideal industries--those that are satisfactory for both production and market analysis--on the implicit assumption that deviations from ideal in practice can be handled as "special cases," for which case-by-case compromises can be effected. That ideal industries have been found in the United States in only 10 percent of the cases is a measure of how far the traditional view of the classification problem is from the empirical reality of actual industry structure.

B. Heterogeneity index

The ECPC has developed a new statistical approach that will assist in determining production-oriented economic groupings. This method is explained in ECPC Report No. 2, "The Heterogeneity

Index: A Quantitative Tool to Support Industrial Classification" [7], which applies the new method to 4-digit manufacturing industries in the United States.

The heterogeneity index is based on the following regularity: When producing units have the same production function and face the same input prices, each producing unit will exhibit the same proportionate expenditure on each productive input (shares of inputs in total cost) as will every other producing unit. When producing units have different production functions, their input expenditures will differ. The heterogeneity index measures the dispersion in relative expenditures on inputs among the establishments in an industry, or in a proposed industry. When the establishments have the same production functions, they will have the same input shares in total costs, and the heterogeneity index will be zero. The value of the index rises as establishment heterogeneity within the industry increases; that is, the index takes on a larger value as establishments with dissimilar production processes are combined into a single category.

The heterogeneity index can be used, in conjunction with other information, to judge how closely existing industries correspond to a production-oriented grouping. It can also be used to evaluate proposals to form new production-oriented industries, or to break apart or combine existing ones.

ECPC Report No. 2 also compares the results from the new heterogeneity index with the judgments that were incorporated into the matrix of ECPC Report No. 1. Note that the matrix judgments were formed before the heterogeneity index was computed, so that the matrix and the index could be used as independent evaluations. The degree of correspondence between these two completely independent evaluations, though not perfect, is both intriguing and promising (see ECPC Report No. 2).

The heterogeneity index is an important new tool that is available for implementing a production-oriented economic concept in a classification system.

C. Services classifications

The three North American countries have agreed to give special attention to classifications for services industries, as well as for high-tech and new and emerging industries. The classification of services poses special difficulties and because of this the ECPC has released a paper (ECPC Issues Paper No. 6, "Services Classifications" [4]) that discusses the application of a production-oriented economic concept to services industries.

The ECPC has been especially challenged by those who have said that our approach may be practical for goods but will not work

for services. We believe the application of production-oriented reasoning to services industries is practical, and ECPC Issues Paper No. 6 discusses practical interpretations of the economic concept. Moreover, the matrix exercise (ECPC Report No. 1) also applied the production-oriented concept to services industries in a pragmatic way, and we believe that this exercise shows that the production-oriented concept can be applied to services.

The task of classifying services industries will, however, be especially difficult. Additional special reports on the classification of services will be released as the work proceeds.

V. Applications of the ECPC's Research Approaches to the Classification Systems of Other Countries

We believe it would be especially rewarding to know the economic concepts that have been incorporated into industry definitions in classification systems outside the United States and Canada. It would also be valuable to test the heterogeneity index on the industry classifications of other countries. Exchanging the results of similar studies carried out on classification systems in use in different countries would provide a good way to determine where--that is, in which classification systems--the best ideas for industry groupings are to be found.

In the past, comparisons of different classification systems have more or less given the result: We do ours this way and we think ours is best, and you do yours that way and you think yours is best. However, we can now do better: Carrying out analysis of classification systems along the lines of ECPC Reports Nos. 1 and 2 and the Statistics Canada study [10] potentially provides a much more productive exchange of information than has been possible in the past. Rather than "splitting the difference" between mutually exclusive classification outcomes, performing some economic analysis on classification systems, of the type incorporated into the Statistics Canada and the two ECPC reports described in this paper, would produce new and valuable information for improving industry classifications.

Moreover, explicit conceptual analyses of classification systems would offer the potential for melding the international desire for comparability in industrial statistics with the goal of improving the available industrial statistics for the needs of users. Rather than setting the two goals against each other, or elevating the one over the other, as has sometimes inadvertently been true in the past, we need to gain wider understanding and support for a new approach: Constructing internationally comparable industrial statistics--where internationally comparable economies exist--that conform to a consistent economic concept provides the worldwide best course for the future of industrial statistics.

REFERENCES

[1] Bureau of the Census, *Proceedings*, International Conference on Classification of Economic Activities, Williamsburg, Virginia: U.S. Department of Commerce, November 6-8, 1991.

[2] Coats, R. H., "The Classification Problem in Statistics," *International Labour Review*, April 1925, pp. 511 and 520.

[3] Economic Classification Policy Committee, Issues Paper No. 1, "Conceptual Issues," *Federal Register*, March 31, 1993, pp. 16991-17000; Economic Classification Policy Committee, Bureau of Economic Analysis (BE-42), U.S. Department of Commerce, Washington, D.C. 20230.

[4] _____, Issues Paper No. 6, "Services Classifications," March 1994. Available from Economic Classification Policy Committee, Bureau of Economic Analysis (BE-42), U.S. Department of Commerce, Washington, D.C. 20230.

[5] _____, *Federal Register* (forthcoming). Available from Economic Classification Policy Committee, Bureau of Economic Analysis (BE-42), U.S. Department of Commerce, Washington, D.C. 20230.

[6] _____, Report No. 1, "Economic Concepts Incorporated in the Standard Industrial Classification Industries of the United States," May 1994. Available from Economic Classification Policy Committee, Bureau of Economic Analysis (BE-42), U.S. Department of Commerce, Washington, D.C. 20230.

[7] _____, Report No. 2, "The Heterogeneity Index: A Quantitative Tool to Support Industrial Classification," May 1994. Available from Economic Classification Policy Committee, Bureau of Economic Analysis (BE-42), U.S. Department of Commerce, Washington, D.C. 20230.

[8] U.S. Department of Commerce, Bureau of Economic Analysis, *Survey of Current Business* 70 (February 1990): 2.

[9] U.S. Executive Office of the President, Office of Management and Budget, *Standard Industrial Classification Manual, 1987*, 705 pages. (For sale by: National Technical Information Service, 5285 Port Royal Road, Springfield, Virginia 22161, Order no. PB 87-100012.)

[10] Young, Kenneth, "The Conceptual Basis of the Standard Industrial Classification," Standards Division, Statistics Canada, February 1994.

REVISING THE UNITED STATES STANDARD OCCUPATIONAL CLASSIFICATION (SOC) SYSTEM

Thomas J. Plewes
U. S. Bureau of Labor Statistics

I. Background

Historically, various United States Federal agencies, primarily the Department of Labor and the Bureau of the Census, have developed their own separate occupational classification systems, designed to meet their own specific statistical and programmatic needs. The lack of comparability between these various sources of occupational information and data led to multi-agency interest in and action to develop a Standard Occupational Classification (SOC) system, beginning in 1966. The SOC, first published in 1977 and revised once in 1980, was intended to provide a mechanism for cross-referencing occupation-related data collected by various economic and social statistics programs in order to maximize the analytical utility of these data.

The major underlying principle of classification in the SOC is work performed, not skills, training, education, licenses or other credentials. More specific occupations are grouped into the most detailed SOC categories based on their similarity in terms of work tasks and activities. Other classification principles include the following: SOC groupings are independent of the work setting, unless it alters the nature of the job; supervisors are identified separately from workers; large or small size is not a determinative factor for separate identification; and comparability to the international standard classification of occupations (ISCO).

The SOC was intended to be comprehensive in coverage, including all occupations for which work is performed for pay or profit, including unpaid farm work. The 1980 SOC was comprised of 664 distinct occupations at the most detailed level. It was not intended to meet all specialized analytical or organizational management purposes, but to serve as a general tool for reconciling various sources of occupational data.

In 1983, the major sources of U.S. occupational employment data -- the establishment-based Occupational Employment Statistics (OES) survey and the Current Population Survey (and Decennial Census) of households -- became more comparable when each adopted a new classification structure based on the SOC.

II. The Need for a New SOC

The SOC, unfortunately, never was implemented fully across all Federal occupation-related data collection efforts. Various

federal agencies continue to use their own distinct occupational classification structures. For example, the Department of Labor's Employment and Training Administration uses the Dictionary of Occupational Titles (DOT); the Department of Education uses its Classification of Instructional Programs (CIP); and the Office of Personnel Management has its own occupational classification structure. As a result, reconciling different occupational data sources continues to be difficult at best. In addition, the 1980 version of the SOC is outdated, as new occupations -- particularly in technical and health-related fields -- have emerged since that time (and are incorporated into some of the current occupational classification structures).

There are other reasons that attention recently has focused on occupational information. Concern with the quality of the U.S. workforce, skill formation issues, and changes in occupational structures due to new technology and shifts to "high-performance" work organizations, all highlight the importance of accurate, timely, and comparable occupational information to support program planning, career guidance, and training development. As such many users and producers of occupational data feel that it is time to re-examine the SOC and to develop a classification structure that meets the occupational information needs of the twenty-first century.

III. Actions to Inform the SOC Revision

In November 1991, the Office of Management and Budget (OMB) designated the Department of Labor as the lead agency to coordinate the development of a new U. S. Standard Occupational Classification (SOC) system by 1997, in time for implementation in the 2000 Census. Since that time, the Bureau of Labor Statistics' Office of Employment and Unemployment Statistics and the Dictionary of Occupational Titles (DOT) staff of the Employment and Training Administration (ETA) have been working together to organize activities aimed at developing information and alternative approaches related to classification principles for the new SOC. These activities have included commissioning contract papers on major occupational classification issues.

An International Occupational Classification Conference was held in June 1993, sponsored by the Bureau of Labor Statistics. The Conference provided a forum for the discussion of new ideas and alternative approaches to occupational classification issues. It included many individuals and agencies directly involved with the occupational classification user community, as well as international occupational experts from numerous countries. The papers, discussions, and ideas generated at the Conference are serving to inform revision activities for the SOC.

Some of the major issue areas addressed at the Conference are described below.

1. New challenges and alternative approaches to occupational classification: Currently, all federal occupational classification systems are based on work performed or job titles. As the pace of occupational change has increased, many people are becoming more concerned with issues of skills transferability between jobs or occupations in order to facilitate transitions in an increasingly volatile economic environment. An important issue raised during the conference is whether a new U.S. SOC should be based primarily on skill type and skill level, rather than work performed.

2. The feasibility and desirability of creating a unified occupational classification structure for government statistical and programmatic purposes: Although some Federal agencies may prefer to maintain their separate classification structures, others feel that net value could be provided to users of occupational data by developing a more unified Federal classification structure. At a minimum, there seems to be consensus that a more unified Department of Labor occupational classification structure is desirable, and movement in this direction has been occurring, even prior to the Conference. In its final report, the Secretary of Labor's Advisory Panel on the Dictionary of Occupational Titles recommended that a revised Dictionary conform to the classification structure of a revised SOC system and, in the interim, conform to the Bureau of Labor Statistics' Occupational Employment Statistics (OES) system.

3. How a revised SOC could meet the needs of users of occupational information who are dissatisfied with the current classification systems: Due to the current system of multiple occupational classification structures, users must obtain important related information -- such as demographic characteristics, industry and geographical distribution, worker attributes and skill requirements, and wages -- from different sources with different underlying classification structures. As a result, the information obtained from one source is not compatible with information derived from another source, leading to frustration on the part of many users. Another source of dissatisfaction lies with the perceived currency and accuracy of current occupational classification structures. Some structures, including the SOC, have not been updated for more than a decade, and therefore, many new occupations that have emerged as a result of new technology and changed forms of work organization are not included in current classification structures.

4. International perspectives on occupational classification and lessons for the U.S. SOC revision: A full day of the conference was devoted to international occupational classification issues. The international experience is important for two reasons: One relates to the international comparability of data, and the other relates to lessons that can be learned from the experience of other countries. A decision to move towards a common international classification system, such as the International Standard Classification of Occupations (ISCO-88), would

inevitably result in a loss of nation-specific occupational detail that many users of national data regard to be critical. In addition, there are questions about the degree to which ISCO-88 is structured on clear, consistent, and appropriate principles. The second reason to examine the international experience is to try to draw lessons from other countries, many of which have recently made substantial revisions to their national occupational classification systems. Issues explored included new approaches to principles of occupational classification (e.g., skill type and skill level); the level of effort and resources required and methodologies used to develop new systems; and the feasibility and desirability of developing a unified national classification system to replace existing disparate ones.

IV. SOC Revision Process

Following the Conference, the Office of Management and Budget established an SOC Revision Policy Committee, chaired by the Bureau of Labor Statistics, with representatives from the Bureau of the Census, the Employment and Training Administration, the Office of Personnel Management, the Defense Manpower Data Center, and, *ex officio*, the Office of Management and Budget (OMB). A Charter for the Committee recently has been approved.

The Policy Committee is charged with an examination of the Federal Government's various occupational classification systems for statistical and administrative uses, and with providing recommendations to OMB on the structure and implementation of a new SOC. The charge to the Committee includes: (1) identifying the major statistical uses of occupational classifications; (2) identifying and developing new concepts, structures, and methodologies to determine what constitutes an occupation; (3) developing a standard classification system based on these concepts; (4) planning the implementation of the new classification system; and (5) ensuring that there is ample opportunity for widespread public participation in the revision process.

The principal use of a revised SOC would be statistical, but it also would serve as a framework for administrative purposes and other occupational classifications. The Policy Committee will evaluate the utility of alternative classification structures in consideration of the following: (1) Ensuring compatibility between the descriptive material of the new Dictionary of Occupational Titles (DOT) and the revised SOC; (2) current public interest in a skills-based classification system; (3) users' needs for historical comparability of data; (4) the expertise of other countries in revising national classification systems; (5) desirability, but not necessity, of compatibility with international occupational classification systems; and (6) the need for all Federal Government occupational classification systems to be part of the SOC framework.

The Policy Committee will adopt processes that ensure ample opportunity for public participation. These processes will involve all stakeholders, including the range of occupational data users, both government and private, as well as data collectors and data providers. The Policy Committee will consider forming a Consultation Group, composed of Federal agencies not represented on the Policy Committee and interested public and private parties (e.g., States, associations, private individuals). Such a group would meet on a flow basis, as necessary, to provide input to the work of the Policy Committee. Notice of the Policy Committee's work will be widespread and will be published in the *Federal Register*, and all interested parties will be given the opportunity to be included on a mailing list.

The conceptual framework for the new SOC is to be completed prior to July 1995 to allow for testing related to the 2000 Census, as well as for the administration of the 1996 DOT National Content Test. The completed occupational classification structure should be available by July 1997 to coincide with development of the 2000 Census.

COMMENTS ON THE REVISIONS OF THE STANDARD INDUSTRIAL AND OCCUPATIONAL CLASSIFICATIONS

Joel Popkin
Joel Popkin and Company

The beginnings of the present effort to revise the Standard Industrial Classification (SIC) originated in the Census Advisory Committee structure in the mid 1980s. At that time, I represented the American Economic Association. At those committee meetings, I recall getting the assignment to comment on papers about how the SIC ought to be revised, and asking myself why I had not drawn a more interesting assignment. I clearly did not recognize that a revolution in economic classification was afoot.

Charles Waite, Associate Director of the Census Bureau, was handing out those assignments, and the papers I commented on were written by his staffers, Pamela Powell-Hill and James Monahan. That was 10 years ago, and I think marks the birth of this much needed and very important current effort to conduct a "clean-slate" revision of the SIC. Following those developments, Charles Waite planned and convened in Williamsburg what turned out to be a seminal international conference on economic classification. About the time the plans for the conference were being initiated, Jack Triplett, Chief Economist of the Bureau of Economic Analysis, presented at the 1990 Census research conference a very important paper illuminating the conceptual issues relevant to the classification of economic activity. Hermann Habermann, then Chief Statistician of the U.S. government, lent support to a continuation and formalization of those efforts.

The project to revise the SIC now has a full head of steam with a target for implementation in the 1997 economic censuses. OMB appointed the BEA lead agency, and Jack Triplett is chairman of the government-wide Economic Classification Policy Committee (ECPC). There are three elements of U.S. leadership in this significant undertaking. The first was the Williamsburg conference itself. The second was the successful negotiation among the countries of NAFTA of an agreement to develop a common, North American industrial classification system (NAICS). The third key element in the pervasiveness of this effort, and its enhanced chances of success, was that the North American plan and the Williamsburg conference were instrumental in prompting Eurostat, the statistical agency of the European Community, to reconsider some decisions it had made about industrial classification and to explore moving in the direction of fundamental rethinking of classification systems that the United States has promulgated.

There are three fundamental kinds of decisions that have to be made in designing an industrial classification system. The first is the selection of the unit of observation. The second is the concept by which individual observations should be grouped. And the third is the hierarchy along which groups should be aggregated.

Decisions about two of those three elements have already been made. The first is that the classification system will retain the establishment as the unit of observation except in cases where its use is not appropriate or feasible. The second is that the underlying concept for classification will embody a production-side approach. Establishments will be grouped together that share the same kind of production techniques and processes. The third issue, that of hierarchy, is still being studied.

Part of my role here today is that of discussant of the SIC revision process that is now underway. Some of you will be familiar with my views and recommendations if you have seen the paper I wrote for the Williamsburg conference.

With respect to the unit of observation, my recommendation was to change from an establishment based system to one which I characterized as focusing on divisions, departments or subsidiaries (DDS) within companies as units of observation. I made that recommendation for three reasons. The first is that the establishment is not as prevalent an economic unit of observation as it once was. That is at least partly due to the advances in telecommunications which permit output to be produced with more inputs obtained from different establishments within the company. That leads me to the second reason I recommended a larger unit of observation such as the DDS. It is that at a higher level of aggregation, the matching of inputs and outputs and the full accounting of all inputs may be more feasible and data collection simplified. The problem posed in using the establishment is that not all inputs can be accounted for, especially some purchased services and inputs of information, technology, and management skills from central offices and other establishments within the company. I felt that by moving the unit of observation to a higher level of aggregation within the company, those inputs could be captured and the activity of separate business units (SBO) or DDSs could be relatively well accounted for and measured. That approach is not new. It is used currently in the Census M3 report on "Inventories, Shipments, New and Unfilled Orders" in which data are collected directly from divisions of companies; and it is also being utilized in the Annual Capital Expenditures Survey (ACES) that the Census Bureau has developed. The third reason I made this recommendation was that it seemed as though it might ease reporter burden to the extent that establishment records are increasingly being consolidated at company or division levels. Nonetheless, the ECPC recommendation to use the establishment, but with a recognition that there may be exceptions, goes some distance to alleviating my concern about the use of the establishment.

With respect to classification concept there were two candidates. The production-oriented approach or the market-oriented approach. Each approach serves many legitimate uses, and both can be justified. I thought it would be inappropriate to recommend multiple classification systems simply because the resources are limited. So I thought it was necessary to recommend one approach. For me, it was the demand approach. The ECPC has adopted the production approach, but also has indicated that it is undertaking work to develop a structure in which outputs can be

classified by market grouping for both goods and services. Such market groupings (commodity product classes) already exist in the government. For example, in the price statistics program, the PPIs classified by stage-of-process indexes reflect a market-oriented measure, while the PPIRs represent the kind of price index that would be used to deflate shipments or outputs and measure productivity. I also thought that the market-oriented system would fit better into the harmonized system being used internationally to collect trade statistics.

As I mentioned, the hierarchy issue is still undecided. One recommendation I have made, described more fully in an article in the November 1993 issue of the Survey of Current Business, is to break the large service sector, which as currently defined accounts for two-thirds of the economy, into two sectors. One part would be called "distribution networks" covering retail, wholesale, transportation, communication and other network suppliers. The other grouping would consist of the traditional kinds of services which tend to be labor intensive--such as personal and business services.

As if undertaking the revision of one classification was not enough, the researchers of the federal statistical system have assumed yet another undertaking--a clean-slate look at the way we classify occupations. This effort was lead by the Bureau of Labor Statistics which in June 1993 also convened an international conference. I gave a paper at that one, too. It stressed the need to define the unit of observation--the job as I see it; then to develop an underlying concept for the grouping of jobs; and finally, a hierarchical framework within which those groupings can be aggregated. Among my conclusions in that paper was that occupational classification would serve more purposes if it could be thought of in a three-dimensional context. Jobs should first be aggregated by both type and skill level. The third dimension, though not as well defined, could be along the line of whether the job involves symbolic logical work, production process work, or in-place personal service, a classification scheme developed by Robert Reich in his book, The Work of Nations. Perhaps information, goods, and services would be another way to view such a classification at higher levels of aggregation. Perhaps, this third dimension would capture, a classification index, which in concert with the other two dimensions, would approximate how employers view or define the labor markets in which they buy factors of production. In any event, if we move in that direction, our occupational classification would resemble a three-dimensional matrix, a Rubik's cube. That would facilitate not only the analysis of markets for certain kinds of occupations, but also provide a reading on the skill level required for those occupations and the kinds of training that individuals might need to reach that skill level.

In closing, I think these two efforts to completely revise the SIC and SOC are major statistical developments with considerable impacts. I am most pleased to see U.S. government statisticians take the lead in achieving progress in these fundamental areas.

Comments on Economic Classification Revisions

Joe Mattey [1]

May 24, 1994

[1]The author is on the staff of the Board of Governors of the Federal Reserve System and is a research associate at the Center for Economic Studies (CES), U.S. Bureau of the Census. These remarks are for presentation at the seminar on classification sponsored by the Council of Professional Associations on Federal Statistics, May 25, 1994, in Bethesda, Maryland. The comments reflect the author's own views, not the official views of the Federal Reserve System or Census Bureau.

My remarks will be devoted to the industrial classification revision plans discussed by Carole Ambler in her presentation of Jack Triplett's paper. My background in this area stems both from my experience using the existing classification system–in, for example, analyzing productivity developments by industry–and from my work with plant-level data as a researcher at the Census Bureau's Center for Economic Studies.

Triplett's paper addresses six issues. First, he argues that the United States has mounted this effort to revise the classification system because users demand it; for example, users find the existing classification system inadequate for productivity studies. Second, Triplett explains that this revision of the SIC, unlike those in the past, is not going to be riddled with compromises among competing uses. The third section explains why a production-oriented concept has been chosen, and the fourth section discusses what could be done to appease those most interested in the classification concept that ran a closest second, the market-oriented concept that groups products according to their degree of substitutability. The fifth section argues that a conceptually-based classification system is practical, and the final section advocates that the research approaches of the Economic Classification Policy Committee be applied to the classification systems of other countries.

I would like to elaborate on several of these issues. First, from the perspective of productivity studies in manufacturing, I believe that the need for an improved classification system largely arises from the difficulties we have in implementing the existing system consistently over time and across surveys. For example, the four-digit SIC classification of an individual manufacturing establishment often differs depending on whether the code has been assigned on the basis of product detail collected by the Census Bureau or on the basis of information available to others who initially identify the birth of new establishments, such as the Social Security Administration, Bureau of Labor Statistics or the IRS. Moreover, even within the Census Bureau's SIC assignment system, the industry affiliation of multi-product plants can switch frequently over time. These classification difficulties (and deterioration in sampling frames for a broader range of reasons) cause published individual industry-level time-series to change too abruptly from year-to-year. Users often cope by modelling productivity at more aggregate levels, following the SIC hierarchy for the aggregation. But the current SIC hierarchy was not designed to preserve similarity of input structures upon aggregation, and the resulting aggregate analyses often do not make much sense. Thus,

for productivity analysis there is a demand for an improved SIC system in two respects; we will be much better off if the new system achieves greater continuity at detailed levels (over time and across data sources) and if the hierarchy for aggregation better preserves similarities of the production process upon aggregation.

Whether this consistency goal will be achieved ultimately is an empirical matter. The production-oriented unifying concepts of the revised classification system offer some promise of greater consistency, particularly if the process of how things are made in a given establishment tends to be more fixed than what an establishment makes. In other words, there is some hope that the new technologically based classification system will reduce the extent of SIC switching because it is easier for, say, a manufacturing plant to alter its product mix among goods that are not close substitutes than it is for that plant to change the basic manufacturing process. Ultimately, then, a fundamental task of classification is to find meaningful characteristics of establishments that are relatively fixed.

As an economist, any discussion of fixed factors of production automatically evokes images of the capital stock in place and also, to a certain extent, the human capital embodied in a firm's employees. The production orientation favored by the committee seems quite natural. My only advice is that when production processes are analyzed, particular attention should be paid to the fixity of the elements when deciding whether they are defining features of the industry.

Triplett discusses how the committee would be likely to proceed in determining the defining features of the industry in the section of his paper on whether a conceptually-based classification system is practical. He mentions three studies that demonstrate how classification decisions could be made. Two of these are "matrix papers" that offer subjective descriptions of the extent to which the existing classification systems in the United States and Canada fit the production orientation. A third paper presents a quantitative heterogeneity index for use as a diagnostic tool.

I have had the opportunity to read drafts of these papers. The overall impression that they leave is that alot of work remains to be done, particularly on achieving a consensus on the defining features of industries. The Canadian paper puts it well in saying: "In the United States, the E.C.P.C. has analyzed part of the SIC. There is an official concordance between the two classifications, so the results could be compared for similarly defined in-

dustries. The initial comparisons showed numerous differences in the way the two countries had applied the concepts. (p. 14)" Given this illustration that in a subjective process of assigning characteristics to industries, experts will differ in their application of concepts, they conclude that it would be desireable to have objective measures.

My final remarks concern the one objective measure of heterogeneity that has received the most attention, the heterogeneity index originally proposed by Frank Gollop in 1986. The second report of the Classification Committee presents this heterogeneity index for selected manufacturing industries. The basic process of using the index for classification starts with a tentative grouping of establishments into various industries. Then, for each industry, a weighted average of the differences in input cost shares among establishments in the tentative industry is computed. The relative sizes of the establishments in terms of, say, shipments, can be used as weights.

Thus far, this index has been calculated using only ten types of inputs, each of which is a very aggregate concept: production workers, other labor, fuel, electricity, purchased services, agricultural materials, mineral inputs, nondurable materials, durable materials and capital. I must confess that when inputs are defined at such an aggregative level, I find the heterogeneity index relatively useless for classification. To see this, one can contrast the results of the heterogeneity index for the fluid milk industry with the subjective process illustrated in the U.S. matrix paper. The latter paper states that "...the physical properties of fluid milk dictate many of the processing methods and the types of machinery and equipment that must be used to handle it (p. 11)." I interpret this as meaning that if a plant has the types of machinery and equipment specially designed for handling fluid milk, than it must be a milk processing plant. In contrast, the quantitative heterogeneity index just looks at the overall cost of capital among plants within the industry, without regard to the type of capital equipment. Similarly, the heterogeneity index as computed just looks at the overall cost of agricultural materials, whether or not these materials have anything to do with the defining features of milk production.

In the case of capital equipment and structures the use of the aggregative data can be defended on the grounds that detailed information is not available. However, detailed information on materials use is available from the Census of Manufactures. In some of my own work with the plant-level microdata, I have gone to the opposite extreme, singling out specific detailed

materials that comprise a large fraction of total materials costs in the industry in question. For example, I analyzed the degree of heterogeneity in the use by fluid milk processors of whole milk from dairy farms.

My own statistical work demonstrates some of the difficulties one encounters when attempting to develop a quantitative index of heterogeneity, and whether or not the Gollop index can be successfully applied depends on how these issues are resolved. For example, not all plants report data on specific materials use. Small plants, in particular, omit information on detailed materials use because the Census forms instruct them to do so if a minimum value threshold is not surpassed. Moreover, in any given industry, the inquiries on specific materials are restricted to only a few pre-selected materials. Which Census form a plant receives depends on the tentative classification of the plant. So, anyone trying to develop a quantitative index of heterogeneity for re-classifying plants faces the problem that the data needed to make such a reclassification might not be collected, exactly because the initial classification was inappropriate.

In summary, the revised classification system has the potential for helping users of the data quite a bit, particularly those interested in production function relationships. However, it seems like alot of work remains to be done to develop the consistency needed to achieve this benefit.