

Part 2 : Chapter 4 - Formulation of Trade  
Indices based on UN COMTRADE for SITC  
Revisions

権利	Copyrights 日本貿易振興機構（ジェトロ）アジア 経済研究所 / Institute of Developing Economies, Japan External Trade Organization (IDE-JETRO) <a href="http://www.ide.go.jp">http://www.ide.go.jp</a>
シリーズタイトル(英 )	I.D.E. statistical data series
シリーズ番号	91
journal or publication title	Trade-Related Indices and Trade Structure
page range	121-129
year	2007
URL	<a href="http://hdl.handle.net/2344/00008942">http://hdl.handle.net/2344/00008942</a>

# Chapter 4

## Formulation of Trade Indices based on UN COMTRADE for SITC Revisions

KUROKO Masato

### Introduction

The Institute of Developing Economies has formulated trade indices (indices for unit value, value, and quantity) each fiscal year from FY2002. In FY2004, the indices were formulated using UN trade statistics (COMTRADE) for SITC Revision 1 (SITC-R1) obtained from the Internet<sup>1</sup>. For the present fiscal year, the Institute set itself the aim of formulating more accurate long-term indices with a minimum of abnormal fluctuations. Therefore, in addition to aggregating COMTRADE data using Kinoshita and Yamada's 20 industrial classifications (see Table 1) as in FY2004, we conducted calculations based on all the SITC revisions rather than solely on SITC-R1, in order to generate trade indices that form long-term time series<sup>2</sup>. This chapter will discuss the methods we have employed and the modifications we have made in formulating these indices.

### 1. Input Data and Formulated Indices

COMTRADE data obtained from the UN website in 2005 was employed as input data in the formulation of the present indices<sup>3</sup>. In the absence of any requirement for further qualification, this online UN trade data will be termed COMTRADE. AID-XT basic data was used for the entirety of the Taiwan

data, and also to supplement any data that was unavailable in COMTRADE<sup>4</sup>.

In FY2004, we employed time series data using SITC-R1 in our formulation of trade indices; we did not employ R2 or R3 data. This fiscal year we formulated indices using data for all the SITC revisions, i.e. indices connected by "multiple SITC revisions".

For each SITC and HS revision, the COMTRADE data set contains data from the period in which the revision was applicable, and data from later periods. The latter is data using the original revision converted by the UN to data for another revision. For example, in the case of Japan, R1 data in the SITC data series covers the reporting years 1962 to 2003. Similarly, R2 covers the years 1976 to 2003, and R3 covers 1988 to 2003. In the HS series, HS-0 covers the years 1988 to 2003, HS-96 1996 to 2003, and HS-2002 2002 to 2003 (see Fig1-1). Only the SITC series has been used in the formulation of the trade indices discussed in this paper. Original, unconverted data was employed, in the case of R1 from 1962 to 1975 and in the case of R2 from 1976 to 1987. These data were employed to enable the formulation of trade indices corresponding more closely with original data (see Fig1-2). From 1988 to 2003, the original data was HS series data, therefore, we used SITC R3 data converted from originally HS data for use in formulating the present indices. This is

Table 1 Kinoshita and Yamada's 20 Industrial Classifications

Number	Description	Number	Description
01	AG Agricultural products	11	PC Petrochemical products
02	MI Mining	12	NM Ceramics
03	FD Foodstuffs	13	IS Iron and steel
04	TX Textile	14	NF Non ferrous
05	AP Apparel	15	MT Metal products
06	LT Leather products	16	MC Machinery
07	WD Lumber and wood products	17	EM Electrical equip. & mchn.
08	PP Paper and pulp	18	TE Transport equip.
09	RB Rubber and plastics	19	PI Precision instruments
10	CH Chemical products	20	MM Miscellaneous pro.

(Source) Kinoshita and Yamada (1993). For a correspondence table with SITC commodity classifications, see Noda and Kuroko (2006).

Fig1-1 Period by each SITC revision in which COMTRADE data exists (Reporting country = JPN)

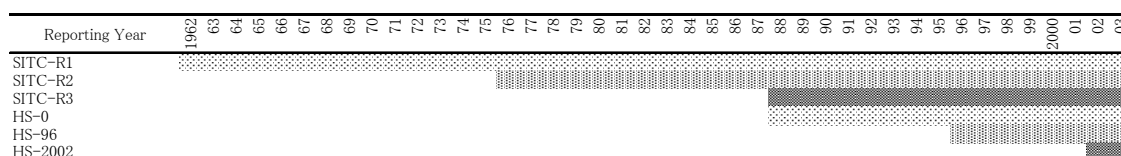


Fig1-2 FY2004's method for formulating Indices using original data of each SITC revision. (Index series use forward and backward direction.)

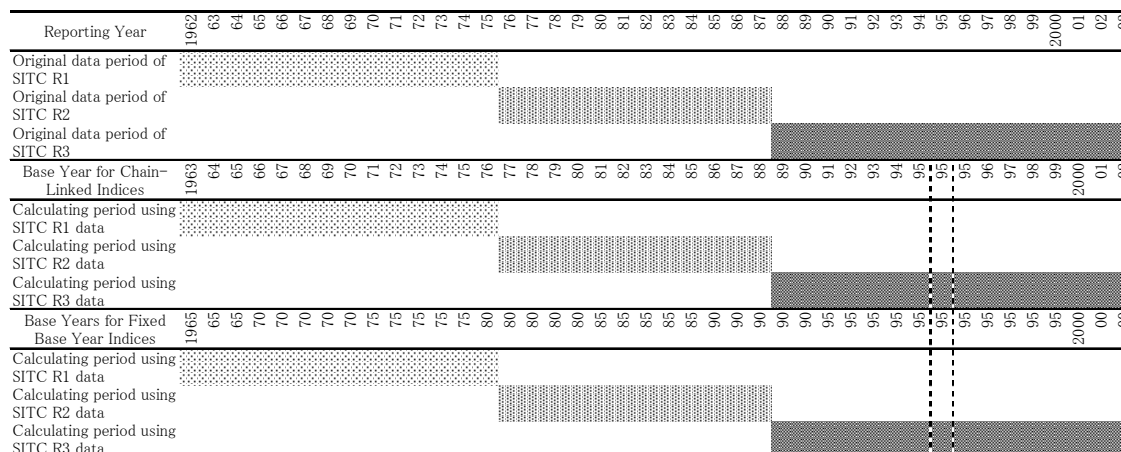
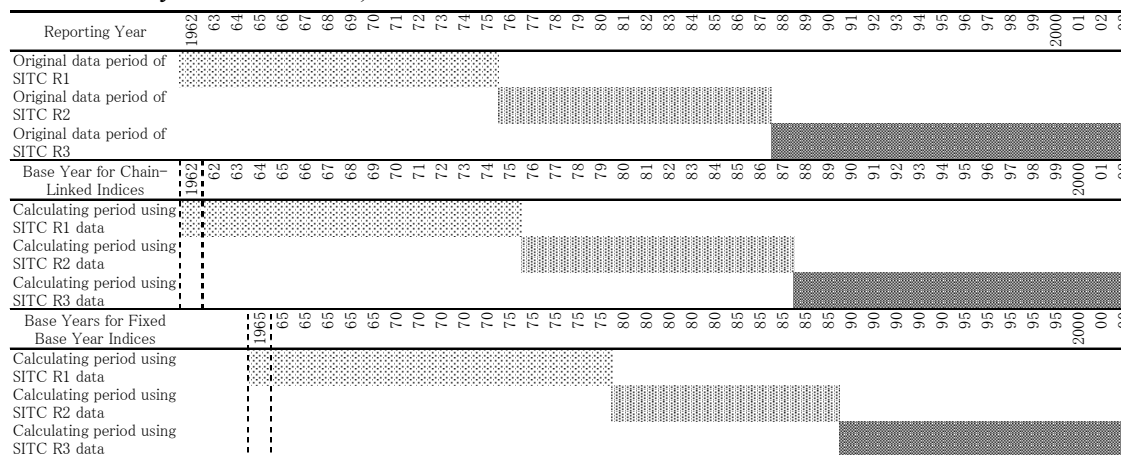


Fig1-3 This fiscal year's method for formulating indices using original data of each SITC revision. (Index series use solely forward direction.)



(Source) Formulated by author.

Table 2 Conditions for Formulation of Present Indices

Item	Conditions
Reporting country	<p>The 31 reporting countries and regions represented in COMTRADE and Taiwan. In the following list the code in brackets is the ISO three-letter country code.</p> <p>Australia (AUS), Austria (AUT), Belgium (BEL), Canada (CAN), China (CHN), Germany (DDR, DEU), Denmark (DNK), Spain (ESP), Finland (FIN), France (FRA), United Kingdom (GBR), Greece (GRC), Hong Kong (HKG), Indonesia (IDN), Ireland (IRL), Iceland (ISL), Italy (ITA), Japan (JPN), South Korea (KOR), Luxembourg (LUX), Malaysia (MYS), Netherlands (NLD), Norway (NOR), New Zealand (NZL), Philippines (PHL), Portugal (PRT), Singapore (SGP), Sweden (SWE), Taiwan (TWN), Thailand (THA), Turkey (TUR), United States (USA)</p>
Import/export category	<p>All imports, exports and re-exports (Re-export value is included in export value in online COMTRADE data).</p>
Trading partner country/country group	<p>World total and country groups. The countries in each group are shown below. These groups are taken from the trading partner country group table (GP_CTRY). Figures in brackets represent UN country codes.</p> <p>Trading partner country group EU: Austria (40), Belgium (56), Belgium-Luxembourg (58), Denmark (208), Finland (246), France (251), Germany (276), Fmr Dem. Rep. of Germany (278), Fmr Fed. Rep. of Germany (280), Greece (300), Ireland (372), Italy (381), Luxembourg (442), Netherlands (528), Portugal (620), Spain (724), Sweden (752), United Kingdom (826)</p> <p>Trading partner country group JPN: Japan (392), Ryukyu Isd (647)</p> <p>Trading partner country group ASIA: China (156), "China, Hong Kong SAR" (344), Indonesia (360), Rep. of Korea (410), Fmr Sarawak (457), Malaysia (458), Fmr Peninsula Malaysia (459), Fmr Sabah (461), Taiwan (490), Philippines (608), Singapore (702), Thailand (764)</p> <p>Trading partner country group US/CA: Canada (124), USA (before 1981) (841), USA (842), US Virgin Isds (850)</p> <p>Trading partner country group ETC: Countries and regions not included in the above (Some countries and regions not included)</p> <p>The following trading partner country categories have been excluded. Figures in brackets represent UN country codes.</p> <p>Br. Antarctic Terr. (80), "Caribbean, nes" (129), "Eastern Europe, nes" (221), "Northern Africa, nes" (290), "CACM, nes" (471), "Africa CAMEU region, nes" (472), "LAIA, nes" (473), "Europe EU, nes" (492), "Oceania, nes" (527), "Neutral Zone" (536), "Other Eurpe, nes" (568), "Other Africa, nes" (577), "Rest of America, nes" (636), "Europe EFTA, nes" (697), Bunkers (837), Free Zones (838), Special Categories (839), US Misc. Pacific Isds (849), "Western Asia, nes" (879), "Areas, nes" (899)</p> <p>World totals include both totals calculated using trading partner country = world data in COMTRADE and aggregated totals using individual trading partner country data.</p>
Types of indices	<p>Unit value, value and quantity indices formulated using Laspeyres, Paasche and Fischer formulas. Indices were first calculated for each of the Kinoshita and Yamada 20 industrial classifications (KY20), and weighted aggregates of these indices were used as general indices.</p>
Base year	<p>Indices were calculated using the following two base year methods.</p> <p>(1) Indices were formulated with every fifth year (1965, 1970, 1975, ..., 2000) as the base year, and years in the five-year period following the base year as the comparison years (For example, when 1993 is the comparison year, the base year is 1990). These were chain-linked to form an index series with 2000 as 100. (Fixed base year method)</p> <p>(2) Indices were formulated with the previous year as the base year (For example, when 1999 is the comparison year, 1998 is the base year). These indices were chain-linked to form an index series with 2000 as 100. (Chain-linked method)</p>

(Source)Formulated by author.

considered to minimize the decline in accuracy due to conversion of commodity classifications.

Kinoshita and Yamada (1993) showed correspondence only between the 20 industrial classifications (KY20) and SITC-R1 and R2, but we added an original correspondence table with R3 in this project in order to enable formulation of indices using multiple revisions. Correspondence with the HS series remains an issue for the future<sup>5</sup>.

When calculating trading partner country = world indices, in addition to using the trading partner country = world total data available in COMTRADE, we aggregated individual trading partner country data and used it in formulating trading partner country = world indices. The use of data for individual trading partner countries increased calculation time, but this method was expected to enable the formulation of indices that correspond more closely to reality.

Table 2 shows the conditions used in the formulation of the trade indices in this project.

## 2. Procedures of Formulation of Indices

Fig2 shows an overview of the procedure to formulate indices in this project<sup>6</sup>.

1. Formulation of conversion tables for Kinoshita and Yamada's 20 industrial classifications (KY20): First, conversion table (SITC05 table) with correspondence between the SITC and the KY20 were formulated based on the correspondence tables in Kinoshita and Yamada (1993), and the correspondence tables with SITC-R3 formulated in this project. SITC classifications that no correspondences could be established with were recorded on SITC06 table, and correspondences of this data were later added to the conversion table<sup>7</sup>.

2. Formulation of COMTRADE original data tables: In this step, the flat files retrieved from COMTRADE were imported without alteration into

CTO<CTRY><CREV> tables for each reporting country (CTRY) and commodity revision (CREV). CTM<CTRY><CREV> tables were formulated by selecting only the most detailed commodity classification (MDCC) data from CTO<CTRY><CREV> tables<sup>8</sup>.

3. Supplementation of missing data with AID-XT data: Data missing in the COMTRADE data series was supplemented with AID-XT basic data.

4. Conversion from the SITC to the KY20: Using the conversion tables for the KY20 industrial classifications formulated in step 1, item of KY20 classification converted from the SITC was added to the tables (TR\_<CTRY>).

5. Formulation of correspondence tables for SITC revisions and reporting years: Correspondence tables expressing the relationship of correspondence between the SITC revisions and reporting years (Fig1) were separately formulated for chain-linked and fixed base year methods (REVYEAR4 table, REVYEARF04 table).

6. Formulation of chain-linked indices (1): The TR\_<CTRY> tables and the correspondence tables with reporting years for each revision formulated using chain-linked methods (REVYEAR4 table) were joined (TR03 table) and aggregated for each KY20 classification (TI01 table). Unit value indices were calculated. These unit value indices calculated using each formula; Laspeyres, Paasche and Fischer. Naturally, unit value indices could not be calculated in cases in which quantity data was missing, because unit value were calculated by value divided by quantity for each trade data. In addition to this, cases in which the rate of change from the previous year was greater than 5× or lower than 1/5 were not employed in calculations. Basket commodity items were also excluded. Basket commodity items include various heterogeneous commodity items that are not necessarily able to be classified to existing

classifications. Here, items with SITC ending with 9 or beginning with 9 (SITC section 9: “Commodities and transactions not classified elsewhere in the SITC”) were regarded as basket commodity items.

7. Formulation of chain-linked indices (2) (Formulation of indices for KY20 classifications): Value and quantity indices were calculated from the TI01 table, and index series with 2000 as the base year were formulated (TI04 table). Coverages were calculated from the TR03 table into TC03 table. The TC03 and TI04 tables were joined to produce final output of indices for each of the KY20 classifications with coverages. How much amount of value and how many numbers of data in the total MDCC data were employed in calculations are important problem evaluating the validity of indices. These were calculated as a coverage for value and a coverage for data and were also recorded on the index tables for the KY20 classifications. Value indices were calculated on the basis of the total MDCC data. Quantity indices were calculated by dividing the value indices by Fischer unit value indices.

8. Formulation of chain-linked indices (3) (Formulation of weighting tables): The weighting of each KY20 classification for each of the same key items (i.e. the same reporting country, partner country, direction of trade and reporting year [base year or comparison year]) was calculated in order to enable formulation of general indices from the indices formulated for the KY20 classifications. In Laspeyres weighting tables (TW55), data was aggregated for each base year, and in Paasche weighting tables (TW58) for each comparison year.

9. Formulation of chain-linked indices (4) (Formulation of general indices): The weighting tables formulated in step 8 and the aggregated table for each KY20 classification (TI01 table) were joined to enable formulation of general indices (TI21 tables). Value indices (TI22 tables), like the index tables for each KY20 classification, were calculated

based on the entire data series. Value indices (TI23 table) were calculated by dividing the value indices by Fischer unit value indices. Index series with 2000 as the base year were calculated (TI24 table), and the TC23 table on which the coverages were determined and the TI24 table were joined to formulate general indices.

10. Formulation of chain-linked indices (5) (Formulation of terms of trade indices): Terms of trade indices were calculated by dividing the export indices formulated in steps 7 and 9 by import indices. Terms of trade indices were also calculated for each of the KY20 classifications as well as general indices.

11. Formulation of fixed base year indices (1) and 12. Formulation of fixed base year indices (2): These steps for fixed base year indices were corresponding steps to step 6 for chain-linked indices. In the case of fixed base year indices, data which exist every reporting country for five years was extracted from the TR\_<CTRY> tables (MDCC data with KY20 classifications), to formulate TR03 table. This table was joined with the revision / reporting year correspondence table (REVYEARF04 table) for fixed base year indices to formulate aggregated table for each KY20 classification (TI01 table). Following this, the procedures for fixed indices were identical to those described for chain-linked indices in steps 7 to 10.

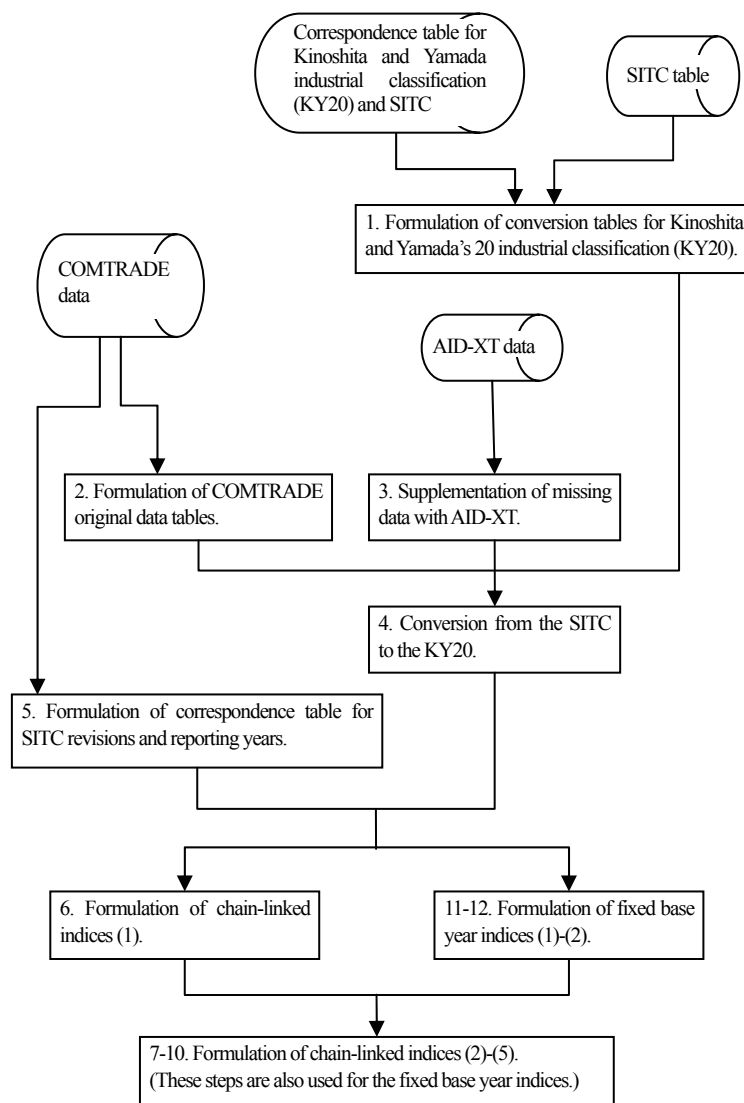
### **3. Changes in Method of Formulating Indices**

Some modifications have been made to the method of calculating indices compared to the FY2004 method.

#### **3.1 Establishment of Uniform Direction of Indices**

The direction of the indices calculated in FY2004

Fig2 Overview of the procedure to formulate indices



(Source)Formulated by author.

differed before and after 1995. Indices before 1995 were directed backwards through time, and indices after 1995 were directed forward. This is indicated by the base year for each index in Fig1-2. 1995 represented a boundary; before 1995 base years were years following the comparison year, and after 1995 base years were years preceding the reporting year. For example, in the case of chain-linked indices, if the comparison year was 1981, the base year was 1982, and if the comparison year was 2001, the base year was 2000. As Fig1-1 shows, because data exists for all revisions up to the most recent year, when

indices are formulated using backward direction, the breaks between revisions in 1975/76 and 1987/88 can be ignored. In the case of chain-linked indices, the base year for the comparison year 1975 is 1976. Because 1976 actually falls within the period of use of R2, base year data is unavailable, and R1 indices should be formulated only between 1962 and 1974. However, because the COMTRADE data includes R2 data converted to R1 data for 1976 and later, as indicated in Fig1-1, 1976 R1 data (converted from R2 data) can be employed to formulate indices for 1975. The same is true of the break between R2 and

R3 in 1987/88. Calculations can also be conducted in the same way for fixed base year indices.

As this indicates, in the case of indices of backward direction, the breaks between revisions are unproblematic. However, it is generally known that different directions of index change the characteristics of indices. For example, Laspeyres indices of backward direction are the inverse of Paasche indices of forward direction<sup>9</sup>. When index calculations were conducted using individual trading partner data with the same base year direction as FY2004, unacceptable discrepancies resulted between the Laspeyres and Paasche indices. For this reason, all the index series in this project all used forward direction in order to enable accurate comparisons.

In the case of chain-linked indices which Fig1-3 typically shows, when the index series solely use forward direction, the base year is always one year before the reporting year. For example, if the comparison year is 1981, the base year is 1980, and if the comparison year is 2001, the base year is 2000. These cases are unproblematic; problems arise at the breaks between revisions. Using the example discussed above, in the case of chain-linked indices, at the 1975/76 break between revisions, the base year for 1975 is 1974, and the base year for 1976 is 1975. In the case of 1975/1974, R1 data is available for both years, and this case is therefore unproblematic. In the case of 1976/1975, however, only R2 data is available for 1976, and indices can therefore not be formulated. The index series must be terminated at 1975. For this reason, as shown in Fig1-3, a one-year lag is introduced, and the break between revisions is considered as occurring in 1976/77 rather than 1975/76. This makes it possible to use R1 data to calculate indices for 1976 and to use R2 data for 1977, thus making the index series continuous. This is also true with regard to the break between 1987 and 1988.

In the case of fixed base year indices, a further lag must be introduced because the indices are formulated using five-yearly base years. For example, because 1975 is also the base year for 1976, R1 data must be used for both 1975 and 1976. 1975 is the base year for the entire five-year period 1976-1980, and R1 data must therefore be used for the entire period. For 1981, the base year is 1980, which falls within the period of use of R2, and R2 data can therefore be used. In the same way, a lag of five years must be used in the case of the break between R2 and R3 (see Fig1-3).

### **3.2 Formulation of Indices by Trading Partner Country Group**

In FY2004, the IDE formulated indices using only data for trading partner country = world total. However, data for trading partner country = *individual country* was employed to formulate aggregated indices for trading partner groups (EU, Japan, Asia, US/Canada, etc.)<sup>10</sup>. This enabled analyses that were not possible using the trading partner country = world total data to be conducted, such as determining whether significant differences existed between indices for each of the trading partner country groups. In addition, assuming that reporting country conducted trade of commodities of different quality depends on its trading partner country, aggregating indices for trading partner groups would render commodity quality uniform, and this could be expected to stabilize fluctuations in the indices. Special categories such as nes (not elsewhere specified) exist as trading partners in the COMTRADE data. These data have been excluded from the formulation of indices<sup>11</sup>.

### **3.3 Changing the Base Year in Fixed-Base-Year Calculations**



In FY2004, when calculating fixed base year indices, commodity sets differ in different comparison years for which same base year is employed. For example, in the case of export unit value indices for US chemical products, indices were formulated for the years 1996 to 2000 with 1995 as the base year, but because the applicable SITC commodities differed with each reporting year, the SITC commodities set making up the content of the chemical industry differed in each reporting year. In the sense that makeup of commodity set differed year-to-year in fixed base year indices, they had identical characteristics to chain-linked indices.

This fiscal year, we employed more exacting criteria in the formulation of fixed base year indices: only commodity items for which trade data was available for all the comparison years having the same base year were used in calculations. In the example cited above, only commodities for which trade data is available for all the comparison years in the five-year period between 1996 and 2000 would be used in the formulation of fixed base year indices for chemical industry. This condition can be assumed to reduce the coverage, but by ensuring that indices for the same industrial classifications in the same comparison year were formulated on the basis of the same commodity set, it would also further minimize fluctuations and highlight the different characteristics of the fixed base year indices as compared to chain-linked indices.

In FY2004, the IDE has applied the criteria of only using data falling between figures of 1/5 and 5× for rate of change in unit value compared to the base year, and this criteria has also been applied this fiscal year.

### **3.4 Increased Connectivity between Index Series**

Chain-linked indices formulated for each year and fixed base year indices formulated for each five years

were linked to form index series. If indices cannot be formulated for only one year, the indices for the following years cannot be linked. This has been particularly true in the case of indices of US machinery-related industry. One solution to this problem is to link the indices treating indices for the year in question as 100 (this assumes that no change has occurred). This method was adopted this fiscal year, and has resulted in increased connectivity between the indices.

However, opinion exists to the effect that the assumption that indices assigned for a year for which they could not be formulated display no change in comparison with the base year is unwarranted. Further study of methods of responding to this problem will be required.

### **Conclusion**

Future issues to be addressed with regard to the formulation of indices can be indicated as follows:

- (1) Quantitative evaluation of the extent to which the modifications made this fiscal year have enabled irregular fluctuations to be controlled;
- (2) Evaluation of indices by means of comparison with indices published by national governments;
- (3) Establishment of methods of evaluating and supplementing missing indices based on other indices, etc.;
- (4) Formulation of indices linking multiple commodity classification revisions, including the HS series;
- (5) Calculation of indices with lower levels of deviation, such as geometrical mean indices, Toernqvist indices, etc.; and
- (6) Use of the trimmed mean method in order to exclude data with significant deviations from the distribution of fluctuation of the indices.

- 1 See Kuroko, M. (2005).
  - 2 See Kinoshita and Yamada (1993).
  - 3 <http://unstats.un.org/unsd/comtrade/>  
There is a considerable difference in format between data available from this URL up to 2005 and data available from 2006. Data in the original format, obtained in 2005, was employed in this chapter.
  - 4 COMTRADE data is missing for the following categories:
- | Reporting country | Reporting year | Direction of trade |
|-------------------|----------------|--------------------|
| Japan             | 1992           | Import, export     |
| Thailand          | 1988           | Export, re-export  |
- 5 See Noda and Kuroko (2006), pp. 11-20, for conversion tables with all SITC revisions, including KY20-SITC R3 correspondence.
  - 6 For a more detailed figure, see Kuroko (2006), pp.106-117.
  - 7 The following classifications exist in SITC-R1, but were not included in the conversion tables (Kinoshita and Yamada (1993)): “0”, “1”, “2”, “3”, “33”, “6”, “7”, “8”. Of these, “33” was excluded because it was contained in “3”. “0”, “1”, “2”, “3”, “6”, “7”, “8” were added to the conversion tables later, and corresponded to KY20 classification “21”.
  - 8 See Noda and Fukao (2005) for considerations on the most detailed classification (MDCC) and correction. Because MDCC data used at stages in this project was uncorrected, the totals for this data may not match commodity totals. The coverages indicated later in the text therefore do not represent the ratio of applied data to the product totals, but the ratio of the applied data totals to the MDCC data totals.
  - 9 See Allen (1975), pp. 59-60.
  - 10 See Table 2 for countries included in country groups.
  - 11 See Table 2 for excluded trading partner categories.

- [1]Allen,R.G. D. 1975. *Index Numbers in Theory and Practice*, Macmillan Press.
- [2]Kinoshita, S., and Yamada, M. 1993. "Estimates on Sectoral Export Price Indices by Country for Global Economic Modeling" , *Economic Research*, No.97, Nagoya University.
- [3]Kuroko, M. 2005. "Classification of trade indices based on UN COMTRADE data connected by SITC-R1", in Y. Noda (eds.), *Trade Indices in East Asian Countries and Regions: Basic Subjects from Compilation to Application, IDE Statistical Data Series No.88*, Institute of Developing Economies, JETRO.
- [4]Kuroko, M. 2006. "Kokuren Boeki Tokei ni motozuku Boeki Shisu no Kaitai", in Noda and Kuroko (eds.), *Choki Jikeiretsu ni okeru Boeki Data to Boeki Shisu no sakusei to oyo* (“Chapter 4. Revised Trade Indices based on UN COMTRADE”, in Noda and Kuroko (eds.), *Compilation and Application of Trade Indices and Trade Data in the long term.*), Institute of Developing Economies, JETRO. ([http://www.ide.go.jp/Japanese/Publish/Report/pdf/2005\\_02\\_04\\_04.pdf](http://www.ide.go.jp/Japanese/Publish/Report/pdf/2005_02_04_04.pdf)).
- [5]Noda, Y. and Fukao, K. 2005. "Evaluation to consistency of compiling world trade matrix: Based on basic trade statistics for new and former AID-XT", in Y. Noda (eds.), *Trade Indices in East Asian Countries and Regions: Basic Subjects from Compilation to Application, IDE Statistical Data Series No.88*, Institute of Developing Economies, JETRO.
- [6]Noda, Y. and Kuroko, M. 2006. "Higashi Ajia shokoku to beikoku ni okeru boueki kanren shisu no mikata", in Noda and Kuroko (eds.), *Higashi Ajia shokoku chiiki no boeki kanren shisu* (“Preface How to use the Trade related Indices of East Asian Countries and Regions and US.”, in Noda and Kuroko (eds.), *Various Trade Indices of East Asian Countries and Regions and US.*), Institute of Developing Economies, JETRO. ([http://www.ide.go.jp/Japanese/Publish/Report/pdf/2005\\_02\\_04a\\_mokuji.pdf](http://www.ide.go.jp/Japanese/Publish/Report/pdf/2005_02_04a_mokuji.pdf))