Chapter 1

Evaluation to Consistency of Compiling World Trade Matrix:

Based on Basic Trade Statistics for New and Former AID-XT

NODA Yosuke and FUKAO Kyoji

The existence of long-term and continuous data for the statistical elements transaction value and quantity, and a lack of significant fluctuations in unit prices obtained by dividing transaction value by quantity, when viewed as a time series, are important considerations in the formulation of long-term time series data for trade statistics and trade indices. In addition, to formulate indices for individual partner countries rather than for world, it is essential to attain an accurate understanding of the changes in the partner country's data over time. This chapter will evaluate the consistency of classification categories, namely commodity, country and quantity unit in new and former basic statistics for AID-XT; Ajiken Indicators of Developing ecomies : eXtended for Trade statistics.

Basic statistics for AID-XT, namely the world trade statistics database compiled and maintained by Institute of Developing Economies: IDE, has two subcategories: former one and new one. Former basic statistics for AID-XT（former AID-XT in short）is composed of UN trade statistics purchased directiy and obtained from UN Comtrade database compiled by the UN Statistics Division, OECD and Taiwan trade statistics. The heterogeneities in the data arising from its formulation by different organizations were rendered consistent using IDE codes. New AID-XT compiled by IDE is composed of UN trade statistics obtained from another UN Comtrade database which the UN Statistics Division has made available online since 2003 and Taiwan trade statistics. While former AID-XT employed OECD trade statistics for OECD member nations, new AID-XT is characterized by its sourcing of data for all countries other than Taiwan from UN trade statistics.

This chapter makes attempt to examine methods of correction for both of AID-XT in as much detail as possible. Y. Noda has previously presented an overview of evaluation of consistency and correction of trade statistics in "Evaluation to Revision to Consistency of Compiling World Trade Matrices". However, that study employed former AID-XT, and evaluation of consistency was limited to transaction value with respect to statistical values and commodity classifications with respect to classification categories. In addition to commodity classifications, the other major classification categories in trade statistics are partner country and unit of quantity. Statistical values further include quantity. In this chapter we will broaden our focus to take in partner country, unit of quantity and quantity

in addition to commodity classification, and will formulate general evaluation and correction procedures for trade statistics.

1. Consistency in trade statistics

Trade statistics are normally broken down into six classification categories, as follows: Reporting country (*rc*), year (*y*), direction of trade (*d*), commodity classification (*c*), partner country (*pc*) and unit of quantity (*qu*), and are formulated employing two statistical values, transaction value (v) and quantity (q), for each of these classification categories. Irrespective of the commodity classification system or their relationship to digit level classification code, at IDE, where commodity classification codes have a transaction value greater than zero in the statistical data actually obtained, the codes which do not have lower strata classification codes are termed the most detailed classification code (*mdcc*).

Terming the commodity classifications composed of *mdcc* as $C = \{Total\ C_1 \cdots C_m\}$ and partner countries $P = \{World\ P_1 \cdots P_n\}$, transaction value data obtained as trade statistics can be expressed as

(1-1) $\qquad v_{ij}(rc,d,y) \quad i \in C, j \in P$

for commodity classification C_i and partner country P_j for each reporting country, year and direction of trade. If we fix reporting country, year and direction of trade, it can be expressed simply as v_{ij}. If the total value of commodities in commodity classification *i* is expressed as *T* and world for partner country *j* is expressed as *W*, v_{TW} is simultaneously total value of commodities and transaction value for partner country world.

Using total commodity value v_{TW} as a standard enables us to evaluate the consistency of the

trade matrix. Commodity classifications in trade statistics contain stratified digit level classification codes, and transaction value data exists for all of these codes at their respective levels. When an SITC system is used as the commodity classification,

(1-2) $\qquad v_{ij}(rc,d,y)[k] \quad i \in C, j \in P$

or simply expressed, $v_{ij}[k]$, where the transaction value expressed by the *k*-digit level classification code is $k = 1,2,3,4, mdcc$. When HS system is used, the k-digit level classification code is expressed as $k = 2,4,6, mdcc$.

To maintain the consistency of this transaction value in the trade matrix, we establish error terms for partner countries, *error of P*, or $e_p[k]$, and *k*-digit level commodity classifications, *error of C* [*k*], or $e_c[k]$. We further employ $e_{c,p}[k]$, the intersection of the *error of P* and *error of C* [*k*] terms as shared error for partner countries and commodity classifications. This enables a consistent world trade matrix to be formulated on the basis of *k*-digit level commodity classifications. Terming total partner countries and total commodity classifications (*k*-digit level classification codes) $v_{\bullet\bullet}[k]$ and following the definition of error of consistency provided by Noda in "Evaluation to Revision to Consistency of Compiling World Trade Matrices", we express the error in the trade matrix originating from commodity classifications

(1-3) $\qquad e_c[k] + e_{c,p}[k] = v_{TW}[k] - v_{\bullet W}[k]$

and from partner countries

(1-4) $\qquad e_p[k] + e_{c,p}[k] = v_{TW}[k] - v_{T\bullet}[k]$

Total error is the sum of error originating in commodity classifications, the error originating in partner countries and shared error (commodity classifications and partner countries), and is expressed as

(1-5) $\qquad e[k] = v_{TW}[k] - v_{\bullet\bullet}[k]$

Table 1 is a summarized trade matrix transaction

Table 1 Summarized trasaction table of trade matrix generated by correlating partner countries to columns and k-digit level of commodity classifications to rows

C P	P : Total of partner countries	Error of P	World
$C[k]$:Total of commodity error of $C[k]$	$x_{\bullet\bullet}[k]$ $e_c[k]$	$e_p[k]$ $e_{c,p}[k]$	$x_{\bullet W}[k]$ $e_c[k]+e_{c,p}[k]$
Total	$x_{T\bullet}[k]$	$e_p[k]+e_{c,p}[k]$	x_{TW}

value table using k-digit level classification codes in commodity classifications.

The consistency evaluation tables for k-digit level classification codes used in the formulation of trade matrices are arranged in the following order for each reporting country and direction of trade: year y, commodity classification system sc, total value of commodities v_{TW}, total error $e[k]$, number of k-digit level classification codes $\{d_1 \cdots d_6\}$, signed absolute error of error originating with commodity classifications $e_c[k]+e_{c,p}[k]$ and error originating with partner countries $e_p[k]+e_{c,p}[k]$, and relative error of v_{TW}. Table 2 shows an example of a consistency evaluation table using data for the US and China from new AID-XT. For the US, consistency is evaluated based on *mdcc* from 1981 to 1990 using SITC for commodity classification; for China, consistency is evaluated based on *mdcc* from 1992 to 1997, with HS as the system of commodity classification.

Consistency has been achieved in commodity classifications when the total transaction value of the lower level classification codes in a stratified classification system composed of codes of various digit levels matches the transaction value of the corresponding higher level classification codes.

What we are terming correction of inconsistency in commodity classifications in this section is, given a situation in which the total transaction value of the lower level classification

codes does not match the transaction value of the corresponding higher level codes, is a process which utilizes the higher level rather than the lower level codes, and ensures that after substitution the total transaction value of all the classification codes matches the total value of commodities.

2. Consistency of partner countries

While it is possible to evaluate the consistency of trade statistics in terms of partner countries, because such evaluation relies on a sum check based on the total value of commodities for each reporting country year and direction of trade and v_{TW} of partner country world, it is not possible to evaluate the conversion error for individual partner countries when the country codes employed in the UN, OECD and Taiwan trade statistics are converted to unified IDE country codes. The consistency of individual partner countries can be evaluated by comparing the total transaction value for imports and exports for each partner country based on totals for commodities in trade statistics formulated by other international organizations with total transaction values in AID-XT base data. For example, total values for partner countries in OECD trade statistics and IMF Direction of Trade (DOT) data are used for comparison when evaluating partner countries in UN trade statistics.

Using the ratio of total value of commodity to

表 11　Hong Kong transaction import for commodity total by partner countries in 1999

g	ide_cc	pc	H	$H^{US\$}$	U	O
	000000	World	1392718	180710649	180710649	180710649
1	118140	Korea Rep.	65432	8490060	8773650	63487
1	117180	Korea Dem. P. Rep.	397	49176	63487	8773650
2	117960	Taiwan	100426	13030669	0	12944869
2	198340	Other Asian nes	0	0	12944869	0
3	154940	Qatar	104	13494	13456	15563
3	155282	Oman	121	15700	15563	58
3	155283	United Arab Em	21229	2754546	274543	287999
7	422960	Dominican Republic	5	649	618	0
7	423395	Dominica	0	0	0	618
8	428940	Neth Antile	0	0	343	0
8	428942	Aruba	.	.	58	343

note: *G* represents commodity group, *ide_cc* IDE county code, *pc* description of country. *H* is obtained from Hong Kong trade statistics of CD-ROM and its unit is million Hong Kong \$, Unit of $H^{US\$}$ based on *H* is converted into 1.000US\$.　*U* and *O* are UN trade statistics and OECD trade statistics with 1,000US\$, respectively.　Symbol . means value less than 1,000US\$.

partner country world enables comparison to trade statistics using the local currency of the formulating country without conversion to US dollars. Table 11 shows different transaction values based on Honk Kong trade statistics, UN and OECD trade statistics for Hong Kong transaction import for commodity total by partner countries in 1999. The table makes sure of the facts that Korea Rep and Korea Dem. P. Rep. are reversed in O.

3. Consistency of units of quantity and quantity

Units of quantity and quantities in basic statistics for AID-XT drawn from UN, OECD and Taiwan trade statistics are converted to unified IDE quantity codes to ensure consistency. Table 13 shows cross references between the major units of quantity used by the various formulating organizations and the IDE quantity unit codes. As this table makes clear, different organizations do not always employ common units of quantity. Therefore, an international evaluation of consistency of units of quantity involving comparison of different formulating organizations is considerably more complex than evaluation of the consistency of units of quantity for each reporting country in data from a single formulating institution.

Only the former method is employed in this chapter, and four standards are employed to evaluate the consistency of units of quantity and quantity. The first of these evaluation criteria is whether or not partner countries are consistent in terms of quantities for each commodity classification and unit of quantity. The second is whether or not there is any change in the same unit of quantity for each commodity classification when considered as a time series. The third is whether or not there are any changes over time in commodity classifications with identical *mdcc*. The fourth criterion is whether there are any changes over time in unit prices when units of quantity are identical.