CHAPTER 4

METHODOLOGY OF THE STUDY
CHAPTER 4

METHODOLOGY OF THE STUDY

So far the conceptual and measurement issues of intra-industry trade have been discussed, now the present chapter broadly deals with the methodologies used in the study to discuss the nature, pattern and determinants on India’s IIT as well as the limitations of the study. The present chapter mainly focused on the system of Indian Trade Classifications (ITC-HS Code), different level of aggregation and their significance, the formula used to calculate Grubel-Lloyd index of calculating the degree of intra-industry trade, disentangling total intra-industry trade into horizontal and vertical intra-industry trade, calculating import-weighted average tariff rate and a brief discussion about panel-data analysis and in the last, the limitations of the study have also been discussed. All these points have been discussed below one-by-one.

4.1 SAMPLING TECHNIQUE (ITC-HS CLASSIFICATIONS)

The data used (secondary data only) in the present study has been taken from ITC (HS) Code. This is called as Indian Trade Classification – Harmonized System of classification, it classifies all the products into different product categories. India uses an eight-digit unique ITC-HS code for different products. Here in this case, all the products are divided into 21 sections and each section is further divided into different chapters. There are total 98 chapters – from HS-01 to HS-98. Each chapter is further divided into four-digit sub-headings and this division continues till the level of eight-digit. This eight-digit code is unique to each product and the trade is done using these codes. For example, at two-digit, chapter HS-85 belongs to “Electrical Machinery and Equipments and Parts thereof”; at four digit the chapter is divided into 48 sub categories like 8501, 8502 …. 8548; at six-digit the category 8501 is further sub-divided in 13 sub-categories like 850110, 850120 and so on; while eight-digit specifies a particular type of product; it is shown as – 85 is for - Electrical Machinery and Equipments and Parts thereof; 8501 is
for — “Electric motors and generators (excluding generating sets)”; 850110 is for — “Motors of an output not exceeding 37.5W” and 850120 is for “Universal AC/DC motors of an output exceeding 37.5W”; at eight-digit 85011011, 85011012, 85011013 are DC motors of different types, as micro motor, stepper motor and wiper motor respectively. This is the way by which each product is given an eight-digit unique code and they are traded and recorded according to their product code. The lists of the products of all chapters at two-digit, i.e., all 98 chapters are given in Annexure I.

All the data have been taken from the sample of all 98 chapters and finally four HS chapters have been selected for detailed study.

4.1.1 DIFFERENCE BETWEEN FOUR-DIGIT AND SIX-DIGIT AGGREGATION LEVEL

As far as different studies on international trade are concerned, it depends on the objective of the study that what level of aggregations is required. Some studies may be conducted at the highest level of aggregation, i.e., two-digit level while some may be conducted in the lowest level of aggregation, i.e., eight-digit. Normally, industry has been considered as either at four-digit or at six-digit level of aggregation. Since, theoretically it has been assumed that the level of aggregation affects the degree of intra-industry trade, thus the present study is based on both 4-digit and 6-digit level of aggregation. In the case of defining industry at the four-digit level, the value of all the products at six-digit level, under the same four-digit level, have been considered; while in the case of defining industry at six-digit level, the value of all the products at eight-digit level, under the same six-digit level, have been considered. Therefore, if industry is defined as at the four-digit level then it means that it is carrying the summation of all the values of six-digit level, at six-digit level the summation of all the values of eight-digit level have been considered.
4.2 MEASURING GRUBE-LLOYD INDEX

In the present study GL-index has been calculated for measuring the degree India's of intra-industry trade. The nature and pattern of IIT has been considered for both four-digit and six-digit level. Although different indices have already been proposed to calculate the level of intra-industry trade, but here GL-index has been used because even today it is the most widely used and most widely accepted index. To obtain the average level of intra-industry trade for a country \( j \), they proposed a weighted average measure of IIT\(_j\), which is shown as:

\[
IIT'_j = \frac{\sum_{i=1}^{n} (X_i + M_i) - \sum_{i=1}^{n} |X_i - M_i|}{\sum_{i=1}^{n} (X_i + M_i)} \times 100 \quad \ldots \quad (4.2.1)
\]

where \( IIT'_j \) is the degree of IIT for a country \( j \), \( X_i \) and \( M_i \) is the export and import value of industry \( i \) of the country, \( n \) is the number of industries. The equation 4.2.1 can be used for both at 4-digit and 6-digit level of aggregation. The equation ranges from 0 to 100, and in this case closure the value to 100 higher will be the degree of intra-industry trade, while on the other hand, closure the value to 0 more will be inter-industry trade.

But the equation 4.2.1 has one limitation for measuring the level of intra-industry trade, that is – it makes no allowance for any imbalance in the country's total trade and is downward biased for measuring the degree of IIT, in case of trade imbalance. The greater the imbalance, the lower will be the value of IIT. Therefore, to avoid this problem, they proposed an alternative adjusted formula, which adjusts for any trade imbalance, which is shown as:

\[
IIT_j = \frac{\sum_{i=1}^{n} (X_i + M_i) - \sum_{i=1}^{n} |X_i - M_i|}{\sum_{i=1}^{n} (X_i + M_i) - \sum_{i=1}^{n} X_i - \sum_{i=1}^{n} M_i} \times 100 \quad \ldots \quad (4.2.2)
\]
The above equation, 4.2.2, is also called as an adjusted equation, and it takes care of trade imbalance, if any, because in case of trade imbalance (surplus or deficit) the unadjusted equation would be biased downwards and true extent of intra-industry trade would be underestimated. The equation has been used to calculate the degree of IIT at both 4-digit and 6-digit level of aggregation.

Both the equations 4.2.1 and 4.2.2 have been used, for calculating the degree of intra-industry trade. Wherever it was required to calculate unadjusted IIT, equation 4.2.1 has been used; and when it is required to calculate adjusted IIT equation 4.2.2 is used. Both of these two equations have been referred directly in the rest of the study.

4.3 DISENTANGLING TOTAL INTRA-INDUSTRY TRADE INTO HORIZONTAL AND VERTICAL COMPONENTS

Total intra-industry trade can be further divided into horizontal intra-industry trade (IIIT) and vertical intra-industry trade (VIIT), as shown in equation 4.3.1:

\[ \text{IIT} = \text{HIIT} + \text{VIIT} \] \hspace{0.5cm} (4.3.1)

In case of HIIT, products differ in their attributes but do not differ in quality or price. Producers in the industry are assumed to use the same factors of production and same techniques. On the other hand, in VIIT, products can differ in terms of quality which leads to difference in prices also. Here producers use different factors of production and different techniques.

To measure, whether IIT is horizontal or vertical, the following equation, 4.3.2, has been used:

\[ 1 - \alpha \leq \frac{U_{\text{exp}}}{U_{\text{imp}}} \leq 1 + \alpha \] \hspace{0.5cm} (4.3.2)

where \( U_{\text{exp}} \) and \( U_{\text{imp}} \) are unit-value of export and import respectively. “\( \alpha \)” is called as Dispersion Factor and is used to differentiate between horizontal and vertical product
differentiation. There is no fix value for $\alpha$, but in most of the cases researchers used 0.15 or 0.25 values. The present study has used 0.15 as dispersion factor, which is the most frequently used value. It means if the export and import unit values difference equal to or greater than 15 percent, then the difference between the two unit values is considered as significant to assume quality differences and are regarded as vertical product differentiation, while on the other hand, if the difference is less than 15 percent then the products would be considered of almost same qualities and be regarded as horizontal product differentiation. Therefore, if the ratio of unit value of export to import is lies between 0.85 (i.e., 1-0.15) to 1.15 (i.e., 1+0.15), then the IIT would be called as horizontal intra-industry trade (HIIT); while on the other hand, if the ratio does not lie in the range of 0.85 to 1.15, the IIT would be called as vertical intra-industry trade (VIIT).

Apart from this, vertical intra-industry trade can be further classified into two types – low-quality VIIT (LQVIIT) and high-quality VIIT (HQVIIT). The determining factor for this will again be the value of ratio of unit value of export to import. If the ratio is less than “1-\alpha” (i.e., here it will be less than 0.85 because dispersion factor, $\alpha$, is taken as 15%), then it will be called as LQVIIT, on the other hand, if the value of the ratio is more than 1+\alpha (i.e., more than 1.15) it will be called as HQVIIT. Low-quality VIIT depicts that a country is importing high-priced product, i.e., high-quality product, and exporting the low-priced product, i.e., low-quality products in such a way that the ratio of unit value of export to the unit value of import is less than or equal to 0.85 (assuming 15% dispersion factor). Similarly in case of high-quality VIIT, a country is exporting high-priced product (i.e., high-quality product) and importing the low-priced product (i.e., low-quality products) in such a way that the ratio of unit value of export to unit value of import is equal to or higher than 1.15 (assuming 15% dispersion factor). In both the cases, the product must belong to the same industry.

As far as calculating unit value of exports and imports are concerned, it is calculated as shown in the equation 4.3.3 –

$$UV_{exp} = \frac{Value_{exp}}{Units_{exp}} \quad \text{and} \quad UV_{imp} = \frac{Value_{imp}}{Units_{imp}} \quad \ldots \ldots \quad (4.3.3)$$
To calculate unit value of export, as shown in equation 4.3.3, the total value of export is divided by the number of units of the product, similarly to calculate unit value of import, total value of import is divided by the total number of units imported, of the product. In the present study, the currency used is USD, and the unit is tonne, kgs or litre, depending on the nature of the product and the availability of the data. Selecting tonne, kgs or litre does not make any difference in calculating the ratio because same notations have been used for both export and import, for a particular product group.

4.4 IMPORT-WEIGHTED AVERAGE TARIFF (IWAT)

A tariff is a tax levied by a country on its import. It is synonymous with the word import duties or custom duties. The basic objective of levying tariff is twofold – to collect revenue or to protect the domestic industries. The former is called as Revenue Tariffs and the latter is called as Protective Tariffs. The motive of imposing revenue tariff is to earn revenue from import and normally it is lower, because low tariff will not result into price escalation of the imported goods into the domestic market and hence leads to higher demand which finally leads to a collection of high revenue from import. On the other hand, protective tariff is imposed to restrict the import in order to protect the domestic industries and therefore the rate of tariff is generally very high. This high tariff rate will lead to price escalation of the imported goods in the domestic market and will increase its price and reduce the demand which helps the domestic industries because the demand of the product of domestic industries will not decrease and they win the war of price competition. Thus it can be said that in general, imposing tariff leads to increase in the price of imported product in domestic market and hence it restricts the imports and international trade. The wave of globalization leads to reforms in tariff structure and countries are reducing their tariff to promote world trade. Therefore a decrease in tariff rate should increase the international trade which may lead to increase in intra-industry trade. Hence, a decrease in tariff rate may leads to increase in the degree of IIT.

In the present study the effect of tariff rate on intra-industry trade have been considered and this could have been done by either calculating a simple average or calculating a weighted average. Here simple average tariff has not been considered because although it measures the overall degree of protection but it does not adjust for the significance of
different products in the trade profile, so a high tariff on an insignificant product may overstate the degree or protection.

Therefore the calculation is based on weighted average tariff because it takes into account each product mentioned in the import profile of a country. However, this method is also not devoid of limitations and its major limitation is it understates the level of protection because a very heavily protected product will be imported less and therefore receive a small weight. The formula used in the present study to calculate import-weighted average tariff (IWAT) is shown below:

$$\text{IWAT}_j = \frac{\sum_{i=1}^{n} (w_i \times m_i)}{\sum_{i=1}^{n} m_i} \quad \ldots (4.4.1)$$

where $w$ stands for weighted average tariff, $m$ represents the import value and $n$ represents the total number of products imported. The same equation, 4.4.1, can be used for both 4-digit and 6-digit level of aggregation.

4.5 PANEL-DATA ANALYSIS

Time series and cross section are two most commonly used tools used for the empirical analysis in economics and other areas of research. In time series data we observe the values of one or more variables over a period of time, for example, analyzing GDP or national income of a country for several time periods (years, quarters, months etc.). Therefore time series analysis can also be called as vertical analysis because in this case same data for different time periods, of one parameter, arranged chronologically and the technique is mainly used for trend analysis. In cross-section data, values of one or more variables are collected for several sample units, or entities, at the same point of time, for example, GDP growth for different Asian countries for a given time period. Therefore, cross-section analysis can also be called as horizontal analysis because in this case same
data for different objects but for the same year is given and the technique is mainly used for comparative analysis.

In panel data the same cross-sectional unit is surveyed over time. In short, panel data have space as well as time dimensions, for example, data regarding GDP, per-capita income and population of different countries (cross-section data) arranged chronologically for several years (time series). The advantage of using panel data is it gives a holistic view about the nature, pattern and determinants of the data under consideration. This technique is also called as pooled data (pooling of time series and cross-sectional data), combination of time series and cross-section data, micropanel data, longitudinal data (a study over time of a variable or group of subjects), event history analysis (like studying the movement over time of subjects through successive states or conditions), and cohort analysis (like following the career path of 1970 graduates of a business school). Therefore, in short, panel data analysis means that movement over time of cross-sectional units.

The advantages of using panel data over time series and cross-section data, as mentioned by Gujarati (2006), are:

a) It is more informative because it combines both time series and cross-sectional data. Moreover, because of combining these two series, it gives more variability, less collinearity among variables, more degree of freedom and more efficiency.

b) It relates to individuals, firms, states, countries etc., over time, there is bound to heterogeneity in these units and the technique takes such heterogeneity explicitly into account.

c) By studying the repeated cross-section of observations, panel data are better suited to study the dynamics of change.

d) It can better detect and measure the effects that simply cannot be observed in pure cross-section and pure time series data.

e) It enables us to study more complicated behavioral model.

f) By making the data available for several thousand units, panel data can minimize the bias that might result if we aggregate individuals of firms into broad aggregates.
Broadly, the model of panel data analysis can be classified into two types: the fixed effect approach and the random effect approach. In fixed approach, we assume either intercept or slope or both of the regression equation is constant, depending on the objective of the study. Fixed approach models can be of different types:

a) The intercept and slope coefficient are constant across time and space and the error term, of a regression equation, captures differences over time and individuals. This is the simplest approach and just estimates the usual OLS regressions.

b) The slope coefficients are constant but the intercept varies over individuals. This is also known as Fixed Effects Model (FEM) or Least Square Dummy Variable (LSDV) model. In this model, dummy variables are used for different parameters or time period. Number of dummy variables should be one less than the total number of parameters used, to avoid dummy-variable trap.

c) Slope coefficients are constant but the intercept varies over individuals as well as time.

d) All coefficients vary across individuals.

Fixed effects model or LSDV model has to be used carefully and few things have to be kept in mind like if we use too many dummy variables then the number of degrees of freedom would decrease and in case of using so many variables there is always the possibility of multicollinearity.

The other type of panel data model is Random Effect Model (REM) or it is also called as Error Component Model (ECM). This model assumes that if we are not sure about the nature of dummy variables then why should we include it in the equation and sacrifice on the degree of freedom. In that case it is better to use an error component in the equation which represents common mean plus an error term, representing a mean deviation for each individual parameter under consideration.

The question that which model should be used in a particular case, FEM or ECM, it depends on the assumption used in the study. Here fundamental differences between the usages of these two models are given below:
a) If $T$ (the number of time series data) is large and $N$ (the number of cross-sectional units) is small, then practically there will be little difference between the result of FEM and ECM, therefore, the choice is based on the computational convenience. But FEM, in this case, may be preferred.

b) When $N$ is large and $T$ is small, then the results of these two models would differ significantly. In that case if we strongly believe that the individual, or cross-section unit in our sample are not random drawings then FEM should be used otherwise ECM would be appropriate.

c) If individual error component (of REM) and one or more regressors are correlated, then FEM should be used.

d) If $N$ is large and $T$ is small, and if the assumptions underlying ECM hold, ECM estimators would be more appropriate than FEM.

It is difficult to discuss the panel data analysis in detail, therefore only the basic concepts of the model has been discussed here. The present study mainly used Fixed Effects Models (FEM) for discussing about the determinants of intra-industry trade of India.

4.6 LIMITATIONS OF THE STUDY

Although the present study deals with different dimensions of India’s intra-industry trade but certainly it is not devoid of limitations. The major limitations of the study are:

a) The present study is based on the overall basis therefore to offer a precise policy guideline we need to do a detailed study of country-wise/product-wise. However the present study provides a framework for understanding the nature and pattern of India’s intra-industry trade and the result of the study can certainly be used as a base for further study.

b) The accuracy of the calculation depends upon the accuracy of the classifications of the products into different HS Code. It has already been discussed in earlier chapters that the products can be classified into an industry on a different basis and to avoid any confusion, ITC HS-Classification has been used here. Therefore
the accuracy of the result will depend on the fact that how accurately the products have been classified.

c) For disentangling HIIT and VIIT, the unit values used is average unit-values and not the exact because it has been calculated by dividing total export (import) value with the total number of units. It may be possible that few products of the concerned HS-Code are of high-value and the other may be of low-value but the calculation used here will average-out these values.

d) Grubel-Lloyd index, which is used in the study, is a share measure, i.e., the share of IIT in gross trade and not an absolute or volume measure. It means a high GL-index does not correspond to high trade volume. Therefore while analyzing the value of GL-index one has to be careful about the fact and decisions should be taken accordingly.