Author - Canadian Industry Statistics Development Team

Publication Date - 1999-04-12

Pitfalls in the use of International Merchandise Trade tabulated on an SIC Basis

General Description

Although international merchandise trade data are collected and tabulated on a commodity basis, some users prefer to use data classified by industry (SIC) so that it is easily compared with other Statistics Canada data. The Industry Overviews use trade data tabulated on an SIC basis.

Statistics Canada produces several versions of the merchandise trade data series tabulated by SIC, using two distinct approaches: primary production based and production based. The primary production approach, the basis used by the Industry Overviews, is normally applied when using data referring to the importers and exporters of the commodities. The production-based approach incorporates concordance tables from the Industry Measures and Analysis Division and International Trade Division of Statistics Canada. Both approaches are crude approximations and should be viewed as such. Prior to analyzing any of the data and trends shown throughout the Industry Overviews, users should be familiar with the nature of the trade data and not overburden it with implications that it was never designed to support.

Primary Production Approach

The primary production approach attributes exports of a specific commodity to whichever industry is normally associated with the majority of its production in Canada. For imports, goods are assigned to the industry that normally is associated with that type of production in Canada.

The industry code assigned to the commodity is in accordance with the producer and not the exporter or importer. In other words, if lumber was exported by a wholesale distributor but manufactured by a lumber mill, then the good is assigned to the SIC for lumber mills (not for wholesale distributors).

Caveats

The concerns covered in this note are not with the quality of the merchandise trade data but with the conversion of the commodity-based data into industry-based data, notably:

- -the inclusion of individual SICs is different between the trade data and other economic series data (i.e. production, employment, etc.);
- -valuation may be different; and/or
- -timing may be different.

The effect of these differential factors can vary among SICs, and also over time and geography.

There are many situations where the establishment produces a mix of commodities, some of which belong to an SIC not assigned to that particular producer. Since there is only one SIC assigned, it is not always possible to be exact in attribution and classification.

Purity of Shipment Data for Establishments

A key measure of the impurity is the Coverage Ratio which shows how much of an industry's major product is also produced by other industries. The most appropriate way of interpreting the coverage ratios is to think of these as the accuracy of the association of the product to its major producer. The higher the coverage ratio, the better the chance of the product being linked to its major producer.

For example, the coverage ratios for Major Group 33 (Electronic and Electronic Products Industries) for the year 1990, show that it is not unusual that over 20% of the exports are classified in the wrong SIC. At the 4-digit level, coverage ratios can vary significantly, often less than 75%. The median, mean (simple average), and minimum values are 90.0, 86.6, and 44.0 respectively. The ratios are significantly improved and reach between 90 to 99% when aggregating the industry codes up to the 2-digit level.

In addition to assigning a product to the wrong industry, another problem may occur that the relationship between commodity-based exports and industry-based exports may be biased. For example, in 1990, the Electronic Computing and Peripheral Equipment Industry (SIC 3361) and the Office, Store and Business Machine Industries (SIC 336) are clear examples of this bias because the exports are much higher than shipments (to both home and abroad) for that year.

The reasons include:

- -wrong imputation to the industry;
- -exports for the year were partly out of stocks owned or held by wholesalers and retailers; and
- -manufacturers and exporters (wholesalers, general distributors) applied inconsistent valuation to the same products.

The proportions in which these factors affect the data are not known but the numbers suggest that the joint effect is not trivial. There are industries which, because of their configuration (produce mostly for exports, have high coverage ratios, and handle their own freight forwarding), the ratio is more meaningful. Examples are the Automobile Industry and the Aircraft Industry. But for others, even the year-to-year variations should be taken as a reliable indicator.

It should be emphasized that the more detailed the level, the less reliable the figures. It is clear, however, that as exports are overstated in some SICs due to over-inclusion, there will be other SICs which are understated. As one aggregates the data to the 3 and 2-digit levels, the inclusion problem diminishes greatly. The effects of timing and valuation problems will, however, tend to remain.

The Principal Problem with the Primary Production Approach

The understanding that the shipments of an industry can contain goods not associated with that industry is the single most critical element to dealing with the inclusion problem.

The primary production approach ignores the fact that the same commodity is produced by two or more industries. Through the use of concordance tables, commodities are attributed to their primary industry of production only, and an exporting industry may be ignored.

Refining the Primary Production Approach

A method of addressing the allocation problem posed by cross-production (the production of one commodity by several industries) is to look at the actual production mix of a particular commodity between industries and then allocate the trade in that commodity on that basis.

This could lead to an improvement over the primary industry of production method, but, it is not necessarily more accurate. As the production mix changes from year-to-year in the input-output tables, the weights for allocating commodities across industries must be changed. Since the input-output tables are typically produced a number of years after the fact, the only ratios currently available may be inappropriate. Furthermore, the input-output table is less detailed in terms of commodity and SIC.

Cross-production is most prevalent at the 3 and 4-digit SIC levels. At the 2-digit or division levels, classification is sufficiently broad that the likelihood of inappropriate assignment is minimal.

For more details on these matters contact:

Statistics Canada
Attn: Intenational Trade Division, Client Services Section
9th Floor - Jean Talon Building
Ottawa, Ontario
K1A OT6

Tel: (613) 951-9647 or 1-800-294-5583
FAX: (613) 951-0117