17th Voorburg Group Meeting on Service Statistics
Nantes – 23-27 September 2002

A contribution for the Session on Classifications

Developing a Demand-based Aggregation Structure for the North American Product Classification System

By Paul Johanis (Statistics Canada)

Developing a Demand-based Aggregation Structure for the North American Product Classification System

1. Statistics Canada, in cooperation with the national statistical agencies of Mexico and the United States, is developing a North American Product Classification. In the first phase of the project, which started in April 1999, the initiative covered the products of Finance (NAICS Sector 52 excluding Insurance), Information (Sector 51), Professional, Scientific and Technical Services (Sector 54), and Administrative and Support and Waste Management Services (Sector 56) of the North American Industry Classification System (NAICS Canada 1997). In the second two-year phase of the project, started in April 2001, the identification of detailed products has continued, covering the products of selected industries in Transportation and Warehousing (Sector 48-49), Educational Services (Sector 61), Health Care and Social Assistance (Sector 62), Art, Entertainment and Recreation (Sector 71) and Accommodation and Food Services (Sector 72). The full coverage of all service products is expected to be completed by late 2005, for implementation in survey programs for reference year 2007. Work on the aggregation structure for this new classification was started during Phase 2 of the project and we are still early in the developmental phase of this activity.

2. The stated objective of the North American partners in this project is to develop a market-oriented and demand-based classification system for products. Grouping and aggregating products according to demand side characteristics is required for the economic analysis of market share, concentration ratios and import competition, among others. Demand-based aggregation structures are currently widely used in expenditure and price statistics. The goal of the three countries is to develop a product classification system that: (a) is not industry-of-origin based, but can be linked to the NAICS industry structure; (b) is consistent across the three countries and encompasses the products of both services and goods industries alike; and (c) promotes improvements in the Central Product Classification System of the United Nations.

3. Since the beginning of the North American Classification System (NAPCS) project three years ago, more than 1,800 detailed products have been identified and defined. The number of products continues to increase with the work of Phase 2. It is now time to start developing the structure under which these products shall be grouped. From the outset of the project, it was agreed that a demand-based aggregation structure should be developed, i.e., grouping together products that satisfy a given need or purpose. To further develop this idea, a tri-lateral working group has been established, which has developed several illustrations of what an eventual structure might look like and dealt with a number of conceptual issues.

4. Preliminary work was done mainly on compiling an inventory of various classification categories for products, demand based and others, used by national and international statistical agencies. This has served as input for general discussions of

the working group on the aggregation structure. A number of working assumptions were adopted from these discussions. For example, it was decided that the structure should cover both goods and services, even though the detailed product identification work only covers services at this point. It was also decided that there would not be a distinction made at the top level of the classification structure between goods and services but that true demand categories would be identified, which integrate both goods and services. It was also agreed that there would not be a class of customer distinction made between businesses and households at the top level of the classification.

5. The relationship of NAPCS to other national and international product and activity classifications was considered and certain working assumptions were accepted in this respect. First, the industry of origin link of detailed NAPCS classes to their producing industries defined in NAICS terms would need to be preserved. Secondly, wherever possible, the detailed classes of NAPCS would be related to one and only one detailed class of the CPC. With the potential convergence of NACE, NAICS and ISIC in 2007, and the opportunity to make revisions to the CPC in the same timeframe, the development of NAPCS could give rise to a fully articulated set of product and activity classifications, which reconcile supply side and demand-based approaches (see Appendix 1).

6. The adoption of NAPCS would be facilitated if the aggregation structure that is designed is similar to structures already in use in statistical programs that currently use demand-based categories, such as expenditure statistics for businesses and households as well as for price indexes. These structures were therefore mined for potential applicability. Findings were presented to the Methods and Standards Committee of Statistics Canada and to its Advisory Committee on Services Statistics. At those presentations, emphasis was placed on the challenge of accommodating in a single structure both intermediate demand (goods and services purchased by businesses as inputs to the production of goods and services) and final demand (by households, businesses and governments).

7. As a contribution to the Tri-lateral working group, Statistics Canada developed several illustrations that used a variety of existing demand-based aggregations as building blocks. The Canadian Consumer Price Index (CPI) was used as the base for constructing these illustrations. (The CPI is also a well known indicator in other countries). The CPI basket is derived from family expenditure surveys that provide information on the spending habits of households; the expenditure data are used as weights in the CPI. The CPI includes goods and services for which it is possible to associate a sum of money that a consumer must pay to purchase a specific quantity and quality of a good or service. The goods and services are organized according to a classification system that is demand-based. Every product has a unique place in the classification. Products are grouped with other items either because they have a common end-use or because they are considered substitutes for each other. These families of products are joined together at different levels in the classification system

to form a hierarchy. The lowest level is called a basic class while the highest level is known as a major component. The goods and services in the CPI are grouped into 181 basic classes and eight major components. These components and classes are useful to categorize household expenditures however they do not cover all of the goods and services available in the economy.

8. For business expenses that are part of intermediate demand, the expense categories of the industrial product price indexes (IPPI), of the raw materials price indexes (RMPI), and of the purchased service inputs (PSI) were considered. The IPPI includes all manufactured commodities in Canada and a few commodities (in a special category: S-99) that are not normally classed as manufactured goods, but are produced by manufacturing industries. In the IPPI, the goal is to combine similar commodities according to where they fit in the stage of processing, who produces them and why they were purchased. There are five (or six) levels of combination which eventually group together all commodities. More than 1,300 different commodities are distinguished and they are called principal commodity groups (PCGs) corresponding to the fifth level. For 50 PCGs, there is a sixth level for more detail. Moving on to progressively more general groupings, the third level (56 different groupings) corresponds to the M-level definitions in the Canadian Input-Output tables, the second level (21 categories) to the S-level definitions of the same tables, and the first level to all commodities.

9. The RMPI cover the price movement of a number of commodities that are used in Canadian industry, and pertain to raw rather than manufactured products. The term "raw material" refers either to a commodity that is sold for the first time after being extracted from nature, or a substitutable recycled product (e.g. metal scrap). Except for category S-99 there is no overlap with the IPPI. The RMPI is subdivided into seven major categories, and at the finest level of detail, there are 62 separate commodities corresponding to individual PCGs or to groupings of PCGs.

10. The PSI categories used in Canadian business surveys cover the expenses for services purchased from other businesses. PSIs include 16 major categories and 50 categories at the lowest level of detail. They are linked to CPC V1 classes. These categories were presented at previous Voorburg meetings in papers dealing with business demand for services.[1]

11. Although these commodity groupings (industrial products and raw materials) and purchased service inputs are normally part of intermediate demand by businesses, it is not always the case. Some industrial goods go directly to final demand, for example

[1] Statistics Canada, Business Demand for Services: Canadian experience, 14th Voorburg Group, Christchurch, New Zealand; Statistics Canada, Purchased Services in Canada, 15th Voorburg Group, Madrid, Spain

lumber, while some finished goods, for example food, go into intermediate demand of businesses and government.

12. Four different options were constructed from these building blocks:

    Option 1A: a single aggregation structure for both final and intermediate demand, using CPI high-level aggregates only as the top level of the classification;

    Option 1B: a single aggregation structure for both final and intermediate demand, adding new high-level aggregates to deal with products that are mostly found in intermediate demand;

    Option 2A: two aggregation structures, one for final demand and another for intermediate demand, without any overlap between the two;

    Option 2B: two aggregation structures, one for final demand and another for intermediate demand, with overlap.

13. In Option 1A, to create a single aggregation structure for both final and intermediate demand using CPI high-level aggregates at the top level of the classification, it was necessary to insert new categories at appropriate levels within the CPI classification to cover expenditure categories not explicitly covered by the CPI. For example, the major component "Transportation" in the CPI is restricted to transportation products purchased by households. It could be used, however, to represent all transportation products. New categories, taken from the transportation expenses listed in the PSIs, would therefore need to be inserted under Transportation. These are "Goods transportation, warehousing and storage" and "Other transportation vehicles", and the corresponding basic classes. Introducing this new detail under the major component Transportation, however, changes the content and meaning of that component. This would cause a break in the time-series statistics that currently use this aggregate, while requiring changes of the labels of other sub-groups, even though their contents are the same, for example, adding the word "passenger" to the sub-headings "private transportation" and "public transportation" to distinguish them from goods transportation.

14. In Option 1B, this issue is dealt with be inserting the missing categories at the major component level of the CPI. In this way, the content and meaning of CPI major components remain the same while new high-level categories are added. For example, in option 1A, the categories "Goods transportation, warehousing and storage" and "Other transportation vehicles" were inserted at the second level while in Option 1B they are inserted as major components #6 and #7 respectively. As a result, while there are more major components, the original major components are not modified as to content and meaning. However, it may be necessary to change the label of a major component (from "5 Transportation" to "5 Passenger transportation") since the previous label does not exactly represent its content (goods transportation not being

included). As a result of this procedure, there are 29 major components and 76 second level categories in Option 1B.

15. In Option 2A, two separate structures were developed: one for intermediate demand and one for final demand. All product groups are adjudged to belong to one or the other, which is what we mean by "no overlap". All CPI components are included in the final demand structure. IPPI components were assigned to either intermediate or final demand based on their intended use. The major commodity groups in the Canadian IPPI are defined as intermediate goods or finished goods. Intermediate goods are further subdivided into first stage intermediate goods, which are those used primarily for the manufacture of other intermediate goods, and second stage intermediate goods, which are those used primarily for the manufacture of finished goods. The ratios of production values of first and second stage intermediate goods and of finished goods within each major commodity groups were used to decide whether to assign the whole major commodity group to the final or intermediate demand structure. As result, the final demand structure includes nine major components and the intermediate demand structure contains 20 top-level aggregates.

16. In Option 2B, two separate structures are created, one for final demand and one for intermediate demand, but the same product can appear in both structures. As a result, there are 16 major components for the final demand structure, 27 major components for intermediate demand, and 14 major components appear in both structures.

17. At its meeting of May 2002, the Trilateral working group determined that Option 1B best illustrated the type of aggregation structure that would be designed for NAPCS (see Appendix 2). It was decided that each country would develop a similar illustration, using the aggregations found in their respective Prices statistics programs. These would be compared for similarities and differences at a subsequent meeting, planned for September 2002.

18. The detailed methodology for constructing Option 1B was shared with the trilateral partners and guided their work. In general, the approach is to start with CPI categories and adding major components taken from the categories of the PSI, IPPI, and RMPI. First, the following PSI categories were added: Professional and business products, Financial products, Construction services, Other government services, and Other products, with their sub-categories. Some fine-tuning will be necessary, for example, in the CPI category "2.2.6 Other owned accommodation expenses", services related to the dwelling such as legal service, appraisal, surveying, and mortgage penalty are included. Those services are also included in the added PSI categories. A similar approach was adopted for the IPPI. The IPPI commodity groups-S and M levels were compared to the new set of CPI/PSI major components. In the event of a complete overlap, the IPPI commodity group-S has not been used. For example, the IPPI-S "23 Autos, trucks, other transportation equipment" is assumed to be covered in CPI major components "5 Passenger Transportation", "7 Other transportation vehicles" and CPI

sub-sub-group "9.1.2 Purchase and operation of recreational vehicles". In other cases, the removal of commodity groups was done at the second (M level) or third level. For this report, no further detailed analysis was done below the fourth level (primary commodity group-PCG) and fifth level. However, this task would be necessary to eliminate some overlaps. For example, in the CPI, "computer equipment and supplies" is part of the major component "7 Recreation, education and reading" whereas in the IPPI commodity group – M level "22 Machinery & equipment", there is a principal commodity group for computing equipment. A decision would be required as to keep computer equipment with "Recreation, education and reading" or to move it somewhere else. Finally, the seven major categories of the RMPI were added as second level to the IPPI-S 99 category as a way to aggregate miscellaneous remaining commodities (fine-tuning still necessary to eliminate overlaps).

19. Each country has now produced an Option 1B and they are being compared for similarities and differences. At the September meeting, additional illustrations were tabled, some using Personal Consumption Expenditure categories as the base rather than CPI categories, others using the broad framework of expenditure based GDP as the basis for the aggregation structure.

20. Also at the September meeting, it was recognized that the eventual outcome of the Trilateral working group's work might not be a single aggregation structure but a number of alternative aggregation structures, constructed from common building blocks. These building blocks would be identified from the detailed product lists developed by the NAPCS subcommittees. These lists currently include hierarchical structures within narrow product areas and the top levels of these structures could be assessed as "building blocks". Each country has now agreed to develop an alternative aggregation structure, which will be tabled and considered as initial proposals at the next meeting of the working group in January 2003. The Canadian contribution will be a structure based on existing Canadian Prices statistics aggregates while the Mexican proposal will be based on Personal Expenditure categories. The U.S. will propose two or three alternatives, of which one will be based on an expenditure based GDP model.

21. Based on the work to date, there is a good expectation that a demand based aggregation structure will be developed and adopted for NAPCS, in the planned timeframes. These call for options to be made available for public consultation during 2003, with a target to finalize the structure (or structures) by the end of 2003. Such a structure will open up new opportunities for reconciling and harmonizing North American statistics beyond activity and production statistics into such areas as prices and expenditure statistics, in a way that takes into account international product classification standards and their evolution over the next five years.