

Chapter 2

Conversion of Taiwan Trade Statistics into UN Standard Form

EBIHARA Etsuo and NODA Yosuke

The Institute of Developing Economies uses an original method to convert the form and content of Taiwan trade data from 1971 onwards into UN standard trade statistics format in order to employ it as basic data in the IDE world trade data system (Ajiken Indicators of Developing economies: eXtended for Trade statistics : AID-XT), which enables international comparison of trade data. The purpose of this chapter is to discuss an attempt to convert Taiwan trade data directly into the format used in the UN COMTRADE database, independently of the AID-XT.

1. Conversion to AID-XT Basic Data

The IDE uses both data from the UN COMTRADE database and trade statistics formulated by the OECD (International Trade by Commodity Statistics: ITCS) as basic data for the AID-XT. In order to do so, the IDE unified classification system is used as a shared classification system for the UN and OECD statistics. These classifications do not necessarily correspond to the UN COMTRADE classifications. The process of formulating AID-XT basic data from Taiwan trade data can be summarized as follows.

(1) All the commodity classifications, partner countries, and units of quantity in the classification

categories included in the original data are extracted for year and import/export category. These commodity classifications, partner countries, and units of quantity are saved temporarily in a working file. If a comparison with correspondence tables for partner countries in Taiwan trade data and IDE unified country codes reveal any that do not correspond, new relationships of correspondence are formulated and added to the correspondence tables. The same procedure is required for units of quantity.

(2) Detailed commodity codes are formulated based on HS, using the commodity codes from the original data up to six digits from the left. Partner countries are converted to IDE unified country codes, and units of quantity are converted to IDE unified units of quantity.

(3) Transaction values and quantities are totaled for year, import/export category, commodity classification, partner country and unit of quantity. At this stage, transaction values are shown in 1,000 New Taiwan dollars. Transaction values in 1,000 New Taiwan dollars are converted to these in 1,000 US dollars.

2. Conversion to UN Standard Format

This section discusses the formulation of

relationships of correspondence between classification categories used in Taiwan trade data and in UN COMTRADE data, which are essential in conforming the former to the latter. Correspondence tables between country codes and units of quantity are required for conversion.

As discussed in section 1, the IDE has formulated correspondence tables between country codes used in Taiwan data and IDE unified country codes for each year. In addition, because a correspondence table between IDE and UN country codes has also been formulated, these have been combined as one table, and a

Taiwan data country code → IDE country code → UN country code
conversion table has been formulated.

The purpose of formulating this correspondence table was naturally to convert Taiwan country codes into UN country codes; however, the IDE country codes act as an intermediary. Therefore, if a country code is not defined in the intermediary, then the conversion of that country code will not link up, and the purpose will not be achieved. The Taiwanese country name and country code tables were checked to find data that had been overlooked, and checks were conducted to determine whether there were any codes that did not link up because they did not appear in the IDE country codes. To cite an actual example: The country code for "Mayotte" is 353 in Taiwanese statistical data, but this code did not appear on the correspondence table between Taiwanese and IDE data. In this case, the correspondence table was corrected, and a string was added. "Other Asia", "Other Europe", "Venezuela" and "Nauru" represented similar cases.

When converted, it is sometimes the case that multiple Taiwanese country codes become the same code in terms of IDE and UN codes. This is because

there are differences between the Taiwanese, UN and IDE classifications. In general, the Taiwanese codes have been considered the most detailed. The following is an example in which only a UN country name and three kind of codes were extracted from the tables.

Malaysia 033 127650 45458
Malaysia 034 127650 45458
Malaysia 035 127650 45458

In this case, 033, 034 and 035 of Taiwanese codes were allocated to 127650 defined by the IDE, and would have been classified as 45458 if conversion had been conducted without further checking. Consultation of explanatory materials concerning Taiwanese trade data obtained by the IDE showed that 034 represented Malaysia: Sabah and 035 represented Malaysia: Sarawak. The correspondence table between UN country codes and IDE unified country codes and country names showed that these were allocated to 45461 and 45457 respectively. The correspondence table was therefore corrected to reflect this. In the case of Belgium: Luxembourg, because Luxembourg is defined as a separate country in the UN codes, codes were allocated as follows: 206 →53058, 237→53442. In the IDE unified country codes, Monaco is classified with France. In converting between codes, it was defined separately, with the following code allocation: 241→90915. Similarly, Liechtenstein was discriminated from Switzerland (236→90916). As this indicates, efforts have been made to respect Taiwanese country code classifications as much as possible, but in cases like India, which Taiwan defines as "India: West Coast " and "India: Others" (018 and 019 respectively), the classifications have been combined into a single "India" classification. The same is true in the cases of "France", "Spain", "USA", "Canada" and "Mexico".

The country codes 502, 528, 820 and 710 did not appear on the conversion table. Each of these was extracted from the original file and checked with data in a print statistical publication. They were found to represent, respectively, "BR T/C in C.A.", "Neth. T/C in C.A.", "Origin Unidentifiable", and "British Pacific Is". These were given new codes and added to the correspondence table.

The units of quantity used in Taiwan and UN trade statistics also differ, and code conversion was therefore also necessary in this case. The tables formulated in section 1 above were edited into one table. Missing codes were supplemented by referring to print statistical materials. A list of units of quantity printed on the first page of the Taiwanese statistical materials was referred to. If a judgment on quantity could still not be made, the relevant records in the original data were compared with the statistical materials in order to find records in which quantity, transaction value, and partner country matched, and in which there was a record of units. The Taiwanese standard in terms of units was also respected, but new codes have been assigned as necessary.

Main quantities are generally recorded in kilograms, but there are cases in which metric tons have been used. In such cases, quantities have been multiplied by 1000 and considered as kilograms.

The Monthly Reports that form the basis of Taiwanese trade statistics do not indicate the HS revision that is employed, making it necessary to specify the revision. First, the number of times that commodity classification codes appeared was checked for each year. From 1989 to 1996, codes are concentrated around HS1988 categories, and the revision was therefore judged as HS1988. From 1998 to 2002, codes are concentrated around HS1996 categories, and the revision was therefore judged as HS1996. From 2004 to 2005, codes are

concentrated around HS2002 categories, and the revision was therefore judged as HS2002. However, because HS1988 and HS1996 categories were almost equally mixed in 1997, this was judged as a transitional period between revisions. On Table 4, this period is shown as Mixed (1). Similarly, 2003 was judged as a transitional period between HS1996 and HS2002, and is shown as Mixed (2).

3. Data Processing for Conversion

This section will discuss the data processing procedures employed to convert Taiwan trade data to UN COMTRADE database data format. The respective formats of the data were compared to determine the categories employed, and the following conversion procedures were decided on.

The country code for Taiwan itself(490) in the UN COMTRADE database was set as the reporting country code. Based on the combination of *D1* and *D2*, the following import/export categories were employed:

$D1=1$ and $D2=0$ →2 (Export)
 $D1=2$ and $D2=0$ →1 (Import)
 $D1=1$ and $D2=1$ →4 (Re-import)
 $D1=2$ and $D2=1$ →3 (Re-export)

With respect to commodity classifications, CCCN codes have been transposed to the → Commodity code column. Up to 1988, CCCN classifications are employed, and from 1989 HS codes are employed. "CN" has therefore been appended to commodity classifications up to 1988, and "HS" to commodity classifications from 1989. In the UN COMTRADE database, the 1988, 1996, and 2002 HS revisions are classified as H0, H1, and H2 respectively. Correspondence with each revision is also required in the case of HS, but because old and new commodity classifications are mixed during periods of transition, HS has been considered as

unified here. Detailing the codes for HS revisions remains an issue for the future.

Three-digit country codes have been converted to five-digit UN country codes using the conversion table included as Annexed Table 1, and input in → Country code (trading partner country). Year (Western calendar) has been set with consideration of the names used for the input files for Taiwan data. Values (CIF or FOB) have been transposed to → Value, main quantities to →Main quantity (Kg), and sub-quantities to →Sub-quantity. Taiwanese unit

codes have been converted to UN unit codes using correspondence table in section 2 and transposed to → unit used for sub-quantities in UN format. If there were no main quantities but there were sub-quantities in kilograms, these quantities have been transposed to Main quantities (Kg), and zero recorded for sub-quantities in the data as standardized to UN format. If there were main quantities in metric tons, the quantity has been multiplied by 1000 to convert it to kilograms, and it has been recorded as the main quantity in the data as standardized to UN format.