

Indices of Comparative Advantage: A Study with Focus on Differences between Gross Trade and Value-Added Trade

A thesis submitted to Indian Institute of Technology Guwahati in partial fulfillment of the requirements for the degree of Doctor of Philosophy



Submitted By
Kaveri Deb

Department of Humanities and Social Sciences
Indian Institute of Technology Guwahati
Guwahati 781039
India

September 2015





Dedicated
to
Maa, Baba & Bingo





Indian Institute of Technology Guwahati

Department of Humanities & Social Sciences

Guwahati - 781 039

(Assam), **INDIA**

Declaration

I hereby declare that the thesis entitled “**Indices of Comparative Advantage: A study with Focus on Differences between Gross Trade and Value-Added Trade**”, is the result of investigation carried out by me in the Department of Humanities and Social Sciences, Indian Institute of Technology Guwahati, India, under the supervision of Dr. Bodhisattva Sengupta, Assistant Professor (Economics), Department of Humanities and Social Sciences, IIT Guwahati.

In keeping with the general practice of reporting observations, due acknowledgements have been made, whenever the work described is based on the findings of other investigators.

(Kaveri Deb)

Research Scholar

Department of Humanities and Social Sciences

IIT Guwahati

September, 2015





Bodhisattva Sengupta
Asst. Professor
(Economics).

Indian Institute of Technology Guwahati

Department of Humanities & Social Sciences

Guwahati - 781 039

(Assam), **INDIA**

Phone: +91-361-2582575

Fax: +91-361-2690762

E-mail: bsengupta@iitg.ernet.in

Certificate

This is to certify that the thesis entitled “**Indices of Comparative Advantage: A Study with Focus on Differences between Gross Trade and Value-Added Trade**”, submitted by Ms. Kaveri Deb for the Doctor of Philosophy in Economics in the Department of Humanities and Social Sciences, Indian Institute of Technology Guwahati, embodies research work carried out under my supervision and guidance. The present thesis, or any part of it, has not been submitted to any other University or Institute for award of any degree or diploma. She has fulfilled all the requirements according to the rules of the institute for submission of thesis.

(Bodhisattva Sengupta)

Supervisor



Acknowledgment

I would like to express my heartfelt gratitude towards all those who have immensely helped me during this long arduous journey and eventually assisted me to emerge as an experienced researcher and a mature individual.

First and foremost, I must express my deepest sense of appreciation and respect for my thesis supervisor, Dr. Bodhisattva Sengupta for providing me the opportunity to work under his supervision. His constant encouragement and constructive criticisms have played a crucial role in writing the thesis and bringing it to its present form.

I would like to thank my doctoral committee members, Prof. Saundarjya Borbora, Dr. Mrinal Kanti Dutta, Dr. Debarshi Das and Dr. Rajshree Bedamatta for carefully reading the reports at various stages of evaluation and providing many useful comments, which have helped in the development of this thesis.

I whole heartedly appreciate the support and encouragement of other faculty members of the HSS department. I would also like to acknowledge the contribution of HSS office staff, Durga Sarma, Bandana Khatanaier and Parag Kalita for helping me with various academic as well as non-academic issues.

Words can hardly express how much I owe to Dr. William R. Hauk, Jr. of University of South Carolina, without whose assistance certain portions of the thesis would have been difficult. I must also express my sincere gratitude and indebtedness to Prof. John H. McDermott for providing me the rare opportunity as a visitor at the Economics Department of Moore School of Business, University of South Carolina, for a few months and allowing me to interact with the faculty members and graduate students. This experience immensely helped me in broadening my horizon of knowledge and boosting my research potential.

I must express my thanks to Prof. Jeffrey M. Wooldridge, Prof. Nityananda Sarkar, Dr. Alex R. Hoen and Dr. Manisha Chakrabarty for various useful comments on different aspects of my doctoral work. I would also like to thank Dr. Debashis Chakraborty for his encouragement and kind words, which motivated me to complete my work as early as possible.

I seize this opportunity to express my gratitude towards my erstwhile mentors, Prof. Deepak Nayyar of Jawaharlal Nehru University and Prof. Partha Basu of Indian Institute of Technology, Kharagpur, for their guidance, which helped me in the initial stages towards forming the necessary foundations for current research.

It would be unfair if I fail to express my gratitude towards Dr. Pahi Saikia and Dr. Bipul Bhuyan without whose persuasions and initiatives, I may not have been a part of this wonderful IITG “HSS family”.

Friends and family form an important part of this long and enduring journey, and without their constant support and encouragement, the completion of this thesis perhaps would have been an impossible task. I would therefore be forever be grateful to my closest friends, Anushree Bhattacharya, Manomohini Dutta, Srobonti Chattopadhyay and Sudatta Banerjee, for being epitomes of mental and spiritual strength and making sure that I do not lose the struggling spirit. May God infuse the taste of such true friendship in everyone. Heartfelt thanks go to Debarati Mitra and Partha Sarathi Mandal for being such close family friends. I cannot be indebted enough to my departmental friends, Pranti, Bandana, Nirmala, Munmi, Mahsina, Gunjan, Baban, Hemanta, Nabanita, Madhuri, Pankaj for many academic and non-academic discussions. I also thank other research scholars for creating such a cordial and friendly environment in the research scholar lab. Finally, I am thankful to my parents, in-laws and other family members, who have been a constant source of love, encouragement and support. Udit, Baba and Api... I cannot express how much grateful I am to you all, for always being at my side and help me battle through all the hurdles in making my dream come true.

(Kaveri Deb)

Contents

	Page No.
Acknowledgement	i
List of Tables	vii
List of Figures	ix
Abbreviations	xi
Abstract	xv
Chapter 1: Introduction	1
1.1 Background of the Study.....	1
1.1.1 Inter-linkage between the Comparative Advantage Theories.....	2
1.1.2 Importance of Inter-Industry Trade in the Presence of Intra-Industry Trade.....	4
1.1.3 Emergence of Revealed Comparative Advantage Indices.....	6
1.1.4 RCA Indices and Global Supply Chains.....	8
1.1.5 Relation between Comparative Advantage, Global Supply Chains and Intra-Industry Trade.....	9
1.2 Gap Areas.....	12
1.3 Research Questions.....	14
1.4 Hypotheses.....	14
1.5 Data and Methodology.....	15
1.6 Structure of the Thesis.....	18
Chapter 2: Review of the Literature	21
2.1 Introduction.....	21
2.2 The Revealed Comparative Advantage Index of Balassa.....	23
2.2.1 Critical Analysis of the RCA Index of Balassa.....	26
2.2.2 Alternative Indices of RCA.....	34

2.2.3	Derivation of Theoretically Consistent RCA Indices.....	44
2.3	Growing Importance of Global Production Chain and the RCA Indices.....	47
2.3.1	Vertical Specialization and Value-Added Trade: Implications for RCA Indices.....	50
2.4	Rational of the Present Study and Conclusion.....	62
Chapter 3: Analyzing the Empirical Distributions of RCA Indices		65
3.1	Introduction.....	65
3.2	Methods and Data.....	71
3.2.1	Cumulative Distributions and Summary Statistics of RCA Indices.....	71
3.2.2	Graphical Analysis of RCA Indices.....	71
3.2.3	Tests of Significance on the Stability of RCA Index Distributions.....	75
3.2.3.1	Wilcoxon's Signed Rank Test.....	75
3.2.3.2	Wilcoxon's Rank Sum Test.....	78
3.2.3.3	Kolmogorov Smirnov Test.....	81
3.3	Discussion of Results.....	83
3.3.1	Discussions on the Summary Statistics and Graphical Analysis.....	83
3.3.2	Discussions of Results for the Tests of significance.....	91
3.3.2.1	Results for Wilcoxon's Signed Rank Test.....	91
3.3.2.2	Results for Wilcoxon's Rank Sum Test.....	94
3.3.2.3	Results for Kolmogorov Smirnov Test.....	96
3.4	Conclusion.....	99
Chapter 4: Consistency of RCA Indices with the Theories of Comparative Advantage		101
4.1	Introduction.....	101
4.2	Testing the Consistency of RCA Indices with the Theories.....	103
4.2.1	Methods and Data.....	104
4.2.1.1	Estimation of Regression Coefficients: OLS and Robust Estimate of Standard Errors.....	113
4.2.1.2	Estimation of Regression Coefficients: Robust Regression.....	116

4.2.2	Discussion of Results.....	118
4.2.2.1	Results for Heckscher-Ohlin Model.....	118
4.2.2.2	Results for Ricardian Model.....	123
4.3	Adjusting Export-Import Based RCA Indices for Trade Barriers.....	127
4.3.1	Methods and Data.....	127
4.3.2	Discussion of Results.....	131
4.4	Conclusion.....	133
Chapter 5: Value-Added Trade and RCA Indices		135
5.1	Introduction.....	135
5.2	Methods and Data.....	137
5.3	Discussion of Results.....	142
5.4	Robustness Check on the Stability of Empirical Distributions of Indices.....	145
5.4.1	Summary Statistics of RCA Indices.....	146
5.4.2	Testing the Stability of RCA Index Distributions.....	149
5.5	Conclusion.....	155
Chapter 6: Application of Value-Added in Export based RCA Indices in the context of China and India		157
6.1	Introduction.....	157
6.2	Methods and Data.....	159
6.3	Discussion of Results and Policy Recommendations.....	161
6.3.1	Comparison of Country Sectoral Ranks over Years.....	167
6.3.2	Comparison of Country Sectoral Ranks between Indices.....	168
6.3.3	Sectoral Exports Identification on the basis of NRCA Index Ranks.....	169
6.4	Conclusion.....	172
Chapter 7: Summary of Findings, Conclusions and Policy Suggestions		175
7.1	Introduction.....	175
7.2	Summary of Findings.....	176
7.2.1	Empirical Distributions of Gross Trade based RCA Indices.....	176

7.2.2	Consistency of Gross Trade based RCA Indices with the Comparative Advantage Theories.....	177
7.2.3	Value-Added Trade and RCA Indices.....	179
7.2.4	Evaluation of China and India's Exports to United States on the basis of Value-Added Adjusted Indices.....	180
7.3	Conclusion.....	182
7.4	Policy Suggestions.....	183
7.5	Limitations of Study.....	184
7.6	Scope for Future Research.....	185
Appendix A		187
Appendix B		199
Appendix C		219
Appendix D		251
Appendix E		265
References		271

List of Tables

Table No.	Title of Table	Page No.
3.1.A	: Empirical Distribution of Balassa Index for Each Year: Entire Sample Analysis	84
3.1.B	: Empirical Distribution of RSCA Index for Each Year: Entire Sample Analysis	84
3.1.C	: Empirical Distribution of ARCA Index for Each Year: Entire Sample Analysis	85
3.1.D	: Empirical Distribution of NRCA Index for Each Year: Entire Sample Analysis	85
3.1.E	: Empirical Distribution of Log of Balassa Index for Each Year: Entire Sample Analysis	86
3.1.F	: Empirical Distribution of RTA Index for Each Year: Entire Sample Analysis	86
3.1.G	: Empirical Distribution of RC Index for Each Year: Entire Sample Analysis	87
3.2	: Results Summary for Wilcoxon's Signed Rank Test for Countries	92
3.2.1	: Results Summary for Wilcoxon's Signed Rank Test for Sectors	94
3.3	: Results Summary for Wilcoxon's Rank Sum Test for Sectors	95
3.4	: Results Summary for Wilcoxon's Rank Sum Test for Countries	96
3.5	: Results Summary for Kolmogorov Smirnov Test for Sectors	97
3.6	: Results Summary for Kolmogorov Smirnov Test for Countries	98
4.1	: Results Summary for Heckscher-Ohlin Theory: With/ without White's Correction	121
4.2	: Robust Regression Results Summary for Heckscher-Ohlin Theory	122
4.3	: Results Summary for Ricardian Theory: With/ without White's Correction	124
4.4	: Robust Regression Results Summary for Ricardian Theory	125

Table No.	Title of Table	Page No.
4.5	: Results Summary for Rank Correlation Test for Heckscher-Ohlin Theory	131
4.6	: Results Summary for Rank Correlation Test for Ricardian Theory	132
5.1	: Estimated Results for Ricardian Theory: Combined Samples	142
5.2	: Results Summary for Ricardian Theory: Sector Specific Samples	143
5.3	: Results Summary for Heckscher-Ohlin Theory: Sector Specific Samples	144
5.4.A	: Summary Statistics for Balassa Index for Each Year: Entire Sample Analysis	147
5.4.B	: Summary Statistics for RSCA Index for Each Year: Entire Sample Analysis	147
5.4.C	: Summary Statistics for ARCA Index for Each Year: Entire Sample Analysis	147
5.4.D	: Summary Statistics for NRCA Index for Each Year: Entire Sample Analysis	147
5.4.E	: Summary Statistics for Log of Balassa Index for Each Year: Entire Sample Analysis	148
5.5	: Results Summary for Wilcoxon's Signed Rank Test for Countries	152
5.5.1	: Results Summary for Wilcoxon's Signed Rank Test for Sectors	152
5.6	: Results Summary for Wilcoxon's Rank Sum Test for Sectors	153
5.7	: Results Summary for Wilcoxon's Rank Sum Test for Countries	153
6.1	: NRCA Index Ranks for Sectors with Revealed Comparative Advantage	162
6.2	: Log of Balassa Index Ranks for Sectors with Revealed Comparative Advantage	163
6.3	: NRCA Index Ranks for Sectors with Revealed Comparative Disadvantage	165
6.4	: Log of Balassa Index Ranks for Sectors with Revealed Comparative Disadvantage	166

List of Figures

Figure No.	Title of Figure	Page No.
2.1	Vertical Specialization	49
2.2	Gross Exports Accounting	59





Abbreviations

APLR	: Average Productivity of Labor Ratio
ARCA	: Additive Revealed Comparative Advantage
CIS	: Commonwealth of Independent States
FDI	: Foreign Direct Investment
FER	: Factor Endowment Ratio
GATT	: General Agreement on Tariffs and Trade
GDP	: Gross Domestic Product
GTAP	: Global Trade Analysis Project
HOV	: Heckscher-Ohlin-Vanek
HS	: Harmonized System
IDE-JETRO	: Institute of Developing Economies - Japan External Trade Organization
IMF	: International Monetary Fund
ISIC	: International Standard Industrial Classification
MISE	: Mean Integrated Squared Error
NRCA	: Normalized Revealed Comparative Advantage
OECD	: Organization of Economic Cooperation and Development
PIM	: Perpetual Inventory Method
PWT	: Penn World Tables
RCA	: Revealed Comparative Advantage
RC	: Revealed Competitiveness
RSCA	: Revealed Symmetric Comparative Advantage
RTA	: Relative Trade Advantage
SITC	: Standard International Trade Classification
TiVA	: Trade in Value-Added
TPP	: Trade Production and Protection
TRAINS	: Trade Analysis Information System
UN COMTRADE	: United Nations Commodity Trade Statistics
UNCTAD	: United Nations Conference on Trade and Development

UNIDO	: United Nations Industrial Development Organization
VAX	: Value-added in Exports
VS	: Vertical Specialization
WIOD	: World Input Output Database
WIOT	: World Input Output Tables





It is better to be roughly right than precisely wrong.

- John Maynard Keynes



Abstract

The Revealed Comparative Advantage (RCA) indices are frequently applied in international trade literature to determine the potential of countries in exporting various products. The indices using post trade data, therefore tries to identify the products that could be traded amongst countries on the basis of the comparative advantage theory of trade. However, reliable usage of such indices for policy analysis must be based on the assurance that their structure reflects the existing theories of comparative advantage. Further, the empirical distributions of such indices must also be considerably stable over time, sectors and countries to justify their considerable usage for comparative analysis. The current thesis focus on examination of these two features of the RCA indices. Five of such indices considered in the thesis are based only on exports, namely – Balassa index, Revealed Symmetric Comparative Advantage (RSCA) index, Additive Revealed Comparative Advantage (ARCA) index, Normalized Revealed Comparative Advantage (NRCA) index and Log of Balassa index. Two other indices considered are based on both exports and imports, namely – the Relative Trade Advantage (RTA) index and Revealed Competitiveness (RC) index. The advantage of exports-and-import based indices over exports-only indices is that they can account for intra-industry trade and re-exports. Consideration of exports-and-import based indices brings to the fore another pressing issue. Imports of countries are often distorted by trade barriers and therefore the phenomenon must be addressed while assessing the performance of RCA indices based on both exports and imports. Thus, the influences of tariff barriers on country imports are also taken into account by making suitable adjustments to the import values. The consistency of such adjusted indices with the theories is further reaffirmed.

An important aspect of the considered research topic is that the originally suggested indices are based on gross trade values. The present pattern of trade is dictated more by the purpose of cost minimization at different stages of production of a product. In this context, the reduction in tariffs and other trade costs on intermediate products, combined by growth in foreign direct investments and revolutions in information technology, played significant roles in disintegration of the production process, and allow the trading nations to reap the benefits that each country has to offer. Thus global production networks are such that, the final product generated includes the value-added not only by the exporting country but also by

several other countries through which the product passed during its production stage. Under such circumstances gross trade data of a country may include considerable double counting, as inputs may need to cross border multiple times. In this scenario, gross trade based RCA indices may fail to reflect the true comparative advantage of a country, and hence need is felt for adjusting the RCA indices of a country for domestic value-added in trade. The thesis therefore explores the implications of such adjustment on the two previously mentioned features of the RCA indices.

Detailed analyses of the RCA indices based on secondary sources of data reveal that it is indeed difficult to identify any particular index which meets both the features of an “ideal” index. Thus none of the indices considered for analysis, are simultaneously consistent with the existing theories of trade and possess stable distributions across all parameters considered. On the basis of analyses of gross trade based indices, RSCA index is found to be more consistent with the theory than any other index. But the distribution of NRCA index is found to be more stable simultaneously over time, sectors and countries, than any other gross trade based index. The data analyses fails to provide any evidence in favor of the exports-and-import based indices over the exports-only indices. While testing the consistency of indices with the theories, the exports-and-import based indices are found to be either at par or inferior to the exports-only RSCA index. Even consideration of tariff barriers on import figures could not produce any significant changes in the consistency of exports-and-import based indices with the theory. Further the stability of the distributions of exports-and-import based indices is not found to be better than the exports based NRCA index. On account of such findings the exports-and-import based indices are not considered for analysis, after adjusting the indices for domestic value-added in trade.

After adjusting the exports-only indices for domestic value-added in exports, the NRCA index retains its position, but Log of Balassa index is found to be more compatible with the comparative advantage theories than any other index. Given the growing significance of global production networks, the conclusion in the thesis is drawn with respect to the analysis of value-added in export based indices. Thus, based on the results of data analyses, it is recommended that the policy makers may make use of both NRCA index and Log of Balassa index, to formulate policies for export expansion among countries.

Results of the thesis have been published in various forms in:

I. Journal Publications:

- i) Deb, K. and W.R. Hauk (2015), “RCA Indices, Multination Production and the Ricardian Trade Model”, *International Economics and Economic Policy*. doi: 10.1007/s10368-015-0317-z.
- ii) Deb, K. and P. Basu (2011), “Indices of Revealed Comparative Advantage and their Consistency with the Heckscher-Ohlin Theory: A Cross Sectional Analysis”, *Foreign Trade Review*, Vol. 46, No. 3, pp. 3-28.

II. Conference Proceedings:

- i) Deb, K. and P. Basu (2012), “Consistency of Revealed Comparative Advantage Indices with the Heckscher-Ohlin Theory”, *Conference Proceedings of Knowledge Globalization Conference, Pune, India, January 5-7*, Vol.6, No.1, pp. 122-131.



Chapter 1

Introduction

1.1 Background of the Study

In the literature on international trade theory a considerable volume of work has been devoted on the determination of factors that drive the pattern of trade between countries. The theories in this area explore the factors behind a particular country exporting one good while importing another from its trading partner. In contrast to Adam Smith's (1776) theory on absolute cost advantages in trade, David Ricardo recognized in 1817 that in a two country – two commodity framework, even if one country has an absolute cost advantage in both the commodities, it is possible for it to engage in gainful trade with the other country if each country specialize in the production of the product in which they have comparative cost advantage.¹ Thus, each country would specialize and export that product in which it incurs not just the lower resource cost of production (the amount of labor required to produce a unit of output), but lower opportunity cost of production compared to the other country. Assuming two goods X and Y, opportunity cost of producing a given amount of X is the amount of Y that could be produced by utilizing the resources required to produce X. Hence, if the resulting output of Y in a country is lower than the output of Y in any other country, the former country would be having comparative advantage in production of X. This idea came to be widely known as the Ricardian comparative advantage theory of trade. In the Ricardian model, comparative advantage is the outcome of the differences in technology or labor productivity between countries, assuming labor to be the only factor of production. Hence, comparative advantage of a country in good Y requires that the country is twice as productive in production of good X relative to the other country but three times as productive in production of good Y relative to the other country (Suranovic, 2015). Later Eli Heckscher (1919) and Bertil Ohlin (1933) focused on the relationships between the composition of a

¹ The idea of comparative advantage was originally proposed by Robert Torrens in 1815 in *Essay on the External Corn Trade*. Later in 1817, David Ricardo formalized the idea in his book *On the Principles of Political Economy and Taxation*.

country's factor endowments and the flow of commodity trade. The Heckscher-Ohlin model thereby identified comparative advantage as the outcome of differences in relative factor endowments between countries. Assuming a two country-two commodity-two factor of production framework, the model showed that a country relatively abundant in one factor of production would specialize and export that product which could be produced by using the country's abundant factor intensively.

Thus, while both of these two theories predicted that trade is driven primarily by the principle of comparative advantage, they made different assumptions about the origin of comparative advantage. Ricardian model focused on differences in labor productivity as the main driver of comparative advantage, without any reference to the relative factor endowments of countries. On the other hand, Heckscher-Ohlin model, as noted above, focused on differences in factor endowments (such as capital, labor, and natural resources) across countries and, for tractability purposes, assumed that factor productivity was roughly the same across countries.

Before moving on to the literature review and definition of the problem, it is essential to elaborate on certain issues to develop the research framework.

1.1.1 Inter-linkage between the Comparative Advantage Theories

The empirical tests on the Ricardian theory (MacDougall, 1951; Balassa, 1963; Golub and Hsieh, 2000) have confirmed the predictions of the theory that relative productivity differences would determine the flow of trade between countries. Empirical tests on the Heckscher-Ohlin model began with the work of Leontief (1953) which failed to produce results in support of the model. Further tests continued with the many goods and many factors extension of the model associated with the work of Vanek (1968), called the Heckscher-Ohlin-Vanek (HOV) model. The model is the factor content version of the Heckscher-Ohlin model as it does not analyze the countries' trade patterns in individual goods but rather compute the amounts of labor, capital, land, and other factors of production present in the production and consumption of a country (Feenstra, 2004).² The HOV model

² The HOV model tries to relate the factor content of net trade of any country to its underlying factor endowments (Feenstra, 2004). In the process the model predicts that a capital abundant country would export capital services and a labor abundant country would export labor services.

has been the subject of great deal of empirical research by Leamer (1980), Bowen *et al.* (1987), Trefler (1993, 1995), Davis and Weinstein (2001) and many more. An important fact that emerges from these researches is that HOV model cannot explain the pattern of trade unless the assumption of identical technologies is withdrawn. Bowen *et al.* (1987) introduce technological differences across countries to test the HOV theorem and found the modified model to be better explained by the data. Trefler (1993, 1995) shows the predictions of the HOV model to improve with the incorporation of productivity differences across countries. Davis and Weinstein (2001) estimated different technology matrices (used to construct factor content of trade) for a group of countries based on the idea that industry specific factor intensities should depend on country factor endowments, and found results in favor of the HOV model. Realization for differing technologies highlights the importance of Ricardian tradition. While validation of the HOV model is a subject of ongoing research, examining the Ricardian theory of comparative advantage and simultaneously studying the implications of traditional Heckscher-Ohlin theory of comparative advantage, can provide important insights into the relevance of the theories. At this stage it needs to be noted that a complete description of comparative advantage needs to account for both the Ricardian and Heckscher-Ohlin theories. Although Ricardian theory has wider acceptance than the Heckscher-Ohlin theory, given the fact that endowment differences could account for productivity differences, the importance of Heckscher-Ohlin theory in explaining the flow of trade cannot be ruled out altogether. As pointed out by Deardorff (1984) the productivity and factor endowment based approaches to comparative advantage are to some extent related. Ricardian model assumes that labor productivity differences are due to technological differences. However, large relative factor endowment differences could enforce complete specialization, and induce productivity differences (as cited in Golub and Hsieh (2000)). Leromain and Orefice (2014) cite the case of Saudi Arabia whose higher productivity in the manufacture of plastic materials is due to its endowment of petrochemicals. Thus, not only technological differences, but also factor endowment differences have important role to play in influencing the flow of trade between countries through their effects on labor productivity. Since differentiating the effects of technology and factor endowments on labor productivity is a difficult task, it seems plausible to study the implications of Ricardian model and Heckscher-

Ohlin model separately which are based on technological differences and factor endowment differences respectively.

1.1.2 Importance of Inter-Industry Trade in the Presence of Intra-Industry Trade

The orthodox trade theories on comparative advantage formed the basis of inter-industry trade or trade in dissimilar products. With the development of new trade theory, the validity of the orthodox trade theories had been questioned. It had been argued that the recent surge of intra-industry trade or trade in similar products can be explained by new trade theories only. However, this idea had been challenged by several theorists and they had demonstrated that even the orthodox trade theories can account for intra-industry trade. The initial works on intra-industry trade models due to Krugman (1979, 1980, 1981), Lancaster (1980), Helpman (1981) and Eaton and Kierzkowski (1984), assumed that products are horizontally differentiated.³ This implies that the varieties of the same product may differ on the basis of their actual or perceived characteristics even though they are of same quality. Thus, consumers' choices among different varieties of such product are determined by their preferences; at the same price some consumers may prefer to buy some variety, while some other consumers may prefer to buy some other variety. Intra-industry trade based on such horizontally differentiated products is typically between countries with similar factor endowments. Such trade cannot be explained by traditional trade theories but rather by the integration of economies of scale and horizontal product differentiation (Faustino, 2008). The other kind of intra-industry trade models introduced by Falvey (1981), Falvey and Kierzkowski (1987) and Flam and Helpman (1987) are based on vertical product

³ Spence (1976), Dixit and Stiglitz (1977), Lancaster (1975, 1979) offered alternative representations of horizontal differentiation. According to Lancaster, consumers differed in their selection of an "ideal variety" of a differentiated good. According to Spence, Dixit and Stiglitz, every consumer would prefer different varieties of a differentiated good. This phenomenon could be termed as the "love of variety" approach. Lancaster (1980) and Helpman (1981) made use of the "ideal variety" approach in their models, while Krugman (1979, 1980, 1981) made use of the "love of variety" approach. Helpman and Krugman (1985) show that although these two approaches have different interpretations, they do not differ much in terms of the generated results. As cited by Greenaway *et al.* (1995) both the approaches reveal that horizontal intra-industry trade can be linked with preference diversity and decreasing costs.

differentiation.⁴ In these models, it is assumed that consumers rank different varieties of the same product on the basis of product quality. Thus, goods differ on the basis of quality, and consumers would prefer one variety over the other if they are sold at the same price. The existing literature accept the fact that intra-industry trade based on vertically differentiated products can be accounted by the traditional theories of comparative advantage (Greenaway and Milner, 1986; Greenaway, Hine and Milner, 1994, 1995; Tharakan and Kersten, 1995; Blanes and Martin, 2000). Hence, in the models of Falvey (1981), Falvey and Kierzkowski (1987), relatively labor abundant countries would specialize and export labor intensive products or the lower quality varieties of a product, and relatively capital abundant country would export capital intensive products or higher quality varieties of the same product. Flam and Helpman (1987) base their model on the essence of Ricardian comparative advantage so that intra-industry trade takes place due to differences in technology between the countries on the production side and differences in income distribution on the consumption side (Ando, 2006). Thus, while one could expect a positive association between vertical intra-industry trade and comparative advantage of countries, the association between comparative advantage and horizontal intra-industry trade could be expected to be negative. Faustino (2008) shows however that comparative advantage theory can positively influence not only vertical intra-industry trade but also horizontal intra-industry trade and the total intra-industry trade (which includes both vertical and horizontal intra-industry trade). His study therefore provides an important contribution to the literature by noting that relative autarkic price differences are a crucial determinant of intra-industry trade, at least during a certain portion of a product's life cycle.⁵ Such findings strengthen the arguments in favor of the relationship between orthodox trade theory and intra-industry trade. In a nutshell, comparative advantage

⁴ Other works on vertical product differentiation include the one by Shaked and Sutton (1984) where intra-industry trade are driven by scale economies and a small number of firms. Thus, while the models by Falvey, Falvey and Kierzkowski and Flam and Helpman assume a perfectly competitive market structure, Shaked and Sutton are more particular about the role of the market structure as they realize the significance of 'natural oligopolies'.

⁵ The author notes following Vernon's product cycle theory, at more mature stages of a product, all types of intra-industry trade tend to be determined by other factors such as foreign direct investment rather than comparative advantage, generating an inverted U shaped relationship between comparative advantage and intra-industry trade.

theory can explain not only inter-industry trade but also forms a part of new trade theories. Hence, the idea of comparative advantage and theories encompassing it, still remain an important strand of international trade theory.

However, as identified in Sanidas and Shin (2010, 2011) the term comparative advantage mentioned so far is used in a traditional sense. Comparative advantage can also be defined in a broader sense. Palley (2008) referring to the idea of technology transfer and production methods between countries point out that comparative advantage in the modern world is not necessarily endowed and could be created. He defends his argument by stating:

“In the 18th century world, trade was driven by the search for exotic spices and raw materials. In that epoch, climate and natural resource endowments significantly determined the pattern of comparative advantage, and little could be done to alter this pattern. In today's economy, comparative advantage is driven by technology, and technology can be importantly influenced by human action and policy” (p. 201).

Irrespective of the sense in which comparative advantage is defined, it can unanimously be considered to be a source that causes international trade to take place.

1.1.3 Emergence of Revealed Comparative Advantage Indices

While theories describing the idea of comparative advantage have been developed, an important question that arises in this context is how to apply this idea in determining the comparative advantage of countries in real world. The issues seemed to have been largely addressed with the development of the Revealed Comparative Advantage (RCA) index of Balassa (1965), which is based on gross export values. As noted by Ballance *et al.* (1987), any RCA index could be used for three purposes. First, they serve as dichotomous measures i.e., on the basis of some value they help in categorizing countries in two groups - those with comparative advantage and those with comparative disadvantage in a product. Second, they can be used as ordinal measures i.e., ranking of sectors with respect to a country or ranking of countries with respect to a given sector. Third, they aid as cardinal measures i.e., help in determining the degree of comparative advantage enjoyed by one country in a commodity with respect to another country. With respect to a country, the usage of an index as a cardinal measure implies quantifying its comparative advantage in a particular sector relative to

another sector. These realizations have prompted the usage of Balassa's RCA index for comparing countries or sectors over time (Hiley, 1999; Richardson and Zhang, 2001; Bender and Li, 2002; Kijboonchoo and Kalayanakupt, 2003; Hinloopen and Marrewijk, 2004; Batra and Khan, 2005; Smyth, 2005; Burange and Chaddha, 2008; Wignaraja, 2011; Karaalp, 2011; Ekram *et al.*, 2013; Hassan, 2013). Apart from these standard uses, Balassa's RCA index and a few of its modifications (to be discussed later) have also been used in econometric analysis for determining the structural changes in trade specialization through the application of Galtonian regression models (Dalum *et al.*, 1998; Laursen, 1998; Frantzen, 2008; Sharma & Dietrich, 2007; Sanidas and Shin, 2010; Sanidas and Shin, 2011).

That research on RCA indices did not stop after the proposed Balassa index, would be evident from literature following the advent of the index. Hillman (1980) and Bowen (1983) theoretically tried to link the RCA indices with the concept of comparative advantage and argued that the index is not in conformity with the theoretical notion of comparative advantage. Studies also tried to evaluate the performance of Balassa index in empirical analysis. For instance, the applicability of the Balassa index towards cross country or cross industry comparison was questioned (Yeats, 1985; Hinloopen and Marrewijk, 2001; Benedictis and Tamberi, 2001; Benedictis and Tamberi, 2004; Hoen and Oosterhaven, 2006; Yu *et al.*, 2009). The index is also asymmetrical in its distribution, thus resulting in non-normally distributed residuals in parametric regression analysis (Dalum *et al.*, 1998; Laursen, 1998). The assumption of normality is not necessary for generating coefficient estimates but for drawing inferences using traditional t and F statistics. Thus asymmetric distribution which underlies the index of Balassa may not generate reliable inferential statistics. To overcome the deficiencies of the index, alternative RCA indices were suggested in the literature. For instance, Dalum *et al.* (1998) and Laursen (1998) suggested the Revealed Symmetric Comparative Advantage (RSCA) index to address the possibility of non-normally distributed residuals associated with the asymmetric distribution of the Balassa index. Hoen and Oosterhaven (2006) with the objective of making the mean of the distribution of Balassa index stable across sectors, formulated the Additive Revealed Comparative Advantage (ARCA) index. Proudman and Redding (1998) made an attempt at fixing the mean of the Balassa index distribution across sectors by normalizing the original index by its cross-sectoral mean. Yu *et al.* (2009) suggested the Normalized Revealed Comparative Advantage

(NRCA) index, which not only has a stable mean across sectors but also across countries. The index is also characterized by stable distributions over time. Vollrath (1991) suggested an index by taking the logarithm of Balassa index which on account of its possible normal distribution could generate normally distributed residuals in parametric regression analysis. With due recognition to the fact that RCA indices based on both demand and supply effects are more consistent with the theoretical concept of comparative advantage, Vollrath (1991) recommended Relative Trade Advantage (RTA) index and Revealed Competitiveness (RC) index. However, the recommendations for alternative RCA indices are not restricted to trade based indices only. As observed by Vollrath (1991) RCA indices can be recommended in many alternative forms by using various combinations and transformations of post trade variables used to determine comparative advantage of countries. Bowen (1983) argued that the structure of the Balassa index, based on gross export figures, is in violation of the comparative advantage theory which essentially is a net trade concept. Accordingly he developed two indices – the Net Trade Intensity index and Production Intensity index which are based on the association between a country's production, consumption and trade. Lafay (1992) made an attempt at analyzing the influence of macroeconomic variables on RCA indices and suggested the Lafay index comprising of exports, imports and total GDP of a country. However, in order to provide more intensive analysis on RCA indices, the discussions in this thesis will be restricted to trade based indices only.

1.1.4 RCA Indices and Global Supply Chains

While the trade based RCA indices mentioned so far incorporate gross trade values, considering growing significance of global supply chains (where a commodity could pass through multiple countries during its production stage) it becomes imperative to study the implications of such value-added trade for the RCA indices. In the presence of global supply chains, the final product generated includes the value-added not only by the exporting country but also by several other countries through which the product passed during its production stage. Given the growing importance of production fragmentation, it is indeed essential to reconsider the RCA indices on the basis of domestic value-added in trade figures. A country's comparative advantage or potential in exporting a particular product relative to other countries can only be understood if its RCA index is adjusted to reflect domestic value-

added in trade. Thus if China exports a \$400 iPhone using \$250 worth of imported contents, then only \$150 should count towards iPhone sector's exports from China and this figure should be included in China's comparative advantage calculation. RCA measures based on value-added in exports have been considered in the literature to address the potential differences between such measures and the RCA measures based on gross export figures (Koopman *et al.*, 2010; Koopman *et al.*, 2014, Meng *et al.*, 2012).

1.1.5 Relation between Comparative Advantage, Global Supply Chains and Intra-Industry Trade

It is necessary to recognize certain issues with respect to the Ricardian and Heckscher-Ohlin trade models, global supply chains and intra-industry trade. Trade in parts and components, which is the essence of value-added trade, brings to the fore the concept of intra-industry trade. Hence it might seem that integrating the concept of RCA index with global supply chains is misleading, as the assumptions underlying the comparative advantage theories and intra-industry trade are different. However this need not be so as will be revealed by the following discussions.

Pittiglio (2014) observed that the primary purpose of both the theoretical and empirical literature on intra-industry trade has been to analyze the exchange of different varieties of products without any reference to intermediate goods. As previously discussed, Krugman (1979, 1980), Lancaster (1980), Helpman (1981), Eaton and Kierzkowski (1984) considered intra-industry trade to exist in horizontally differentiated products between countries with similar factor endowments. Such trade was explained by economies of scale in the literature. On the other hand, Falvey (1981), Falvey and Kierzkowski (1987) insisted on vertical product differentiation, which could take place among countries with dissimilar factor endowments, and could therefore be explained by the Heckscher-Ohlin relative factor endowment model. Unlike horizontal intra-industry trade, which relied on monopolistic competition and increasing returns to scale, the proponents of vertical intra-industry trade based their models on perfectly competitive markets and constant returns to scale.

However, Grubel and Lloyd (1975) argued that the simultaneous exchange of goods within the same industry between two countries need not be restricted to final products only, but

also to exchange of final goods for intermediate inputs or intermediate goods for other intermediate goods (as cited by Pittiglio p. 469). In the more recent literature, Jones *et al.* (2002), Ando (2006) and Türkcan (2011) have argued that international fragmentation of production may lead to the inter-country exchange of intermediate goods within the same industrial classification. According to Pittiglio (2014) this exchange of intermediate goods may take three forms – the exchange of horizontally differentiated intermediate goods, the exchange of vertically differentiated goods, and vertical specialization which involves value-addition by the associated countries. Arndt (1997) as well as Feenstra and Hanson (1997) suggest that vertical specialization involving value addition incorporates the ideals of Heckscher-Ohlin model, whereby production fragmentation is undertaken to exploit the factor cost differences across countries. Deardorff (2001), Jones and Kierzkowski (2001) also exhibit that international fragmentation of production could be due to factor price differences caused by technology differences as in Ricardian model and or factor endowment differences as in Heckscher-Ohlin model. Türkcan and Ates (2011), by dividing total intra-industry trade in automobile parts between USA and 29 OECD countries into horizontal and vertical components, found that a substantial portion of this trade is vertical. The extent of this vertical intra-industry trade, which in effect is the outcome of the international fragmentation of production, is positively related to inter-country differences in market size, factor endowments and outward foreign direct investment. The fact that multinational corporations emerge to take advantage of cross country differences in factor rewards has also been recognized by Helpman (1985). As observed by Kinoshita and Campos (2006), the location of foreign direct investments associated with the international fragmentation of production is closely linked with the comparative advantages of the host country, which the home country in turn has to take into account to increase the profitability of investments.

Thus the fragmentation of production process is primarily undertaken to reap the benefits of cheaper factor costs or differences in technology. In this scenario, vertical specialization can be explained by the theories on vertical intra-industry trade, but not by the theories on horizontal intra-industry trade. Since the assumptions underlying the inter-industry trade theories of Ricardo or Heckscher-Ohlin are not contradictory to those underlying the vertical intra-industry trade models, it may not be unusual to link the comparative advantage theories with the theories on vertical specialization. Even if the importance of horizontal intra-

industry trade is realized, Davis (1995) shows that increasing returns to scale is not essential for intra-industry trade to occur and some of the characteristics of intra-industry trade highlight the importance of Ricardian determinants of trade. When intra-industry trade is defined to be the exchange of goods with similar factor intensity, trade becomes sensitive to even minor technical differences.⁶ Thus, even when the markets are perfectly competitive and returns to scale are constant, Davis shows that intra-industry trade could be explained by the traditional Ricardian theory, i.e., technical differences could account for intra-industry trade. The disaggregation of intra-industry trade in intermediate goods into horizontal differentiation, vertical differentiation and vertical specialization have further reinforced the importance of Ricardian or Heckscher-Ohlin determinants of comparative advantage in the context of intra-industry trade and the prevalence of classical and neo-classical assumptions of perfect competition and constant returns to scale. Thus, inter and intra industry trade may not be two completely different strands of literature, and integrating the two concepts through the adjustment of RCA indices for domestic value-added in exports is an important step in the literature. Further, as noted by Ando (2006), the exchange of products under vertical specialization need not be restricted to products within the same industrial classification. Value addition could result in intermediate products or in final goods with different industrial classifications. As for example, computer chips could be used not only in computers but also in cars, micro-vans, cameras etc. (Jones and Kierzkowski 2001). Thus, although Arndt (1998) and Deardorff (2001) insist on the importance of intra-product specialization to define fragmentation of production process, as argued, production fragmentation could also involve inter-product specialization. Under such circumstances, vertical specialization would go beyond the boundary of intra-industry trade and would more be at par with the concept of inter-industry trade. Hence, the adjustment of RCA indices for domestic value-added in trade

⁶ Goods with similar factor intensity are characterized by excellent substitution possibilities. Moreover, as inherent in the intra-industry trade literature, with number of goods being substantially large relative to the number of factors, it is possible to expand the production of some goods at the expense of others without raising the marginal opportunity costs. Both of these features of intra-industry trade underline the essence of Ricardian determinants of trade such that with linear transformation curves, trade could take place on the basis of technical differences (Davis 1995).

need not necessarily imply trying to determine inter-industry trade through the instruments of intra-industry trade (Deb and Hauk, 2015).

1.2 Gap Areas

The discussions in the previous section are chiefly aimed at laying the ground for the identification of possible areas of research. The first possible issue which demands exploration in greater details is determination of the consistency of the RCA indices suggested in the literature with the comparative advantage theories. While many RCA indices have been proposed to address one or more shortcomings of the Balassa's index, empirically examining and comparing their consistency with the theories on comparative advantage continues to remain largely an unexplored area. Such an exercise will help in the determination of an index which incorporates the basis of the theory into its structure and which can therefore be reliably used by the policy makers for identifying comparative advantage of countries in various products. In this context, it is worth noting that since many of the RCA indices are based on both exports and imports, they can be expected to incorporate the distortionary impacts of the trade barriers imposed on imports. Hence, adjusting the RCA indices for trade barriers on imports can evidently be justified. One can argue that even exports are subjected to trade barriers, so one should suitably adjust export values too. But Vollrath and Vo (1990) found export performance of countries are more likely to be influenced by relative efficiencies and relative factor endowments than by policy induced market distortions. They observed imports on the other hand to be influenced more by government policies such as shifts in domestic monetary growth. These findings provide support for adjustment of import values and non-adjustment of export values for trade barriers. Thus, construction of RCA indices with relevant adjustments to import data and comparing their performance with RCA indices without adjustments to imports, will permit an examination of whether distortionary impacts of trade barriers on imports are significant enough to produce misleading results about the comparative advantage of countries.

However, establishing a theoretically consistent index may not be sufficient as these indices are also used for cross sector, cross country and cross time comparisons. In this respect, stability of the index distribution is an important feature which must also be tested. An index can be identified as "ideal" if it features a stable distribution over sectors, countries and time,

and additionally be empirically consistent with the comparative advantage theories. Although the existing literature have tried to analyze the empirical distribution of Balassa's RCA index, examining the distributions of other RCA indices is definitely a significant contribution to the literature.

The next important issue demanding attention arises from the growing significance of value-added trade. All of the RCA indices originally suggested are based on gross exports or imports. In the context of rapidly changing world trade scenario, one might wonder how far these indices portray the true picture. As discussed, today most countries of the world are involved in intermediate goods production with a good being passed from one country to another with value added by each country until a final product emerges. Under such circumstances, it is not really the gross exports or imports of a country that one should be concerned with to determine a country's comparative advantage, but rather the value-added by a particular country in its exports or imports. Although some of the literature have tried to re-estimate Balassa's RCA index using value-added approach, estimating the other RCA indices identified in the literature using similar methodology, and comparing their performance continues to be an un-attempted task.

It is important to note at this outset the scope of the present research. The existing empirical research on Ricardian and Heckscher-Ohlin models determine whether flow of trade is according to the respective models using observed data (MacDougall, 1951; Balassa, 1963; Golub and Hsieh, 2000; Leontief, 1953; Leamer, 1980; Bowen *et al.*, 1987; Trefler, 1993, 1995; Davis and Weinstein, 2001). The observed trade figures by themselves do not provide any information on the sector specific comparative advantage or disadvantages of countries. The current study uses indices which provide information on the comparative advantage and disadvantage of countries in various products. Ad hoc usages of such indices are based on the assumption that these reflect the Ricardian or Heckscher-Ohlin notions of comparative advantage. It seems worthwhile in this context to empirically verify this assumption. The objective of the current study is therefore not to empirically verify the predictions of the theory but to examine whether different RCA indices are compatible with the theories of comparative advantage.

1.3 Research Questions

On the basis of the identified areas of research, the following research questions could be formulated:

1. Among all the exports-only and export-and-import based RCA indices, which index has sufficiently good structural features and can generate stable distributions over time, sectors and countries? In this context analyzing the empirical distributions of the RCA indices is of considerable significance.
2. Which export based RCA index generates reasonably good results when empirically analyzed to determine its consistency with the Ricardian and Heckscher-Ohlin theory of comparative advantage?
3. Do the export-and-import based indices perform better than the exports-only indices in terms of their consistency with the theories?
4. Do adjustments to import data in order to account for trade distortions improve the performance of the export-and-import based indices?
5. Do the previously established results change with adjustment of RCA indices in order to account for domestic value-added in trade?

1.4 Hypotheses

Since the objective of the present study is primarily to determine the consistency of various RCA indices with the Heckscher-Ohlin and Ricardian theories on comparative advantage, following two hypotheses can be outlined:

H₁: RCA indices in labor intensive commodities would rise with a rise in ratio of relative labor force to relative capital stock. RCA indices in capital intensive commodities would fall with a rise in ratio of relative labor force to relative capital stock.

H₂: RCA indices would rise with a rise in the ratio of average labor productivity of a country in a particular sector relative to all considered sectors, to the average labor productivity of the world in that particular sector relative to all sectors.

The above specified alternative hypotheses in the context of Heckscher-Ohlin theory (H₁) and Ricardian theory (H₂) are tested against the following null hypothesis:

H_0 : *There is no relation between the RCA index values and relative factor ratio or average productivity ratio.*

1.5 Data and Methodology

The present study is based on secondary sources of data. World Bank's Trade Production and Protection (TPP) Database by Nicita and Olarreaga (2007) is the primary source of data for exports and imports, value-added, number of employees and output. The database contains data on several indicators for about 100 developed and developing countries from 1976 to 2004 on 28 manufacturing sectors based on 3-digit International Standard Industrial Classification of all Economic Activities (ISIC) revision 2 codes. Working with manufacturing sector rather than any other sector avoids the possibility of biases that could arise from the geophysical and geographical characteristics of countries, as noted by Benedictis and Tamberi (2004). The TPP database is also the source of input-output coefficient matrices necessary for the adjustment of RCA indices for domestic value-added in exports. The country specific data on labor force and capital stock have been collected from the Extended Penn World Tables version 4 by Marquetti and Foley (2011). The database provides data on about 176 countries for the years 1963 to 2009, subject to the fact that not all countries report data for every year considered. Marquetti and Foley make use of the Penn World Table (PWT) version 7.0 to arrive at the estimates for labor force and capital stock. The estimate for capital stock is obtained by Perpetual Inventory Method (PIM) using data on investment series calculated from the variable real investment share of GDP present in PWT 7.0. The PIM assumes that capital stock for each period is a function of some predetermined capital stock, investment and depreciation rate. Although the method is subject to criticism due to the consideration of common rate of depreciation across countries and over-valuation of estimated capital stock in case of developing countries, it is the most widely used approach towards arriving at comparable capital stock estimates across countries.⁷ The data on import

⁷ Pritchett (1996) suggests that in developing countries for each public dollar invested, less than 50 cents of effective capital would be created due to inefficient public sector. The PIM method on the other hand assumes away any such inefficiency involved. As a result, the public capital stock series for developing countries estimated using PIM, may be overvalued. This problem is however negligible in case of developed countries (Arestoff and Hurlin, 2008).

demand elasticities and tariff rates required for the adjustment of import values for tariff barriers, are acquired respectively from Kee *et al.* (2008) and Nicita and Olarreaga (2007). To our knowledge, Kee *et al.* (2008) is the only source of data for import demand elasticities. Nicita and Olarreaga (2007) obtain the data on tariff from the United Nations Conference on Trade and Development's (UNCTAD) Trade Analysis Information System (TRAINS) database, which is a comprehensive source of data on trade statistics. But the original data being based on 6 digit Harmonized System (HS) codes, Nicita and Olarreaga (2007) converted them into 3 digit ISIC codes using concordance tables. Finally to study the comparative advantages of India and China in United States with the help of more recent dataset, World Input Output Database (WIOD) has been utilized. As TPP database could provide data only till the year 2004, WIOD with data on more recent years is useful. WIOD provides data on input-output tables and several other indicators for about 40 countries and rest of the world from 1995 to 2011, although at higher levels of product aggregation compared to the TPP database. However, for this part of analysis in the thesis, primary products are considered in addition to manufactures, in order to provide a broad over view on the possible exports of China and India to United States.

For analyzing the empirical distribution of the RCA indices both graphical methods and tests of significance are used. Stability of the distributions of RCA indices over time are studied graphically using cumulative distributions and Kernel density plots. The graphical analyses are supplemented by Wilcoxon's signed rank test to determine the statistical significance of the changes in distributions over time. Time stability of the distributions ensures application of the indices for time series analysis. Further, given the extensive usage of the RCA indices as an ordinal or cardinal measure, the statistical significance of the shifts in the distributions of the RCA indices over countries and sectors are analyzed using Wilcoxon's rank sum test and Kolmogorov Smirnov test. The reliance on non-parametric methods for stability analysis stems from the possibility of non-normally distributed residuals for most indices, which implies unreliable inferences in case of parametric estimation methods.

In order to test the consistency of the RCA indices based on gross trade values with the Ricardian and Heckscher-Ohlin theories on comparative advantage, cross sectional bivariate regressions with robust estimate of standard errors (in presence of heteroskedasticity) or with

conventional Ordinary Least Squares standard errors (in absence of heteroskedasticity) are considered. The indicator used to measure Heckscher-Ohlin comparative advantage is the ratio of labor force of a country relative to the world labor force, to the stock of capital of the same country relative to the world stock of capital. The indicator used to measure Ricardian comparative advantage is the ratio of average labor productivity of a country in a particular sector relative to all considered sectors, to the average labor productivity of the world in that particular sector relative to all sectors. Given the possibility of outliers violating the assumption of normally distributed residuals, the OLS regression results are complemented by robust regression results. In case of RCA indices based on both exports and imports, the effects of tariff barriers on the performance of such indices are also examined. In order to factor out the effects of tariff barriers on imports, the import values are first adjusted using the data on import demand elasticities and applied ad-valorem tariff and the reconstituted index values are tested for their consistency with the comparative advantage theories using non-parametric Spearman's rank correlation technique. Adoption of non-parametric techniques for this part of analysis is due to the reduced sample size resulting from the adjustment of import figures. The details of the methodology adopted in this context are discussed in chapter 4.

To understand the significance of value-added trade for RCA indices, the latter based on gross export values are adjusted using the input-output coefficients matrices and reveal the extent of domestic value-added in RCA indices. The details of the adopted methodology are discussed at length in chapter 5. The adjusted RCA indices are then tested for their consistency with the theories on comparative advantage using unbalanced panel data approach with robust estimate of standard errors. In order to test the consistency of the indices with Ricardian theory, three different kinds of regression are considered for each index - (1) Random effects regressions with separate fixed effects for country and year by incorporating all sectors, countries and years in the sample. (2) Random effects regressions with country-year fixed effects with the same sample (3) Random effects regressions with only year fixed effects for separate sectors in the sample. In order to determine the consistency of the RCA indices with the Heckscher-Ohlin theory, only the sector specific regressions with year fixed effects are considered.

1.6 Structure of the Thesis

The thesis is comprised of seven chapters including the present one.

Chapter 2 is a review of the available literature on RCA indices and global supply chains. The chapter discusses the structure of the alternative RCA indices and their respective deficiencies. Also, in the event of the growing incidence of value-added trade, the relevant literature are examined and their implications for the existing RCA indices are recognized in this chapter.

Chapter 3 empirically reviews the structure of different RCA indices given the current database. This chapter particularly focuses on the stability of empirical distribution of each index, as stable distributions permit meaningful comparison of the index values over time, sectors and countries.

Chapter 4 addresses the second, third and fourth research questions simultaneously. The consistency of gross export based RCA indices with the comparative advantage theories of Heckscher-Ohlin and Ricardo is studied. RCA indices based on both gross exports and imports is further analyzed to note any significant differences in their performance from that of RCA indices based only on exports. A re-analysis of compatibility of various indices with the theories of comparative advantage, after adjustment of import values for trade barriers, is also a part of chapter 4.

Recognizing the importance of global value chains, chapter 5 adjusts the RCA indices to incorporate domestic value-added in trade in place of gross trade figures, and eventually examines whether this adjustment has any effect on the results presented in chapter 4.

The analysis of several alternative RCA indices in chapters 2, 3, 4 and 5, reveal that it is indeed difficult to select any one particular index which could be employed for comparability purposes and which at the same time would empirically be consistent with the theories on comparative advantage. The choice accordingly becomes restricted to two particular indices - the NRCA index of *Yu et al.* (2009) and the Log of Balassa index of Vollrath (1991). These two indices are employed in chapter 6 to study the comparative performance of two countries – India and China while exporting commodities to United States, and the differences in

interpretations of the two indices are analyzed. The chapter therefore provides some information of the sectors based on which strategic policies for export expansion can be undertake in the respective countries.

Finally, chapter 7 contains the summary of findings, conclusions, policy suggestions, limitations of the study, and presents an overview of the possible avenues for further research.





Chapter 2

Review of the Literature

2.1 Introduction

One of the oldest propositions in the pure theory of international trade is the Law of Comparative Advantage according to which, the pattern of trade is determined by the opportunity cost of producing goods by the countries. If each country specializes in production and trade of that commodity in which it incurs lower opportunity cost, the total world output of all traded commodity will increase.¹ This proposition is central to both the Ricardian comparative cost theory and Heckscher-Ohlin factor proportions theory, as long as comparative advantage is determined by differences in relative autarky prices (Deardorff, 1980).

While comparative advantage has been conceptualized theoretically to explain the pattern of international trade, an important issue in this regard is to quantify comparative advantage for empirical analysis. In the process of devising measures to identify the comparative advantage of countries in real world, one frequently comes across several problems. The first problem is, while the theory requires measurement of pre-trade prices, data available for measuring comparative advantage pertains to the post-trade situation. Moreover, the data are subjected to distortionary impacts of government intervention and imperfect information (Vollrath, 1991). The second problem encountered during the empirical measurement of comparative advantage, arises out of commodity aggregation. It is possible for a country to have comparative disadvantage in a composite commodity and comparative advantage for a disaggregated commodity within the composite group (Vollrath, 1991). Suitably addressing these concerns is a challenge in empirical analysis. However in the context of the first

¹ This is accompanied by an increase in world welfare. The theory is based on some general assumptions. Goods are assumed to be homogenous, transport costs do not exist, labor is assumed to be homogenous within a country but heterogeneous between countries, labor is mobile between industries within a country but immobile between countries, labor is fully employed and perfectly competitive goods and factor markets exist.

problem, Benedictis and Tamberi (2004) observe that even if autarkic relative prices were available, they might fail to reflect the true comparative advantage of countries. They justify their argument on the basis of literature on overlapping-generations model (Galor, 1992; Mountford, 1998) and on oligopoly theory (Cordella, 1998), and note that trade between countries might not always occur on the basis of their autarkic relative price positions. Thus, developing and using some measure based on post-trade data to determine the comparative advantage of countries, does not seem to be a serious drawback. With regard to the second problem, devising measures which are consistent across commodity disaggregation is also a matter demanding attention. This phenomenon, as subsequent discussions will reveal, cannot be addressed by existing methods for measuring comparative advantage. Therefore assessing the comparative advantage or disadvantage of a country in disaggregated sectoral groups on the basis of its comparative advantage or disadvantage in aggregated sectoral groups, may not be the right approach.

Ballance *et al.* (1987) develop a theoretical framework to demonstrate the relationship between the theoretical concept of comparative advantage and the measures of comparative advantage that are practically feasible. They explain their framework through the following diagram:

$$EC \rightarrow CA \rightarrow TPC \rightarrow RCA$$

According to the above expression, economic conditions (EC) of various countries play a role in determining the comparative advantage (CA) of each country. The comparative advantage of each country then acts as a deciding factor towards country specific trade, production and consumption (TPC). Indices developed by making use of TPC and other post trade variables are employed to indicate the CA of countries and are termed as RCA indices or the Revealed Comparative Advantage indices.

The relationship between CA and TPC holds in a two country - two sectors - two factors of production framework. But as noted by Drabicki and Takayama (1979), the situation is far more complicated by the presence of many countries, sectors or factors and as such the straightforward linkage between CA and TPC may cease to hold. However, Deardorff (1980) show that net exports and relative autarkic prices are negatively correlated under conditions

such as presence of arbitrary number of goods and countries, as well as presence of policies and barriers to trade. Thus, even though CA may not be strictly measurable in accordance to the theory, indices based on post trade data would after all be able to guide us towards the underlying pattern of comparative advantage (Ballance *et al.*, 1987).

Ballance *et al.*'s (1987) theoretical framework implies that for determination of comparative advantage or disadvantage, RCA indices are employed. By comparing national sectoral shares of actual output or trade with their international analogues, these RCA indices help in identifying the comparative advantage or disadvantage of countries in various products.

As discussed in chapter 1 (p. 7-8), the RCA indices can broadly be classified under the following three headings: (1) Exports-only indices (Balassa, 1965; Dalum *et al.*, 1998; Laursen, 1998; Hoen and Oosterhaven, 2006; Yu *et al.*, 2009); (2) Exports-and-Import indices (Vollrath, 1991); (3) Trade-cum-production indices (Bowen, 1983; Lafay, 1992). This chapter presents a review of the literature on the first two categories of index. However, as recognized in chapter 1 (p. 8-9), discussions on RCA indices will be incomplete without noting the implications of global supply chains on the RCA indices. In this context a review of the literature on value-added trade must also be presented in this chapter. The rest of the chapter is therefore divided into three more sections. Section 2.2 presents a review of the literature, beginning with the RCA index of Balassa and its various modifications suggested thereafter. The literature on global supply chains and their implications on RCA indices are discussed in section 2.3. The rational of present study and conclusions are presented in section 2.4.

2.2 The Revealed Comparative Advantage Index of Balassa

Although the concept of Revealed Comparative Advantage Index was pioneered by Bela Balassa in 1965, Liesner in 1958 made a preliminary attempt at quantifying comparative advantage using post trade data. Liesner used relative export performance in order to assess bilateral comparative advantage between Britain and one of its European competitors, for a single commodity, while exporting to Europe. The index proposed by Liesner can be expressed in the following form:

$$\frac{X_a^i/X_a^e}{X_a^d/X_a^e} \quad (2.1)$$

Here X equals export value, a denotes any specified commodity, i , d and e denote respectively Britain, any specified European competitor of Britain, and European countries collectively. However, this index considers only a single commodity in its formulation. Therefore the identification of a country's comparative advantage or disadvantage among a set of countries is on the basis of that commodity only. But, in order to understand the pattern of comparative advantage, it is essential to evaluate in addition, a country's export of a particular commodity relative to other commodities.

By adjusting Liesner's index, Balassa (1965) developed the Revealed Comparative Advantage (RCA) index for measuring comparative advantage. His objective was to analyze the effects of trade liberalization resulting from Kennedy Round of GATT. Balassa comments "*comparative advantage appears to be the outcome of a number of factors, some measurable, others not, some easily pinned down, others less so. One wonders, therefore, whether more could be gained if, instead of enunciating general principles and trying to apply these to explain actual trade flows, one took the observed pattern of trade as a point of departure....*"(p. 116-117). He further acknowledged, comparative advantage could be revealed through the exchange of merchandise between countries, as goods flow is expected to reflect differences in relative costs of production between countries (Balassa, 1977). Thus a stronger relative trade performance of a country in a commodity, can be considered to reflect its greater comparative advantage in production of that commodity.

The original RCA index of Balassa can be defined in the following form:

$$\frac{X_a^i/X_a^c}{X_m^i/X_m^c} \quad (2.2)$$

Here m denotes all manufactured goods and a to any one of the manufactured goods. i denotes any one of the following developed countries: Belgium, Canada, France, Italy, Japan, Luxembourg, Netherlands, Sweden, United Kingdom, United States and West Germany. c refers to all of these 11 industrial countries. The index is thus expressed as the ratio of a country's share in 11 countries' exports of a particular product, to its share in the 11 countries' exports of all manufactured goods. A value of the index greater (less) than unity

signifies revealed comparative advantage (disadvantage) in product a by country i . The comparative advantage neutral point is at unity. The noteworthy fact about Balassa's index is that, unlike the Liesner's index, it compares the export performance of a single commodity with that of the manufactured exports as a whole. Thus it permits an evaluation of relative export performance of a country not only with respect to other countries or a group of countries, but also with respect to other commodities or a group of commodities.

Identifying the fact that Balassa's index is restricted in terms of both commodity and country coverage, it was later modified to include all countries and all traded commodities (Vollrath, 1991). Hence, the index takes the following form:

$$\frac{X_a^i/X_t^i}{X_a^w/X_t^w} \quad (2.3)$$

Here a refers to any specific commodity (not necessarily manufactures); t refers to all traded commodities (both manufactures and non manufactures); i and w to any country and the world respectively. Country i is said to 'reveal' comparative advantage in a particular commodity a if the value of the index exceeds unity. The country will 'reveal' comparative disadvantage in commodity a if the value of the index is less than unity. In the thesis henceforth, while referring to the Balassa's RCA index, the above version of the original index will be considered.

It is to be noted, in classical and neo-classical framework, comparative advantage is measured by considering two country – two commodity structure. For this purpose, while computing the comparative advantage of country i using any index, the reference group of countries should be entire world less country i i.e., rest of the world. Similarly, the reference group of commodities should be all commodities excepting product a . But, problem arises when using such indices for inter-country or inter-commodity comparison. The reference group of countries and commodities will change from country to country and from commodity to commodity respectively. Therefore, the utility of such indices for inter-country or inter-commodity comparison is questionable. To overcome such difficulties, the country under consideration is included in the reference group for the index of Balassa and for all other export based RCA indices considered subsequently. The commodity under consideration is also included in reference group of commodities.

The RCA index of Balassa has achieved more importance among the applied international trade economists than any measure based on net exports because it provides a broader view of the concept of specialization of countries (Brasili *et al.*, 2000). As will be evident from the structure of the index, it compares the performance of a country in exporting a product with that of all countries, whether or not they export that particular product to the considered country. But in case of measures incorporating net exports, the comparison is with respect to all those countries that export that product to the considered country. It is in this respect Brasili *et al.* (2000) project RCA index of Balassa (and for that matter all exports-only RCA indices) as an all inclusive measure of overall specialization of countries.

2.2.1 Critical Analysis of the RCA Index of Balassa

With the RCA index of Balassa assuming immense popularity among researchers, several attempts have been made in the past to link the index to the theory of comparative advantage. Hillman (1980) theoretically tried to explore the relation between Balassa's RCA index and comparative advantage as indicated by pre-trade relative prices. He argued, while comparing the RCA indices for two commodities with respect to a country, the values of the indices are independent of the comparative advantage as indicated by pre-trade relative prices. However, while comparing the RCA indices for two countries with respect to a commodity, Hillman showed that the correspondence between RCA index values and the pre-trade relative prices would hold only under certain circumstances. If the share of a country's exports in the world exports of particular commodity is very small (closer to zero) and if the share of that commodity in the concerned country's total exports is also very small (closer to zero), then Balassa's RCA index is consistent with the comparative advantage as indicated by the pre-trade relative prices. However, at higher levels of commodity aggregation, the share of a commodity in the country's total exports could be closer to unity. Under such circumstances, Balassa's RCA index and the comparative advantage as indicated by the pre-trade relative prices would be inconsistent. Hence, finer the level of commodity disaggregation, greater is the possibility that Balassa's RCA index will be consistent with the theory. Following Hillman's argument, application of the index with highly disaggregated data will be able to produce more theoretically consistent results than with aggregated data. In fact, given the method of product aggregation, inter industry trade can be more observed at finer levels than

at higher levels of commodity aggregation. At higher levels of product aggregation, intra-industry trade might be observed. For instance, considering International Standard Industrial Classification of All Economic Activities (ISIC) code, a country might be exporting a product classified under one 4 digit ISIC code and importing another product classified under a different 4 digit ISIC code. But if both the products are under the same 3 digit sectoral classification, the country will appear to be engaged in inter-industry trade on the basis of 4 digit classifications but intra-industry trade on the basis of 3 digit classification. Hence, the argument forwarded by Hillman is supported by the methods of sector classification. Therefore, Balassa's index can be expected to reflect the underlying inter-industry trade theory when higher levels of sectoral disaggregation are taken into account.

Among other similar studies, Bowen (1983) showed that an index of revealed comparative advantage based entirely on trade flows is not valid. He used the probabilistic framework assumed by Kunimoto (1977) to put forward his arguments. Kunimoto (1977) theoretically interpreted the differing values of the revealed comparative advantage index. He proposed that deviations of the indices from unity “*measure the extent to which actual trade deviates from the trade expected in a world in which factors influencing the direction (but not the level) of a country's trade are absent*” (Bowen, 1983, p. 464). Following Kunimoto, Bowen represented the Balassa's index of revealed comparative advantage as:

$$\frac{X_a^i}{E(X_a^i)} \text{ where } E(X_a^i) = \left(\frac{X_t^i}{X_t^w} \right) X_a^w \quad (2.4)$$

It is the ratio of actual exports of commodity a by country i , to exports of commodity a by country i expected in a world where, world exports of commodity a are distributed among countries in proportion to the share of their total exports in world total exports. Thus, $E(X_a^i)$ represents expected trade in a hypothetical world where factors influencing the direction of country's trade are absent, which in effect implies absence of any relative advantages. But, Bowen argued, formulation of such a hypothetical world is based implicitly upon the assumption that each commodity is exported by all countries, which is not a reasonable assumption. If simultaneous exports and imports of every product by every country are ruled out, the index becomes undefined.

Bowen, however, recognized the fact that it is possible for every country to export every commodity if the presence of intra-industry trade is recognized. But, according to him, it would be unjustifiable to assume that such intra-industry trade would occur even at very high levels of commodity disaggregation.² His argument can be questioned as intra-industry trade or trade based on variety of a product can exist even at finer levels of commodity disaggregation, although as argued previously, the instances of inter-industry trade might be relatively more. But whether the deviation of actual trade from expected trade of intra-industry type represents comparative advantage is a matter of consideration.

Although Hillman (1980) and Bowen (1983) express their skepticism about the theoretical foundation of the Balassa index, Yeats (1985) in an empirical study examined the consistency of the index with the factor proportions theory of Heckscher-Ohlin and came up with contradictory results. According to the Heckscher-Ohlin theory, richer countries with relative abundance of capital should specialize in the production and export of capital intensive products, while poorer countries with relative abundance of labor should specialize in the production and export of labor intensive products. To establish the consistency of Balassa's RCA index with the factor proportions theory, he thus regressed the average per capita income (Y/N) of five countries recording highest RCA values for an industry, on the value added per worker (L) in that industry. Following the works of Lary (1968) and Tuong and Yeats (1980), he assumed that higher value added per worker would imply more capital per labor and therefore more capital intensive industries. Since such industries are typically located in rich countries, higher value added per worker would be associated with higher per capita income. Similarly, lower value added per worker would imply less capital per worker and hence would indicate labor intensive industries. Labor intensive industries would be associated with poor countries and therefore lower per capita income. Thus, a positive

² As discussed in chapter 1 (p. 8), Bowen suggested a Net Trade Intensity Index to determine the comparative advantage of countries and overcome the deficiencies of trade based indices. Net trade of a country in his formulation, is identified to be the difference between its production and consumption of a product and to measure the comparative advantage of a country in a product, its net trade is scaled with respect to expected production. However, his index is not considered for analysis in the thesis because it concentrates more on the validity of trade based RCA indices, and thus avoiding Bowen's production based RCA index seems quite justified.

association could be expected between the value added per worker and the average per capita income. With data on 40 manufacturing industries and 47 countries, his estimated regression equation took the following form:

$$\left(\frac{Y}{N}\right)_i = -2522.96 + 75.79 \times L_i \quad \bar{R}^2 = 0.51 \quad (2.5)$$

(6.22)

Figure in parenthesis denote the t value, i represents the i th industry. Thus, the estimated coefficient establishes a positive and significant association between value added per worker and per capita income. Yeats further fitted another regression equation by incorporating a dummy variable (D) to classify industries as natural resource based and non-resource based. The modified regression was fitted to determine whether resource endowments alter the relation between labor intensity and per capita income. The modification did not alter the findings of the previous model as it produced modest improvement in the explanatory power of the model. The modified estimated regression equation took the following form:

$$\left(\frac{Y}{N}\right)_i = -2118.24 + 69.21 \times L_i + 11.29 \times D_i \quad \bar{R}^2 = 0.57 \quad (2.6)$$

(5.87)

(1.86)

Thus, Yeats tried to establish a link between the values of the Balassa's RCA index and the predictions of factor proportions theory and came to the conclusion that the RCA index of Balassa is at par with the theory. However, his results need to be analyzed in detail as he does not consider the RCA index values directly into the regression.

Problems with the index of Balassa in empirical estimation have also been recognized in the literature. Yeats (1985) for instance, pointed out that the traditional method of ranking industries in a country according to the value of the Balassa's RCA index may fail to indicate that the country could be a leader in a particular industry as compared to other countries. Yeats ranked countries according to the RCA values in a particular industry to check the consistency of this ranking with the ranking of industries in a particular country according to their RCA values. The inconsistency between the two rankings led Yeats to conclude that Balassa's RCA index does not accurately rank industries according to a country's real

comparative advantage. Yeats pointed out that such inconsistency in ranking arises out of the fact that different industries have different distributions of country index values. The phenomenon highlights the responsiveness of the index to the size of a sector or country. Hence, the index might signify stronger comparative advantage for countries with smaller market share in the world export market i.e., smaller $\frac{X_t^i}{X_t^w}$, and for commodities with smaller market share in the world export market i.e., smaller $\frac{X_a^w}{X_t^w}$. Accordingly the distributions might differ across countries or across commodities.³

The other important shortcoming of the Balassa's index is its asymmetric distribution. Laursen (1998) and Dalum *et al.* (1998) argue that the asymmetric property of the Balassa's index makes it unsuitable for econometric analysis. The lower bound of the index is zero. The upper bound of the index can assume any value greater than unity and will generally tend to infinity.⁴ Laursen comments, with such a skewed distribution, the error term in any regression analysis incorporating the RCA index will be non-normal and hence, traditional *t* and *F* statistics cannot be used reliably. The asymmetric distribution also ensures that values above unity are given greater weights than values below unity. Thus for a highly skewed

³ If in an industry, the country index values are highly concentrated around unity, then the country with the greatest comparative advantage in the industry can have relatively low RCA index value. If on the other hand, in another industry, the country index values are widely distributed around unity, then the country which does not have greatest comparative advantage relative to other countries might have very high index value (Yeats 1985). If an industry has large share in the world export market then the RCA values will be low and hence will be concentrated around unity. If on the other hand an industry has smaller share in world export market, the RCA values will be widely spread around unity. This represents the responsiveness of Balassa's index to the size of a sector. Hence there might arise inconsistency between industry-country ranking and country-industry rankings. Similar arguments can be presented with respect to the responsiveness of the distribution of the index to the size of a country.

⁴ If the reference group of countries consists of the world less country *i*, the lower and the upper limits of the distribution will be 0 (rest of the world completely specializes in product *a*) and undefined (country *i* completely specializes in product *a*) respectively, with 1 being the comparative advantage neutral point. But if the reference group is the entire world, the limits of the distribution will be as mentioned in the text. Use of entire world as the group, ensures comparability of the index values across country. The limits of the distribution of the Balassa index are derived in A.1.1 of the Appendix A.

distribution, the arithmetic mean of observations across sectors (for a country) or across countries (for a sector) will be above 1. Hence, if one tries to determine the changes in a country's average specialization through the movement of mean of the distribution, one will be influenced relatively more by the sectors with value greater than unity (Benedictis and Tamberi, 2001). Thus, mean might not be a good indicator for the index of Balassa. This asymmetry in the distribution is to some extent, the result of the size bias of Balassa's index. To cite an example, if two countries have the same share in the world exports of a product $\left(\frac{X_a^i}{X_t^w}\right)$ but different shares in the world export market of all products $\left(\frac{X_t^i}{X_t^w}\right)$, the country with larger share in the world export market will be less specialized in the product. It follows, smaller countries will have higher values for the index and more skewed distributions (Benedictis and Tamberi 2004).

Hoen and Oosterhaven (2006) also noted the difficulties of the index due to its size bias. They argued the distribution of the index will vary with the level of aggregation of the commodity. At more and more detailed sectoral classification, the denominator of the index becomes smaller and smaller, which alters the extreme points of its distribution. The minimum point remains fixed at zero while the maximum point of the distribution increases, producing more and more skewed distributions for increasing commodity disaggregation. The resulting unstable distribution across commodity categorization will produce variable mean, making comparison of index values across commodity difficult.⁵ Hoen and Oosterhaven attributed this problem to the ratio form of the index. To provide support for their claim, using the export data for Netherlands and Poland for the year 1997 they demonstrated that the distributions of the index varied with varying sectoral classifications. The standard deviation and the maximum values for the RCA index increased with higher sectoral disaggregation. The minimum value however remained fixed at zero, which implied increasing mean value (always above 1) for detailed sectoral classifications. A fluctuating mean value over the years 1992 to 1996 for a group of EU member countries was also documented by Hinloopen and Marrewijk (2001). As in the case of Hoen and Oosterhaven's analysis, they found estimated mean value to be greater than unity in all cases. Their analysis of the distribution of the index however revealed its stability over time but not over countries.

⁵ The mean of the distribution for Balassa index is derived in A.2.1 of Appendix A.

Thus, they conclude that the index values are comparable across time but their comparison across countries must be done with care. Yu *et al.* (2009) also recognized the problems associated with the variability of mean for the index of Balassa.

It is interesting to note however, despite the problems associated with the RCA index of Balassa, its extensive application in studies incorporating cross-country or cross-commodity comparability can be found. Richardson and Zhang (2001) analyze the variations in USA's trade patterns across time, sectors and regions by making use of the Balassa's RCA index. They explore USA's export patterns across 38 major trading partners after disaggregating them into regional groups (excepting China and Japan, who are treated separately) between 1980 and 1995 at different levels of commodity aggregation. They find variations in the pattern of US comparative advantage across regional groupings and that these differences have altered during the period studied. They also find discrepancy in US comparative advantage across product aggregation, for US exhibits comparative disadvantage in some product group but comparative advantage in disaggregated products of the same product group and vice versa. Although the first kind of discrepancy could be due to the sensitivity of the Balassa index to sector size, it is not possible to explain the second kind of discrepancy by the structural deficiencies of the index.

Bender and Li (2002) examine the changes in revealed comparative advantages of East Asian, South East Asian and Latin American countries over the period 1981 to 1997. Analyzing the RCA index values they conclude that South East Asian and Latin American economies have gained comparative advantage in certain sectors, while East Asian economies have lost comparative advantage in some of those sectors over the period considered, and that the former two groups of countries have become more competitive by revealing comparative advantage in similar commodities.

Hinloopen and Marrewijk (2004) use Balassa's RCA index to study the changing pattern and competitiveness of China, Hong Kong and Taiwan. They rank the sectors according to the index values in the three countries to analyze the changes in their export performance over the years 1970 to 1997. They note among the three countries analyzed, Taiwan has undergone maximum extent of structural transformation by developing comparative advantage in technically advanced products over agricultural commodities.

Batra and Khan (2005) used the RCA index to determine the differences in the patterns of comparative advantage between India and China for the years 2000 and 2003, and found the pattern to vary at different levels of commodity disaggregation. Those sectors categorized on the basis of two digit Harmonized System (HS) classification, placed among the top ten according to the RCA index values, were unable to retain their ranks, when their corresponding six digit commodity classification was considered.

In a similar study, Wignaraja (2011) compared the RCA index values for China and India for the year 2009 at two digit HS level. Based on the calculated index values, he concluded China has more sectors with an index value greater than unity, compared to India, and is undoubtedly a relatively better performer in manufactured exports, particularly in medium and high technology products.

In another study by Burange and Chaddha (2008) the RCA index was applied to analyze the structure and change in comparative advantage of India in various products, at different levels of product disaggregation over the period 1996 to 2005. They also note that items with revealed comparative advantage at (HS) six digit level of commodity disaggregation need not come from the sectors with HS two digit level of commodity aggregation and with RCA values greater than unity. For instance, although machinery and equipments do not find any position among the two digit level sectors with comparative advantage, items such as tyre-moulding machinery, engines, turbines, track laying tractors, cycle brakes and pedals and others, emerge as sectors with comparative advantage when commodity classification at six digit level is considered. Obviously, whether such results are indeed due to the comparative advantage of India in such disaggregated sectors or due to the sensitivity of the index to the sector size, is a matter requiring further examination and can some extent be answered by analyzing the consistency of the Balassa index with the theory.

In his study on Ireland, Smyth (2005) explored the changes in its pattern of comparative advantage over the period 1997 and 2002 for different product level disaggregation, and found a shift in comparative advantage of the economy from indigenous industries to industries requiring high technology. The study does however document greater values for the RCA index at higher levels of commodity disaggregation than at lower levels and can be due to the responsiveness of the index to the size of a sector.

To apply the index for cross-country comparison, Karaalp (2011) compares Turkey's comparative advantage or disadvantage in 16 product groups, with the Commonwealth of Independent States (CIS) countries between 1996 and 2008 while exporting to the world. Based on the RCA index values, his analysis revealed substantial amount of similarity in the export potentials of Turkey and CIS countries.

Ekram *et al.* (2013) considering 11 sectors in textiles and apparel industry for the years 2000 to 2009, used Balassa's RCA index to determine the sectors in which Bangladesh and China reveal comparative advantage while exporting to the world market. They then use Spearman's rank correlation coefficient to examine the sectors in which the two countries compete with or complement each other. A positive value for the rank correlation coefficient indicates competition, while a negative value indicates complementarity among the said nations in the considered product category.

Hassan (2013) applied Balassa's RCA index to compare the performance of Pakistan in exporting agricultural products vis-à-vis some major Asian competitors during the years 2001 to 2010. Thus, considering four product groups – rice, fish and fish preparations, fruits and vegetables, and meat and meat preparations, he compared and contrasted the export potential of Pakistan with leading Asian competitors - China, India, Thailand and Vietnam, on the basis of calculated index values.

Thus, usage of Balassa's RCA index for cross-country or cross commodity comparison even today, signifies the ignorance about the underlying deficiencies of the index which makes it unsuitable for such applications. Hence, it seems quite logical to state that the results of these studies need to be interpreted carefully.

2.2.2 Alternative Indices of RCA

The previous section enumerates the problems associated with the Balassa's RCA index and the evidence provided in support. With due recognition to the problems associated with the RCA index of Balassa for empirical estimation, a number of alternative RCA indices have been proposed in the literature.

Identifying the asymmetrical distribution of the index of Balassa, Dalum *et al.* (1998) and Laursen (1998) suggests a simple modification of the index in order to make its distribution symmetric. The suggested index, known as the Revealed Symmetric Comparative Advantage (RSCA) index takes the following form:

$$RSCA = \frac{\frac{X_a^i/X_t^i}{X_a^W/X_t^W} - 1}{\frac{X_a^i/X_t^i}{X_a^W/X_t^W} + 1} \quad (2.7)$$

The lower limit of the distribution of this index is -1 with upper limit tending to +1. The comparative advantage neutral point will be close to zero.⁶ Country *i* will reveal comparative advantage in product *a*, if the value of the index is positive, and reveal comparative disadvantage if the value of the index is negative. This transformation of the Balassa index ensures that the values above zero have the same weight as the values below zero. Laursen (1998) attempted to demonstrate the difference in the estimated regression coefficients and thereby test for the normality of the error terms by incorporating Balassa's RCA index and the RSCA index in separate Galtonian regression⁷ equations of following form:

$$RCA_{ij}^{t_2} = \alpha_i + \beta_i RCA_{ij}^{t_1} + \varepsilon_{ij} \quad (2.8)$$

In the above equation, t_1 and t_2 represent the initial and final year respectively; *i* and *j* represent the sector and country respectively. Regressions were run for 19 countries by incorporating data on 19 sectors for the years 1971 to 1991. The regression equation (2.8) was fitted to test the null hypothesis that the level of specialization of countries as measured

⁶ If the reference group of countries consists of the rest of the world, the lower and the upper limits of the index will be exactly equal to -1 (rest of the world completely specializes in product *a*) and +1 (country *i* completely specializes in product *a*) with zero being the comparative advantage neutral point. The limits of the distribution of the RSCA index are derived in A.1.2 of the Appendix A.

⁷ The concept of Galtonian regression was first adopted by Hart and Praise (1956) from Galton's (1889) methodology while studying the heights of fathers and sons and analyzing size wise concentration. Since then, the Galtonian regression method has been extensively employed in the literature to examine technological specialization patterns (Cantwell 1989), convergence of productivity levels over time (Hart 1995) and changes in the structure of trade specialization using RCA indices (Dalum *et al.*, 1998; Laursen, 1998; Frantzen, 2008; Sharma & Dietrich, 2007; Sanidas and Shin, 2010; Sanidas and Shin, 2011).

by the RCA indices, remains unchanged over time. Laursen applied the Jarque Bera test for normality of the error terms from the estimated regression to find that the hypothesis of normality of error terms was rejected for 2 out of 19 regressions in case of RSCA index, and for 8 out of 19 regressions for Balassa's index at 10% level of significance. Based on these findings Laursen concluded that compared to Balassa's index, his symmetric index could generate reliable results when used in econometric analysis. However, Benedictis and Tamberi (2001) in their paper replicated Laursen's exercise using more recent database. They found that at higher and higher levels of commodity disaggregation, the hypothesis of normality of residuals is progressively rejected for the RSCA index.⁸ They also claim that the interpretation of the values of the index is not easy.⁹ Moreover, like the Balassa index, the RSCA index too does not feature a stable distribution with a stable mean, as will be evident from considering the sum of the index values for a country across sectors, or for a sector across countries. The mean of the distribution for the index is derived in A.2.2 of the Appendix A. This phenomenon raises concern over the applicability of the index for cross country or cross commodity comparability. With a variable mean across countries for a particular sector, if a country gains comparative advantage in that sector, it will be difficult to identify another country which must have lost comparative advantage in the same sector, so as to maintain a constant mean across countries. Hence, comparing countries to determine the extent of comparative advantage or disadvantage of a country in a sector with respect to another country becomes difficult. With a variable mean across sectors for a particular country, if the country gains comparative advantage in some sector, it does not follow that it must have lost comparative advantage in some other sector, in order to maintain a constant mean across sectors. Hence, comparing sectors to determine the extent of comparative advantage or disadvantage of the country in one sector compared to another becomes difficult.

⁸ At 2 digit level of commodity aggregation, only in 4 out of 13 cases, the hypothesis of normality of error term was rejected. However, at 3 digit level of commodity aggregation, only 3 passed the normality test and at 4 digit level, none passed the test.

⁹ According to Benedictis and Tamberi (2001) the index only shows the proportional difference in a country's sectoral quota with respect to the world.

It is to be further noted, being a monotonic transformation of the Balassa index, the RSCA index tends to reveal comparative advantage or disadvantage in the same sectors as the Balassa index. Further it tends to preserve the same sectoral ordering with respect to a country or same country ordering with respect to a sector. But whether the performance of the RSCA index is determined by its biasness to the size of a sector or a country or by the inherent comparative advantage or disadvantage of a country, can only be revealed on studying its consistency with the comparative advantage theory.

To overcome the difficulties associated with the distribution of the Balassa index, Hoen and Oosterhaven (2006) suggested a modification of the index so as to ensure a symmetric stable distribution across commodities for countries. They replaced the ratio form of the index with the deviation form. Their index takes the following form:

$$ARCA = \frac{x_a^i}{x_t^i} - \frac{x_a^w}{x_t^w} \quad (2.9)$$

The authors term it as the Additive RCA (ARCA) index. They insisted that country i should be deducted from reference group of countries, in which case the lower and the upper limits of the distribution of the index will be exactly equal to -1 and +1 respectively. Zero will be the comparative advantage neutral point. However, in that case, the reference group will be changing from country to country, posing difficulty for inter-country comparison, as pointed out earlier. Hence, with entire world as the reference group, the lower and the upper limits of the distribution will tend towards -1 and +1 respectively. If the index takes a value greater (less) than zero then country i will reveal comparative advantage (disadvantage) in product a . An approximate value of zero denotes comparative advantage neutral point. The limits of the distribution of the index are derived in A.1.3 of Appendix A.

As evident from the structure of the index, the arithmetic mean of its distribution across sectors is equal to zero. The mean is stable irrespective of the number of countries with respect to which the comparative advantage of a country is analyzed and irrespective of the sectoral classification adopted. The mean of the distribution of the index is derived in A.2.3 of Appendix A. Using the export data for Netherlands and Poland for the year 1997 Hoen and Oosterhaven exhibit that ARCA has a much more stable distribution than Balassa's index for different sectoral classifications. The standard deviation of the distribution

gradually decreased with increased sectoral classification and the median would equal zero or the estimated mean value of the index. However, Yu *et al.* (2009) point out that while the sum of the Additive RCA scores of all commodities for a particular country is constant and equals zero, the sum of the Additive RCA scores for a particular commodity across all countries is not constant. This in effect implies that the mean value of a commodity's Additive RCA scores is not characterized by a stable value of zero. As a result, while it is possible to compare a country's comparative advantages in different commodities, application of the index for cross country comparison with respect to a sector is doubtful.

The ARCA index being a monotonic transformation of the Balassa index with respect to a sector, behaves similar to the Balassa index in identifying the comparative advantage or disadvantage of a country in different sectors, in ranking countries with respect to a sector but not in ranking sectors with respect to a country. However, as mentioned in the case of RSCA index, only a study on the consistency of the ARCA index with the theories on comparative advantage will be able to establish whether the index is influenced by the sectoral size or by the underlying comparative advantage of countries.

In an attempt to overcome many of the deficiencies of Balassa's index and other indices, Yu *et al.* (2009) by utilizing the probabilistic framework of Kunimoto (1977), came up with a new index for measuring comparative advantage. In this framework, comparative advantage of country i in product a , is measured by deviation of actual exports of a by i , from its expected exports of the same product, in a world of no relative advantages. The expected exports define the comparative advantage neutral level. The deviation of actual exports from expected exports is further normalized by the world total exports. The Normalized Revealed Comparative Advantage (NRCA) index of country i in product a is thus expressed in the following form:

$$NRCA = \frac{x_a^i}{x_t^w} - \frac{x_t^i}{x_t^w} \frac{x_a^w}{x_t^w} \quad (2.10)$$

If the index is positive (negative), then country i reveals comparative advantage (disadvantage) in product a .

Yu *et al.* (2009) argue that their index satisfies most of the desirable properties of the RCA index for empirical analysis. The index is symmetric about zero with the lower and the upper limits of the distribution being $-1/4$ and $+1/4$ respectively. The limits of the distribution are derived in A.1.4 of Appendix A. The sum of a commodity's (country's) normalized RCA scores over all countries (commodities) equals zero i.e., the mean value is constant and stable. The mean of the distribution of the index is derived in A.2.4 of Appendix A. According to the authors, given the stability of the mean and thereby, the distribution of the index, it is possible to compare the indices across countries and commodities. While all the previously discussed indices can be sensitive to the size of the country or the sector or both, the NRCA index can possibly circumvent the problem, due to its deviation form and normalization of the index by total world exports (Deb and Basu, 2011). In fact, Yu *et al.* compared the RCA scores for Balassa's index and their NRCA index in case of Hawaii's agricultural exports to United States. Their calculations confirmed the fact that Balassa's index is sensitive to the size of the sector or share of a product in the export market. The NRCA index in contrast exhibited large scores for products with large share in the export market. They also found comparable RCA scores could be established for the indices of Balassa and NRCA only in 15 out of 60 cases, thus exhibiting the difference in ranking of products by the two indices. The proponents of the index also proclaim its usage for temporal comparability due to the stability of its mean over time.

However, for the index, normality of the distribution is not guaranteed by the symmetry in its distribution. The same is true in case of ARCA index. In fact, Hoen and Oosterhaven (2006) show that the ARCA index does not fit into the specifications of a normal distribution, due to high kurtosis. As a result, just like Balassa's index, some caution must be employed in case of these two indices, when analyzing the results of parametric tests that assume normally distributed variables.

Vollrath (1991) modified the Balassa RCA index by considering its logarithmic transformation. This monotonic transformation of the Balassa index also has the potential to address the asymmetrical distribution of Balassa's index. However, Vollrath defined the rest of the world and the rest of the commodities exported as the reference group of countries and commodities respectively. To avoid having the reference group of countries or commodities

change from country to country or commodity to commodity, all countries and all commodities exported will be considered as the reference groups in the thesis. Thus the suggested RCA index, henceforth to be referred to as Log of Balassa index, can be represented in the following form:

$$\ln \left(\frac{X_a^i / X_t^i}{X_a^w / X_t^w} \right) \quad (2.11)$$

The new index measures on a logarithmic scale, the extent to which exports of commodity a in country i 's total exports differs from exports of the same commodity from world's total exports. Hence, unlike the RSCA index which also addresses the problem of asymmetry in Balassa index, the Log of Balassa index can be provided an easier economic interpretation. The index is symmetrically distributed about the value zero, which is the comparative advantage neutral point. The upper and lower limits of the distribution of the index are $+\infty$ and $-\infty$ respectively. The limits of the distribution of the index are derived in A.1.5 of Appendix A. Being a monotonic transformation of the Balassa's index, the current index is identical to the Balassa's index while ranking sectors with respect to a country or ranking countries with respect to a sector. The index also shares the difficulties in comparability of the index values across sectors or across countries due to variations in the mean of the index distribution, which has implications for the stability of the distribution. The mean of the index distribution is derived in A.2.5 of Appendix A. The other deficiency of the index is that it cannot be defined in case the export of a product by a country is zero (Laursen, 1998). Nevertheless, the greatest advantage of the index arises from its logarithmic form, which could ensure that the estimated residuals from any regression with the index as the dependent variable will be normally distributed. Logarithmic transformation by inducing symmetry in distribution of Balassa index helps in overcoming the problems associated with skewed distributions (Dalum *et al.*, 1998). In other words, the t and the F statistics will be more reliable in case of Log of Balassa index. As noted, because Log of Balassa index is simply the logarithmic transformation of Balassa index, it is expected to perform very similar to the Balassa index in identifying sectors or countries with comparative advantage or comparative disadvantage. But as noted in case of RSCA index and ARCA index, whether much of its performance is due to its country or sector size sensitivity, or due to inherent comparative

advantage or disadvantage of a country in a product, can only be determined on analyzing the consistency of the index with the theories on comparative advantage.

The RCA indices mentioned so far share one characteristic – they all use gross exports as the primary variable in constructing the index.¹⁰ In the current international economy, a country quite often exports and imports the same product. Hence, indices based solely on export values may fail to reflect the true comparative advantage of countries in the presence of large amounts of intra-industry trade and re-exports. Vollrath (1991) suggests that indices which incorporate both demand and supply effects are more in line with the idea of comparative advantage as recognized in the theory, than any other index. This inevitably calls for an examination of RCA indices based both on exports and imports. But it is to be kept in mind that incorporation of imports for the determination of comparative advantage comes with some costs. Balassa (1977) argued that data on imports could be highly distorted due to the incidences of subsidies, quotas and other import restrictions.¹¹ As a result RCA indices based on import data might not reflect the true comparative advantage of countries. Moreover, following the argument by Brasili *et al.* (2000), any measure based on exports and imports

¹⁰ As mentioned in chapter 1, another index of revealed comparative advantage based on gross export figures, have been suggested in the literature – the index by Proudman and Redding (1998). The new index suggested by Proudman and Redding normalizes the original Balassa index by its cross sectoral mean in an effort to stabilize the mean value of the index across time. This would help to establish the comparability of the index values across sectors within an individual country (Yu *et al.*, 2009). But in their proposed index, they do not address one of the chief concerns of the present research – the symmetry of the distribution of the index and hence do not seem to be a substantial improvement over the Balassa's index. For that reason the RCA index by Proudman and Redding has not been separately considered for analysis in this chapter. They do however make the Balassa's index free from country-size effect by the way of normalizing Balassa's index with its arithmetic mean (over sectors).

¹¹ Balassa (1965) considered a relative export-import measure of comparative advantage, defined in the following form:

$$\frac{X_a^i/M_a^i}{X_m^c/M_m^c}$$

where M refers to imports and other notations are very similar to those stated in expression (2.2). But later Balassa (1977) himself rejected this measure recognizing the fact that country wise imports could be influenced by differing incidences of subsidies, quotas and other import restrictions. Hence, comparative advantage computed on the basis of export-import ratio could generate biased results (as cited in Vollrath (1991)).

might not be a comprehensive measure of specialization of countries, as it compares the performance of a country in exporting a product only with respect to those countries who export that particular product to the considered country. Despite such criticisms associated, it seems worthwhile to analyze some of the RCA indices based on both exports and imports, and study the differences in their performance with RCA indices based solely on exports.

Vollrath (1991) suggested two export and import based indices, one without logarithmic transformation of the variables, which he termed as the Relative Trade Advantage (RTA) index, and the other with logarithmic transformation, termed as Revealed Competitiveness (RC) index. The two indices were expressed in the following respective forms:

$$\frac{X_a^i/X_n^i}{X_a^r/X_n^r} - \frac{M_a^i/M_n^i}{M_a^r/M_n^r} \quad (2.12)$$

$$\ln \frac{X_a^i/X_n^i}{X_a^r/X_n^r} - \ln \frac{M_a^i/M_n^i}{M_a^r/M_n^r} \quad (2.13)$$

Here M stands for imports. Superscript r refers to world minus country i , while subscript n refers to all traded commodities minus commodity a . Country i will have a comparative advantage in product a if either of the above indices take a value greater than zero. It will conversely have a comparative disadvantage in product a if the indices take a value less than zero. Thus, RTA index determines comparative advantage or disadvantage of countries by measuring the difference between export intensity and import intensity with respect to product. RC index adopts the same procedure to determine the comparative advantage or disadvantage of countries, but the exports and imports intensity are now measured on a logarithmic scale.

According to Vollrath, the RCA index (2.13) is extremely sensitive to small values of exports or imports and it also requires a country to be simultaneous exporter and importer of the same product. The index becomes undefined in positive direction in case a country does not import a product, and in case a country does not export a product, the index becomes undefined in negative direction. Hence, the limits of the distribution are undefined. Vollrath argues, RCA index (2.12) might be preferable, at lower levels of commodity aggregation. However, at higher levels of commodity aggregation, index (2.13) may be preferred as it tries to treat supply and demand dimensions of comparative advantage equally (Vollrath, 1991).

It is to be noted however, in Vollrath's original specifications of RTA and RC indices, rest of the world and rest of the commodities exported were considered to be the reference group of countries and commodities respectively. This, as discussed, creates difficulty in using the indices for cross-country or cross-commodity comparison, due to the change in reference group from country to country and from commodity to commodity. For that reason, Vollrath's indices will be modified to include the entire world and all commodities to be the reference group of countries and commodities respectively. Thus, the following versions of the RTA and RC indices will be considered henceforth in this thesis:

$$\frac{X_a^i/X_t^i}{X_a^w/X_t^w} - \frac{M_a^i/M_t^i}{M_a^w/M_t^w} \quad (2.14)$$

$$\ln \frac{X_a^i/X_t^i}{X_a^w/X_t^w} - \ln \frac{M_a^i/M_t^i}{M_a^w/M_t^w} \quad (2.15)$$

Both the indices in conformity with their original versions are symmetric about the value zero. The upper and lower limits of their distributions tend towards $+\infty$ and $-\infty$ respectively. The limits of the distributions of each index are derived in A.1.6 and A.1.7 of Appendix A. The RC index because of its logarithmic form, can have normally distributed estimated residuals from regressions with the index as the dependent variable. Both the indices however might not fare well for cross country or cross commodity comparison due to their unstable mean across countries or across commodities. The means of the distributions for each index are derived in A.2.6 and A.2.7 of Appendix A. But it must be acknowledged that, for both the RTA and RC indices, susceptibility to the size of the sector or of the country will be indeterminate compared to the exports-only indices, because of their relative demand and relative supply dimensions. Whether a small country will have a large magnitude of RCA index will depend upon the relative strength of the demand and supply. Similar arguments will be applicable for size of commodities.

A discussion of the alternative RCA indices thus tends to suggest that it is indeed difficult to select any particular index which is perfect from every respect. While the NRCA index addresses some of the shortcomings of the Balassa's index, it falls short with respect to the feature of a normal distribution. Although RSCA index, Log of Balassa index and RC index do have the potential to generate reliable regression estimates, they suffer from the problem

of unstable distribution due to unstable mean. As discussed, an unstable distribution prevents the indices from being used for reliable cross country or cross commodity comparison. NRCA index however does possess this essential feature and hence arguably can be used for cross country or cross commodity comparison, although it may not be the index which can generate reliable inferential statistics. In fact, arguments in support of the above claim have also been provided by Sanidas and Shin (2010) who make similar observations and argue that none of the indices can be considered perfect.

2.2.3 Derivation of Theoretically Consistent RCA Indices

Although the RCA indices considered in subsection 2.2.2 address one or more short comings of the RCA index of Balassa in empirical analysis, an important issue that has prompted additional research is concern about consistency of RCA indices with the comparative advantage theories. Considerations about the theoretical foundations of the Balassa index prompted Leromain and Orefice (2014) to propose a new RCA index consistent with the Ricardian theory of comparative advantage. Using the econometric framework adopted by Costinot *et al.* (2012) they estimate that part of export flows of a country which is only due to the intrinsic productivity level in a given industry of that country. While doing so, they controlled for all other determinants of export performance of a country through importer-exporter fixed effects and importer-industry fixed effects.¹² Thus assuming a Ricardian model with one factor of production (labor) and k industries in a perfectly competitive setting, they considered exports of k from country i to country j to be determined in the following manner:

$$\ln x_{i,j,k} = \delta_{i,j} + \delta_{j,k} + \theta \ln(z_{i,k}) + \varepsilon_{i,j,k} \quad (2.16)$$

$\delta_{i,j}$ captures country pair fixed effects, $\delta_{j,k}$ captures importer-industry fixed effects, and $\theta \ln(z_{i,k})$ represents exporter-industry specific technological differences. $z_{i,k}$ is the intrinsic or fundamental productivity level of country i in industry k and hence it is exporter-industry

¹² Importer-exporter fixed effects as defined by Leromain and Orefice incorporate physical distance between the countries, prevalence of colonial ties, common language etc. Importer-industry fixed effects incorporate variations in policy barriers, demand shocks, bias in the tastes of consumers, etc.

specific. It represents the deterministic component of labor productivity. θ assumed to be constant across all countries and industries, is a measure of intra-industry productivity heterogeneity. It represents the distinctive differences in technological know-how across varieties within any industry. It is the random component of labor productivity, drawn from countries, industries and varieties (Costinot *et al.*, 2012). $z_{i,k}$ captures the factors reflecting the cross-country variations in productivity such as climate, infrastructure and institutions, and is of primary interest to construct an index based on the Ricardian comparative advantage. The authors estimate $z_{i,k}$ which is assumed to be unknown, by estimating technological differences through exporter-industry fixed effect, $\delta_{i,k}$ from the following regression equation:

$$\ln x_{i,j,k} = \delta_{i,j} + \delta_{j,k} + \delta_{i,k} + \varepsilon_{i,j,k} \quad (2.17)$$

Where $\delta_{i,k} = \theta \ln(z_{i,k})$. Hence the estimates of $z_{i,k}$ could be obtained by calculating $e^{\delta_{i,k}/\theta}$.¹³

The authors then use estimated $z_{i,k}$ to construct an RCA index which is very similar to Balassa's ratio form of RCA index. The only difference between the two indices is that the export values are now adjusted to reflect mainly the part of exports accounted by the intrinsic productivity levels of each country in each sector. Because of the presence of importer-exporter fixed effects and importer-industry fixed effects in the regression equation (2.17), the estimated $z_{i,k}$ has been cleaned from all other determinants of export performance. Thus, the index proposed by Leromain and Orefice and derived from the proposed index by Costinot *et al.* (2012), can be represented in the following form.¹⁴

¹³ Costinot *et al.* (2012) estimated a value of 6.53 for θ , and Leromain and Orefice made use of that value in their paper.

¹⁴ Costinot *et al.* (2012) referred to the estimated $\frac{z_{i,k}/z_{i,k'}}{z_{i',k}/z_{i',k'}}$ as the relative revealed measure of productivity and

claimed that it could be used as an alternative to Balassa's RCA index. In their formulation, the comparison of revealed productivity is between two countries i and i' and two commodities k and k' . The authors claim that contrary to Balassa's index, their relative revealed measure of productivity has substantial theoretical foundation, as evident from the manner in which $z_{i,k}$ is estimated. However, as argued by Leromain and Orefice

$$\frac{z_{i,k}z}{z_i z_k} \quad (2.18)$$

z is the average of $z_{i,k}$ coefficients across all countries and all sectors, z_i is the average of $z_{i,k}$ coefficients across all sectors for a particular country, and z_k is the average of $z_{i,k}$ coefficients across all countries for a particular sector. A value greater than unity for the above index implies that, country i has a higher productivity level in sector k compared to the expected ($z_i z_k$), given the worldwide average level of productivity (z). In the process of calculating the index the authors used the export data on various product groups for about 20 exporting and 76 importing countries for the years 1995 to 2010. Eventually, Leromain and Orefice demonstrate the superiority of the empirical distribution of their RCA index compared to Balassa's RCA index. The distribution of their index has a lower dispersion and skewness, better symmetry about unity and greater stability across time compared to Balassa's index. Their RCA measure also turned out to possess better ordinal ranking properties than Balassa's measure. Their methodology is significantly different from those adopted in case of previously mentioned RCA indices and most importantly their suggested index incorporates the ideals of Ricardian comparative advantage in its structure. However, their adopted methodology suffers from one particular drawback and realization of that drawback casts doubt on the reliable usage of this RCA index. They assume a time invariant θ , which implies that the idiosyncratic differences in the technological know-how across varieties within any industry do not change over time. The authors themselves recognize that by making the above assumption, importance of big technological shocks that industries may experience over time is assumed away. If technological shocks do influence θ upward, then the computed value of fundamental productivity, $z_{i,k}$, will be low compared to the case where θ remains constant over time. This follows from the fact that in the expression for calculation of $z_{i,k}$, θ enters exponentially. Hence, RCA index values computed using the values on

(2014) the measure is dependent upon country-sector benchmark (i' and k') with respect to which the productivity of country i in sector k is analyzed. Although this methodology adheres to the strict two countries and two commodities framework of Ricardo, it does not yield in understanding the relative performance of any country among many. Comparing the performance of a country with respect to that of a reference group of countries helps in determining the position of the considered country among many, which is more relevant for policy purposes. In this sense, perhaps the approach adopted by Leromain and Orefice (2014), where they incorporate a group of countries as the reference group or the benchmark group, deserves better credibility.

fundamental productivity must be analyzed with some caution during the event of technological shocks. Further, although the index of Leromain and Orefice has better empirical distribution compared to the index of Balassa, it may still be more asymmetrically distributed than many symmetrically distributed indices identified in this chapter. In this respect, the index may be less appropriate in many exercises for which symmetrical distribution of the index is necessary.

2.3 Growing Importance of Global Production Chain and the RCA Indices

The previous section documents the comparative advantage of countries based on gross exports or imports values. The RCA indices mentioned so far take into account the relative gross exports or imports of a country to determine whether that particular country has comparative advantage or disadvantage in a product. However, in the event of changing nature of trade flows throughout the world over the last four to five decades, not only in terms of quantity but also quality, raises concern over the relevance of trade indices based only on gross exports or imports. Hummels *et al.* (1998) consider the internationalization of production to be the primary reason for this change in trade flows. As Feenstra (1998) notes, trade in intermediate goods, such as parts and components, now constitute the largest share of world trade and very often final consumer goods sold in one country are made up of components processed in one or more countries. Meng *et al.* (2012) tried to map the evolution of trade in goods by different end use categories over the period 1995 to 2010.¹⁵ Their analysis reflected the fact that trade in intermediate goods are the main drivers of growth in world goods trade, followed by household consumption and capital goods.¹⁶ The phenomenon of higher volume of intermediate goods being produced in subsequent stages in

¹⁵ End use categorization classifies goods as intermediate goods, final consumption goods and capital goods.

¹⁶ The recent change in global merchandise trade can be explained by several factors. First, higher rate of reduction in tariffs and other trade costs on intermediate products compared to final products over the last 20 years, have contributed to greater movement of parts and components all over the world (Miroudat, 2009). Second, the growth of foreign direct investment (FDI) boosted the trade in intermediate products (Miroudot, 2009). Third, induced by the first and second factors, intra-firm trade has increased and has contributed to increased intermediate goods trade (Yi, 2003). Finally, domestic market oriented reforms by countries like China has enabled other countries and multinational enterprises to involve China in their global production network (Meng *et al.*, 2012).

different countries and then exported to other countries for further processing, typically describes the situation of global value chain which has alternatively been termed as “vertical specialization”, or “fragmentation of production” or “outsourcing” or “global supply chains” (Meng *et al.*, 2012, p. 2). Hummels *et al.* (2001) in their paper, provide a clear understanding of the term “vertical specialization”. They note that in order to define a process as “vertical specialization”, the good must be produced by atleast two countries with value added by each country during its successive production stages. Imported inputs must be used by atleast one of the countries participating in the production network and some of its output generated must be exported.

The authors also provide a diagrammatic exposition of the phenomenon of vertical specialization chain involving three countries.

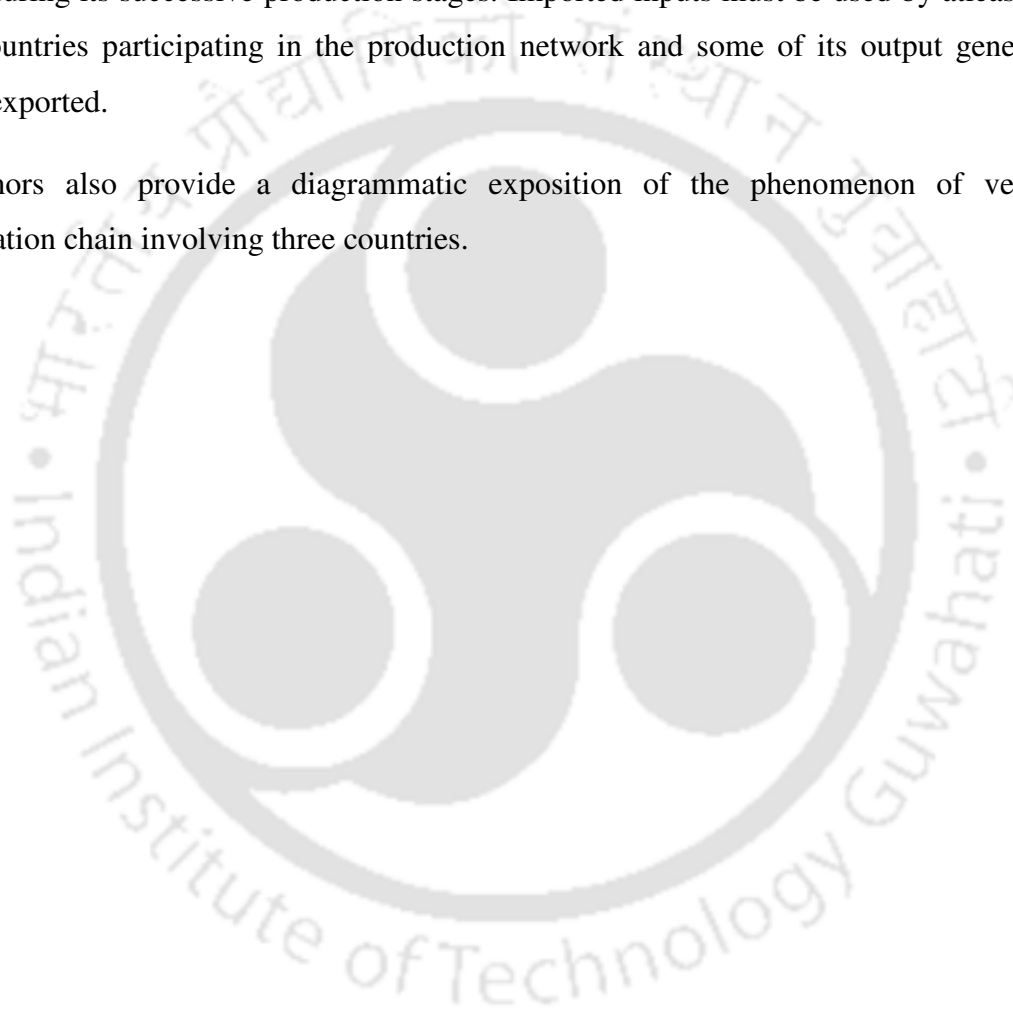
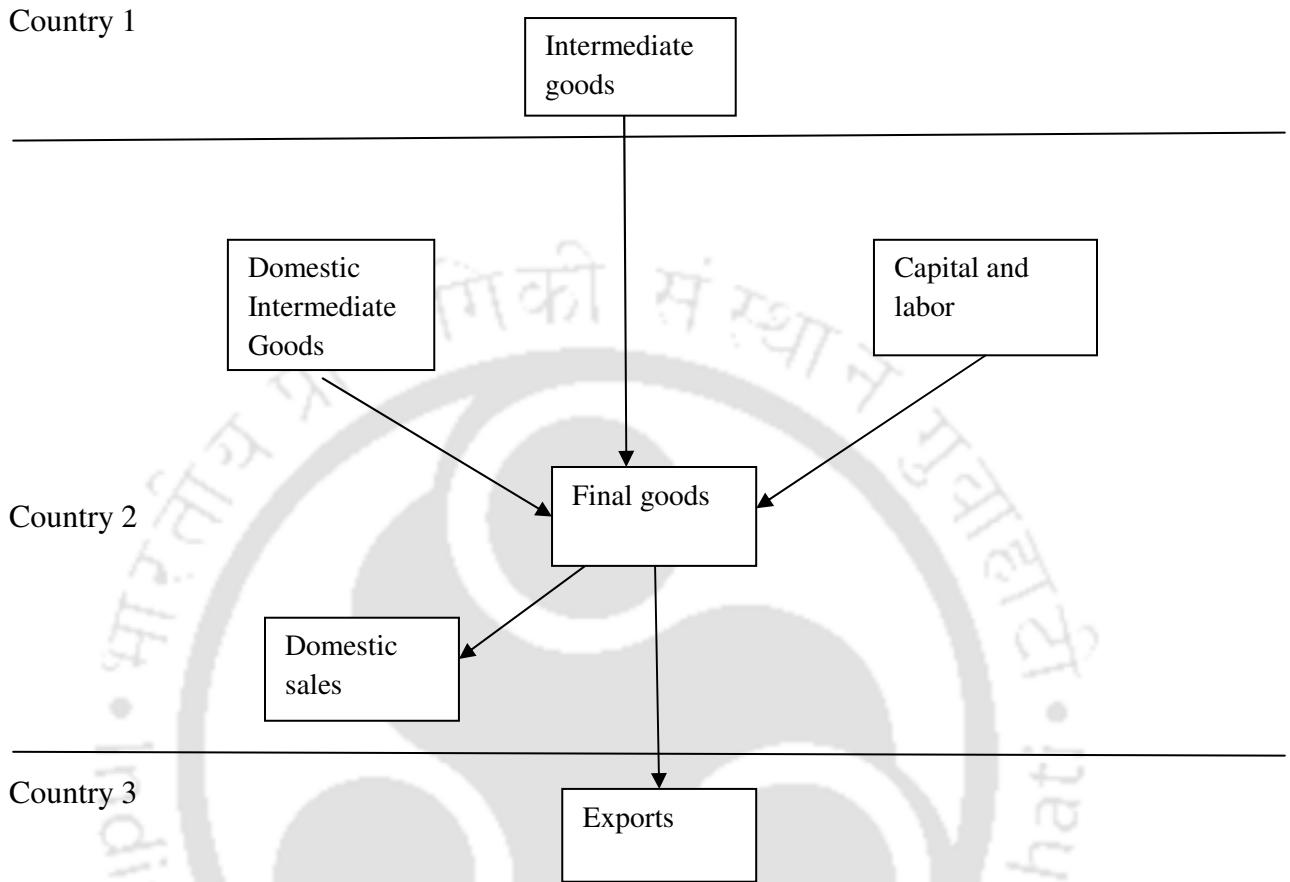


Figure 2.1: Vertical Specialization



Source: Hummels *et al.* (2001, p.78)

In the above diagram, country 1 manufactures intermediate goods and exports them to country 2. Country 2 again combines these imported intermediate goods with domestic intermediate goods and other factor inputs like labor and capital to produce a final good. Some of these final goods are for domestic consumption while the rest are exported to country 3.

According to Jones and Kierzkowski (2001) the phenomenon of production fragmentation is similar to creative destruction in Schumpeterian tradition. Gains from specialization can be more effectively exploited if production process is disintegrated and distributed among countries with differing technologies and factor prices. Thus, underlying comparative advantage of countries is a determining factor behind the production fragmentation process.

The authors also note that technical progress in services sector is significantly responsible for the growing success of global production networks. Lower transport costs and improved telecommunication services have facilitated internationally connected production blocks.

It is however to be noted, global production chains do not only incorporate the activities within multinational firms but also activities by other firms who are not directly associated with a multinational company. Such firms are usually under some contractual agreement with the multinational firm. Thus, one needs to look beyond the activities of a multinational firm. To define global production networks, one need to consider all imported intermediates used in the production of the multinational firm, and final goods sold under the brand name of the company, in addition to the imports directly by the firm itself (Feenstra, 1998).¹⁷

The realization of the phenomenon of global value chains has made gross export or imports as incomplete measures of RCA. There is a growing belief that measures based on value-added trade will be better able to reflect the potential of a country in exporting a particular product. Thus, several literature have developed measures of comparative advantage based on value-added to exports figures, arguing that the comparative advantage exhibited by such measures differ widely than the typical RCA measures based on gross export figures (Koopman *et al.*, 2014, Meng *et al.*, 2012).

2.3.1 Vertical Specialization and Value-Added Trade: Implications for RCA Indices

In this subsection, a brief review of the literature on vertical specialization and corresponding value-added trade, and hence the latter's contribution in redefining the RCA indices is provided.

Hummels *et al.* (2001) provided the first empirical measurement of a country's participation in vertically specialized trade. They define vertical specialization of country k and sector i as,

¹⁷ Certain studies have tried to determine the effects of increased activities of multinational firms on the domestic factors such as wages and employment. Lawrence (1994) found the changes in imports by US multinational firms are too meager to influence changes in domestic wage and employment. In another study by Krugman (1995), FDI flows through multinational firms were found to be minimal to effect employment and wage changes. Thus focusing on the multinational firms alone may not provide the true impact of outsourcing.

$$VS_{ki} = \left(\frac{\text{imported intermediates}}{\text{gross output of } i} \right) \times (\text{Exports of } i \text{ by } k) \quad (2.19)$$

Thus VS or vertical specialization is simply foreign value-added or imported input content in exports. Imported input content in exports is generated on multiplying the share of imported input in gross output by the exported amount. The authors allow for the fact that imported inputs can be used not only directly, but also indirectly in the production of an export good. Thus the production of a domestic input which is directly used to produce exports, might require imported inputs. Further, a domestic input directly used to produce exports might require another domestic input, which necessitates imported inputs in its production. Similar other stages can be identified where the imported input is embodied in exports at the fourth, fifth or sixth stage. Considering two sectors 1 and 2, the direct and indirect imports embodied in exports of country k can be presented in terms of input requirement matrices:

$$(1 \quad 1) \left[\begin{aligned} & \left(\begin{pmatrix} a_{11}^M & a_{12}^M \\ a_{21}^M & a_{22}^M \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a_{11}^M & a_{12}^M \\ a_{21}^M & a_{22}^M \end{pmatrix} \begin{pmatrix} a_{11}^D & a_{12}^D \\ a_{21}^D & a_{22}^D \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) \\ & + \left(\begin{pmatrix} a_{11}^M & a_{12}^M \\ a_{21}^M & a_{22}^M \end{pmatrix} \begin{pmatrix} a_{11}^D & a_{12}^D \\ a_{21}^D & a_{22}^D \end{pmatrix} \begin{pmatrix} a_{11}^D & a_{12}^D \\ a_{21}^D & a_{22}^D \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \dots \right) \end{aligned} \right] \quad (2.20)$$

a_{ij}^M represents imported input from sector i used to produce one unit of sector j 's output. Hence the (2×2) matrix containing a_{ij}^M , $i, j = 1, 2$, denote the imported input-output coefficient matrix. a_{ij}^D represents domestic input from sector i used to produce one unit of sector j 's output. Hence the (2×2) matrix containing a_{ij}^D , $i, j = 1, 2$, denote the domestic input-output coefficient matrix. X_1 and X_2 are the amount of exports from sector 1 and 2 respectively. We proceed by first interpreting the terms within third bracket of the above expression. In the first term, each element of the (2×1) vector obtained by multiplying the vector of exports by the rows of the imported coefficient matrix, gives imported input from each sector used to produce exports. The second term represents imported inputs indirectly used to produce exports. Thus multiplying the vector of exports by rows of the domestic coefficient matrix gives domestic input from each sector used to produce exports. Multiplying the resulting vector by imported coefficient matrix gives imported inputs from each sector needed to produce required domestic inputs. The third term also represents the

imported inputs used indirectly to produce exports, when imported inputs are embodied in domestic inputs required to produce those domestic inputs employed in generating exports. Infinite such possibilities can be possible. Multiplying the sum of the resulting vectors within third bracket by a unity vector of the order (1×2) , generates the total imports used as inputs in producing exports by country k or vertical specialization of country k . That total imports divided by the total exports by country k or the sum of X_1 and X_2 , is defined to be the VS share of total exports of country k by Hummels *et al.*

Thus generalizing expression (2.20) in the context of N sectors, the vertical specialization by country k will be:

$$\begin{aligned} & u(A^M X + A^M A^D X + A^M A^D A^D X + A^M A^D A^D A^D X + \dots) \\ & = uA^M(I - A^D)^{-1}X \end{aligned} \quad (2.21)$$

u is the $(1 \times N)$ vector of 1s, A^M is the $(N \times N)$ imported coefficient matrix, I is the identity matrix, A^D is $(N \times N)$ domestic coefficient matrix, X is the $(N \times 1)$ vector of exports.

$(I - A^D)^{-1}$ captures the domestic output requirement to produce exports.

The VS share of total exports country k will thus be represented as:

$$\frac{VS_k}{X_k} = \frac{uA^M(I - A^D)^{-1}X}{X_k} \quad (2.22)$$

X_k is the total exports of country k .

The methodology of Hummels *et al.* has been applied by the National Research Council (2006) to determine the foreign content in the exports of United States. However, Koopman *et al.* (2010, 2014) questioned the validity of two basic assumptions made by Hummels *et al.* which limits the applicability of their measure in the real world. These two assumptions required for their measure to accurately reflect foreign content in exports include:

- 1) Imported inputs applied for the manufacture of export goods and goods meant for domestic final consumption, must be used in similar intensity. However, in the presence of export processing zones, it will be quite logical to expect that the

applicable technology will differ from that applied in case of goods meant for domestic final demand.

- 2) All imports have 100% foreign content. This assumption will be violated if several countries exports intermediate goods. In that case, a country can import intermediate goods, add value and export them to be processed further by the country from which it had previously imported the intermediate goods. Or, a country can import intermediate goods which contain value that had been previously added domestically at an earlier stage in the production process.

To address the first assumption Koopman *et al.* (2012) compute the shares of foreign and domestic content in a country's exports in the presence of processing exports, and establish Hummels *et al.*'s measure as a special case of their advocated general measure. They treat processing exports of a country as a category distinct from normal exports, and allow for differences in the use of imported and domestic intermediates for processing exports, goods meant for normal exports and goods meant for final domestic sales. Assuming N sectors, they derive the foreign content share in a country's total exports as:

$$TVSS = uA^{MD}(I - A^{DD})^{-1} \frac{E - E^P}{te} + uA^{MD}(I - A^{DD})^{-1}A^{DP} \frac{E^P}{te} + uA^{MP} \frac{E^P}{te} \quad (2.23)$$

A^{MD} is the $(N \times N)$ imported input-output coefficient matrix, with its each element a_{ij}^{MD} giving the amount of imported inputs from sector i used to produce one unit of sector j 's output domestically. A^{DD} is the $(N \times N)$ domestic input-output coefficient matrix when domestic inputs are used to produce goods meant for domestic consumption or normal exports. A^{MP} and A^{DP} are the $(N \times N)$ input-output coefficient matrices when imported and domestic inputs respectively are employed in processing exports. u is a $(1 \times N)$ unity vector. E and E^P are the $(N \times 1)$ matrices of total exports and processing exports respectively. The difference between the two vectors gives the vector of normal exports. te is a scalar indicating the country's total exports. The first term on the right hand side of the expression (2.23) represents the import content (both direct and indirect) in normal exports of the country. The second term represents indirect import content in processing exports and the

third term represents direct import content in processing exports. Koopman *et al.* show that if $A^{DD} = A^{DP}$ and $A^{MD} = A^{MP}$ i.e., input-output coefficient matrices do not differ in cases of normal exports and processing exports, or if processing exports do not exist, i.e., $E^P = 0$, then the expression (2.23) reduces to:

$$TVSS = uA^{MD}(I - A^{DD})^{-1} \frac{E}{te} \quad (2.24)$$

which is equivalent to Hummels *et al.*'s formula for vertical specialization as a share of exports. The authors after constructing relevant input-output tables, used their new formula to estimate the share of foreign value-added in the manufactures export of China in 1997 and 2002 and found that to be about 50%, which is almost double the amount implied by the measure of Hummels *et al.*

Daudin *et al.* (2011) tried to overcome the limitations of the second assumption by computing not only the import content of exports (VS as per Hummels *et al.*'s definition) but also the share of exports used as inputs to produce further exports (which Hummels *et al.* define as VS1 but refrain from its computation). The authors also defined VS1* as that part of VS1, which comes back to the country of origin after being processed abroad. The authors define vertical trade of a country as being twice the sum of VS and VS1* and total value-added trade as being total standard trade minus the sum. The authors found through their calculations, that the patterns of inter country value-added trade in various sectors is quite different from that of standard trade. For the purpose they use GTAP (Global Trade Analysis Project) versions 5, 6 and 7 which provides data on final demand and input-output tables for each region.

Johnson and Noguera (2012) also made an effective attempt at addressing the second assumption of Hummels *et al.* They formulate the value-added export ratios or the VAX ratios to compute and examine the value-added content of trade at the bilateral level, using input-output and bilateral trade data. The authors in their analysis allow for the fact that a country's imports are not 100% foreign sourced. With a two country framework they exhibit their VAX ratio to represent domestic content in exports due to Hummels *et al.* (2001) only in case, a country does not export intermediates to another country to be exported back to it in processed form. This implies the country only exports finished products to the other

country by using imported intermediates from the latter, as necessitated by Hummels *et al.*'s vertical specialization measure. They also compute the bilateral VAX ratios in a three country framework carefully distinguishing between the cases where country 2 imports intermediates from country 1 to be processed and exported as final products to 1, and the case where a country 3 exports intermediates to country 2 to be processed and exported as finished products to country 1. Thus the authors allow for the possibilities where a country imports final products with domestic contents in them, and a country exports intermediate products to a country to be processed and exported as final product to a third country. In their description the authors thus establish a situation where a country trades with another country indirectly or in value-added terms, but not directly. The authors using data from GTAP database version 7.1, found through their calculations US exports to Canada in value-added terms to be 40% smaller and US-China deficit to be 30-40% smaller compared to corresponding measures based on gross exports.

All of the literature mentioned so far recognized the importance of value-added trade as compared to gross trade in the rapidly changing world trade scenario, and demonstrated the deviations of value-added exports from gross exports. As already pointed out, Koopman *et al.* (2010, 2014), and Meng *et al.* (2012) recomputed the RCA indices after taking into account the importance of value-added trade. Koopman *et al.* (2010, 2014) decomposed each country's gross exports into separate components of value-added trade. In the process, their constructed framework incorporated almost all the measures of value-added and vertical trade suggested in the literature. In order to arrive at their decomposition, they assume the gross output of every country must be absorbed as intermediate or final good by the country itself or by its trading partners. Hence assuming two countries r and s and two sectors 1 and 2, the framework adopted by Koopman *et al.* (2010) can be expressed as:

$$X_r = A_{rr}X_r + A_{rs}X_s + Y_{rr} + Y_{rs}$$

and

$$X_s = A_{sr}X_r + A_{ss}X_s + Y_{sr} + Y_{ss}$$

X_r and X_s are the (2×1) gross output vector of countries r and s respectively. A_{rs} is the (2×2) matrix of input-output coefficients with each element a_{ij} representing inputs from

sector i of country r being used to produce one unit of sector j 's output in country s . It follows A_{rr} , A_{sr} and A_{ss} being (2×2) input-output coefficient matrices, can be similarly defined. Y_{rs} is the (2×1) vector of final demand in country s for the goods of country r . Y_{rr} , Y_{sr} and Y_{ss} being (2×1) vector of final demands can be similarly defined. Hence in block matrix form the previous two equations can be represented as:

$$\begin{bmatrix} X_r \\ X_s \end{bmatrix} = \begin{bmatrix} A_{rr} & A_{rs} \\ A_{sr} & A_{ss} \end{bmatrix} \begin{bmatrix} X_r \\ X_s \end{bmatrix} + \begin{bmatrix} Y_{rr} + Y_{rs} \\ Y_{sr} + Y_{ss} \end{bmatrix}$$

which implies,

$$\begin{bmatrix} X_r \\ X_s \end{bmatrix} = \begin{bmatrix} I - A_{rr} & -A_{rs} \\ -A_{sr} & I - A_{ss} \end{bmatrix}^{-1} \begin{bmatrix} Y_r \\ Y_s \end{bmatrix}$$

which implies,

$$\begin{bmatrix} X_r \\ X_s \end{bmatrix} = \begin{bmatrix} B_{rr} & B_{rs} \\ B_{sr} & B_{ss} \end{bmatrix} \begin{bmatrix} Y_r \\ Y_s \end{bmatrix}$$

Y_r and Y_s as defined by Koopman *et al.* (2010) are the (2×1) vectors of total final demand for the goods of country r and s respectively. B_{rr} is a (2×2) matrix giving the amount of gross output of country r required to generate a unit increase in final demand for the goods of country r . B_{rs} , B_{sr} and B_{ss} can be similarly defined.¹⁸ To determine the domestic value-added in exports, a (1×2) vector of domestic value-added by country r in each sector, V_r can be introduced. Thus, each element of V_r represents the share of direct domestic value-added in total output of each sector (Koopman *et al.*, 2010). It follows $V_r B_{rr}$ is a (1×2) vector with each element giving the domestic value-added by country r in producing one more unit of each of the domestic products. Defining V_s to be a (1×2) vector of domestic value-added coefficients of country s , $V_s B_{sr}$ denotes the (1×2) matrix of domestic value-added by country s in producing one more unit of each of the domestic products to be absorbed in country r , namely, foreign value-added for country r . The authors assume that since all value-added must either be domestic or foreign, the following relation must hold:

$$V_r B_{rr} + V_s B_{sr} = u \quad (2.25)$$

¹⁸ B_{rr} , B_{rs} , B_{sr} and B_{ss} can be solved by applying the algebra of inverse of a partitioned matrix.

u is the (1×2) unity vector.

Hence, from expression (2.25) it follows, total gross exports from country r to country s is the sum of domestic value-added in exports and foreign value-added in exports:

$$V_r B_{rr} E_{rs} + V_s B_{sr} E_{rs} = u E_{rs} \quad (2.26)$$

Where E_{rs} is a (2×1) vector of exports from country r to country s . Rearranging and collecting terms, the left hand side of equation (2.26) can be broken down into various value-added and double counted components so that the gross exports of country r to country s can be expressed as:

$$u E_{rs} = V_r B_{rr} Y_{rs} + V_r B_{rr} A_{rs} X_{ss} + V_r B_{rr} A_{rs} X_{st} + V_r B_{rr} A_{rs} X_{sr} + V_s B_{sr} E_{rs} \quad (2.27)$$

The first term on the right hand side gives domestic value-added by country r in the exports of final goods to country s . Second term on the right hand side gives domestic value-added by country r in the exports of intermediate inputs to country s to be used by latter to produce domestically needed products. $A_{rs} X_{ss}$ gives the total intermediate goods requirement from country r to produce the gross output of country s and absorbed in country s . Hence $V_r B_{rr} A_{rs} X_{ss}$ gives domestic value-added by country r in producing those intermediates. The third term on the left hand side gives domestic value-added by country r in the exports of intermediate inputs to country s to be used by the latter to produce goods for a third country represented by t . The fourth terms gives domestic value-added by country r in the exports of intermediate inputs to country s to be used by latter to produce goods for country r . As explained, the last term is the foreign value-added in exports of country r .¹⁹ The detailed steps on arriving from expression (2.26) to (2.27) are provided in the Appendix A.3.

Therefore, considering N sectors and many countries, the gross exports of country r , E_{r*} (i.e., $u \sum_s E_{rs}$, where u and E_{rs} are now $(1 \times N)$ and $(N \times 1)$ vectors respectively) can be represented by the following expression:

¹⁹ As common in input-output literature, the authors continue to assume that the value-added shares in a product are same for any country irrespective of whether the product is used as an intermediate good or as a final good.

$$\begin{aligned}
E_{r*} = & V_r B_{rr} \sum_{s \neq r} Y_{rs} + V_r B_{rr} \sum_{s \neq r} A_{rs} X_{ss} + V_r B_{rr} \sum_{s \neq r} \sum_{t \neq r, s} A_{rs} X_{st} \\
& + V_r B_{rr} \sum_{s \neq r} A_{rs} X_{sr} + \sum_{s \neq r} V_s B_{sr} E_{r*}
\end{aligned}
\tag{2.28}$$

The vector of exports and gross output are now of the orders $(N \times 1)$, A_{rs} is the $(N \times N)$ input-output coefficient matrix and $V_r B_{rr}$ and $V_s B_{sr}$ are the $(1 \times N)$ vector of domestic value-added and foreign value-added in one additional unit of each of the N sectors' output respectively.

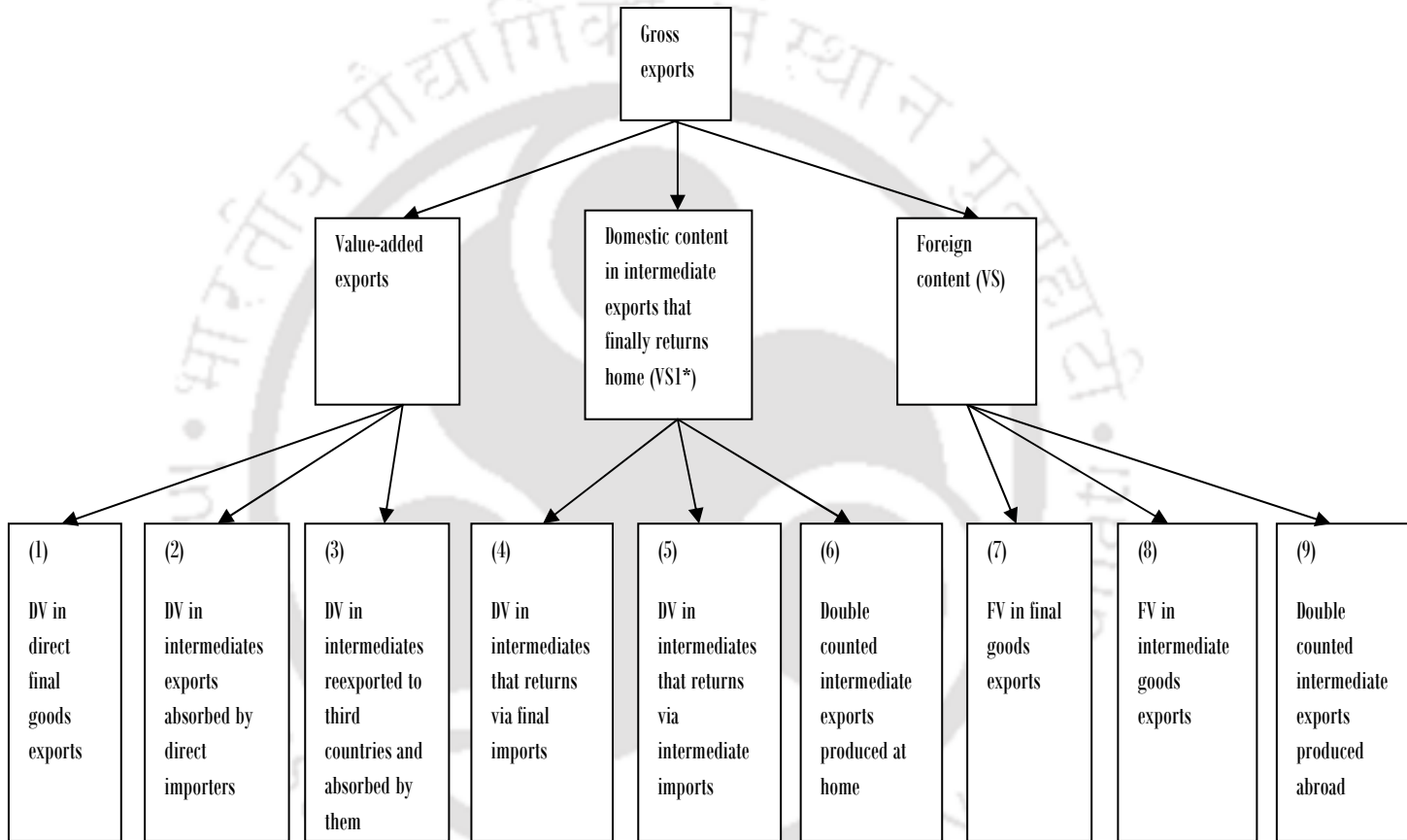
Koopman *et al.* (2010) consider the sum of first three terms on right hand side of equation (2.28) as country r 's value-added exports to the world, while the sum of first four terms as domestic content in country r 's gross exports. Further, the sum of third and fourth term constitutes the VS1 in gross exports, as defined by Hummels *et al.* (2001). A portion of the fourth term, which constitutes domestic value-added in intermediates that return home via final goods imports, represents VS1* computed by Daudin *et al.* (2011).²⁰ Sum of the first three terms divided by gross exports represent Johnson and Noguera's VAX ratio. The authors further define the sum of fourth and fifth terms to be the double counted portion of trade in official trade statistics, as they incorporate movement of products that cross national borders at least twice. The fifth term actually represents the foreign value-added in the country r 's gross exports. However, Koopman *et al.* allows for adjustment for domestic content returned home while calculating the foreign value-added in gross exports. Hence, they incorporate the possibility for unrestricted intermediate goods trade among countries, or waiving the assumption that country r 's imports have 100% foreign content. It is in this respect, Koopman *et al.*'s formulation differs from the formulation adopted by Hummels *et*

²⁰ The definition of VS1* by Daudin *et al.* (2011) differs from the definition adopted by Koopman *et al.* (2010, 2014). While Daudin *et al.* takes into account the import of only final goods with domestic value-added, Koopman *et al.* also considers intermediate goods imports in the calculation of VS1* (Koopman *et al.*, 2014). In this sense Daudin *et al.*'s specification is a subset of Koopman *et al.*'s specification. Although Koopman *et al.* (2010) are not so specific about this difference, Koopman *et al.* (2014) make the distinction clear.

al. (2001) as the latter does not allow for more than one country exporting intermediate goods.

Thus, Koopman *et al.* (2014) provided a detailed decomposition of each value-added and double counted component in expression (2.28). The authors provide a complete description of gross exports accounting through the following box diagram:

Figure 2.2: Gross Exports Accounting



Source: Koopman *et al.* (2014, p. 482).

DV in the above boxes stand for domestic value-added and FV represents foreign value-added. As explained by Koopman *et al.* (2014), value-added exports by a country equal (1) + (2) + (3). Domestic content in a country's exports is represented by (1) + (2) + (3) + (4) + (5) + (6). The authors offer a distinction between the country's domestic value-added in exports and value-added exports in aggregate and acknowledge that the two measures would be equal in case imports do not contain any domestic value-added. (7) + (8) + (9) is the vertical

specialization or VS. (3) + (4) + (5) + (6) is part of VS1. (4) is the VS1* as computed by Daudin *et al.* (2011). (4) to (9) are sources of multiple counting in official trade statistics.

Koopman *et al.* (2010, 2014) construct a global Inter Country Input Output (ICIO) table based on GTAP database version 7 and trade data from United Nations Commodity Trade Statistics (UN COMTRADE) database, for decomposing gross exports into the individual components as described. Later the authors provide a comparative analysis of Balassa's RCA indices computed on the basis of gross exports and domestic value-added in gross exports for several countries and several sectors. Their analysis revealed contradictory pictures for some sectors, e.g., India has comparative advantage in sectors such as finished metal products and business services on the basis of gross exports data, while comparative disadvantage in the same sectors on the basis of domestic value-added in gross exports data. For each sector, the authors also noticed substantial changes in the country rankings. Tables on their comparative study are presented in the Appendix A.4. However, the validity of country ranking based on Balassa's RCA index continues to remain a questionable issue due to its responsiveness to the size of a country or sector, as discussed in section 2.2.1, and hence should be studied with due concern.

Meng *et al.* (2012) constructed alternative RCA indices based on trade in value added (*TiVA*), after making a distinction between value-added in final goods and value-added in intermediate goods. Assuming two countries r and s , $TiVA^{rs}_j$ signifies country r 's value-added in sector j induced by country s 's total final demand when international input-output table is available. On similar notations, $TiVA^{rs}_i$ represents country r 's value-added in all sectors to fulfill the final demand for product i by country s . The authors define, $TiVA^{rs}$ to be the sum of value-added contained in country r 's exports of final products to country s , and cross border value-added in terms of trade in intermediate goods. Summing $TiVA^{rs}$ over all countries s , $TiVA^r$ can be obtained. Using sector level (j) and product level (i) *TiVA*, the authors define two alternative RCA indices. The first of these RCA indices identify the “comparative advantage of country r 's sector j in global value chain” (p. 9) and is represented as follows:

$$RCA^r_j = \frac{TiVA^r_j / \sum_j TiVA^r_j}{\sum_r TiVA^r_j / \sum_r TiVA^r} \quad (2.29)$$

$TiVA_j^r$ in the expression (2.29) is actually $TiVA_j^{rs}$ summed over all countries.

The second kind of RCA index identifies “country r ’s entire comparative advantage in providing value-added to fulfill other countries’ final demand for product i ” (p. 9) and is represented as follows:

$$RCA_i^r = \frac{TiVA_i^r / \sum_i TiVA_i^r}{\sum_r TiVA_i^r / \sum_r TiVA^r} \quad (2.30)$$

$TiVA_i^r$ in the expression (2.30) is actually $TiVA_i^{rs}$ summed over all countries.

Thus by differentiating final goods from intermediate products, the RCA indices proposed by Meng *et al.* (2012), establish a link between a country’s value-added and other countries’ final demands. The authors used the international input-output tables constructed by means of OECD’s (Organization of Economic Cooperation and Development) and IDE-JETRO’s (Institute of Developing Economies - Japan External Trade Organization) input-output tables and OECD bilateral trade data for estimating their indicators. On comparing their proposed RCA measure with the Balassa’s RCA index based on gross export figures, the authors noted important differences in ranking of some countries in certain sectors. For instance in the year 1995 China’s position in office, accounting and computing machinery was 9th based on traditional RCA measure while 15th based on $TiVA$ based RCA. The computed RCA values depicted a move from comparative advantage to comparative disadvantage. In motor vehicles, trailers and semi-trailers Korea’s position based on traditional RCA measure was 16th (comparative disadvantage) but increased to 8th (comparative advantage) when $TiVA$ based RCA measure was used.²¹ Similar changes were noticed in case of other countries also.

However Meng *et al.* in their determination of domestic value-added components, do not consider the presence of a third country with which country r though not directly, can trade indirectly. Thus country r can export intermediates to s to be processed and exported to a third country. Under such circumstances, value-added by country r in its indirect exports to that third country should be taken into consideration. Therefore, when determining the position of country r relative to other countries in exporting a product to the world, and thereby formulating RCA indices on the basis of domestic value-added in exports by country

²¹ The comparison pertains to 30 countries – 29 countries and the rest of the world (treated as one).

r , like Koopman *et al.* (2010), it seems quite reasonable to consider the sum of first three terms from equation (2.28) as an indicator of domestic value-added in exports.

The existing literature has therefore recognized the growing importance of internationalization of production and some of them have effectively suggested the re-computation of the RCA indices based on value-added trade rather than gross trade.

2.4 Rational of the Present Study and Conclusion

The above analysis of the existing literature reveals a few issues which require further exploration. It is important to realize that the trade based RCA indices are frequently used in the literature to determine the export potential of countries and these indices suffer from the intrinsic problem of valid theoretical foundation. Attempts to establish the consistency of Balassa's RCA index with the theory of comparative advantage have failed to produce any conclusive results. The approach has largely been theoretical with no reference to empirical analysis. Although empirical analyses are largely dependent on observed data, which can be distorted due to various policy interventions and country specific characteristics, empirically examining the consistency of RCA indices with the theories on comparative advantage helps to understand the relevance and justification of these indices in real world. Existing literature falls short in this context. The objective of this current research is therefore trying to determine the consistency of RCA indices identified in the literature with the theories on comparative advantage as imagined by Ricardo or Heckscher-Ohlin, by analyzing observed data. Since the current thesis deals with the trade based RCA indices, the purpose of the current research is to examine the consistency of these RCA indices with the theories of comparative advantage. A comparative analysis of various RCA indices will help in determining one or more RCA indices that have the comparative advantage theory ingrained in their structure. Although Yeats (1985) made an attempt in this regard by empirically analyzing the consistency of Balassa's index with the factor proportions theory, the limitation of his study has been highlighted in the text. Further, Leromain and Orefice (2014) have suggested theoretically consistent RCA index, but the deficiencies associated with their index have also been recognized. In this context, the trade based RCA indices, are too not perfect from every respect. They can reflect comparative advantage or disadvantage on the basis of calculated values, due to factors other than productivity differences or factor endowment

differences, such as, trade barriers, historical trade relationships, demand shocks generated within countries and tastes and preferences. Such problems are inherent to any RCA index based on observed data, and therefore each of them will fall short with respect to some criterion. Thus the values generated by any RCA index need to be interpreted with some caution. Hence, in this context, restricting the analysis in this thesis to trade based RCA indices and trying to examine the consistency of these indices with the inherent comparative advantage of countries like factor endowment differences or productivity differences empirically, is certainly an important step towards establishing theoretically consistent indices. However as noted in chapter 1 (p. 13), research must not stop at identifying theoretically consistent indices, but must also move forward and analyze the empirical distributions of those RCA indices. A theoretically consistent index with stable distribution over sectors, countries and time, can be identified as an “ideal” index and will facilitate reliable usage of the index for comparative analysis.

In the context of the above mentioned arguments, it is also essential to realize the importance of production fragmentation. The phenomenon of fragmentation of production, which gives rise to value-added trade, has noteworthy consequences for the trade based RCA indices. These indices should be adjusted for domestic value-added in trade to factor out the contributions of other countries in the production process of any particular product and understand the true export potential of a country in that product. Although the importance of such adjustment has been realized in the literature through adjustment of Balassa index for domestic value-added in trade, extending the concept for other indices is a non treaded area which requires further exploration. Since, this thesis is primarily concerned about the comparative performance of alternative trade based RCA indices with the theories of comparative advantage, adjusting all such indices for value-added in trade and examining the changes in their performance will form a crucial extension of the literature and constitute a robustness check on the conduct of indices. Thus addressing the above issues will help in filling such gaps in the existing literature and aid in recognizing a reliable index for policy purposes.



Chapter 3

Analyzing the Empirical Distributions of RCA Indices

3.1 Introduction

The Revealed Comparative Advantage (RCA) index originally proposed by Balassa (1965) finds its application towards understanding the possibilities of gainful exchange of products amongst trading nations. The index has however been a subject of discourse and potential improvements of the index have been consequently suggested in the literature, to address many deficiencies that the index suffers from.

As also noted in chapter 1 (p. 6-7), Ballance *et al.* (1987) recognized that the usefulness of any RCA index can be divided into three categories. First, the indices help in identifying those countries which reveal comparative advantage in a particular sector and those countries which do not. Second, based on the values of the indices, it is possible to generate cross country ranking with respect to a particular sector, or cross sector ranking with respect to a particular country. Third, they facilitate quantification of the extent of comparative advantage of a country in a sector compared to some other country or the extent of comparative advantage of a country in a sector compared to any other sector. The second application hints towards the usage of the indices as an ordinal measure while the third application implies usage of the indices as a cardinal measure. These two applications of the RCA indices will be dealt with in this chapter.

According to Yeats (1985) usage of the index as a cardinal measure is more effective than its usage as an ordinal measure. This is because, ranks by themselves do not provide any account of the magnitude of differences. A cardinal measure on the other hand by permitting a measurement on the magnitude of differences, is more useful for commercial and public policy formulations. Yeats then showed that the Balassa index is not suitable either as a cardinal or as an ordinal measure. Despite this fact, the original index of Balassa found its

application in various studies to infer the presence of strong or the weak sectors of a country in comparison to other countries, or to identify the comparative advantage or disadvantage of a country in a commodity compared to other commodities (Batra and Khan, 2005; Smyth, 2005; Wignaraja, 2011; Karaalp, 2011; Ekram *et al.*, 2013; Hassan, 2013). Further, the index also found its application in determining the changes in pattern of comparative advantage of countries over time (Hiley, 1999; Kijboonchoo and Kalayanakupt, 2003). These applications seem to ignore the arguments of Yeats and rely on the assumption that the underlying distribution is stable either over countries or over products or over time, as the case may be. However, concerns have been raised in the existing literature about the stability of the index distributions across countries or across products. For instance, as noted in chapter 2 (p. 31-32), Hoen and Oosterhaven (2006) and Yu *et al.* (2009) argue that Balassa's RCA index is characterized by an unstable mean due to its variable upper bound but fixed lower bound. The variability in the upper bound can be attributed to the sensitivity of the index to size of the sector or the size of a country, so that smaller countries (sectors) tend to have more asymmetric distributions across sectors (countries) respectively. If the upper bound keeps on changing from country to country or from sector to sector, the average or the mean value will also keep on changing. The phenomenon of changing mean value will be evident from the computation of arithmetic mean of index values across countries for a particular commodity or across commodities for a particular country. Evidently, the arithmetic mean will fluctuate from sector to sector or country to country, and evidences of such fluctuating mean value (above 1) have been documented by Hoen and Oosterhaven (2006) and Hinloopen and Marrewijk (2001) in their respective studies. Thus, skewness or asymmetry in the distribution of the index brought about by varying upper bound is certainly responsible for an average index value above unity or above the comparative advantage neutral point. If the underlying distribution is always symmetrically distributed about the value zero, usage of the index as a cardinal measure becomes feasible. For in that case, as explained in chapter 2 (p. 36), if a country gains comparative advantage in a sector, some other country must lose comparative advantage in that particular sector to retain an average value of zero. Under such circumstances, it is possible to identify the comparative advantage or disadvantage of one country compared other countries in exporting a particular product. However, if the computed average is not stable and not equal to zero, then if one country gains comparative

advantage in a product, it will not be possible to say for certain that it has gained comparative advantage compared to other countries, as it does not necessarily follow that some other country has lost comparative advantage in that same product. Hence, comparing countries for a particular product becomes quite unfeasible. Similar observations can be made while comparing sectors for a country. Therefore, the index fails to function properly as a cardinal measure.

A closer look at the phenomenon of unstable mean reveals another issue. An unstable mean will imply an unstable distribution. If the distribution of index values across countries for a particular commodity differs from the distribution of index values across the same group of countries for another commodity, then the same value for the index of Balassa may have different interpretations for different commodities. Accordingly, using Balassa's index to identify comparative advantage or disadvantage of a country in a commodity in comparison to the other commodity must be done with care. This implies, ranking of sectors with respect to a country can be problematic. Thus, the index may not be suitable as an ordinal measure. If on the other hand, the distribution of index values across sectors for a particular country differs from the distribution of index values across the same group of sectors for another country, then again the same value for Balassa's index may imply different things for different countries. Therefore, using Balassa's index to identify a country's position relative to other countries with respect to a sector is not an easy task. Hence, ranking of countries with respect to a sector is difficult. These arguments relate to the observations by Yeats (1985) who takes note of the fact that different sectors may have different country distributions which adds to the infeasibility of direct comparison of index values for a country across sectors. Similar difficulties in case of cross country comparison for a sector were noted by him.

Thus, not only the index fails to act as a cardinal measure, its use as an ordinal measure is also questionable.

On similar grounds, if the distribution of index values across sectors for a country differs from year to year, then problem may arise in interpreting the index values over time. As observed by Leromain and Orefice (2014) the concept of comparative advantage is ex-ante in nature as it is based on pre-trade relative prices. Using pre-trade prices, inferences are drawn

about post-trade scenario. In this regard if RCA indices are used as a proxy for comparative advantage, they must be expected to be sticky over time. Benedictis and Tamberi (2001, 2004), Hinloopen and Marrewijk (2001) have analyzed the stability of Balassa's distribution over time. Benedictis and Tamberi noticed the distribution to be stable over time for France but unstable for Italy, Germany and Japan. Hinloopen and Marrewijk by grouping the individual observations on twelve members of European Union found the distributions to be considerably stable both over months and years. Stability of index distribution over time through stability of the mean also ensures reliable applicability of the index for time series analysis, as recognized by Yu *et al.* (2009).

As also mentioned in chapter 2 (p. 31), due to the asymmetric distribution of Balassa index, using its arithmetic mean to identify the performance of the average sector or the average country can be doubtful. For distributions skewed to the right, the arithmetic mean tends to assign more weight to the sectors with index value exceeding unity than to sectors with index value less than unity (Benedictis and Tamberi, 2001). As a result, a possible interpretation of the mean value will be - a country has comparative advantage in the average sector (on computing the mean across sectors for a country) or an average country has comparative advantage in any sector (on computing the mean across countries for a sector). According to Benedictis and Tamberi (2001) such interpretations are of little significance. Evidently, for rightly skewed distributions, the mean will be greater than the median or the middle most value, and in fact median will be a better indicator of the average sector or country. Unlike arithmetic mean, the median is not influenced by extreme values. A low median (less than unity) in case of a country will imply that the country has large share of sectors with comparative disadvantage, and a high median (greater than unity) will imply large share of sectors with comparative advantage for the country (Benedictis and Tamberi, 2004).

Keeping aside the arguments regarding the usage of mean or median as an indicator of average and given the extensive usage of the Balassa index for cross country or cross commodity or cross year comparison, it seems justifiable to study the empirical distribution of Balassa's index and analyze its stability. However, since Balassa's index has been modified in later literature to overcome many deficiencies of the index, an analysis of the empirical distributions of those indices must also not be overlooked.

As already specified in chapter 1 (p. 8) and chapter 2 (p. 23), the RCA indices to be taken up for analysis in this thesis can be divided into two broad categories –

- (1) The exports based indices i.e., the indices constructed with export values only.
- (2) The exports and import based indices i.e., the indices constructed by including both export and import values.

The five exports based indices – RCA index of Balassa, the Revealed Symmetric Comparative Advantage (RSCA) index by Dalum *et al.* (1998) and Laursen (1998), the Additive Revealed Comparative Advantage (ARCA) index by Hoen and Oosterhaven (2006), the Normalized Revealed Comparative Advantage (NRCA) index by Yu *et al.* (2009) and the Log of Balassa index based on Vollrath (1991) to be considered, take the following forms –

$$\text{Balassa's RCA Index} = \frac{X_a^i/X_t^i}{X_a^w/X_t^w}$$

$$\text{RSCA Index} = \frac{\frac{X_a^i/X_t^i}{X_a^w/X_t^w} - 1}{\frac{X_a^i/X_t^i}{X_a^w/X_t^w} + 1}$$

$$\text{ARCA Index} = \frac{X_a^i}{X_t^i} - \frac{X_a^w}{X_t^w}$$

$$\text{NRCA Index} = \frac{X_a^i}{X_t^w} - \frac{X_t^i}{X_t^w} \frac{X_a^w}{X_t^w}$$

$$\text{Log of Balassa Index} = \ln \frac{X_a^i/X_t^i}{X_a^w/X_t^w}$$

The two RCA indices incorporating both exports and imports – the Revealed Trade Advantage (RTA) index and Revealed Competitiveness (RC) index based on Vollrath (1991) take the following forms:

$$\text{RTA Index} = \frac{X_a^i/X_t^i}{X_a^w/X_t^w} - \frac{M_a^i/M_t^i}{M_a^w/M_t^w}$$

$$RC\ Index = \ln \frac{X_a^i/X_t^i}{X_a^w/X_t^w} - \ln \frac{M_a^i/M_t^i}{M_a^w/M_t^w}$$

In all of the index specifications, i denotes any particular country and a denotes any particular sector. t is the sum of all sectors and w is the world total. However, while calculating the index values in this and subsequent chapters, t and w will include the sum of all sectors and sum of all countries respectively on which data will be available. The terms ‘sector’, ‘commodities’ and ‘products’ are used synonymously in this and subsequent chapters. Therefore X_a^i is the exports of product a by country i , X_t^i is the exports all products by country i , X_a^w is the exports of product a by all countries and X_t^w is the exports of all products by all countries. M_a^i , M_t^i , M_a^w and M_t^w can be similarly be defined for imports.

As evident from the structures of the specified indices, the computed arithmetic mean across sectors or across countries will be stable and equal to the comparative advantage neutral point of zero, only in case of the NRCA index. The ARCA index generates a stable arithmetic mean equal to zero, only across sectors, but not across countries. The means for each index are derived in the Appendix A.2, as also specified in chapter 2. Because of stable arithmetic means across sectors and across countries, Yu *et al.* (2009) argue that the NRCA index is suitable for comparison of index values across time. It seems important in this context to analyze whether the empirical distributions of NRCA index are stable over countries or sectors or time through stable means, and to what extent other RCA indices falls short of the NRCA index. The analysis will thus help in ascertaining the applicability of the indices as ordinal or cardinal measures. Such a comprehensive study on the empirical distributions of RCA indices is lacking in the existing literature. Previously Benedictis and Tamberi (2001, 2004) and Hinloopen and Marrewijk (2001) have attempted to analyze the empirical distributions of Balassa’s RCA index. But their studies do not provide an intensive account of the statistical significance on stability of index distributions over time, over sectors and over countries, in the manner as presented in this chapter.

After a discussion on applicable methods and data in Section 3.2, the results are analyzed in Section 3.3. Section 3.4 concludes the chapter.

3.2 Methods and Data

The methods of analyzing the RCA indices and the data used for the purpose are described in this section. The following subsections elaborate on each method.

3.2.1 Cumulative Distributions and Summary Statistics of RCA Indices

Before analyzing the stability of the empirical distributions of each RCA index, a preliminary analysis on the cumulative distributions and summary statistics on each index are presented for each of the years – 1998, 1999 and 2000. For each year the sectoral observations on the considered countries are grouped together. Similar analysis has been carried out by Hinloopen and Marrewijk (2001), although they made use of monthly data for the years 1992 to 1996 for 12 member states of European Union while exporting to Japan. In their paper, they thus report the cumulative distributions and summary statistics for the index of Balassa only, for each month, each year, as well as for the whole period. The analysis in this chapter differs from their study in a few respects. First, the countries considered are heterogeneous group of countries and they are exporting to the world (or the group of countries on whom data are available) as a whole, instead of any specific country. As noted by Hinloopen and Marrewijk, this approach pertains to comparing countries at different levels of development with each facing differing transport costs. However, consideration of all countries permits a more comprehensive analysis on the export performance of any country with respect to many countries. Consideration of similar countries exporting to a single country tends to provide limited insights into the relative export potential of any country. Further when the reference group comprises of the world, the effect of transport cost may not be significantly relevant as every country is exporting to its immediate neighbors as well as to distant trading partners. The second difference between the studies is, only yearly data can be accessed for analysis in this chapter and relevant results are reported for each considered year only. Third, the analysis in this chapter includes not only Balassa's index but also several other indices proposed to overcome one or more deficiencies of Balassa's index.

3.2.2 Graphical Analysis of RCA Indices

The distribution of the indices are also graphically analyzed using cumulative distribution plots and the corresponding probability density functions represented by Kernel density plots,

for a number of countries for each particular year. Graphical analysis therefore provides a preliminary guide to the stability of index distributions over years. As already discussed, stability of index distributions over time is required to provide valid interpretations to the index values. The methodology of graphically illustrating the stability of index distributions through cumulative distribution plots or Kernel density plots is based on Hinloopen and Marrewijk (2001) and Benedictis and Tamberi (2001, 2004). Hinloopen and Marrewijk (2001) plotted the cumulative distributions for the index of Balassa for 12 member nations of European Union grouped together. Benedictis and Tamberi (2001, 2004) plotted the Kernel density distributions for France, Germany, Italy and Japan. The current study, in order to provide a more intensive evaluation of the distribution plots will like to make use of country specific samples. Consideration of a group of countries, as done by Hinloopen and Marrewijk, fails to reveal the variations in distribution plots due to country heterogeneity. Thus the country sample of Benedictis and Tamberi will be considered in this chapter. Apart from the mentioned set of developed countries, two other emerging developing economies are also considered in this chapter. These two countries are China and India. The consideration of these two developing countries rather than any other developing country is driven by their significant shares in the world population and GDP growth rates.¹ With respect to the selected sample, it must also be noted that China, France, Germany, Italy and Japan were placed among the top ten countries with largest shares in the exports of manufactures by all countries, according to the data for 1998, 1999 and 2000.² India although does not figure among the top ten, being a labor abundant economy as China, it is often considered to be a competitor of China in exports of many labor intensive manufactures

¹ The two countries' competitive population size and GDP growth rates have been noted in the literature. The shares of two countries' in the world population have been estimated to be the largest. As per the World Development Indicators of World Bank, the respective shares of China and India are 18.9% and 17.6% for the year 2014. According to the World Economic Outlook April 2015, released by International Monetary Fund (IMF), India would be the fastest growing economy in the world by recording faster growth rates than China in 2015, as well as in 2016. India's GDP growth rates for 2015 and 2016 have been projected to be 7.5% against China's 6.8% (2015) and 6.3% (2016). In 2014 however China was the fastest growing economy with its growth rate pegged at 7.4% against 7.2% by India.

² The shares were computed using TPP database. This measure could to some extent be influenced by size of countries. Nevertheless it can be accepted as a ready measure for rough comparison of countries.

(Dimaranan *et al.*, 2007).³ Further, there are speculations on the part of Chinese media that with slowing Chinese economy, India might emerge as the next “factory of the world” (PTI, 2015, August 18).

The cumulative distributions plotted for 3 separate years for each country and each index helps to study the shifts in the empirical distributions of the indices. Kernel density plots apart from exhibiting the shifts in distributions over time throw additional light towards the degree of asymmetry in data. Kernel density estimation is aimed at estimating the probability density function of a continuous random variable. The basic objective for the plot is to make conjectures about the population distribution on the basis of finite sample data. It is a non-parametric approach because it does not assume anything about the underlying distribution of the variable. The formulation of Kernel density estimator is attributed to Rosenblatt (1956). Assuming n independent observations $x_1, x_2, \dots, x_n$ from a random variable X , the Kernel density estimator $\hat{f}_h(x)$ for the estimation of the density value $f(x)$ at the point x is given by,

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x_i - x}{h}\right) \quad (3.1)$$

In the above formulation, h is the smoothing parameter, also known as the bandwidth or window-width. $K(\cdot)$ is the Kernel function satisfying the condition $\int K(x)dx = 1$. The range of the data is divided into overlapping intervals of width h and height $\frac{1}{nh}$ with the intervals centered at the data points, and then the estimates of the density at the centre of the intervals are generated (StataCorp., 2013). The Kernel estimator is the sum of the density curves generated at each data point (Ullah, 1988). The Kernel $K(\cdot)$ determines the shape of the density curves at each data point and bandwidth h determines the width of those curves. For $\hat{f}_h(x)$ to be a probability density function, $K(\cdot)$ must be non-negative throughout and satisfy the condition $\int K(x)dx = 1$ (Ullah 1988).⁴

³ It is to be noted, any direct comparison of the current Kernel density plots for the index of Balassa with that of Benedictis and Tambari will not be feasible, as the latter consider the Balassa index in a log scale while plotting the country specific distributions.

⁴ Kernel density plots are similar to histograms. However, unlike histograms, Kernel density plots are continuous. According to Ullah (1988) Kernel density plots are at times problematic due to their sensitivity to the choice of bandwidth or smoothing parameter, which are also fixed across data points. Smaller band widths

The Epanechnikov Kernel function, which is the default function in Stata is used for the estimation of densities and is most efficient in minimizing the mean integrated squared error (StataCorp., 2013).⁵ Along with the default Kernel function, default bandwidth calculated by Stata is used.⁶ The formulae for Epanechnikov Kernel function and default bandwidth utilized by Stata can be represented as follows:

$$K(z) = \begin{cases} \frac{3}{4} \left(1 - \frac{1}{5} z^2\right) / \sqrt{5} & \text{if } |z| < \sqrt{5} \\ 0 & \text{otherwise} \end{cases} \quad (3.2)$$

and

$$h = \frac{0.9m}{n^{1/5}} \quad \text{where } m = \min \left(\sqrt{\text{variance}_X}, \frac{\text{interquartile range}_X}{1.349} \right) \quad (3.3)$$

As already defined, X is the variable for which the Kernel density will be estimated and n is the number of observations on X . Evidently the bandwidth is not a function of the Kernel function but the data (StataCorp., 2013). Since the Kernel density plots are not compared

may result in unsmooth densities, while larger band widths may lead to increased smoothness of the density function. Fixing of bandwidth results in slight bumps towards the tails of the distributions and attempts in smoothing the tails by altering the bandwidth may distort the middle part of the distributions (Ullah, 1988). An effective solution is to incorporate varying bandwidths into the formula for density estimate. Kernel density estimator in that case will be known as the Variable Kernel Estimator. However, in the present analysis, we refrain from such considerations and try to analyze the shifts in index distributions for a country over time using constant bandwidth. Of course, trying to analyze the stability of distributions over time using variable bandwidth can be explored in future.

⁵ The mean integrated squared error is defined as $MISE(\hat{f}_h) = \int E\{\hat{f}_h(x) - f(x)\}^2 dx$. $MISE$ must be minimized to obtain $\hat{f}_h(x)$ and the Epanechnikov Kernel function ensures that the $MISE$ is minimized.

⁶ The choice of the Kernel function is not integral to the estimation of Kernel density function as the choice of the bandwidth (StataCorp., 2013). Epanechnikov (1969) compared the efficiency of various Kernels with respect to his Kernel function and found that there is not much variance in the results (as quoted in Ullah, 1988). In the present analysis, the density estimates were also obtained by using Gaussian Kernel functions and default bandwidth. The generated density plots were found to be very similar to those obtained by using Epanechnikov Kernel function.

across countries for any particular index, need is not felt for keeping the bandwidth same for every country for each index, as done by Benedictis and Tamberi (2004).

3.2.3 Tests of Significance on the Stability of RCA Index Distributions

The present subsection elaborates on the methodology for analyzing the statistical significance of shifts in index distributions over time, over sectors and over countries.

3.2.3.1 Wilcoxon's Signed Rank Test

While a graphical analysis of the data through cumulative distribution and Kernel density plots helps in determining the changes in empirical distributions of the considered RCA indices over time, the data must further be analyzed to understand the statistical significance of those changes. Wilcoxon's signed rank test can be considered to reinforce the results corresponding to cumulative distribution plots and Kernel density plots, by providing evidences on statistical significance on the shifts of empirical distributions over time. A two tailed Wilcoxon (1945) signed rank test is a non-parametric test and tests for any difference in distributions between two samples, provided that these two samples are not independent, but matched or paired or they constitute repeated measurements on a single sample. For each pair of responses the test computes the difference and ranks the absolute value of differences. All the positive and negative differences are then grouped separately and sum of the ranks for positive differences constitute the Wilcoxon's signed rank test statistics. Conover (1999) provides a formal enumeration of the method. Consider a bivariate sample of size n' $(X_1, Y_1), (X_2, Y_2), \dots, (X_{n'}, Y_{n'})$ defined on two random variables X and Y . The test procedure begins by first computing the difference between each pairs of observations on each variable,

$$D_i = Y_i - X_i \quad i = 1, \dots, n' \quad (3.4)$$

The absolute value of the difference is then considered,

$$|D_i| = |Y_i - X_i| \quad i = 1, \dots, n' \quad (3.5)$$

Any pairs of observations for which the difference is zero, are excluded from the sample and the remaining are considered for analysis. Let the remaining number of pairs be denoted by n such that $n \leq n'$. In the next step, ranks from 1 to n are assigned to $|D_i|$. A rank of 1 is

assigned to the pair (X_i, Y_i) with smallest $|D_i|$, and a rank of n is assigned to the pair (X_j, Y_j) ($i \neq j$) with the largest $|D_i|$. In case of pairs with same $|D_i|$, the average rank is assigned. The average is calculated on the basis of the ranks that would have been assigned in the absence of ties. If suppose, three pairs of observations have the same magnitude of difference, which incidentally is also lower than the magnitudes of differences for any other pairs of observations in the sample, then to calculate their ranks, three of them should first be allotted the consecutive ranks 1, 2 and 3 and then the average of the ranks $\frac{(1+2+3)}{3}$ must be allotted to each pair.

Let the rank for each pair of observations (X_i, Y_i) with their associated sign be denoted by R_i , and be called the signed rank. Assuming no ties in the sample, the test statistic is defined as,

$$W = \sum R_i \text{ (where } D_i \text{ is positive)} \quad (3.6)$$

In the presence of ties, the test statistic will be defined as:

$$W' = \frac{\sum_{i=1}^n R_i}{\sqrt{\sum_{i=1}^n R_i^2}} \quad (3.7)$$

The test statistic in this case considers the sum of all signed ranks with their associated + or – sign depending on whether D_i is positive or negative. However, for the considered sample in the present analysis, no ties were observed. Hence, the discussions can proceed by considering the test statistic W .

The two tailed null and alternative hypotheses for the test can be specified as:

$$H_0: E(D) = 0$$

$$H_1: E(D) \neq 0$$

The null hypothesis therefore implies that the difference between pairs of observations on the two variables has a mean value of zero. Since the test assumes that D is symmetric, a zero value for mean implies zero value for the median. A zero value for the mean of the

distribution of the difference implies that the distributions on X and Y have same means, which hints towards identical distributions for the two variables.⁷

Rejection of null hypothesis, which conforms to unequal distributions, will be evident from the values of the test statistics falling in defined the critical region of the distribution of differences when the null hypothesis is true. Benedictis and Tamberi (2004) use the signed rank test for determining the statistical significance of changes in the probability density functions for the index of Balassa between the years 1986 and 1996 in case of Italy, France, Germany and Japan. By considering the distribution of Balassa's index for a particular country at two different points in time, they essentially restrict themselves to matched pairs of observations. This chapter also makes use of the Wilcoxon's signed rank test to study the statistical significance of the shifts in country distributions over time, but the analysis offered in this chapter differs from that of Benedictis and Tamberi in two respects. First, Wilcoxon's signed rank test statistics is computed not only for France, Germany, Italy and Japan, but also for China and India. Second, the statistical significance of shifts in distributions for each country over three years is evaluated not only for the index of Balassa but also for other RCA indices suggested in the literature.

Apart from analyzing the stability of index distributions over years for separate countries, Wilcoxon's signed rank test is also applied to examine the stability of the index distributions over years for separate sectors. Benedictis and Tamberi did not inspect this with Balassa's index. For analyzing the stability of index distributions over time for separate sectors, 7 sectors with following ISIC codes - 321 (manufacture of textiles), 322 (manufacture of wearing apparel except footwear), 323 (manufacture of leather products, leather substitutes and furs), 324 (manufacture of footwear), 351 (manufacture of industrial chemicals), 352 (manufacture of other chemical products) and 385 (manufacture of professional, scientific and measuring equipments) are selected.⁸

⁷ If two distributions are identical, then the positive differences will be symmetrically distributed as the negative differences. Thus the assumption of symmetrically distributed D is required to formulate the null hypothesis (Siegel, 1956; Fraser, 1957).

⁸ The selected sectors confirm to those used in chapter 4 while testing the consistency of the RCA indices with the Heckscher-Ohlin theory.

3.2.3.2 Wilcoxon's Rank Sum Test

The stability analysis is not only restricted to time but can also be extended to countries and sectors. Since comparison over sectors or over countries involve unmatched or non-paired data, Wilcoxon's signed rank test will not be suitable. In this respect Wilcoxon (1945) rank sum test is quite helpful in determining the statistical significance of shift in distributions across countries over sectors or shifts in distributions across sectors over countries. As already discussed, stable index distributions over sectors or countries provides evidences in support of the usage of indices as cardinal or ordinal measures.

Wilcoxon's rank sum test also known as the Wilcoxon-Mann-Whitney test is thus used in this chapter for determining changes in the distribution of RCA indices between any two sectors or between any two countries. Wilcoxon's rank sum test is a non-parametric alternative to 2 sample t test on unmatched or non-paired data, i.e., when the samples drawn are independent. While t test assumes that the underlying population should be normally distributed, Wilcoxon's test assumes away any such requirement on the form of the distributions and it is entirely based on the order of the observations in the samples. This characteristic will particularly be helpful given the fact that not many of the RCA indices are normally distributed. Hence, the statistical inferences will not be biased. The test however assumes that the data are drawn from a continuous distribution. The same assumption also holds for the non-parametric signed rank test previously mentioned. The rank sum test thus tests for the hypothesis that the two independent samples are drawn from two populations which share similar distributions. However, it is to be noted, there can be differences in distributions in the two populations from which the samples come, due to differences in standard deviations or medians or skewness and kurtosis. But the rank sum test might accept the null hypothesis, as it only compares the rank sums of the two groups. Thus, in order to validate the null hypothesis, one needs to assume that the two distributions share the same shape. For in that case changes in distributions will be reflected by shifts in location which permits a comparison of the actual rank sums with their expected or mean value. Effectively, then the test will also permit analysis of the differences in means or medians (Hart, 2001; Campbell, 2006). Thus, deviations from the null hypothesis that the Wilcoxon's rank sum test tries to

test are the location shifts of the distributions, which under the assumption of identically shaped distributions implies testing for differences in means or medians.

Conover (1999) provides a formal description of the test. Consider two random samples drawn from two populations. Let $X_1, X_2, \dots, X_n$ be the random sample of size n from population 1, and $Y_1, Y_2, \dots, Y_m$ be the random sample of size m from population 2. The observations are then combined into a single sample and ranked in an ascending order from 1 to $n + m = N$. Let $R(X_i)$ and $R(Y_j)$ represent the ranks assigned to X_i and Y_j for all i and j . In case, several observations in the combined sample have similar values, or are tied, the average rank is assigned to each such observation. The average rank is calculated in the same manner as in the case of Wilcoxon's signed rank test. Assuming no ties, the test statistics can be expressed as:

$$T = \sum_{i=1}^n R(X_i) \text{ or } \sum_{j=1}^m R(Y_j) \quad (3.8)$$

In the presence of ties the test statistic is expressed as:

$$T' = \frac{T - n \frac{N+1}{2}}{\sqrt{\frac{nm}{N(N-1)} \sum_{i=1}^N R_i^2 - \frac{nm(N+1)^2}{4(N-1)}}} \text{ or } \frac{T - m \frac{N+1}{2}}{\sqrt{\frac{nm}{N(N-1)} \sum_{i=1}^N R_i^2 - \frac{nm(N+1)^2}{4(N-1)}}} \quad (3.9)$$

R_i^2 is the sum of the squares of ranks on all N observations. However, for the sample analyzed in this chapter, tied observations were not found. Hence, it is relevant here to proceed by considering the test statistic T .

The two tailed null and alternative hypotheses can be stated as:

$$H_0: F(x) = G(x) \text{ for all } x$$

$$H_1: F(x) \neq G(x) \text{ for some } x$$

$F(x)$ and $G(x)$ are the distribution functions corresponding to X and Y respectively.

As per the null hypothesis, the rank sum for the two samples will be same, provided the two samples are of the same size. If the two samples share similar population distributions, then

the test statistic will be equal to the mean or the average ranks for each of the two groups. Thus, in case of the two samples of same size, i.e., $n = m$, the test statistic is given by,

$$T \approx \frac{N(N+1)}{4} \quad (3.10)$$

If two samples are of different sizes, i.e., $n \neq m$, the test statistic is given by,

$$T \approx \frac{n(N+1)}{2} \text{ or } \frac{m(N+1)}{2} \quad (3.11)$$

depending on whether X or Y is considered for rank sum.

In case the null hypothesis gets violated the test statistic or the rank sum will deviate from the mean. Hence, the null hypothesis will be rejected if observed rank sum or the value of test statistic is significantly large or small with respect to the distribution of rank sums, when the null hypothesis is true.

It is to be noted, the U test separately developed by Mann and Whitney (1947) is an extension of the Wilcoxon's (1945) test. While Wilcoxon focused on samples of equal sizes, Mann and Whitney explored the possibilities for data with unequal sample sizes as well. However, Mann and Whitney's test and Wilcoxon's rank sum test are equivalent in the sense that one is a linear combination of the other. With two random samples $X_1, X_2, \dots, X_n$ and $Y_1, Y_2, \dots, Y_m$, each of sizes n and m respectively, Mann and Whitney's U statistic is the number of pairs (X_i, Y_j) for which $X_i < Y_j$, and the Wilcoxon's rank sum T statistic as defined, is the sum of the ranks of Y when the observations under X and Y are arranged in ascending order. If suppose in the simplest possible case, the observations corresponding to Y are smaller than the observations corresponding to X , then for every possible shift of Y_n to the right of X_i for all i , T will increase by the number of pairs for which $Y_n > X_i$. The obtained measure represents the U statistics. Hence, the relation between the two statistics can be expressed as:

$$T = \frac{m(m+1)}{2} + U \quad (3.12)$$

Thus, given U , the value of the T statistics can be uniquely determined (Higgins, 2004).

Considering RCA index values across countries for any two sectors, acceptance of the null hypothesis will imply that the distributions across countries for the two sectors are stable. Similarly, while comparing the two countries, acceptance of the null hypothesis will imply the distributions across sectors for two countries are stable. These findings will provide evidence in support of the fact that the index can reliably be used as an ordinal measure or for ranking of sectors (countries) with respect to a country (sector). Further, since the rank sum test analyze changes in distributions through changes in means or medians, stability of the index distribution will imply stability of mean across sectors or across countries. Hence, the index can be used as a cardinal measure. Similar attempts to analyze sectoral or country mean have not been made in the existing literature.

For analyzing the stability of index distributions over countries, the set of six countries previously mentioned are considered, namely, China, France, Germany, India, Italy and Japan. For analyzing the stability of index distributions over sectors, the previously mentioned seven sectors with ISIC codes 321, 322, 323, 324, 351, 352 and 385 are considered.

3.2.3.3 Kolmogorov Smirnov Test

To perform a robustness check on the stability of RCA index distributions over sectors or over countries, the Kolmogorov Smirnov test is considered. This test pertains to unpaired samples and hence is not suitable for analyzing the empirical distributions over time. Smirnov (1939) devised a non-parametric test to determine whether distribution functions associated with two different populations from which two samples are drawn, are identical (Conover, 1999). This test is also known as the Kolmogorov Smirnov two sample test, as it is similar in spirit to the one sample Kolmogorov test for comparing an empirical distribution function with the corresponding hypothesized distribution function. For both tests, the statistics are a function of the vertical distance between two distribution functions. The Kolmogorov Smirnov test computes the test statistics by considering the maximum difference between two cumulative distributions. Conover (1999) provides a formal explanation of the test by considering two independent random samples of sizes n and m , represented by $X_1, X_2, \dots, X_n$ and $Y_1, Y_2, \dots, Y_m$ respectively. $F(x)$ and $G(x)$ are assumed to be their respective unknown distribution functions. $S_I(x)$ can be defined to be the empirical

distribution function corresponding to the random sample $X_1, X_2, \dots, X_n$, and $S_2(x)$ can be defined to be the empirical distribution function corresponding to the random sample $Y_1, Y_2, \dots, Y_m$. The test statistic as Conover (1999) notes, is equivalent to the greatest vertical distance between two empirical distribution functions, and can be expressed as:

$$T = \sup_x |S_1(x) - S_2(x)| \quad (3.13)$$

In the above expression “*sup*” denotes supremum. Thus, T measures the supremum of the absolute value of the difference between $S_1(x)$ and $S_2(x)$ over all x .

The test statistic T presented above corresponds to two tailed test. The test statistic can also be differently specified for one tailed tests. However, in the present analysis, it is more important to determine whether the two distributions differ rather than whether one distribution falls short of or exceeds the other distribution. For that reason, it is of minor significance to go into the details of the test statistics for one tailed tests, and concentrate instead on the test statistic corresponding to two tailed test. Thus, the null and the alternative hypotheses corresponding to the test statistic T can be specified as:

$H_0: F(x) = G(x)$ for all x from $-\infty$ to $+\infty$

$H_1: F(x) \neq G(x)$ for at least one value of x .

Evidence against null hypothesis or unequal distributions, will be provided by the value of test statistic which exceeds or falls short of the critical value according to the distribution of the test statistics when the null hypothesis is true. As pointed out by Lehmann (2006) although the test is based on actual observations, it is similar to an analysis of ranks. This is so because, ranking all the observations will not change the maximum difference between the cumulative distributions.

The Wilcoxon’s rank sum test as previously discussed, detects location shifts in distributions. This feature of the test, along with the assumption of identically shaped distributions signifies changes in mean or median. The Kolmogorov Smirnov test detects shifts in distributions due to changes in mean, standard deviation, presence of outliers, differences in skewness or kurtosis, number of modes etc. Hence, if the test reports statistically significant changes in distributions, it can be due to any one or more factors noted.

The necessary data on exports and imports for 28 manufacturing sectors classified on the basis of 3 digit International Standard Industrial Classification of All Economic Activities (ISIC) revision 2 for the years 1998, 1999 and 2000 are collected from the Trade, Production and Protection (TPP) Database, 1976-2004 by Nicita and Olarreaga (2007). Selection of these three particular years is primarily governed by the availability of the largest number of data on indicators required for the construction of average productivity ratio (to be used in chapter 4) and RCA index values, keeping in mind that for each country in each year, the sectoral observations remains the same. Consideration of years beyond the three selected years tends to reduce the sample size.⁹ Although the observations on average productivity ratio will not be of relevance in this particular chapter, retaining the three selected years for analysis in this chapter helps to maintain parity with the analysis in the later chapter. The summary statistics for the data used in this chapter are provided in Table B.1 of Appendix B. The detailed descriptions on the 28 sectors are provided in the Table B.2 of Appendix B.

3.3 Discussion of Results

The results of the analysis are discussed in this section.

3.3.1 Discussions on the Summary Statistics and Graphical Analysis

For each of the seven indices, Tables 3.1.A to 3.1.G give detailed information on the cumulative distributions and the summary statistics for each of the years 1998, 1999 and 2000. The sample for each index in each year includes 60 countries and each country has 28 sectoral observations.¹⁰

⁹ Years are selected to keep the number of observations as large as possible. In the present chapter, the numbers of countries are about 60 for most of the sectors in each year. However, for the sectors with ISIC codes 332, 354 and 361, only 59 observations exist, on account of missing observations on exports and imports for Nepal in all 3 years.

¹⁰ Nepal reports missing observations on sectors with ISIC codes 332, 354 and 361 for all 3 years considered.

Table 3.1.A
Empirical Distribution of Balassa Index for Each Year: Entire Sample Analysis

Summary Statistics/ Year	1998	1999	2000
p-1	0.0004	0.001	0.001
p-5	0.022	0.028	0.024
p-10	0.067	0.069	0.068
p-25	0.252	0.253	0.250
p-50	0.659	0.647	0.652
p-75	1.335	1.349	1.363
p-90	2.889	2.910	2.819
p-95	4.228	4.346	4.486
p-99	10.040	10.312	11.941
Maximum Value	104.720	113.179	112.484
Mean	1.287	1.320	1.367
Standard Deviation	3.270	3.486	3.773
Skewness	20.314	21.092	18.121
Kurtosis	605.506	640.876	476.159
No. of observations	1677	1677	1677

Table 3.1.B
Empirical Distribution of RSCA Index for Each Year: Entire Sample Analysis

Summary Statistics/ Year	1998	1999	2000
p-1	-0.999	-0.997	-0.998
p-5	-0.957	-0.946	-0.953
p-10	-0.874	-0.870	-0.873
p-25	-0.598	-0.596	-0.600
p-50	-0.205	-0.215	-0.211
p-75	0.143	0.148	0.154
p-90	0.486	0.489	0.476
p-95	0.617	0.626	0.635
p-99	0.819	0.823	0.845
Maximum Value	0.981	0.982	0.982
Mean	-0.207	-0.203	-0.202
Standard Deviation	0.484	0.484	0.488
Skewness	0.178	0.198	0.170
Kurtosis	2.133	2.148	2.128
No. of observations	1677	1677	1677

Table 3.1.C
Empirical Distribution of ARCA Index for Each Year: Entire Sample Analysis

Summary Statistics/ Year	1998	1999	2000
p-1	-0.161	-0.163	-0.169
p-5	-0.119	-0.118	-0.110
p-10	-0.033	-0.034	-0.034
p-25	-0.013	-0.013	-0.014
p-50	-0.003	-0.003	-0.003
p-75	0.005	0.005	0.005
p-90	0.035	0.034	0.035
p-95	0.076	0.078	0.079
p-99	0.307	0.307	0.315
Maximum Value	0.705	0.789	0.729
Mean	0.00001	0.00001	0.00001
Standard Deviation	0.070	0.070	0.069
Skewness	3.590	3.545	3.362
Kurtosis	30.182	31.381	29.350
No. of observations	1677	1677	1677

Table 3.1.D
Empirical Distribution of NRCA Index for Each Year: Entire Sample Analysis

Summary Statistics/ Year	1998	1999	2000
p-1	-0.002	-0.002	-0.002
p-5	-0.001	-0.001	-0.001
p-10	-0.0003	-0.0003	-0.0003
p-25	-0.0001	-0.0001	-0.0001
p-50	-0.00001	-0.00001	-0.00001
p-75	0.00001	0.00001	0.00001
p-90	0.0002	0.0002	0.0002
p-95	0.0008	0.0008	0.0008
p-99	0.003	0.003	0.003
Maximum Value	0.008	0.009	0.008
Mean	<0.00001	<0.00001	<0.00001
Standard Deviation	0.001	0.001	0.001
Skewness	2.262	2.322	2.252
Kurtosis	36.507	40.212	40.405
No. of observations	1677	1677	1677

Table 3.1.E
Empirical Distribution of Log of Balassa Index for Each Year: Entire Sample Analysis

Summary Statistics/ Year	1998	1999	2000
p-1	-6.527	-5.904	-6.132
p-5	-3.671	-3.520	-3.631
p-10	-2.648	-2.627	-2.662
p-25	-1.364	-1.367	-1.383
p-50	-0.413	-0.434	-0.425
p-75	0.295	0.301	0.319
p-90	1.061	1.068	1.036
p-95	1.442	1.469	1.501
p-99	2.307	2.333	2.480
Maximum Value	4.651	4.729	4.723
Mean	-0.674	-0.652	-0.672
Standard Deviation	1.610	1.569	1.642
Skewness	-1.275	-1.127	-1.324
Kurtosis	6.433	6.034	7.307
No. of observations	1668	1673	1672

Table 3.1.F
Empirical Distribution of RTA Index for Each Year: Entire Sample Analysis

Summary Statistics/ Year	1998	1999	2000
p-1	-4.404	-4.027	-4.736
p-5	-1.351	-1.504	-1.401
p-10	-0.987	-1.014	-1.040
p-25	-0.601	-0.608	-0.610
p-50	-0.201	-0.200	-0.197
p-75	0.381	0.371	0.349
p-90	1.720	1.745	1.692
p-95	3.373	3.420	3.471
p-99	9.035	9.412	10.440
Maximum Value	61.144	69.308	66.721
Mean	0.227	0.249	0.268
Standard Deviation	2.673	2.835	3.154
Skewness	9.552	11.065	10.339
Kurtosis	180.080	228.632	173.340
No. of observations	1677	1677	1677

Table 3.1.G
Empirical Distribution of RC Index for Each Year: Entire Sample Analysis

Summary Statistics/ Year	1998	1999	2000
p-1	-6.362	-5.881	-5.798
p-5	-3.420	-3.274	-3.414
p-10	-2.358	-2.254	-2.351
p-25	-1.096	-1.148	-1.140
p-50	-0.313	-0.304	-0.322
p-75	0.416	0.396	0.385
p-90	1.398	1.384	1.360
p-95	2.092	2.103	2.009
p-99	4.140	3.858	3.917
Maximum Value	6.949	6.877	7.035
Mean	-0.423	-0.418	-0.446
Standard Deviation	1.722	1.656	1.691
Skewness	-0.608	-0.503	-0.760
Kurtosis	6.283	6.081	6.921
No. of observations	1668	1673	1672

p-1 to 99 in Tables 3.1.A to 3.1.G are the percentile points and they give detailed information of the cumulative distributions of the RCA indices. A value of 0.252 corresponding to the p-25 point for the index of Balassa in the year 1998 signifies that 25% of the observations have index values below 0.252 and 75% of the observations have index values above 0.252. Although the percentiles appear to vary between the years, particularly towards the upper ends of the distributions, the differences do not seem to be very significant. Thus, the cumulative distributions can be regarded as stable over time for the Balassa index. Similar observations can be made about the cumulative distributions of the other indices.

Considering the summary statistics for the index of Balassa in Table 3.1.A, the computed arithmetic mean is above the comparative advantage neutral point of unity (approximately) and is seen to fluctuate over time.¹¹ However, the differences in its magnitudes are quite small. The levels of skewness and kurtosis are well beyond the demarcated values of 0 and 3 for a symmetrical normal distribution in case of all the three years. In fact, apart from being skewed, a large kurtosis implies that a sizeable portion of the observations for the index is

¹¹ A little deviation of the calculated arithmetic mean from the value of unity is expected as the reference group in the index calculation includes all countries and all commodities.

towards the end of the distribution. A mean value of 1.287 greater than median value of 0.659 for the year 1998 implies the distribution is skewed to the right, and this holds true for the other years also. The median also does not seem to fluctuate to a significant extent. In case of RSCA index in Table 3.1.B, the arithmetic means are different from the comparative advantage neutral point of zero (approximately), but they seem to be quite stable over the years.¹² The extent of asymmetry in the distribution and the degree of “peakedness” are also not high. In fact they are closer to the prescribed value of 0 and 3 for a normal distribution. The means and the medians although not equal to each other, do not seem to differ to a significant extent. Thus the empirical distribution of the RSCA index is almost at par with its theoretical distribution. For the ARCA index in Table 3.1.C, the mean does not significantly differ from the comparative advantage neutral point of zero (approximately) for each year. The median is also observed to be quite stable over the years. The measures on skewness and kurtosis, though not so large as that of Balassa index, are certainly greater than that of RSCA index, and are quite different from the prescribed values of 0 and 3 for a normal distribution. Almost similar conclusions emerge while analyzing the summary statistics of the NRCA index in Table 3.1.D, although it seems to be more symmetrically distributed compared to the ARCA index. Greater asymmetry in the empirical distribution of the ARCA index is probably due to the clubbing of countries and sectors in a particular year. The theoretical structure of the ARCA index ensures that its empirical distribution across sectors will be symmetric, but the same need not follow in case of its empirical distribution across countries. Both Log of Balassa and RC indices in Tables 3.1.E and Tables 3.1.G respectively report minor fluctuations in arithmetic mean and median values over time. The computed means are also different from the comparative advantage neutral point of zero to some extent.¹³ However, the numbers of observations differ between the years for the two indices. Taking into consideration the levels of skewness and kurtosis and also minor differences between the means and medians for each of the two indices, the empirical distributions of both Log of Balassa and RC indices can be conceived to be a little close to a normal distribution.

¹² Deviations to some extent from the value of zero are expected as the reference group includes all countries and all commodities.

¹³ The comparative advantage neutral point is approximately zero in case of Log of Balassa index. But equal to zero in case of RC index.

However, their performance in this respect is still short of the RSCA index. Like most indices, the RTA index in Table 3.1.F records minor fluctuations in the mean and median values over time, which implies overall stability in its distributions. The levels of skewness and kurtosis recorded by the index are higher than any other index, excepting the index of Balassa. Additionally, the mean values for each year being greater than the median, confirms the fact that in general the distributions are skewed to the right. Thus the empirical distribution of the RTA index is not symmetrically distributed. Therefore an analysis of Tables 3.1.A to 3.1.G suggests that over time the empirical distributions are quite stable for the RCA indices. The empirical distribution of RSCA index seems to be most well behaved, for not only it is stable over time, but approximately normal. The Log of Balassa index and RC index also presumably fit into the category of normal distribution but certainly not as well as the RSCA index. The NRCA index, although characterized by a stable symmetrical empirical distribution, cannot be considered to be normally distributed. Among the indices analyzed, only the ARCA index and NRCA index report means substantially close their comparative advantage neutral points of zero as expected theoretically.

The findings corresponding to each index in Tables 3.1.A to 3.1.G relate to all countries and all sectors for each particular year. The cumulative distribution plots and the Kernel density plots across 28 sectors to be considered subsequently, are specific to a single country, and represent the shifts in distributions over 1998, 1999 and 2000 in case of each index. As previously mentioned, the countries sampled for analysis include China, France, Germany, India, Italy and Japan. The cumulative distribution plots for each country are presented in the Figures B.1 to B.6 in the Appendix B.

Contrary to findings in Table 3.1.A and as also claimed by Hinloopen and Marrewijk (2001) variability in cumulative distributions of Balassa index to certain extent are observed for the considered sample. Only for France, the cumulative distributions for the Balassa index are similar to the extent that they are almost indistinguishable for each year. The RSCA index has quite noticeable variations in the distributions over the years for India. For other countries the distributions seem to vary over time, but the differences do not appear to be as large as noticed in case of India. The ARCA index exhibits noticeable stability of distributions over the years, particularly for Germany, France and Japan. Similar

observations can be made about the NRCA index. For other countries both the indices report very insignificant variations in distributions over time. Log of Balassa index do not seem to perform well in generating stable distributions over the years, as observed in case of China, Germany India, and Japan. For Italy and France, the index however exhibits better performance. The distributions for RTA index can also be distinguished from year to year except for France. Similarity in distributions of RC index over 3 years is only observed for Italy, and over 1998 and 1999 for France. Thus among all the 7 indices analyzed, the cumulative distribution plots hint towards the fact that only the ARCA and NRCA index values for a country can reliably be compared over time, as they feature stable distributions across countries analyzed. The cumulative distributions presented in Tables 3.1.A to 3.1.G however represent all of the indices to be largely stable over time.

The Kernel density plots are illustrated in the Figures B.7 to B.12 in the Appendix B for the same set of countries over the same set of years. For each index in the Figures B.7 to B.12, the comparative advantage neutral point can be considered to depict a fixed demarcation value and shifts in density functions with reference to that demarcation point can be noted to determine the stability of the density functions. For all countries, excepting Germany and Japan, the density plot for the Balassa index is skewed to the right. The shifts in density plots to the right of the demarcation point of 1 over the years though not so explicit, but to a certain extent are observed for Germany, India and Italy. In case of India and Italy the shifts are particularly apparent for the year 2000 and for Germany for the year 1999, although there is some amount of ambiguity in the shifts of the plots. Tests of significance to be considered in the next section will be able to determine statistical significance of the shifts. The shift of the plots to the right is indicator of an increase in the number of sectors with index value greater than 1. The plots for the RSCA index are largely symmetric in case of France. Shifts in the density functions for RSCA index over the years to the right of the demarcation point of zero, are most visible for India, and to some extent for Germany. The ARCA and NRCA indices do not exhibit any significant shifts in the density plots. They also seem to be largely symmetric around the demarcation point of zero, with exceptions in case of Japan and Italy. The Log of Balassa index generates an apparently symmetrical probability density function only in case of France, for in other cases some amount of skewness towards the left is observed. Shifts in the density functions to the right of the demarcation point zero are

apparent for India. However, the shape of the function for 1998 differs to some extent from that of the years 1999 or 2000 for India. The RTA index features some amount of asymmetry in its density plots particularly for China, Japan and Germany. Such observed asymmetry in the empirical distribution of the RTA index was also noted in Table 3.1.F, although theoretically the index is supposed to be symmetrically distributed about the value zero. Shifts in density plots towards left and towards right are observed for Germany and for Japan respectively. The RC index is characterized by asymmetry in its density functions and the shifts do not seem to be significant over time.

Thus the graphical analysis produces some contradictory results as compared to the results emerging from Tables 3.1.A to 3.1.G. The differences are expected as the Tables 3.1.A to 3.1.G relate to all countries and all sectors together, while the graphical analysis relates to a particular country and hence are more relevant for case by case policy analysis. In this respect, both NRCA and ARCA indices being generally characterized by stable symmetrical empirical distributions across sectors is no anomaly. Both the Kernel density plots and summary statistics in Tables 3.1.C and 3.1.D however agree to the fact that the distributions of both the indices are far from being normal. Nevertheless, for cross country, cross sector or cross time comparisons, distributions of RCA indices need not be normal. They must however be stable. Asymmetrical empirical distributions of RSCA index, Log of Balassa index and RC index in case of some Kernel density plots although not apparent from Tables 3.1.B, 3.1.E and 3.1.G, is a noteworthy fact and can be attributed to country specific behavior of data.

3.3.2 Discussion of Results for the Tests of Significance

Statistical significance of the shifts in empirical distributions of RCA indices over time, sectors and countries are analyzed in this subsection.

3.3.2.1 Results for Wilcoxon's Signed Rank Test

Since the Wilcoxon's signed rank test permits comparison of only two samples at a time, the 3 years are considered in a group of two, in order to run the test. For each country in each year, the number of sectors is 28. Shifts of the country distributions across sectors over time are analyzed using the signed rank test. Table 3.2 reports the numbers of cases where the null

hypothesis of equality of distributions can be accepted for each country as well as for all countries together. Based on the combination of years, for each country, number of cases analyzed are 3. The reported numbers of cases where the null hypothesis gets accepted for each country in Table 3.2 are out of those 3 cases. Based on the combination of years and number of countries, the total numbers of cases analyzed are 18. The reported numbers of cases where the null hypothesis gets accepted for all countries together are out of those 18 cases. The detailed results for the test for each country and each index are provided in Table B.3 in the Appendix B.

Table 3.2
Results Summary for Wilcoxon's Signed Rank Test for Countries

Country/ Index	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
China	3	3	3	3	3	3	3
Germany	3	3	3	3	3	3	3
France	3	3	3	3	3	3	3
India	1	1	3	3	1	2	3
Italy	1	2	3	3	2	3	3
Japan	3	3	3	3	3	3	3
Total no. of accepted cases for null hypothesis	14	15	18	18	15	17	18

Note: Reported numbers of cases for each index are out of 3 for each country, and out of 18 for all countries together.

As evident from Table 3.2, maximum cases of rejection of the null hypothesis are observed for India followed by Italy. But for all other countries the test reports stable distributions over time for each of the indices considered. The results therefore differ to some extent from the drawn inference from cumulative distribution plots and Kernel density plots. The signed rank test seems to project that many of the differences noticed with cumulative distribution plots and Kernel density plots are not in fact statistically significant. Particularly in case of cumulative distribution functions, shifts in distributions are observed for RTA and RC indices for most countries, but the signed rank test posit such shifts to be statistically insignificant. Shifts in cumulative distribution functions are also noted with Log of Balassa index for China, Germany and Japan and to no extent for Italy. But the signed rank test reveals the shifts to be statistically insignificant for China, Germany and Japan, but

significant for Italy. On the basis of Kernel density plots for Balassa index, shifts in distributions are noted for Germany, India and Italy. The signed rank test however reports the shift to be statistically insignificant for Germany but significant for India and Italy. Shifts in Kernel density plots were also noted in case of Germany for RSCA index, and Germany and Japan for RTA index. But the signed rank test does not show the shifts to be statistical significant. For India however, there seems to be some parity in the graphical analysis and the test of significance. Both graphical analysis and signed rank test agree to the fact that ARCA index and NRCA index are the most stable over time. RC index is also found to be as stable as the ARCA index or NRCA index on the basis of signed rank test, but the same is not supported by cumulative distribution plots, although to some extent by Kernel density plots. This can be attributed to the fact that the shifts in Kernel density plots are more in line with the shifts in means of two distributions. Since the signed rank tests studies the shifts in means of two distributions, the results of Kernel density plots and signed rank test can be expected to coincide. The shifts in the cumulative distribution plots are however not exclusive to changes in means and therefore cumulative distribution plots may not be at par with the signed rank test. According to Table 3.2 RTA index emerge to be the second most stable index over time. The RSCA index and Log of Balassa index do not perform as good as the previously mentioned indices. Balassa index seem to be generating the largest number of cases where the null hypothesis gets rejected. Hence, although Yu *et al.* (2009) argue the distribution of the NRCA index to be time invariant, based on the analysis in this section, the ARCA index and RC index are found to be equally good substitutes for the NRCA index. However, the comparison of Balassa index values over time must be done with care.

To test the stability of the index distributions over years for separate sectors using the signed rank test, 7 sectors with ISIC codes – 321, 322, 323, 324, 351, 352 and 385, are considered. For each of the selected sectors the index distributions over the years are analyzed. For each sector in each year the numbers of countries are 60. Table 3.2.1 reports the numbers of cases where the null hypothesis of equality of distributions can be accepted for each sector as well as for all sectors together. Based on the combination of years, for each sector, number of cases analyzed are 3. The reported numbers of cases where the null hypothesis gets accepted for each sector in Table 3.2.1 are out of those 3 cases. Based on the combination of years and number of sectors, the total numbers of cases analyzed are 21. The reported numbers of cases

where the null hypothesis gets accepted for all sectors together are out of those 21 cases. The detailed results for the test for each sector and each index are provided in Table B.4 in the Appendix B.

Table 3.2.1
Results Summary for Wilcoxon's Signed Rank Test for Sectors

ISIC Code/ Index	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
321	3	2	2	3	2	2	0
322	3	3	3	3	3	3	0
323	3	2	3	3	3	3	2
324	1	1	3	3	1	1	1
351	3	3	3	3	3	3	3
352	2	2	2	3	2	3	3
385	3	3	0	2	3	3	3
Total no. of accepted cases for null hypothesis	18	16	16	20	17	18	12

Note: Reported numbers of cases for each index are out of 3 for each sector, and out of 21 for all sectors together.

The findings of Table 3.2.1 are to some extent different from that of Table 3.2. NRCA index is clearly found to be superior to any other index. Thus, it features the most stable distribution over time for different sectors. Balassa and RTA indices have the second most stable distributions over time. Although ARCA and RC indices are found to be as good as the NRCA in Table 3.2, the same does not hold in Table 3.2.1.

3.3.2.2 Results for Wilcoxon's Rank Sum Test

To evaluate the stability of index distributions over sectors and over countries the Wilcoxon's rank sum test is used. Tables 3.3 and 3.4 report the number of cases where the null hypotheses of equality of distributions over sectors and over countries respectively for each index are accepted. Since, the Wilcoxon's rank sum test permits testing the distributions of only two samples at a time, the test is performed by considering two sectors or two countries at a time. As considered in Table 3.2.1, in Table 3.3, results corresponding to only 7 sectors with ISIC codes 321, 322, 323, 324, 351, 352 and 385 are presented. The sectors

being considered in a group of two for the test, the number of reported cases for acceptance of null hypothesis, are out of a total of 21 cases for each index. For each sector, the numbers of countries are 60. In Table 3.4, results corresponding to 6 countries – China, Germany, France, India, Italy and Japan are presented. The countries being considered in a group of two for the test, the numbers of reported cases for acceptance of null hypothesis, are out of a total of 15 cases for each index. For each country the numbers of sectors are 28. Results are reported in Tables 3.3 and 3.4 for the year 2000 only. The detailed results for the test for each index are reported in Tables B.5 and B.6 in the Appendix B.

Table 3.3
Results Summary for Wilcoxon's Rank Sum Test for Sectors

Summary/ Index	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
No. of accepted cases for null hypothesis	9	9	8	17	9	10	9

Note: Reported numbers of cases are out of a total of 21 cases for each index.

In Table 3.3, the NRCA index is found to be most stable over sectors. Since the Wilcoxon's rank sum test studies the shifts in the distributions due to shifts in the mean or the median (assuming the distributions are of identical shapes), the results corresponding to NRCA index in Table 3.3 is consistent with the observation that the arithmetic mean of the index values across countries is stable between sectors with the value being fixed at zero. Hence, stability of the mean serves dual purpose – as a cardinal measure and as an ordinal measure. It is not only possible to reliably determine the extent of comparative advantage of one country over another in a sector, it is also possible to rank different sectors for a country reliably. The other indices report larger cases for rejection of the null hypothesis for equality of distributions. Since it has been previously claimed that none of the indices other than the NRCA index, generate stable mean across countries with respect to a sector, larger cases of rejection of null hypothesis for those indices should not come as a surprise.

Table 3.4
Results Summary for Wilcoxon's Rank Sum Test for Countries

Summary/ Index	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
No. of accepted cases for null hypothesis	10	10	14	13	10	12	8

Note: Reported numbers of cases are out of a total of 15 cases for each index.

As per the results reported in Table 3.4, the ARCA index is found to be most stable over countries. Since, theoretically the arithmetic mean of the ARCA index values across sectors for a particular country is zero and stable over countries, the results corresponding to ARCA index is as per expectations. NRCA index marginally falls short of the ARCA index in terms of stability of empirical distributions. Thus, taking into consideration the results presented in Table 3.4, the ARCA index is reliable for comparing index values across sectors within a country, i.e., the index serves as a cardinal measure with respect to a country. Also, the position of a country relative to any other country with respect to a sector can be determined with ARCA index. This implies reliability of the index towards ranking of countries in a particular sector. Thus, the ARCA index may serve the purpose of an ordinal measure with respect to a sector. NRCA index may feature a stable mean of zero across sectors, but might not be as reliable as the ARCA index as a cardinal measure with respect to a country or as an ordinal measure with respect to a sector. As observed in Table 3.4, larger cases of rejection of the null hypothesis are noted for the other indices and this directly follows from the fact that the arithmetic mean of these remaining indices vary from country to country.

3.3.2.3 Results for Kolmogorov Smirnov Test

The Wilcoxon's rank sum test results corresponding to Tables 3.3 and 3.4 are compared with the Kolmogorov Smirnov test results reported in Tables 3.5 and 3.6. Table 3.5 reports the numbers of cases for acceptance of the null hypothesis of equality of distributions over sectors for each index. Table 3.6 reports the number of cases for acceptance of the null hypothesis of equality of distributions over countries for each index. Hence, the results reported in Tables 3.5 and 3.6 can be compared with those reported in Tables 3.3 and 3.4 respectively. As in Tables 3.3 and 3.4, the number of reported cases for acceptance of null

hypotheses in Tables 3.5 and 3.6 are out of total of 21 and 15 cases respectively, for each index. Detailed results for the test are presented in Tables B.7 and B.8 in the Appendix B.

Table 3.5
Results Summary for Kolmogorov Smirnov Test for Sectors

Summary/ Index	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
No. of accepted cases for null hypothesis	6	6	2	14	6	8	4

Note: Reported numbers of cases are out of a total of 21 cases for each index.

Compared to Table 3.3, fewer cases for acceptance of the null hypothesis are reported in Table 3.5 for all indices. However, the NRCA index still continues to be featured by the most stable distribution over sectors. Hence, both the Wilcoxon's rank sum test and Kolmogorov Smirnov test agree to the fact that NRCA index is most reliable as an ordinal measure with respect to a country. On referring to Tables B.5 and B.7 in the Appendix B, it will be interesting to note that the value of the test statistic corresponding to each test is same for the Balassa index, RSCA index and the Log of Balassa index for each group of sectors. Since, both the rank sum test and Kolmogorov Smirnov tests are based on the ranks of the observations, the test results are unaffected by changes in scale of the variables. The RSCA and Log of Balassa indices being quasi logarithmic and logarithmic transformations of the Balassa index respectively, involve changes in scale of the variable and that will not alter the rank sum or the maximum difference between the cumulative distributions of the indices. Although in general, compared to Table 3.3, an increase in the number of cases where it is possible to reject the null hypothesis is observed in Table 3.5, on considering the Tables B.5 and B.7 in the Appendix B, one cannot ignore a few cases where the test statistic for Kolmogorov Smirnov test is statistically insignificant, but significant in case of Wilcoxon's rank sum test.¹⁴ This can be attributed to the fact that compared to Wilcoxon's rank sum test, the Kolmogorov Smirnov test is based on many parameters that contribute to deviations from the null hypothesis. Therefore, Kolmogorov Smirnov test has less power to detect a change in distribution exclusively due to a shift in the mean or median but more power to detect

¹⁴ For example in the cases of Balassa index (321-352), RSCA index (321-352), Log of Balassa index (321-352), RTA index (321-323 and 324-352) RC index (322-324).

changes in the shape of the distribution (Lehman 2006). Following this argument, the cases where the null hypothesis of equality of distribution gets rejected in case of both Kolmogorov Smirnov and Wilcoxon's rank sum tests are probably those cases for which, changes in population distributions are not only due to shifts in location but also the shape. The cases where the null hypothesis is rejected by Wilcoxon's test but accepted by Kolmogorov Smirnov's test are those where the population distributions have almost similar shapes but differ with respect to their location, which implies differing means or medians. And finally, the cases where the null hypothesis is accepted by Wilcoxon's test but rejected by Kolmogorov Smirnov's test are probably those where population distributions have differing shapes but similar means or medians.

Table 3.6
Results Summary for Kolmogorov Smirnov Test for Countries

Summary/ Index	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
No. of accepted cases for null hypothesis	10	10	14	12	10	12	6

Note: Reported numbers of cases are out of a total of 15 cases for each index.

In Table 3.6, the numbers of accepted cases for null hypothesis of equality of distributions are same as in Table 3.4 for all but the indices of NRCA and RC. The ARCA index continues to remain the most stable index over countries as in Table 3.4. Thus, the stability of its distribution ensures the index can be used as an ordinal measure with respect to a sector. Considering the Tables B.6 and B.8 in the Appendix B it will be possible to note that the null hypothesis is rejected by the Wilcoxon's test but accepted by the Kolmogorov Smirnov's test while analyzing the equality of population distributions in case of China and Japan for the indices of Balassa, RSCA, Log of Balassa and RTA indices. Similar observations are also noted for the country combinations Germany and Italy (for the indices of Balassa, RSCA and Log of Balassa), India and Japan (for RTA index), and Italy and Japan (for the indices of RTA and RC). As argued in the previous paragraph, this can be attributed to the potentiality of Kolmogorov Smirnov test to identify more changes in shape of the distributions rather than their locations.

Although the conclusions that emerge from both the Wilcoxon's rank sum test and Kolmogorov Smirnov's test do not differ, it seems changes in the distributions of indices due to changes in the mean will be more relevant for the study under consideration. Changes in mean value for a particular index have important implications for the usage of the index as a cardinal or as an ordinal measure. In this respect shifts in distributions exclusively due to changes in arithmetic mean must be considered more important than shifts in distributions which can be due to changes in mean or standard deviation or skewness or kurtosis. Hence, preference can be assigned to results corresponding to the Wilcoxon's rank sum test which reports NRCA index to be most stable over sectors and ARCA index to be most stable over countries. However, since the NRCA index only marginally falls short of the ARCA index while analyzing the stability of the index distributions over countries using Wilcoxon's rank sum test, one can consider NRCA index as good as the ARCA index.

3.4 Conclusion

The empirical distributions of RCA indices are analyzed in this paper with the primary objective of identifying an index which can reliably be used as a cardinal and ordinal measure over time. The country and sector specific results, which are more relevant for policy analysis, uphold the superiority of NRCA index over all other indices taken up for examination in this paper. Thus the index can be used as an ordinal measure for ranking sectors with respect to a country or for ranking countries with respect to a sector. The index can also be used as a cardinal measure for comparing countries with respect to a sector or for comparing sectors with respect to a country. Further, its empirical distribution is time stable which ensures the usage of the index for time series analysis. The RCA index of Balassa and all other subsequently suggested modifications of Balassa index could not compete with the NRCA index in these respects. However, one major drawback of the NRCA index is its non-normal distribution which although not required for usage of the index as a cardinal or ordinal measure, is definitely essential in case of parametric tests necessitating normally distributed variables.

The conclusions drawn in this chapter are dependent upon the dataset used. An extension of dataset both in terms of time, sectors and countries will enable a more detailed analysis on the empirical distributions of RCA indices.

On the basis of discussions in this chapter, if one has to select one RCA index which can consistently be used for over countries, over sectors and over time comparison, the choice will be the NRCA index of Yu *et al.* (2009). However, to be identified as an “ideal” index, the NRCA index must also be consistent with the theories of comparative advantage, which will be taken up for analysis in later chapters.



Chapter 4

Consistency of RCA Indices with the Theories of Comparative Advantage

4.1 Introduction

The Revealed Comparative Advantage (RCA) indices have gained importance over time as a tool for understanding the potential of gainful trade amongst countries on the basis of their respective comparative advantages. Stated alternatively, various RCA indices are expected to identify products or sectors in which any particular country will be incurring lower opportunity costs compared to other countries. Balassa (1965) justified the usage of trade based measures on the ground that they tend to reflect the underlying relative cost differences of countries, and hence will be able to depict the comparative advantage of countries in various commodities. Balassa's proposition invited analytical literature on the theoretical foundation of the RCA index of Balassa. For instance, Hillman (1980) and Bowen (1983) argued that the index is not well integrated into the theory. While Hillman attempted to determine the consistency of the index with the pre-trade relative prices, Bowen questioned the assumptions on which the structure of the index is based. These studies however do not explore the extent to which Balassa's index reflect the comparative advantage theories of Ricardo or Heckscher-Ohlin by making use of real life data. Given the fact that Balassa's index has been modified in various attempts to overcome the certain deficiencies of the index, it is indeed an important task to explore the relevance of Balassa index and all such alternative RCA indices with the comparative advantage theories. The objective for such an analysis is to identify one or more RCA index which is well integrated into the theory and can be reliably used by the policy makers for making policy recommendations.

As pointed out in chapter 2 (p. 28-29, 44-47), some attempts have been made in the past to link the RCA indices with the comparative advantage theory empirically. Yeats (1985) for instance, established the consistency of the Balassa index with the factor proportions theory using the data on 40 manufacturing industries and 47 countries. Leromain and Orefice (2014)

suggested an index which incorporates the Ricardian productivity differences in its structure. But the deficiencies of these studies, as recognized in chapter 2 (p. 29, 46-47), provides impetus to consider the trade based indices from a different angle and analyze their consistency with the theories on comparative advantage.

Examining the consistency of RCA indices with the factor productivity differences or factor endowment differences is quite justified. A good export performance of a country can be due to various factors that are not directly linked to comparative advantage of countries (Leromain and Orefice, 2014). Factors such as formal or informal trade barriers, historical trade relationships, internal demand shocks and differences in preferences, can be identified under this category. Leromain and Orefice cites the case of India whose export of tea to United Kingdom could be due to its past colonial ties with the country and this factor is definitely unrelated to the productivity differences between the two countries. In this context, one should try and analyze how much of the export performance of a country measured by its RCA index, is due to the Ricardian or Heckscher-Ohlin comparative advantage of countries. If a significant association can be established between the RCA index values and the factor productivity differences or factor endowment differences, then it will be possible to conclude that RCA indices are consistent with theories of comparative advantage and hence the usage of the term “Revealed Comparative Advantage” is justified. However as recognized in chapter 1 (p. 12), import figures could be distorted due to the presence of trade barriers. In this context, it seems necessary to study the behavior of RCA indices based on both exports and imports, after they have been adjusted for such barriers, and not merely draw conclusions on the basis of unadjusted RCA indices.

As in chapter 3, the RCA indices to be studied in this chapter are, the RCA index of Balassa, the Revealed Symmetric Comparative Advantage (RSCA) index by Dalum *et al.* (1998) and Laursen (1998), the Additive Revealed Comparative Advantage (ARCA) index by Hoen and Oosterhaven (2006), the Normalized Revealed Comparative Advantage (NRCA) index by Yu *et al.* (2009), and the Log of Balassa index, the Revealed Trade Advantage (RTA) index and the Revealed Competitiveness (RC) index based on Vollrath (1991). The indices take the following forms –

- 1) *Balassa's RCA Index* = $\frac{X_a^i/X_t^i}{X_a^w/X_t^w}$
- 2) *RSCA Index* = $\frac{\frac{X_a^i/X_t^i}{X_a^w/X_t^w} - 1}{\frac{X_a^i/X_t^i}{X_a^w/X_t^w} + 1}$
- 3) *ARCA Index* = $\frac{X_a^i}{X_t^i} - \frac{X_a^w}{X_t^w}$
- 4) *NRCA Index* = $\frac{X_a^i}{X_t^w} - \frac{X_t^i X_a^w}{X_t^w X_t^w}$
- 5) *Log of Balassa Index* = $\ln \frac{X_a^i/X_t^i}{X_a^w/X_t^w}$
- 6) *RTA Index* = $\frac{X_a^i/X_t^i}{X_a^w/X_t^w} - \frac{M_a^i/M_t^i}{M_a^w/M_t^w}$
- 7) *RC Index* = $\ln \frac{X_a^i/X_t^i}{X_a^w/X_t^w} - \ln \frac{M_a^i/M_t^i}{M_a^w/M_t^w}$

Index numbers 1 to 5 are export based indices while the last two are based on both exports and imports. In all the above index specifications, i denotes any particular country and a denotes a particular sector. t is the sum of all sectors and w is the world total. Therefore X_a^i is the exports of product a by country i , X_t^i is the exports all products by country i , X_a^w is the exports of product a by all countries and X_t^w is the exports of all products by all countries. M_a^i , M_t^i , M_a^w and M_t^w can be similarly defined for imports.

This chapter is henceforth divided into a few more sections and subsections. Section 4.2 elaborates on testing the consistency of RCA indices with the comparative advantage theories. The data and methods of analyses are discussed in subsection 4.2.1. The results of analyses are provided in subsection 4.2.2. Section 4.3 discusses the implications on the performance of RCA indices when the import values are adjusted for trade barriers. The data and methods of analysis are discussed in subsection 4.3.1, while results are presented in subsection 4.3.2. Section 4.4 concludes the chapter.

4.2 Testing the Consistency of RCA Indices with the Theories

The objective of this chapter is to test for the consistency of various RCA indices with the Heckscher-Ohlin and Ricardian theories on comparative advantage. In the process, the chapter tends to identify one or more of such indices which incorporate the basis of the

theories in their structure. After a discussion on the applicable methodology and data utilized, the results will be analyzed to determine such indices.

4.2.1 Methods and Data

As already discussed, the Heckscher-Ohlin theory emphasized the differences in relative factor endowments of countries to be the reason for inter-industry trade on the basis of comparative advantage. The Ricardian theory on the other hand, based its argument on the differences in technology or relative factor productivity of countries. To test the consistency of the indices with the Heckscher-Ohlin theory, the relative endowments of factors of production will be considered. Accordingly, with two factors of production labor and capital, if the ratio of labor force of a country to the world labor force exceeds the ratio of capital stock of the same country to the world capital stock, the country will be considered to be labor abundant, and will exhibit comparative advantage in labor intensive commodities. On the other hand, if the ratio of labor force of a country to the world labor force falls short of the ratio of capital stock of the same country to the world capital stock, the country will be deemed to be capital abundant, thereby exhibiting comparative advantage in capital intensive commodities.

Mathematically, country i is labor abundant relative to the world w if,

$$\frac{L^i}{L^w} > \frac{K^i}{K^w} \quad (4.1)$$

Or else, country i is capital abundant relative to the world w if,

$$\frac{L^i}{L^w} < \frac{K^i}{K^w} \quad (4.2)$$

Based on expressions (4.1) and (4.2), a single expression for factor endowment ratio for country i can be expressed as:

$$FER^i = \frac{L^i/L^w}{K^i/K^w} \quad (4.3)$$

Country i is labor abundant and will be exporting labor intensive products if FER^i exceeds unity. Country i will be capital abundant and export capital intensive products if FER^i falls

short of unity. In expressions (4.1) to (4.3), L and K are used to denote the labor force and capital stock respectively. The superscripts i and w represent the country i and world respectively.

The expressions (4.1) to (4.3) can be derived from Leamer (1980) who made use of the Heckscher-Ohlin-Vanek (HOV) theorem to arrive at a testable hypothesis about the relative factor contents in production and consumption. Assuming two factors of production L and K , the HOV theorem for each factor can be expressed mathematically as:

$$F_L^i = L^i - s^i L^w \quad (4.5)$$

$$F_K^i = K^i - s^i K^w \quad (4.6)$$

F_L^i and F_K^i are the computed labor and capital contents of trade respectively for country i . s^i is the country i 's share of world Gross Domestic Product (GDP). L^i , L^w , K^i , K^w have been previously defined in the context of expressions (4.1) – (4.3). According to the HOV theorem, country i will be abundant in labor if $F_L^i > 0$. From expression (4.5), labor abundance for country i therefore implies

$$\frac{L^i}{L^w} > s^i \quad (4.7)$$

That is, country i 's endowment of labor, relative to the world endowment, exceeds the share of country i in world GDP (Feenstra, 2004). Similarly, country i will be capital abundant if $F_K^i > 0$. Therefore from expression (4.6) it follows,

$$\frac{K^i}{K^w} > s^i \quad (4.8)$$

That is, country i 's endowment of capital relative to the world endowment exceeds i 's share in the world GDP.

Leamer (1980) thereby deduced that if labor-capital ratio of country i exceeds that of the world, then the HOV theorem will imply that the labor-capital ratio contained in production exceeds the labor-capital ratio contained in consumption of country i . Leamer's statement can be expressed as,

$$\frac{L^i}{K^i} > \frac{L^i - F_L^i}{K^i - F_K^i} \quad (4.9)$$

The left hand side of expression (4.9) is the labor capital endowment ratio which can alternatively be interpreted as the labor capital ratio utilized in production (Feenstra 2004). The right hand side depicts the labor capital ratio utilized in consumption. Using expressions (4.5) and (4.6) and rearranging terms, from expression (4.9) the following relation can be deduced:

$$\frac{L^i}{K^i} > \frac{L^w}{K^w} \quad (4.10)$$

From expression (4.10), expression (4.1) can be derived. Using the above interpretation of the HOV theorem, Leamer proved that “Leontief paradox” did not exist, as US in 1947 was indeed capital abundant and exporting capital intensive goods.¹ Thus, Leamer upheld the predictions of the Heckscher-Ohlin model. Expression (4.10) therefore is a valid representation of the Heckscher-Ohlin model as Leamer applies the relation to provide evidence in favor of Heckscher-Ohlin model. Hence, usage of factor endowment ratio to characterize Heckscher-Ohlin model is not unreasonable.

To test the consistency of the RCA indices with the Ricardian theory of comparative advantage, the average labor productivity ratio is considered. If the ratio of average labor productivity of a country in particular sector relative to its average labor productivity in all sectors, exceeds the average labor productivity of the world in that particular sector relative to its average labor productivity in all sectors, the considered country will be exhibiting comparative advantage in the considered sector. If on the other hand, the ratio of average labor productivity of a country in a particular sector relative to its average labor productivity in all sectors, falls short of the average labor productivity of the world in that particular sector relative to its average labor productivity in all sectors, the country will be exhibiting comparative disadvantage in the considered sector.

¹ Leontief (1953) using 1947 data found that the capital labor ratio entering US exports was lower than the capital labor ratio entering US imports. Based on the assumption that US is capital abundant, Leontief’s results contradicted the Heckscher-Ohlin theorem. For that reason Leontief’s findings came to be known as the “Leontief paradox.”

Mathematically, country i will have comparative advantage in sector a relative to the world w if,

$$\frac{APL_a^i}{APL_t^i} > \frac{APL_a^w}{APL_t^w} \quad (4.11)$$

Or else, country i will be having comparative disadvantage in sector a relative to the world w if,

$$\frac{APL_a^i}{APL_t^i} < \frac{APL_a^w}{APL_t^w} \quad (4.12)$$

Expression (9) or (10) can be combined into a single expression for average productivity of labor ratio for country i in sector a , and can be expressed as,

$$APLR_a^i = \frac{APL_a^i / APL_t^i}{APL_a^w / APL_t^w} \quad (4.13)$$

Here APL_a^i is the average product of labor in sector a in country i , APL_t^i is the average product of labor across all sectors in country i , and APL_a^w and APL_t^w can similarly be defined for the world or all countries together. If $APLR_a^i$ takes a value greater than unity, country i will have comparative advantage in sector a . If $APLR_a^i$ takes a value less than unity, country i will be having comparative disadvantage in sector a .

The expressions (4.11) to (4.13) follow from the standard two country - two commodity framework assumed by Ricardo. Assuming two countries A and B and two sectors X and Y, the labor requirement matrix can be represented as,

country/ sector	X	Y
A	a_{LX}	a_{LY}
B	b_{LX}	b_{LY}

a_{LX} and a_{LY} are the units of labor required to produce a unit of products X and Y respectively in country A. b_{LX} and b_{LY} are the labor requirements for a unit of X and Y respectively in country B. Corresponding to the above matrix specification, the opportunity cost of producing X and Y in country A and B can be represented as,

<i>country/ sector</i>	<i>X</i>	<i>Y</i>
<i>A</i>	a_{LX}/a_{LY}	a_{LY}/a_{LX}
<i>B</i>	b_{LX}/b_{LY}	b_{LY}/b_{LX}

Country A will be having comparative advantage in sector X if its opportunity cost of producing X is less than that of B's. Mathematically this criterion can be expressed as,

$$\frac{a_{LX}}{a_{LY}} < \frac{b_{LX}}{b_{LY}} \quad (4.14)$$

Rearranging the terms, expression (4.14) can represent the relation between the average labor productivity of the two countries in each sector, as expressed below.²

$$\frac{APL_X^A}{APL_Y^A} > \frac{APL_X^B}{APL_Y^B} \quad (4.15)$$

It is possible to revert back to expression (4.11) from expression (4.15) by considering appropriate changes to the country and commodity notations.

Marginal product of labor instead of average product, might be a better determinant of comparative advantage. However, data limitations prevent us from directly observing the marginal product. Nevertheless, for the present study, it is assumed that these two measures are correlated with each other. It may also be noted that other studies, for instance, MacDougall (1951), Balassa (1963) and Golub and Hsieh (2000) have used output per worker to represent labor productivity while empirically analyzing the Ricardian model. Hence, using average product to represent Ricardian model is nothing new.

Productivity in this chapter is measured as value-added per worker following Balassa (1963) and Golub and Hsieh (2000). Value-added is the gross output excluding intermediate inputs other than labor. Value-added in an industry is the contribution of that industry to sectoral output or Gross Domestic Product (Cobbold, 2003). For measuring labor productivity, value added in production has greater appeal than gross output. In the presence of outsourcing, changes in estimated labor productivity based on gross output, may be due to substitution

² From expression (4.14), $\frac{1/a_{LX}}{1/a_{LY}} > \frac{1/b_{LX}}{1/b_{LY}} \Rightarrow \frac{APL_X^A}{APL_Y^A} > \frac{APL_X^B}{APL_Y^B}$

between labor and intermediate inputs, and not due to changes in technology or efficiency that must get reflected through changes in labor productivity (Cobbold, 2003). For instance, if outsourcing reduces labor inputs and increases intermediate inputs in production leaving gross output constant or increased, labor productivity measured by gross output per unit of labor input, will increase. On the other hand, outsourcing reduces value-added, and with labor input reduced due to input substitution, labor productivity based on value-added does not increase. Due to growing importance of outsourcing, labor productivity based on value-added in production rather than on gross output, would be a better approximation of the contributions of labor in industry output. Since, Ricardo conceived of labor productivity differences due to differences in technology, value-added measure of productivity is more relevant in this thesis than gross output based measure of productivity.

The perceived relationship between RCA index and *FER* can be stated as: *RCA index values for labor intensive commodities will increase with the values of FER. RCA index values for capital intensive commodities will decrease with an increase in the values of FER.*

With reference to the Heckscher-Ohlin theory, more the labor force of each country relative to the world labor force exceeds (falls short of) the capital stock of the same country relative to the world capital stock, more (less) will it produce and export labor intensive products, and can have higher revealed comparative advantage (disadvantage) in such products. Thus, higher the values of *FER*, greater will be the values of RCA index of countries in any labor intensive product. Following similar arguments, higher the values of *FER*, smaller will be the values of RCA index of countries in any capital intensive product.

The perceived relationship between RCA index and *APLR* can be stated as: *RCA index values will increase with the values of APLR.*

Considering the Ricardian theory, more the average labor productivity of a country in a particular sector relative to its average labor productivity in all considered sectors exceeds (falls short of) the average labor productivity of world in that particular sector relative to its average labor productivity in all sectors, higher will be the country's revealed comparative advantage (disadvantage) in such a sector. Thus, higher the values of *APLR*, greater will be the values of RCA index of countries in any sector.

For each index and each sector, the country observations on RCA index are regressed upon the corresponding country observations on *FER* to determine the consistency of each RCA index with the Heckscher-Ohlin theory. Similarly, for each index and each sector, the country observations on RCA index are regressed upon the corresponding country observations on *APLR*, to determine the consistency of each RCA index with the Ricardian theory. The sector specific cross sectional regression equations for the Heckscher-Ohlin model and the Ricardian model accordingly take the following respective forms,

$$RCA^i = \alpha + \beta FER^i + \varepsilon^i \quad (4.16)$$

$$RCA^i = \alpha + \beta APLR^i + \varepsilon^i \quad (4.17)$$

ε^i is the random error term. It represents all other factors that affect RCA index but are not considered explicitly in the regression equation. i represents the i th country.

For regression equation (4.16), the null and the alternative hypotheses can be expressed as follows:

$$H_0: \beta = 0$$

$$H_1: \beta > 0 \text{ for labor intensive products;}$$

Or

$$H_1: \beta < 0 \text{ for capital intensive products.}$$

The alternative hypotheses are one sided based on the assumed relation between RCA index and *FER*. The null hypothesis is thus tested against alternative hypotheses using one tail test of significance.

For the regression equation (4.17), the null and alternative hypotheses can be expressed as follows:

$$H_0: \beta = 0$$

$$H_1: \beta > 0$$

The alternative hypothesis is again one sided based on the assumed relation between the RCA index and *APLR*. The null hypothesis is thus tested against alternative hypothesis using one tail test of significance.

In order to calculate the values of RCA index and *APLR*, the country specific data on exports, imports, value-added and number of employees for 28 manufacturing sectors classified on the basis of International Standard Industrial Classification of All Economic Activities (ISIC) revision 2 for the years 1998, 1999 and 2000 are collected from the Trade, Production and Protection (TPP) Database, 1976-2004 by Nicita and Olarreaga (2007). As also mentioned in chapter 3 (p. 82-83), selection of these 3 particular years is primarily governed by the availability of the largest number of data on indicators required for the construction of average productivity ratio and RCA index values, keeping in mind that for each sector in each year, the country observations remain same. Consideration of years beyond the 3 selected years reduces the sample size. The country specific data on labor force and capital stock for computing the values of *FER* have been collected from the Extended Penn World Tables version 4 by Marquetti and Foley (2011). The database provides data on net fixed standardized capital stock at 2005 purchasing power parity (PPP) rates. PPP rates for two countries are expressed as the ratio of prices in their national currencies for the same amount of goods and services, intended to equalize the purchasing power of two currencies (OECD, 2015). To obtain capital stock estimates at current prices, the following formula has been used:

$$\text{Capita stock at current prices} = \text{Capital stock PPP (constant 2005 prices)} \times \frac{\text{GDP PPP (current prices)}}{\text{GDP PPP (constant 2005 prices)}} \times \text{PPP coversion factor to market exchange rate ratio}$$

The data on GDP and PPP conversion factor to market exchange rate ratio are obtained from World Development Indicators accessed online in March 2013. PPP conversion factor to market exchange rate ratio represents the amount of dollars required to purchase the amount of goods and services in the domestic market that a unit of dollar can buy in United States (The World Bank, 2013). The summary statistics for the data used are presented in the Tables C.1 and C.2 in Appendix C. The countries included for testing the consistency of the indices with the Heckscher-Ohlin theory is provided as a note towards the end of Table C.1. The

countries included for testing the consistency of the indices with the Ricardian theory is provided in the Table C.2.1 in Appendix C.

With the available data, the sectoral RCA index values and *APLR* values for each country in each sector are calculated for each year. The values for *FER* are simultaneously calculated for each country in each year. For each country, the sectoral RCA index values and *APLR* values are then averaged over 3 considered years. The values for *FER* for each country are also averaged over 3 years. In regression equation (4.16) accordingly, for any sector, the country specific average RCA index values are regressed upon average *FER* values. In regression equation (4.17), for any sector, the country specific average RCA index values are regressed upon average *APLR* values.³ Consideration of average value for variables, instead of years specific values in case of cross sectional regressions, smoothes out any years specific shocks that could exist (Kee *et al.*, 2009).

Apart from allowing the dependent variable to be linearly related to the independent variable as represented by the regression equations (4.16) and (4.17), the variables are also permitted to be non-linearly related with each other by regressing the RCA index values upon the log of factor endowment ratio or factor productivity ratio. Regressions in semi log form will help in ascertaining the importance of non-linear relationship between the variables. Obviously the interpretation of the β coefficients will differ for two different forms of regressions. In case of regressions in linear form, β will measure the absolute change in index values with a unit change in factor endowment ratio or factor productivity ratio. For regressions in semi-log form, β will represent the absolute change in index values with a relative change in factor endowment ratio or factor productivity ratio. Regardless of its interpretation, the sign of β is of primary concern in this chapter.

³ With mere 3 years of data, more cross country variations are expected than within country variations. That is why instead of relying on panel data analysis, cross sectional regressions are considered here. Moreover, since the comparative advantage models attempt to analyze the differences in export potential of countries at a point in time, cross sectional regressions rather than time series regressions will be more relevant here.

4.2.1.1 Estimation of Regression Coefficients: OLS and Robust Estimate of Standard Errors

The estimation techniques applicable include Ordinary Least Squares (OLS) coefficient estimates with or without robust estimate of standard errors and robust regression techniques. The OLS with robust estimate of standard errors are explained in this subsection, while the robust regression techniques are explained in the subsection 4.2.1.2.

In usual OLS regressions, the errors are assumed to be identically and independently distributed or homoskedastic. If this assumption gets violated or the errors are heteroskedastic, OLS estimators will not have minimum variance in the class of unbiased estimators. As a result the confidence interval can be larger and the t and F statistics can be smaller leading to the acceptance of false null hypothesis.⁴ To deal with this problem, Huber (1967) and White (1980) independently developed a procedure which estimates standard errors without relying on the assumption of homoskedastic errors. In other words, the estimated standard errors will be robust to the presence of heteroskedastic errors. To demonstrate the estimation procedure, a multivariate regression model in matrix notation can be considered,

$$Y = X\beta + u \quad (4.18)$$

corresponding to the regression equation,

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + u_i \quad (4.19)$$

Y is a $(n \times 1)$ vector of predictand variable values, X is a $n \times (k + 1)$ matrix giving n observations on k predictor variable values, u is a $(n \times 1)$ vector of errors. n is the sample size.

In the presence of heteroskedasticity, the variance of u_i is not constant across observations but not correlated with u_j ($i \neq j$).

⁴ The width of the confidence interval is proportional to the standard error of the estimator (Gujarati and Sangeetha, 2007).

Therefore the variance covariance matrix for u as defined by Johnston and DiNardo (1997) is,

$$E(uu') = \begin{pmatrix} \sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n^2 \end{pmatrix} = \Sigma$$

$$\text{Or } E(uu') = \sigma^2 \begin{pmatrix} \sigma_1^2/\sigma^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n^2/\sigma^2 \end{pmatrix} = \sigma^2 \Omega$$

Under homoskedasticity, $\Omega = I$, where I is the $(n \times n)$ identity matrix.

White (1980) proposed to estimate the regression model (4.19) by OLS and then use the estimated residuals to estimate Σ , that is,

$$\hat{\Sigma} = \begin{pmatrix} \hat{u}_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \hat{u}_n^2 \end{pmatrix}$$

The variance of β can now be estimated as:

$$\text{var}(\hat{\beta}) = (X'X)^{-1}X'\hat{\Sigma}X(X'X)^{-1} \quad (4.20)$$

Standard errors based on the above specified procedure are known as the robust standard errors or White Huber standard errors. Robust estimate of standard errors are most commonly used and using them along with conventional standard errors serves as a way to ascertain the presence or absence of heteroskedasticity. In fact, in the absence of heteroskedasticity, robust standard errors will converge to conventional OLS standard errors in large samples. Thus, robust standard errors will be appropriate even when errors are characterized by homoskedasticity. However, since the sample size is not very large (although large enough for parametric regression analysis), robust standard errors are considered only when the null hypothesis of homoskedasticity is rejected through Breusch-Pagan / Cook-Weisberg test for heteroskedasticity in Stata. The test assumes that the error variance corresponding to

regression equation (4.19) can be expressed as a function of non stochastic variables Z 's such that,

$$\sigma_i^2 = \alpha_0 + \alpha_1 Z_{i1} + \alpha_2 Z_{i2} + \cdots + \alpha_m Z_{im} \quad (4.21)$$

The null hypothesis for the test is:

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_m = 0$$

This implies $\sigma_i^2 = \alpha_0$, which is a constant, and hence the errors are homoskedastic. The null hypothesis is tested against the alternative hypothesis of a positive association between the error variance and the predicted value of Y . The test statistic as proposed by Breusch and Pagan (1979) is one half of the explained sum of squares from the estimated regression equation:

$$g_i = \alpha_0 + \alpha_1 Z_{i1} + \alpha_2 Z_{i2} + \cdots + \alpha_p Z_{im} + v_i \quad (4.22)$$

where $g_i = \frac{\hat{u}_i^2}{\hat{\sigma}^2}$, $\hat{u}_i$ is the residual estimated from the regression equation (4.19) using OLS, and $\hat{\sigma}^2 = \frac{\sum \hat{u}_i^2}{n}$.

v_i is the residual corresponding to the regression equation (4.22). The test statistic is asymptotically distributed as χ^2 with $(m + 1 - 1 = m)$ degrees of freedom when the null hypothesis is true. It is possible to reject the null hypothesis if the computed value of the test statistic exceeds the critical χ^2 value at the considered level of significance.

The results for the Bresuch-Pagan / Cook-Weisberg test are reported in the Tables C.5 and C.8 in the Appendix C. In case of acceptance of the null hypothesis, conventional OLS standard errors are considered. It is to be noted, in case of regression with robust estimate of standard errors, the coefficient estimates and coefficients of determination remains unchanged, only the standard error changes. As a result the t values change and hence the p values, or the probability of committing Type I error (error committed in rejecting the true null hypothesis) change. The smaller the p value, the greater the confidence with which the null hypothesis can be rejected.

4.2.1.2 Estimation of Regression Coefficients: Robust Regression

The consideration for robust regression technique arises from the concern about the presence of outliers, which can severely affect the regression outcomes. Robust regressions are iteratively reweighted OLS approach. In OLS technique, the sum of the squares of residuals is minimized in order to obtain a predicted value of the dependent variable as close as possible to the actual value. But the process of squaring residuals tends to excessively weigh the observations with large residuals and this phenomenon can effectively distort the estimation of parameters (Verardi and Croux, 2009). Because the OLS estimations can artificially inflate the variance of the estimates, it might lower the chances for rejecting the null hypothesis or increase the chances for committing type II error. Hence OLS estimates may be unreliable in the presence of outliers (Edgeworth, 1887). In fact outliers violate the assumption of normally distributed residuals and hence presence of outliers can aid in generating unreliable regression results (Alma, 2011). In this context robust regression or regression estimates that are robust with respect to the presence of outliers is an effective way of dealing with the problem. As argued by Hamilton (2013, p. 222) “*Robust regression aims to achieve almost the efficiency of OLS with ideal data and substantially better-than-OLS efficiency in less ideal situations such as nonnormal errors.*” Robust estimation process deletes any observation which can significantly influence the regression coefficients or whose exclusion will change the regression coefficients (also known as the influential observation), and then down-weights the observations with large absolute residuals. It estimates standard errors and tests hypotheses using a method that does not assume normality.

As identified, with usual OLS, the coefficient estimates are obtained by minimizing the sum of squared residuals. Hence, considering the regression model (4.18) in matrix form, to obtain $\hat{\beta}$ we,

$$\text{Min } \sum u'u \quad (4.23)$$

However, in robust regression approach, originally proposed by Huber (1964), the coefficient estimates are obtained by minimizing the sum of a less rapidly increasing function of

residuals. For practical implementation of the method, Verardi and Croux (2009) simplified the objective function to imply minimization of weighted squared residuals, that is,

$$\text{Min } \sum W u' u \quad (4.24)$$

W is an $(n \times n)$ diagonal matrix of weights w_i . The weights are intended to reduce the effects of large residuals on coefficient estimates. In Stata, to obtain robust regression estimates, the regression model (4.18) is first estimated using usual OLS approach and Cook's distance is calculated. Cook's distance is a measure of influence, calculated on the basis of the difference between regression coefficient estimates and their value in case the i th observation is deleted (McDonald, 2002). As mentioned previously, an observation is said to be influential if its exclusion from the sample alters the regression coefficient.⁵ Higher the value of Cook's distance, the more is the degree of influence.⁶ Stata follows the guideline by Chatterjee *et al.* (2000) according to whom, a value of Cook's distance greater than unity is influential and should be deleted from the sample. This method therefore serves as a preliminary measure to exclude extreme values from the sample. An iteratively reweighting technique is then followed to obtain estimates of the regression coefficients until the estimated coefficients converge. To begin the iteration procedure, an initial estimate of the weight W^0 must be formed which is acquired by estimating the regression model (4.18) by OLS method. The weights are defined to be a function of the residuals, which in turn are determined by the estimated regression coefficient, $\hat{\beta}^0$. At the first iteration $I=1$, weighted least squares are used to minimize (4.24) and obtain $\hat{\beta}^1$. Using $\hat{\beta}^1$ the residuals are estimated and they yield in calculating W^1 . The newly calculated weights are then used in iteration $I=2$

⁵ See Stata Data Analysis Examples: Robust Regression. UCLA: Statistical Consulting Group. From <http://www.ats.ucla.edu/stat/stata/dae/rreg.htm> (accessed April 2014).

⁶ Cook's distance for the i th observation is calculated as:

$$D_i = \frac{(\hat{\beta}_{(i)} - \hat{\beta})' (\hat{\beta}_{(i)} - \hat{\beta})}{p \text{var}(\hat{\beta})}$$

$\hat{\beta}_{(i)}$ is the coefficient estimate obtained by excluding the i th observation. $\hat{\beta}$ is the coefficient estimate obtained by including all observations. p is the number of regression coefficients. See Rodríguez (2014).

to minimize (4.24) and obtain the estimate for W^2 . At each iteration I, the estimated β is therefore:

$$\hat{\beta}^I = (X'W^{I-1}X)^{-1}X'W^{I-1}Y \quad (4.25)$$

The iteration process stops when $\hat{\beta}^I - \hat{\beta}^{I-1} \approx 0$.

In Sanidas and Shin (2011), robust estimation techniques have been applied on Galtonian regressions to check for any inconsistency with the results generated from OLS techniques. However, the robust estimation technique has been criticized from certain perspectives. For instance Rousseuw and Zomeren (1990) point out that it might not be able to identify certain outliers if they occur in a cluster. But as argued by Hamilton (2009), any analyst might not altogether fit a linear model upon noticing such behavior of the data in diagnostic plots. For the applicable data, no such phenomenon was however visible so it seems robust estimation technique can be applied.

There is another problem with the robust regressions. They assume that the errors are independently and identically distributed, which implies that the standard errors, t or F tests and confidence intervals will not be robust to heteroskedasticity or correlated errors. However, regression results with White's heteroskedasticity corrected robust estimate of standard errors already take care of this issue. Thus by considering regressions with robust estimate of standard errors and iteratively reweighted robust regressions, it is possible to address simultaneously the issues of heteroskedasticity and influential outliers, and analyze changes in estimations under different circumstances.

4.2.2 Discussion of Results

The results for the regression analyses are discussed in this subsection. While the results corresponding to Heckscher-Ohlin hypotheses are presented in subsection 4.2.2.1, the results corresponding to Ricardian hypothesis are presented in subsection 4.2.2.2.

4.2.2.1 Results for Heckscher-Ohlin Model

The regression results with or without robust estimate of standard errors and robust regression estimates corresponding to the Heckscher-Ohlin predictions are reported in Tables

4.1 and 4.2 respectively. Results relating to 4 labor intensive sectors – 321 (manufacture of textiles), 322 (manufacture of wearing apparel except footwear), 323 (manufacture of leather products, leather substitutes and furs), 324 (manufacture of footwear) and 3 capital intensive sectors – 351 (manufacture of industrial chemicals), 352 (manufacture of other chemical products), 385 (manufacture of professional, scientific, measuring and controlling equipments not elsewhere classified, and of photographic and optical goods) are reported.⁷ Corresponding to each RCA index for each form of regression, only the number of significant estimated β coefficient with expected positive sign (for labor intensive sectors) and negative sign (for capital intensive sectors), and the coefficient of determination or R^2 averaged over 7 sectors are reported in both the tables. R^2 represents the extent to which the sample regression line fits the data and it measures the proportion of total variation in the dependent variable explained by the model. For the kind of bivariate regressions considered, R^2 is equivalent to the square of the Pearson's correlation coefficient. It is to be noted, Stata itself does not report the R^2 in case of robust regression estimates. Thus, to compute the measures of model fitness with Stata, one has to make use of the *rregfit* program written by

⁷ Classification of sectors into labor intensive or capital intensive has been adopted from the Trade and Development Report 1996 by UNCTAD. The classification is based upon the mix of different skills, technology, capital and scale requirements at the final product stage. Thus the report classifies, textiles, clothing, leather products, footwear to be labor intensive and resource based industries, requiring low levels of skill, technology and capital. The industries may also be considered to make use of indigenous skills and technologies. On the other hand, chemical and chemical products, scientific, photographic and optical goods require the highest levels of skill, technology, capital and scale, and can therefore be considered capital intensive sectors. The report however classifies sectors on the basis of 3 digit Standard International Trade Classification (SITC) revision 2. Hence, a correspondence between SITC and ISIC codes had to be established before classifying the ISIC codes into labor or capital intensive sectors. However, as recognized by UNCTAD, higher skill do not always imply higher capital intensity, because many high technology and capital intensive sectors e.g., electronics involve low skill labor intensive intermediate stages in production. Nevertheless, skill, technology and capital are in general observed to be complementary to each other. Using R & D intensity (R & D expenditure in relation to output) to identify the technology content of an industry for selected OECD countries, Hatzichronoglou (1997) identified the sectors with 2 digit ISIC code 32 as low technology sectors and sectors with 3 digit ISIC codes 351, 352, 385 as medium to high technology sectors. Assuming, medium to high technology sectors are associated with more capital intensity than low technology sectors, the classification adopted by UNCTAD and Hatzichronoglou do not seem to significantly disagree with each other.

UCLA Statistical Consulting, the results of which are shown in Table 4.2. Being unimportant for the present analysis, the estimates of the intercepts are not reported. The detailed results on the estimated β coefficient with associated p value and the R^2 for each sector and each index corresponding to Tables 4.1 and 4.2, are reported in Tables C.6 and C.7 in the Appendix C.

It is to be noted, wherever applicable, the arguments in favor of any particular index will be based on both the number of significant β coefficients with expected signs and the magnitude of average R^2 . It follows, any index will be considered to score over others or be more consistent with the theory, if it reports greater number of significant coefficients with expected sign and simultaneously a higher R^2 .

Before proceeding to the discussion of results, an important issue should be pointed out. The coefficient of determination or the R^2 values between the dependent and the independent variables for the linear form are not strictly comparable with the corresponding R^2 values for the log-linear forms of regression (Maddala, 1988). This limitation is not so important when we are comparing the indices on the basis of the number of significant coefficients. But it is more relevant when we are comparing the strength of association between the variables (Deb and Basu, 2012). However, it should be noted that in the present analysis, in addition to the linear form of regression, semi log form of regression is considered, where the dependent variable is regressed upon the log of independent variable. For meaningful comparison of R^2 values corresponding to two regression equations, the dependent variables must be of the same form, while the independent variables can assume different forms in different regressions (Gujarati and Sangeetha, 2007). Hence, a rough and ready comparison between the R^2 s of two alternative regression specifications for each index is presented here.

Table 4.1
Results Summary for Heckscher-Ohlin Theory: With/ without White's Correction

Regression Form	Summary	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
Linear	No. of significant coefficients	5	5	5	1	5	4	6
	Average R^2	0.1394	0.1464	0.1397	0.0188	0.1686	0.1059	0.2071
Semi log	No. of significant coefficients	4	7	4	2	7	5	6
	Average R^2	0.1417	0.2168	0.1418	0.0584	0.1979	0.1218	0.2793

Note: The results are based on a sample size of 60 for each sector. Reported numbers of significant coefficients for each index are out of a total 7 cases.

In Table 4.1 the export based indices can be considered first. For regressions in linear form, based on the number of significant coefficients with expected signs and the magnitude of R^2 averaged over 7 sectors, Log of Balassa is most consistent with the theory. It is followed by the RSCA index. For regressions in semi log form, RSCA index is found to be more consistent with the theory than Log of Balassa index. NRCA index is the worst performer among all for both forms of regressions. Balassa index and ARCA index perform at par for both regressions in linear and semi log form. However, results corresponding to Balassa index, ARCA index and NRCA index can be questioned due to their non-normal distributions, which could aid in generating non-normal residuals. Since compared to the linear forms of regressions, an increase in the magnitude of R^2 s for the semi log form of regressions are observed, results corresponding to the regressions in semi log form might be assigned greater preference. This implies, given the dataset and the estimation techniques, RSCA index can be accepted to be most consistent with the Heckscher-Ohlin theory of comparative advantage.

Among the indices based on both exports and imports, RC index is found to be performing better than the RTA index, both in terms of the number significant coefficients with expected signs and the magnitude of R^2 averaged over 7 sectors. This phenomenon is observed for both regressions in linear and semi log form. Considering the regressions in linear form, the RTA index performs poorly compared to the exports-only indices of Balassa, RSCA, ARCA and Log of Balassa. For regressions in semi log form, the RTA index continues to perform poorly compared to RSCA index and Log of Balassa index. However, with respect to other

exports-only indices, some contradiction in the number of significant coefficients and magnitude of R^2 would be observed. Based on the number of significant coefficients and the magnitude of average R^2 , the RC index performs better than any of the exports-only indices for regressions in linear form. It however performs almost at par with the RSCA index for regressions in semi log form.⁸ Again, by adhering more importance to the results corresponding to the regressions in semi log form for reasons cited in previous paragraph, both RC and RSCA indices can be projected to be most consistent with the theory.

Table 4.2
Robust Regression Results Summary for Heckscher-Ohlin Theory

Regression Form	Summary	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
Linear	No. of significant coefficients	7	7	7	1	6	5	7
	Average R^2	0.0961	0.1584	0.0964	0.0158	0.1501	0.0765	0.1928
Semi log	No. of significant coefficients	7	7	7	4	7	6	6
	Average R^2	0.1199	0.2178	0.1200	0.0507	0.1872	0.1052	0.2260

Note: The results are based on a sample size of 60 for each sector. Reported numbers of significant coefficients for each index are out of a total 7 cases.

When robust regression techniques are applied, in case of regressions in linear form, the indices of Balassa, RSCA, ARCA and RC report the highest number of significant coefficients with expected signs. The performance of Log of Balassa index is marginally poor compared to these indices. It is followed by the RTA index, and the NRCA index is the worst performer among all. If the magnitude of R^2 averaged over 7 sectors is considered in addition to the number of significant coefficients, RC index will be the best performing index, and it will be followed by RSCA index and Log of Balassa index in succession. For regressions in semi log form, an increase in the number of significant coefficients for Log of Balassa index is observed which tends to establish the index at par with the export based

⁸ For the regressions in semi log form, with RSCA index and Log of Balassa index recording similar number of significant coefficients, since the R^2 for RSCA index exceeds the R^2 for Log of Balassa index, RSCA index is considered to be better performing compared to Log of Balassa index. The magnitude of R^2 for the RC index exceeds that of the RSCA index, but the number of significant coefficients for the former only marginally falls short of the RSCA index. On that ground it is being argued that the RC index performs almost at par with the RSCA index.

indices of Balassa, RSCA and ARCA. A simultaneous decrease and increase in the number of significant coefficients are observed for the RC index and RTA index respectively. But again when the magnitudes of average R^2 s for different indices are considered in addition to the number of significant coefficients, RC index can be argued to be performing almost at par with the RSCA index, with both performing better than the Log of Balassa index.⁹ Hence, through robust regressions in linear form, RC index is found to be most consistent with the Heckscher-Ohlin theory, and through robust regressions in semi log form both RSCA index and RC index are found to be most consistent with the theory. Since the magnitudes of R^2 s in case of regressions in semi log form tend to be larger compared to regressions in linear form for all indices studied, results for the semi log form can again be preferred, which implies that both RC index and RSCA index are most consistent with the theory.

Thus by controlling for heteroskedasticity, among the exports only indices, RSCA index is most consistent with the Heckscher-Ohlin predictions, given the dataset. By controlling for influential outliers and apparently addressing any possibility for non-normality in the estimated residuals, RSCA index is again found to be a good performer among all exports only indices. Overall, it may be generalized that RSCA index establishes its reliability across estimation techniques considered. Log of Balassa index establishes itself to be the second most reliable export based index in terms of its consistency with the Heckscher-Ohlin theory. The other three export based indices of Balassa, ARCA and NRCA are not so consistent with the theory. RC index based on both exports and imports, by performing at par with the RSCA index, could not establish its superiority over the RSCA index.

4.2.2.2 Results for Ricardian Model

Tables 4.3 and 4.4 report the regression results corresponding to the Ricardian predictions. Results are reported for 25 sectors only.¹⁰ As in the case of Tables 4.1 and 4.2, corresponding

⁹ The R^2 for RC index exceeds that of RSCA index, but the number of significant coefficients for the RC index only marginally falls short of RSCA index. Hence on that basis, it is being argued that the two indices perform almost at par.

¹⁰ The results corresponding to the sectors with ISIC (rev.2) codes 353, 354 and 372 are not included in the table as the sample size for each sector is below 30. However, these sectors are considered within the reference group of commodities.

to each RCA index for each form of regression, only the number of significant estimated β coefficient with expected positive sign and the coefficient of determination or R^2 averaged over 25 sectors are reported in Tables 4.3 and 4.4. The detailed results on the estimated β coefficient with associated p value and the R^2 for each sector and each index corresponding to Tables 4.3 and 4.4, are reported in Tables C.9 and C.10 in the Appendix C.

Table 4.3
Results Summary for Ricardian Theory: With/ without White's Correction

Regression Form	Summary	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
Linear	No. of significant coefficients	7	12	7	1	11	7	10
	Average R^2	0.0669	0.0910	0.0699	0.0175	0.0876	0.0721	0.0707
Semi log	No. of significant coefficients	9	13	9	1	11	8	10
	Average R^2	0.0698	0.1033	0.1519	0.0169	0.1061	0.0930	0.1012

Note: The sample size for each sector varies between 46 and 32. Reported numbers of significant coefficients for each index are out of a total 25 cases.

According to the results corresponding to the exports-only indices reported in Table 4.3, RSCA index has the largest number of significant coefficients with expected positive sign for both forms of regressions. The second largest number of significant coefficients is observed for the Log of Balassa index for both forms of regressions. For the linear form of regressions the magnitude of R^2 averaged over 25 sectors is largest for the index of RSCA, followed by Log of Balassa index. However when the RCA indices are regressed upon the log of the average productivity ratio, the R^2 of RSCA index will be ranked as the third after ARCA index and Log of Balassa index. Thus for regressions in linear form, the RSCA index is most consistent with the Ricardian theory on the basis of both the number of significant coefficients and the magnitude of average R^2 . For regressions in semi log form, although RSCA index performs better than any of the export based indices in terms of the number of significant coefficients, some amount of contradiction is noticed while comparing the magnitude of R^2 . However, it can be argued, since the magnitude of R^2 for the ARCA index is only marginally greater than that of Log of Balassa index or RSCA index, while the number of significant coefficients for the RSCA index exceeds that of Log of Balassa index or the ARCA index by a greater extent, RSCA index can be considered to be more consistent

with the Ricardian theory. Thus RSCA index is the best performing index among all export based indices, considering both forms of regressions.

Among the indices based on both exports and imports, RC index is found to report larger number of significant coefficients than the RTA index for both the regressions in linear and semi log form. The R^2 averaged over 25 sectors is higher for the RC index than for RTA index in case of regressions in semi log form. For regression in linear form however, RTA index reports a higher magnitude for R^2 compared to RC index. But since the difference in R^2 is only marginal, while the difference in the number of significant coefficients is large, RC index can be considered to be more consistent with the theory than the RTA index. On comparing the performance of RC index with those of the exports-only indices, one will notice that it falls short of the RSCA and Log of Balassa indices in terms of the number of significant coefficients and the magnitude of R^2 for both linear and semi log forms of regressions. However on the basis of similar criteria, it performs better than the indices of Balassa, ARCA and NRCA for regressions in linear and semi log forms.¹¹ Since as argued in the previous paragraph, RSCA index is the best performing index among all exports based indices, it is possible to conclude this discussion by noting the superiority of RSCA index even over the exports-and-import indices considered.

Table 4.4
Robust Regression Results Summary for Ricardian Theory

Regression Form	Summary	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
Linear	No. of significant coefficients	14	16	14	3	16	10	11
	Average R^2	0.0598	0.1086	0.0597	0.0084	0.0948	0.0574	0.0583
Semi log	No. of significant coefficients	12	16	12	3	16	9	12
	Average R^2	0.0523	0.1174	0.0522	0.0069	0.1076	0.0552	0.0819

Note: The sample size for each sector varies between 46 and 32. Reported numbers of significant coefficients for each index are out of a total 25 cases.

¹¹ For regressions in semi log form however the magnitude of R^2 for ARCA index marginally exceeds that of RC index, but the number of significant coefficients for ARCA index falls short of RC index.

The robust regression results do not differ much than the results reported in Table 4.3. Although compared to Table 4.3, an improvement in the number of cases where we accept the alternative hypothesis of positive association between the variables, is observed for each index, the RSCA index emerge to be the best performing export based index when the magnitude of R^2 averaged over 25 sectors is considered in addition to the number of significant coefficients, in case of both forms of regressions. Log of Balassa is the second best performing export based index. Among the exports and imports based indices, although the RC index continues to perform better than the RTA index, it could not out-perform the RSCA index or the Log of Balassa index based only on exports.

Hence, by controlling for heteroskedasticity in the data, among the exports only indices, RSCA index is argued to be most consistent with the Ricardian predictions. By controlling for influential outliers and apparently addressing any possibility for non-normality in the estimated residuals, among exports only indices, RSCA index is again found to be consistently a good performer for both linear and semi log forms of regressions. Hence, RSCA index (by performing well across regression forms and estimation techniques considered) emerges to be most consistent with Ricardian predictions. Log of Balassa is the second most reliable index in terms of its consistency with the theory. Balassa, ARCA and NRCA indices are not as good as the other two export based indices. Among the exports-and-import based indices, RC index though better than the RTA index, is in fact inferior compared to not only RSCA index but also Log of Balassa index.

Thus summarizing, on the basis of both regressions with (or without) robust estimate of standard errors and iteratively reweighted robust estimation techniques on export based RCA indices, RSCA index is found to be most consistent with the Heckscher-Ohlin theory. RSCA index is again selected to be the index most consistent with the Ricardian theory, based on regressions with (or without) robust estimate of standard errors and iteratively reweighted robust regression technique on export based indices. Among the exports and import based indices, though RC index in most cases is found to be more consistent with the Heckscher-Ohlin and Ricardian theories of comparative advantage than the RTA index, its performance is either at par or inferior to that of the RSCA index, considering alternative estimation techniques. Hence, it can be generalized that RC index could not exceed the performance of

the RSCA index, which has been found to be the best performing index among all exports only indices, given the dataset. Although the RC index incorporates both exports and imports in its calculation and therefore could address the growing incidence of re-exports and intra-industry trade, it cannot be established as an index more consistent with the theories on comparative advantage than the RSCA index, on the basis of the present analysis.¹²

4.3 Adjusting Export-Import Based RCA Indices for Trade Barriers

In the exports-and-import RCA indices taken up for analysis in this thesis, each country's level of relative imports is either subtracted from the level of relative exports or the level of relative exports is divided by the level of relative imports. One potential weakness of this approach is that imports of a product by a country can be subjected to the adverse effects of trade barriers. Hence, this calls for suitably addressing the concern about distortionary impacts of trade barriers on exports-and-import indices.

4.3.1 Methods and Data

In order to address the problem of trade barriers, each country's sector specific import data are adjusted and estimate the import values in the absence of tariffs.¹³ The method followed is a replication of that used by Kee *et al.* (2009) and Hauk (2012) in their articles on

¹² To establish the validity of results reported in Tables 4.1 and 4.3, the residuals normality tests have been performed and the results are reported in Tables C.11.1 and C.11.2 in Appendix C.11. The tests reveal that the regression results corresponding to RSCA index, Log of Balassa index and RC index are more dependable than that of other indices.

¹³ There is one issue which needs to be addressed at this stage. Non-tariff barriers (NTBs) may be as important as tariffs in determining trade levels in many countries. While Kee *et al.* (2009) estimate ad-valorem equivalents (AVE) for several countries' NTBs, their dataset is limited in terms of countries and sectors covered and incorporating their data will greatly restrict the sample size. Although in the literature, different methodologies exist to estimate the incidence of NTBs on imports (Nogués *et al.*, 1986; Andriamananjara *et al.*, 2004; Bradford, 2003; Leamer, 1990; Harrigan, 1993), Kee *et al.* (2009) argue that not all of the proposed measures could be directly compared to tariffs, which readily helps in comprehending the effects of trade barriers on imports. Even if it were possible, Kee *et al.* point out that the dataset would be limited compared to their dataset in terms of the country and tariff line coverage. As realized by Deardorff and Stern (1997), evaluating the tariff equivalents of NTBs is the ideal approach to understand the effects of NTBs across industries and countries. Hence, the computation of AVEs for NTBs by Kee *et al.* (2009) has adequate support in the existing literature.

calculating Trade Restriction Index (TRI) or the related Overall Trade Restriction Index (OTRI). The TRI or OTRI was originally proposed by Anderson and Neary (1994, 1996, 2003) to provide a comprehensive measure of trade restrictiveness.¹⁴ Their study aimed at estimating the level of uniform ad-valorem tariff that would have similar effects on the welfare or import levels of a country as the existing structure of protection. The present study however attempts to estimate the level of imports in the event of elimination of all tariff barriers in a given sector. The following formula will be used to adjust the import values for each country in each sector:

$$M_{i,a,t}^* = \exp [\log M_{i,a,t} - e_{i,a} \log(1 + \tau_{i,a,t})] \quad (4.26)$$

$M_{i,a,t}^*$ is the estimated adjusted level of imports for sector a in country i at time t , $M_{i,a,t}$ is the reported unadjusted level of imports, $e_{i,a}$ is the import demand elasticity for the sector, and $\tau_{i,a,t}$ is the applied ad-valorem tariff barrier for the sector.¹⁵ The derivation of the formula is presented in the Appendix C.12.

The advantage of making this adjustment is that the adjusted values of imports are presumably a better reflection of a country's comparative advantage than unadjusted values that are distorted by tariffs. The disadvantages of this adjustment are twofold. First, the adjusted import data are themselves estimates. Hence there can be some measurement errors involved. Second, the use of adjusted import values greatly restricts the sample size due to unavailability of data on tariffs across economic sectors. As a result, in order to present meaningful comparison between (i) exports-and-adjusted-import indices; (ii) exports-only indices; (iii) exports-and-unadjusted-import indices, the latter two categories of indices are calculated with reduced sample size and the results are presented in Tables 4.5 and 4.6.

The indices are analyzed for their consistency with the Heckscher-Ohlin and Ricardian theories on comparative advantage. It is to be noted, parametric tests can be unreliable with a sample size of less than 30. So instead of resorting to parametric regression analysis, a non-parametric correlation test – Spearman's Rank correlation test for the data is conducted and

¹⁴ A survey on the works of Anderson and Neary on TRIs can be found in Anderson and Neary (2005).

¹⁵ $e_{i,a}$ in Kee *et al.* (2008) has been constructed to be a time invariant measure.

results are analyzed.¹⁶ Spearman's Rank correlation between two variables instead of using the actual values of the variables, ranks the observations on the variables in an increasing order and uses the difference of ranks to calculate the correlation coefficient. Reliance on the order of observations rather than the actual values makes the test independent of the form of distributions.

To arrive at the technical representation of the concept, Conover (1999) considers a bivariate random sample of size n $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$. The ranks on the observations X_i and Y_i are then obtained as $R(X_i)$ and $R(Y_i)$ respectively. An observation with a rank 1 has the minimum value among all observations, while an observation with rank n has the maximum value among all. The measure of correlation as given by Spearman (1904), can then be expressed as,

$$\rho = 1 - \frac{6 \sum_{i=1}^n [R(X_i) - R(Y_i)]^2}{n(n^2 - 1)} \quad (4.27)$$

The expression (4.27) is based on the fact that there are no ties among the observations on a variable. In the presence of ties among a number of observations the average rank is assigned to each such observation.¹⁷ The measure of Rank correlation is then given by:

$$\rho = \frac{\sum_{i=1}^n R(X_i)R(Y_i) - n\left(\frac{n+1}{2}\right)^2}{\left(\sum_{i=1}^n R(X_i)^2 - n\left(\frac{n+1}{2}\right)^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^n R(Y_i)^2 - n\left(\frac{n+1}{2}\right)^2\right)^{\frac{1}{2}}} \quad (4.28)$$

which is equivalent to the Pearson's product moment correlation coefficient formula calculated on the basis of ranks and average ranks (Conover 1999).¹⁸ However, in case of the considered samples for analysis in this subsection, no tie was observed.

¹⁶ The reliability of parametric tests is dependent upon the form of the distribution of variables, which needs to be normal. However with a sample size of less than 30 the potential of a non-normal distribution increases. Non-parametric tests are however not dependent upon the form of the distribution of variables and hence can be relied upon in case the sample size falls below 30.

¹⁷ The average is calculated on the basis of ranks assigned to each observation in the absence of ties. The procedure is same as in the case of ties noticed for Wilcoxon's signed rank test and rank sum test defined in chapter 3.

¹⁸ Pearson's product moment correlation coefficient is given by:

Thus, for each index and each sector, the country observations on RCA indices, *FER* and *APLR* are arranged in an increasing order of magnitude and assigned ranks from 1 to the total number of observations in each sample. The Rank correlation coefficients are then computed using the formula given by expression (4.27).

The null and alternative hypotheses for testing the consistency of RCA indices with the Heckscher-Ohlin model can be stated as:

$$H_0: \rho = 0$$

$$H_1: \rho > 0 \text{ for labor intensive products;}$$

Or

$$H_1: \rho < 0 \text{ for capital intensive products.}$$

The alternative hypotheses are one sided based on the assumed relation between RCA index and *FER*. The null hypothesis is tested against alternative hypotheses using one tail test of significance.

The null and alternative hypotheses for testing the consistency of the RCA indices with the Ricardian model can be expressed as:

$$H_0: \rho = 0$$

$$H_1: \rho > 0$$

The alternative hypothesis is again one sided based on the assumed relation between the RCA index and *APLR*. The null hypothesis is tested against alternative hypothesis using one tail test of significance.

$$r = \frac{\sum_{i=1}^n X_i Y_i - n \bar{X} \bar{Y}}{(\sum_{i=1}^n X_i^2 - n \bar{X}^2)^{\frac{1}{2}} (\sum_{i=1}^n Y_i^2 - n \bar{Y}^2)^{\frac{1}{2}}}$$

For $X_i = R(X_i)$, $\bar{X} = \overline{R(X)} = \left(\frac{n+1}{2}\right)$. Similarly for Y_i . Hence ρ and r are equivalent.

The index reporting the largest number of significant correlation coefficients with expected signs across sectors, will be considered to be most consistent with the theories on comparative advantage.

The data on import demand elasticities and tariff rates are acquired from Kee *et al.* (2008) and Nicita and Olarreaga (2007) respectively.¹⁹ The summary statistics for the data used is provided in Tables C.3 and C.4 in the Appendix C. The countries included for testing the consistency of the indices with the Heckscher-Ohlin theory is provided as a note towards the end of Tables C.3. The countries included for testing the consistency of the indices with the Ricardian theory is provided in Table C.4.1 in Appendix C.

4.3.2 Discussion of Results

The Spearman's rank correlation results corresponding to the Heckscher-Ohlin theory and the Ricardian theory are reported in Tables 4.5 and 4.6 respectively. The tables report only the number of significant coefficients with expected signs. The detailed results on correlation coefficient and associated p-values corresponding to each sector and each index are presented in Tables C.12.1 and C.12.2 in the Appendix C.12.

Table 4.5
Results Summary for Rank Correlation Test for Heckscher-Ohlin Theory

Summary/ Index	Balassa	RSCA	ARCA	NRCA	Log of Balassa	Unadjusted RTA	Unadjusted RC	Adjusted RTA	Adjusted RC
No. of significant coefficients	5	6	5	2	5	5	5	5	5

Note: The sample size is 29 each for sectors 321, 322, 323, 351, 352 and 385, and 25 for sector 324. Reported numbers of significant coefficients for each index are out of total 7 cases.

¹⁹ Import weighted average applied tariff rates are considered for the analysis. The tariff data from Nicita and Olarreaga (2007) only constitute the ad-valorem component of tariff schedule and hence it is not ad-valorem equivalent of specific duties. But the authors acknowledge that, it does not constitute a significant omission from the database as the tariff rates are reported only for manufacturing sectors where specific duties are very rare. The analysis was repeated with simple average applied tariff rates provided by Nicita and Olarreaga, but the results did not differ from that presented in Tables 4.5 and 4.6. Both the methods of measuring trade restrictiveness are however subject to limitations. As Kee *et al.* (2009) points out, in case of import weighted average tariff rates, goods subject to high tariffs would have low import weights. And in case of simple average tariff rates, the reported measure would be biased downwards due to low tariffs on economically meaningless goods.

Looking at the number of significant coefficients with expected signs in Table 4.5, one can infer that adjustment of imports for trade barriers does not improve the performance of either RTA or RC index. In fact both of these exports-and-import indices, are found to be as good as the exports-only indices of Balassa, ARCA and Log of Balassa, in terms of their consistency with the Heckscher-Ohlin theory. Among all, RSCA index is found to generate the largest number of significant coefficients with expected signs. Compared to the parametric regression tests, this test tends to project the superiority of exports-only RSCA index over exports-and-import based RC index. Thus on the basis of this non-parametric test, which assumes away any requirement on the form of index distributions, RSCA index turns out to be most consistent with the Heckscher-Ohlin model. The exports-and-import indices even after adjustment on imports, could not surpass its performance.

Table 4.6
Results Summary for Rank Correlation Test for Ricardian Theory

Summary/ Index	Balassa	RSCA	ARCA	NRCA	Log of Balassa	Unadjusted RTA	Unadjusted RC	Adjusted RTA	Adjusted RC
No. of significant coefficients	10	10	10	6	10	11	12	11	12

Note: The sample size for each sector varies between 19 and 15. Reported numbers of significant coefficients for each index are out of total 28 cases.

A glance at the Table 4.6 will reveal a picture which is slightly different than Table 4.5. Here although adjustments of imports for trade barriers do not seem to add anything to the performance of RTA and RC indices (in terms of the number of significant coefficients), they turn out to be more consistent with the Ricardian theory of comparative advantage than any of the exports only indices. On the basis of the results reported in Tables 4.6 therefore, RC index is most consistent with the Ricardian theory. These findings are in contradiction with those of the corresponding parametric regression estimates. However, this is just one case. On the basis of deductions from Tables 4.1, 4.2, 4.3, 4.4 and 4.5, the exports-and-import based RC index is found to be performing either at par or inferior to the exports-only RSCA index. Given the fact that the aim of this research is to identify indices which would significantly be consistent with both the Heckscher-Ohlin and Ricardian interpretations of comparative advantage on the basis of various estimation techniques, it would be justifiable enough to state that the exports-and-import indices are not a significant addition to the

literature on comparative advantage indices. Adjustment of imports to address the issue of trade restrictions might not be an effective way of improving the reliability of such indices.²⁰

4.4 Conclusion

In this chapter the RCA indices studied are of two types. Five of them are based only on exports, namely, the indices of Balassa, RSCA, ARCA, NRCA and Log of Balassa, and two are based both on exports and imports, namely, the indices of RTA and RC. The indices are analyzed to determine their consistency with the Heckscher-Ohlin and Ricardian theories of comparative advantage. RSCA index is found to be the index most consistent with the theory, as the hypothesized relationships between the variables is found to be significant in most cases. This implies that the RSCA index incorporates the basis of comparative advantage theory into its structure. After RSCA index, Log of Balassa index is the second exports-only index whose structure seems to reflect the underlying theories of comparative advantage. Balassa and ARCA indices do not seem to be as consistent with the theory as the RSCA and Log of Balassa indices, particularly on the basis of the estimated regression equations. On the basis of rank correlation results, they do however seem to perform almost at par with the RSCA and Log of Balassa indices. This is primarily induced by the fact that each index is a monotonically increasing transformation of the other index, with respect to a sector. Since, rank correlation evaluates the monotonic relationship between two continuous variables, it could report two variables to be perfectly correlated with each other, even if the rate of change in two variables are not consistent with each other. Under such circumstances however, product moment correlation based on actual values of the variable could report less than perfect correlation. Analysis in chapter 3 reveals the NRCA index to be the one featuring most stable distribution with a stable mean. However the analysis conducted in the current chapter fails to project the index as an “ideal” index since it performs very poorly in terms of its consistency with the comparative advantage theories.

Among the two exports-and-import indices, which can address the incidences of intra-industry trade and re-exports, RC index performs better than the RTA index. However, the

²⁰ One can obviously point out that this report only addresses tariff barriers and in a sense is limited in scope. This is in fact a limitation of the study and adjustment of imports for both tariff and non-tariff barriers upon availability of suitable data, provides additional scope for future research.

RC index could not establish its superiority over the RSCA index. Even adjustment of imports to address the concern over trade barriers could not seem to improve the performance of exports-and-imports indices to a significant extent.

Thus, an analysis on gross trade based indices reveals the fact that none of the selected indices can be identified to be the “ideal” index. Considering the “ideal” index to be the one which features most stable distribution and be most consistent with the theories on comparative advantage, our analyses in this regard fails to provide any definitive conclusion. However, it must be admitted that the results are dependent upon the data used for analyses. The next chapter tries to re-analyze the performance of RCA indices after adjusting the indices for domestic value-added in trade. Given the growing incidences of production fragmentation, re-examining the adjusted RCA indices will effectively be a robustness check on the performance of RCA indices.

Chapter 5

Value-Added Trade and RCA Indices

5.1 Introduction

The phenomena of global supply chains have received increasing importance over the years. As defined in chapter 2 (p. 47-48), the occurrence of a higher volume of intermediate goods being produced for later production stages in different countries, and then exported to other countries for further processing, describes the development of a global value chain. Global value-added supply chains are characterized by increasing inter-linkages in production process across countries, and hence inputs need to cross borders multiple times. This can lead to considerable amount of double counting in the gross trade data and as a result, gross exports may exaggerate the level of domestic value-added in exports (Johnson, 2014).

Johnson and Noguera (2014) estimated world value-added exports to be 76% of the gross exports in the year 2008, while in early 1970s it ranged around 85% of gross exports. They consider this decline in the percentage share as an indicator of increasing double counting in gross trade data. Johnson and Noguera (2014) also recognize the fact that the observed decline in shares originates only after 1990. Johnson (2014) argues that this event could be explained by liberalization of trade in emerging markets, increase in European Union members, ratification of more and more regional trading arrangements and information technology revolution. According to Johnson, an interplay of all these factors resulted in reduced international trade costs, followed by replacement of domestic input suppliers by foreign input suppliers, and eventual reduction in the global value-added content of trade.

As discussed in chapter 2 (p. 50), growing incidences of value-added trade provide enough indication towards adjusting the gross exports or imports values, and eventually reconstruct the RCA indices. As rightly pointed out by Koopman *et al.* (2014), there is supposed to be differences between the products a country exports and the products it actually produces. A country might appear to be a dominant exporter in a particular sector based on gross exports values, but in reality may be contributing very little to that sector in value-added terms.

These basic differences should be kept in mind while analyzing the comparative advantage of countries.

The RCA indices considered in the previous chapters are based on gross measures of exports and imports, and hence do not incorporate the essence of value-added trade into their structure. The analyses presented in chapter 4 do not generate any support for the fact that introduction of imports in addition to exports in the indices significantly improves the results. In other words, the analyses do not project the exports-and-import based indices as being more consistent with the theories of comparative advantage than the indices based only on exports.¹ Thus in the present chapter, the exports-only RCA indices are first recalculated by introducing domestic value-added in exports in place of gross exports. The adjusted exports-only RCA indices are then empirically analyzed to note any changes with respect to the established results in chapter 4. It should however be noted, Koopman *et al.* (2014) and Meng *et al.* (2012) re-analyze the RCA index values after incorporating domestic value-added in exports. But they limit themselves to Balassa's index only. In this chapter, domestic value-added in exports is applied to the five export based RCA indices considered in this thesis, namely, Balassa index, Revealed Symmetric Comparative Advantage (RSCA) index, Additive Revealed Comparative Advantage (ARCA) index, Normalized Revealed Comparative Advantage (NRCA) index and Log of Balassa index. The consistencies of the adjusted RCA indices with the Heckscher-Ohlin and Ricardian model are then examined.

However, the objective in this thesis is ultimately to determine one or more index which is structurally superior in terms of stable distributions, and when empirically examined will be consistent with the theory of comparative advantage. It therefore seems unfair to conclude this chapter without exploring the value-added RCA indices for their stability over time, sectors and countries.

The data and methods of analysis are discussed in section 5.2. The results are presented in section 5.3. Section 5.4 provides an analysis on the stability of empirical distributions of RCA indices after they have been adjusted for domestic value-added in exports. This will

¹ Obviously we do observe a few cases where RC index or both RTA and RC indices are found to be performing better than the exports only indices. But as argued in chapter 4 (p. 132), those few cases do not hinder us from concluding that the exports and import indices perform no better than the exports only indices.

permit a comparison of the behavior of value-added based RCA index distributions with those of the gross export based RCA index distributions. Finally the conclusions are drawn in section 5.5.

5.2 Methods and Data

In order to adjust the gross export values for domestic value-added in exports, the conceptual framework developed by Hummels *et al.* (2001) as introduced in chapter 2 (p. 50-52), will be used in this chapter. Hummels *et al.* (2001) derived the expression for foreign value-added in exports and their idea was applied by Koopman *et al.* (2012) to arrive at the expression for domestic value-added share in exports. In this chapter, with minor modifications to Koopman *et al.*'s expression for domestic content or domestic value-added, the following expression for domestic value-added in exports (VAX) will be used:

$$VAX = \hat{A}_V(I - A^D)^{-1}X \quad (5.1)$$

Assuming N sectors, A_V is a $(1 \times N)$ row vector of domestic value-added coefficients with its each element a_j^V representing sector j 's ratio of value-added to gross output. $\hat{A}_V$ is the $(N \times N)$ diagonal matrix corresponding to the row vector A_V . I is the $(N \times N)$ identity matrix. A^D is a $(N \times N)$ domestic input-output coefficient matrix. Its each element a_{ij}^D represents the amount of domestically produced input from sector i used to produce one unit of sector j 's output. X is the $(N \times 1)$ vector of gross exports. $(I - A^D)^{-1}$ captures the total domestic output requirement from each sector i to produce exports from sector j . The expression (5.1) has been derived in details in Appendix D.1.

Before proceeding with the analysis, it should be noted that the phenomenon of global value chain has been extensively studied in the literature and the idea of Hummels *et al.* (2001) have been suitably modified to deal with real life complications, such as presence of export processing zones and the possibility of multiple countries exporting intermediate goods (Daudin *et al.*, 2011; Johnson and Noguera, 2012; Koopman *et al.*, 2012). But the present study is more about analyzing the changes in the performance of RCA indices with the incorporation of domestic value-added in exports, and in this context examining the destination of a country's exports and ultimate usage of those exports in the destination

country as an intermediate or final good, is not of much relevance (Deb and Hauk, 2015). This chapter therefore continues to rely on the established methodology of Hummels *et al.* while calculating the domestic value-added in a country's exports and in the process, ignore the presence of processing trade and end use of a country's exports. Thus, this chapter continues to assume that the technology applied to produce goods meant for domestic consumption and goods meant for exports are the same. Instead of relying on gross exports, the RCA indices for each sector are therefore adjusted to take into account how much of a particular sector's domestic output is actually embedded in the country's exports. In effect, each element in the column vector VAX measures, for a country the amount of domestic content from a sector embedded in its exports of N sectors (Deb and Hauk, 2015).²

Unlike the analyses in chapter 4, the analysis in this chapter does not restrict itself to cross sectional regressions. Thus for testing the consistencies of the adjusted RCA indices with the Heckscher-Ohlin and Ricardian theories, unbalanced panel regression analysis with heteroskedasticity corrected robust estimates of standard errors, is resorted to.

In order to test the consistency of the indices with Ricardian theory, three different possibilities are considered and the results generated are compared thereby. First, random effects regressions with separate fixed effects for country and year for each index are run by incorporating all sectors, countries and years. Second, random effects regressions with country-year fixed effects for each index with the same sample are run. Third, random effects regressions with only year fixed effects for separate sectors for each index are run. Country fixed effects take care of the fact that different countries may have different time invariant features which need to be estimated separately, so that we can factor out the effects that these country specific characteristics may have on the relationship between RCA index and productivity ratio. Similarly, year fixed effects factor out the variations across years for reasons such as technological changes, wars, conflicts etc. Country-year fixed effects take

² This formulation however differs to some extent from the domestic value-added concept of Koopman *et al.* (2012). They basically enquire about how much of the domestic content from various sectors could be found in a country's final exports of a product. But in order to determine the comparative advantage of a country in a sector, it seems more relevant to enquire about that sector's domestic content in exports of various products from a country (Deb and Hauk, 2015).

care of country specific characteristics varying over time. In addition, the random effects component allows for the fact that for each sector in any particular country, the residual term will be correlated across years (country-sector or the combination of a country and a sector, is defined to be the cross sectional unit for the panel) (Deb and Hauk, 2015).³ Random effects model allows for the fact that the residual term of the cross sectional unit will be random and uncorrelated with the predictor or independent variable.

The three different forms of regressions to be considered can be represented as follows:

$$RCA_{i,j,t} = \alpha + \beta APLR_{i,j,t} + \mu_i + \pi_t + \varepsilon_{i,j,t} \quad (5.2)$$

$$RCA_{i,j,t} = \alpha + \beta APLR_{i,j,t} + \gamma_{i,t} + \varepsilon_{i,j,t} \quad (5.3)$$

$$RCA_{i,t} = \alpha + \beta APLR_{i,t} + \pi_t + \varepsilon_{i,t} \quad (5.4)$$

where $RCA_{i,j,t}$ is the relevant RCA index for sector j in country i in year t , $APLR_{i,j,t}$ is the average labor productivity ratio for sector j in country i in year t , μ_i is a country fixed effect, π_t is the year fixed effects, $\gamma_{i,t}$ is the country-year fixed effects, and finally $\varepsilon_{i,j,t}$ is a residual term. As described, regression equation (5.4) is sector specific and hence j is fixed, while i and t are allowed to vary. In all of the regression specifications, as previously mentioned, the residual term is allowed to be correlated across time with other observations in the same country-sector.

The interpretation of the coefficient of average productivity ratio for the regression equations (5.2) and (5.3) will however differ from the regression equation (5.4). For regression equations (5.2) and (5.3), a positive coefficient represents, within each country how much an increase in productivity ratio causes the RCA index to increase. Due to the presence of many countries and years, the estimated coefficient effectively is an average across all such country responses. Thus for regression equations (5.2) and (5.3) positive coefficients indicate whether the Ricardian model and the relevant RCA index provide similar predictions about the sector in which countries on average have comparative advantage (Deb and Hauk, 2015). For the

³ If a sector is an outlier in a country for any particular year, it is expected that it will remain an outlier in all other years. Use of random effects model has also been supported by the Hausman Test for random effects vs fixed effects model.

sector specific regression (5.4), a positive coefficient implies both RCA index and Ricardian theory provide similar predictions about the country which is most likely to export a product (Deb and Hauk, 2015). The results corresponding to the regressions are supplemented by the non-parametric Spearman's rank correlation test.

In order to determine the consistency of the RCA indices with the Heckscher-Ohlin theory, only the sector specific regressions with year fixed effects i.e., regression equation (5.4) will be considered. Obviously the independent variable will in that case be the relative factor ratio in place of average productivity ratio. For this part of analysis, the sectors must be divided into labor intensive and capital intensive sectors, and for each of them the alternative hypothesis or the expected sign of association between variables will differ.⁴ For that reason it will not be possible to combine all sectors which include a combination of capital intensive and labor intensive sectors and run panel regressions with either country and year fixed effects or with country-year fixed effects.⁵ The regressions results are supplemented by the non-parametric Spearman's rank correlation test.

β and the coefficient of rank correlation are expected to be positive for the Ricardian model. In case of Heckscher-Ohlin model, both are expected to be positive for labor intensive sectors and negative for capital intensive sectors. Considering the regression results, conclusion in favor of any particular index are drawn on the basis of significant β coefficients with expected signs and the magnitudes of R^2 . Considering the correlation results, conclusions are drawn on the basis of significant rank correlation coefficients with expected signs.

As in chapter 4, the necessary data on exports, value-added, output and number of employees for 28 manufacturing sectors, classified on the basis of 3 digit ISIC (revision 2) have been collected for each country-year in our dataset from the Trade, Production and Protection (TPP) Database, 1976-2004, by Nicita and Olarreaga (2007). The data on country wise labor

⁴ As discussed in chapter 4, there would be a positive association between the RCA indices and relative factor ratio for labor intensive commodities, and a negative association between the same variables for capital intensive commodities.

⁵ Sectors need to be combined in order to have sufficiently large sample otherwise with separate fixed effects for country and year or country-year fixed effects, the degrees of freedom would be quite small.

force and capital stock are from Extended Penn World Tables, version 4, by Marquetti and Foley (2011). In order to adjust gross export values for domestic value-added, this chapter also makes use of the input-output coefficient matrices from the Nicita and Olarreaga (2007). Because the input-output matrices combine several ISIC sectors so that there are only 17 manufacturing sectors for each country, the same had to be done to the data on exports, value-added and the number of employees in order to have compatibility between the two sets of data. In addition, it is to be noted, the input-output tables are based on GTAP (Global Trade Analysis Project) version 4 for which the base year is 1995. This study assumes that the coefficient matrices do not vary much between years and hence applies the same coefficient matrix for all considered years for any country. These data are thus used to construct an unbalanced panel for the years 1991 to 2001 with the number of countries varying between the years due to the fact that data are not available for all 17 sectors for each country in a specific year. All of the 17 sectors are incorporated while testing the consistency of RCA indices with Ricardian predictions. The summary statistics for the data used are presented in the Tables D.1 and D.2 in the Appendix D. The countries included in the sample are provided as a note towards the ends of Tables D.1 and D.2.

While testing for the consistency of indices with Heckscher-Ohlin predictions, the 7 sectors analyzed in previous chapter are initially considered. However, since the input-output database combines several sectors, the relevant data on sectors 352 (manufacture of other chemical products) and 355 (manufacture of rubber products) had to be combined with sector 351 (manufacture of industrial chemicals), and sector 385 (manufacture of professional and scientific equipments etc.) had to be combined with sector 383 (manufacture of electrical machinery apparatus, appliances and supplies). While 351 and 352 are high skill and technology intensive sectors, 355 is chiefly a medium skill and technology intensive sector, as per the classification adopted by UNCTAD Trade and Development Report, 1996. Similarly while sector 385 is defined to be a high skill and technology intensive sector, sector 383 includes some medium skill and some high skill intensive products. In the analyses in previous chapter, the identified high skill and technology intensive sectors have been considered to be capital intensive sector. But clubbing of sectors for the purpose of present analysis can dilute the high skill content of sectors and hence the results corresponding to the capital intensive sectors can be biased to some extent. Also since the input-output tables do

not contain any data on sector 324, it had to be dropped. Hence in this chapter the compatibility of RCA indices with Heckscher-Ohlin theory is explored with effectively 5 sectors – 321, 322, 323, 351/352/355 and 383/385, of which the first 3 are labor intensive and the last two are assumed to be capital intensive.

5.3 Discussion of Results

Table 5.1 and Table 5.2 report the results for Ricardian hypothesis, and Table 5.3 reports the results for Heckscher-Ohlin hypothesis. The estimated regression coefficients (β) and the rank correlation coefficients (ρ) with corresponding p-values and the R^2 s from regression equations (5.2) and (5.3) for each index are presented in Table 5.1. In Table 5.2, only the number of significant coefficients with expected positive sign and the R^2 averaged over 17 sectors corresponding to regression equation (5.4) are reported for each index. The same table also reports the number of significant positive coefficients corresponding to the rank correlation test for each index. Table 5.3 is exactly similar in content to Table 5.2, except that the R^2 s for each index are now averaged over 5 sectors. The p values corresponding to the regression coefficients in each table take into account heteroskedasticity corrected robust estimates of standard errors. The detailed results corresponding to Tables 5.2 and 5.3 are provided in Tables D.3 and D.4 in the Appendix D.

Table 5.1
Estimated Results for Ricardian Theory: Combined Samples

RCA Index/ Results	Regression results with country FE and year FE		Regression results with country-year FE		Rank correlation results
	β	R^2	β	R^2	
Balassa	2.0356** (<0.001)	0.2696	2.0538** (<0.001)	0.2766	0.4152** (<0.001)
RSCA	0.1363** (<0.001)	0.1702	0.1355** (<0.001)	0.1721	0.4152** (<0.001)
ARCA	0.0201** (<0.001)	0.1601	0.0207** (<0.001)	0.1604	0.4365** (<0.001)
NRCA	0.0003** (<0.001)	0.0100	0.0003** (<0.001)	0.0100	0.2846** (<0.001)
Log of Balassa	0.3854** (<0.001)	0.1785	0.3868** (<0.001)	0.1810	0.4152** (<0.001)

Note: values in parenthesis are p values for 1 tail when the alternative hypothesis is $p > t$. ** represents significant at 1%; * represents significant at 5%. Each regression and correlation results are based on sample size of 1156.

As per the regression results in Table 5.1, the index of Balassa is better than the Log of Balassa index due to a higher magnitude of R^2 but equal number of positive and significant coefficients. The rank correlation tests however reveal all the indices to be equally compatible with the Ricardian predictions. Hence considering its consistently good performance in case of correlation and regression analyses, Balassa index can be selected to be the most reliable index. Log of Balassa index is the second most reliable index.

Table 5.2
Results Summary for Ricardian Theory: Sector Specific Samples

Summary/ Index		Balassa	RSCA	ARCA	NRCA	Log of Balassa
Regression	No. of significant coefficients	8	11	12	4	12
	Average R^2	0.1946	0.1457	0.1779	0.0340	0.1613
Rank correlation	No. of significant coefficients	12	12	11	10	12

Note: Each regression and correlation results are based on sample size of 68. The reported numbers of cases are out of a total of 17 cases analyzed.

In Table 5.2, on the basis of the magnitude of R^2 averaged over 17 sectors and the number of positive and significant regression coefficients, ARCA index outperforms the Log of Balassa index. However the difference between the two indices is marginal. Thus the two indices can be considered to be almost at par with each other. Although the index of Balassa records relatively large magnitude of average R^2 compared to either ARCA index or Log of Balassa index, the number of significant coefficients recorded by the former is considerably small relative to the latter two indices. The indices of Balassa, RSCA and Log of Balassa however together report the largest number of significant rank correlation coefficients with expected positive sign. Hence there seem to be some inconsistency between the correlation and regression results of Table 5.2. On the basis of regression analysis, ARCA index and Log of Balassa index can be considered to be most consistent with the Ricardian theory. On the basis of correlation analysis, Balassa index, RSCA index and Log of Balassa index are more consistent with the Ricardian theory than ARCA index. In this respect, results corresponding to Log of Balassa index can be preferred as it performs reasonably well under both estimation techniques and does not generate contradictory results.

Table 5.3
Results Summary for Heckscher-Ohlin Theory: Sector Specific Samples

Summary/ Index		Balassa	RSCA	ARCA	NRCA	Log of Balassa
Regression	No. of significant coefficients	2	1	3	0	3
	Average R^2	0.1834	0.0754	0.1715	0.0124	0.1413
Rank correlation	No. of significant coefficients	5	5	5	4	5

Note: Each regression and correlation results are based on sample size of 68. The reported numbers of cases are out of a total of 5 cases analyzed.

In a test of the consistency of RCA indices with the Heckscher-Ohlin predictions using unbalanced panel regressions, Table 5.3 will reveal that both the indices of ARCA and Log of Balassa have the highest number of significant coefficients with expected signs. Like in the case of Table 5.2, Balassa index reports higher magnitude of R^2 but lower number of significant coefficients compared to either ARCA index or Log of Balassa index. In fact, since the difference in the number of significant coefficients and magnitude of R^2 is only marginal for these three indices, they can be considered to be almost equally consistent with the theory. But on the basis of non-parametric rank correlation test, all excepting the NRCA index, have significant coefficients for all sectors studied, thus implying general compatibility of Balassa, RSCA, ARCA and Log of Balassa indices with Heckscher-Ohlin theory. Overall, considering the results corresponding to correlation and regression analyses, all the three indices of Balassa, ARCA and Log of Balassa emerge to be substantially compatible with the theory.

On the basis of alternative tests of consistency of the RCA indices with the theories on comparative advantage, contradictory results emerge. The contradiction is not so much in case of rank correlation analysis, but in case of regression analysis. Two points need to be noted in order to arrive at a definite conclusion about any particular index. First, in case of samples with all countries and all sectors (in Table 5.1), Log of Balassa index although marginally falls short of the Balassa index, it is still better than all other indices considered in this chapter. Second, for sector specific samples, the performance of Log of Balassa index is consistently good across the theories on comparative advantage. Thus given the reasonable performance of Log of Balassa index, it can be identified to be the index which integrates considerably well with the theory. Hence, contrary to the findings in chapter 4, after adjustment of export based RCA indices for domestic value-added in exports, Log of Balassa

index is a reasonable indicator of the underlying comparative advantage of countries. The cross sectional analysis in chapter 4 established RSCA index as the most reliable among all RCA indices. On the basis of panel regression results, domestic value-added in export based RSCA index fails to retain its position.⁶

5.4 Robustness Check on the Stability of Empirical Distributions of Indices

On the basis of the analysis on stability of the empirical distributions of RCA indices presented in chapter 3, NRCA index was considered to be reliable for cross sectors, cross countries and cross time comparison. However, it is important to note that the indices analyzed in chapter 3 were based on gross trade figures and hence the results based on gross trade figures might not be relevant for indices based on value-added in export figures. For that reason, it seems worthwhile to reanalyze the RCA indices for stability of empirical distributions, after they have been adjusted for domestic value-added in exports. The analysis in this section is therefore a robustness check on the performance of RCA indices for inter-country, inter-sector and inter-year comparison. In the previous section Log of Balassa index is argued to be most compatible with the theories on comparative advantage on the basis of both regression and correlation results. If the same index also emerges to be most stable over time, sectors and countries, among all the five indices analyzed, then given the growing

⁶ There is one notable feature about the data used in the present chapter. For about 12% of the observations on sectoral domestic value-added in total exports, the calculated values exceed the observations on sectoral domestic value-added. In principle, it is impossible for domestic value-added in exports to exceed the total value-added for a given sector. There are two potential explanations for this discrepancy in data. One could be that the assumption about input-output matrices for exports being similar to the input-output matrices for domestic production (as assumed by Hummels *et al.* (2001) and addressed by Koopman *et al.* (2012)) does not always hold. The other possibility could be mis-measurement in some of the variables that have been used to calculate value-added in export numbers. The results presented in Tables 5.1, 5.2 and 5.3 do not include such implausible observations. However, for the interest of the reader the results corresponding to Tables 5.1, 5.2 and 5.3, are presented in Tables D.5, D.6 and D.7 in the Appendix D by including all such implausible observations. The results are found to be different and contradictory. On basis of Tables D.5 and D.6, Log of Balassa index can be considered to be the most consistent with the Ricardian theory (considering both regression and correlation results). On the basis of Table D.7, Balassa index can be considered to be most compatible with the Heckscher-Ohlin theory (considering both correlation and regression results). Log of Balassa index however is found to fall short of not only Balassa index and ARCA index but also of RSCA index.

importance of value-added trade, Log of Balassa index can be projected to be the “ideal” index that can be reliably used by policy makers.

5.4.1 Summary Statistics of RCA Indices

Before beginning with the analysis on the stability of the distributions of the RCA indices, it will be worthwhile to study the summary statistics of their distributions. This would enable a comparison between the distributions of the domestic value-added in exports based RCA indices with their gross exports based counterparts reported in chapter 3. Although the years and the number of observations considered in this chapter differ from those considered in chapter 3, the following analysis would permit a rough comparison between the performance of indices under two different circumstances.

From the panel dataset considered in this chapter, three particular years – 1993, 1994 and 1995 are selected based on the largest number of common set of countries for all the three years, to be considered for analyzing the stability of the index distributions.⁷ Beyond 1995, the number of countries for each year tends to decrease. However for presenting the summary statistics, besides the common set of countries, the countries not common to each year are also considered. Hence for the year 1993, the countries included are Argentina, Japan, Korea, Portugal, Turkey, Uruguay and United States. In 1994, all of the above countries excepting Uruguay are considered. Instead of Uruguay data are available on Peru. In 1995, apart from all the countries considered in 1993, additional data are available on Egypt. Hence the number of countries in each of the years 1993 and 1994 are seven, but in the year 1995 eight countries are present. Thus, not only the sample size differs between years but also the sample content differs, although the differences are not large. For each country, there are observations on 17 sectors. Tables 5.4.A to 5.4.E contain the summary statistics for each index for the three considered years. Only the mean, median, standard deviation, skewness and kurtosis are reported for each index in each year.⁸

⁷ The common set of countries for 1991 and 1992 will be 9. But it is preferable to consider years as recent as possible.

⁸ The cumulative distributions are not considered additionally as the reported parameters in Tables 5.4.A to 5.4.E are sufficient indicators of stability of distributions through stability of means.

Table 5.4.A
Summary Statistics for Balassa Index for Each Year: Entire Sample Analysis

Summary Statistics/ Year	1993	1994	1995
Mean	2.002	1.877	2.081
Median	0.838	0.853	0.977
Standard Deviation	3.301	3.049	3.250
Skewness	2.891	2.968	2.939
Kurtosis	10.818	11.469	11.631
No. of Observations	119	119	136

Table 5.4. B
Summary Statistics for RSCA Index for Each Year: Entire Sample Analysis

Summary Statistics/ Year	1993	1994	1995
Mean	-0.047	-0.060	-0.011
Median	-0.088	-0.079	-0.012
Standard Deviation	0.455	0.456	0.463
Skewness	0.358	0.261	0.103
Kurtosis	2.433	2.495	2.353
No. of Observations	119	119	136

Table 5.4.C
Summary Statistics for ARCA Index for Each Year: Entire Sample Analysis

Summary Statistics/ Year	1993	1994	1995
Mean	<0.00001	<0.00001	<0.00001
Median	-0.004	-0.003	-0.001
Standard Deviation	0.074	0.079	0.075
Skewness	0.915	1.188	0.142
Kurtosis	7.756	9.681	7.008
No. of Observations	119	119	136

Table 5.4.D
Summary Statistics for NRCA Index for Each Year: Entire Sample Analysis

Summary Statistics/ Index	1993	1994	1995
Mean	<0.00001	<0.00001	<0.00001
Median	-0.00003	-0.00003	-0.000002
Standard Deviation	0.003	0.003	0.003
Skewness	0.084	0.129	0.263
Kurtosis	7.139	7.168	8.228
No. of Observations	119	119	136

Table 5.4.E
Summary Statistics for Log of Balassa Index for Each Year: Entire Sample Analysis

Summary Statistics/ Year	1993	1994	1995
Mean	-0.082	-0.145	-0.018
Median	-0.177	-0.159	-0.023
Standard Deviation	1.179	1.243	1.215
Skewness	0.416	-0.119	0.087
Kurtosis	3.251	4.170	3.213
No. of Observations	119	119	136

A look at the Tables 5.4.A to 5.4.E will reveal that fluctuations in arithmetic mean over the years are more apparent for Balassa index, RSCA index and Log of Balassa index than that of other indices, although all of them reported quite stable means in chapter 3. The ARCA index and NRCA index consistent with chapter 3 continue to report stable arithmetic means over the years analyzed. Thus, the empirical distributions of ARCA index and NRCA index can be expected to be more stable compared to any other index. Further analysis of Tables 5.4.A to 5.4.E suggest that except the RSCA index, all other indices report lower average levels of skewness and kurtosis compared to those reported in Tables 3.1.A to 3.1.E in chapter 3. For RSCA index, the average levels of skewness and kurtosis are in general higher in Table 5.4.B than in Table 3.1.B. Nevertheless as in chapter 3, the distribution of RSCA index is closer to normal distributions as it has skewness closer to 0 and kurtosis closer to 3. Similar observations can be made about the distribution of Log of Balassa index in Table 5.4.E, and in fact it turns out to be even more close to a normal distribution than in chapter 3 (p. 87). Averaging over three years, as in chapter 3 the levels of skewness reported by ARCA index is greater than that reported by RSCA, NRCA and Log of Balassa indices. While in chapter 3 RSCA index reported the least average skewness, after adjustment of indices for sectoral domestic value-added in exports, NRCA index reports the least average skewness among all indices analyzed. The Balassa index still continues to be characterized by a skewed distribution as evident from the values on skewness, kurtosis, mean and median. Compared to the means reported in Tables 3.1.B to 3.1.E in chapter 3, RSCA index, ARCA index, NRCA index and Log of Balassa index report means much closer to zero which is the comparative advantage neutral point, thus providing easier interpretation of the average

sector or country.⁹ The mean for the index of Balassa is however greater than those reported in Table 3.1.A in chapter 3. The difference between the means and medians for the RSCA index for the year 1995 is very insignificant. Such findings tend to hint towards symmetrical distribution of RSCA index. For the other years the differences are a little more apparent. The figures on skewness for each year tend to match with the magnitude of differences between means and medians. The difference between the mean and the median for the Log of Balassa index is more apparent for the year 1993, however for the years 1994 and 1995 the differences are minor. Consistent with the measures on skewness, the differences between means and medians are more visible for the ARCA index than for the NRCA index. Standard deviation is lowest for the NRCA index. Among the two apparently normally distributed indices of RSCA and Log of Balassa, the latter is characterized by lower standard deviation, thus implying relatively lower deviations of observations from the mean. Hence, after the adjustment of export based RCA indices for domestic value-added in exports, the RSCA index and Log of Balassa index still continue to be more normally distributed than any other index considered. The ARCA and NRCA indices report more stable empirical distributions over the years than any other index. But two features of ARCA and NRCA indices as also claimed in chapter 3, remains relevant in this chapter. Firstly, symmetry in distribution is more observed in case of the NRCA index than ARCA index. Secondly, both of their distributions are far from normal.

5.4.2 Testing the Stability of RCA Index Distributions

The analysis in section 5.4.1 relates to the behavior of the empirical distributions of RCA indices by incorporating all countries and all sectors in a sample. As in chapter 3, the empirical distributions of RCA indices can now be analyzed for country specific and sector specific samples. The five RCA indices based on domestic value-added in exports are therefore examined using Wilcoxon's signed rank test and Wilcoxon's rank sum test.¹⁰ The tests have been explained in details in chapter 3 (p. 74-80). Wilcoxon's signed rank test is

⁹ The comparative advantage neutral point is zero for NRCA index but approximately zero for RSCA, ARCA and Log of Balassa indices due to inclusion of all sectors and all countries within the reference group.

¹⁰ Graphical analyses are not considered in this chapter as the tests of significance on stability of distributions could be considered to be more robust than the shifts in graphical plots.

used for studying the statistical significance of shifts in empirical distributions of RCA indices over time. Wilcoxon's rank sum test is used for examining the statistical significance of shifts in empirical distributions of the RCA indices over countries and sectors for any particular year. On the basis of the null hypothesis for Wilcoxon's signed rank test, the distributions for each index are similar or stable over time analyzed. On the basis of the null hypothesis for Wilcoxon's rank sum test, for each index, the sample groups of countries or sectors share the similar population distributions, or in other words, the distributions for each index are stable over countries or sectors analyzed. In chapter 3, Wilcoxon's rank sum tests have been complemented by Kolmogorov Smirnov tests. But as recognized in chapter 3 (p. 98), changes in distributions due to changes in mean or median of the distribution is more relevant for the thesis than changes in distributions due to changes in either mean or median or standard deviation or skewness or kurtosis. For that reason the relevant conclusions in chapter 3 for similarity of distributions for two independent samples are based on the rank sum test. Hence, two independent samples in this chapter are analyzed for similarity of distributions with the rank sum test only for each index.

As already noted, the three years - 1993, 1994 and 1995 are selected and the set of six countries common to all three years are – Argentina, Japan, Korea, Portugal, Turkey and USA. As previously noted, for each country, there are observations on 17 sectors. For the Wilcoxon's signed rank test for countries, the three years are considered in a group of two for each of the six countries and for each index. For Wilcoxon's signed rank test for sectors, three years are considered in a group of two for each of the five sectors and for each index. The five sectors considered are similar to those selected for analyzing the consistency of RCA indices with Heckscher-Ohlin model in the previous section. These sectors are with ISIC codes 321 (manufacture of textiles), 322 (manufacture of wearing apparel except footwear), 323 (manufacture of leather products, leather substitutes and furs), 351 (incorporates manufacture of industrial chemicals, chemical and rubber products) and 383 (incorporates manufacture of electrical machinery, apparatus, appliances and supplies, as well as professional, scientific and measuring equipments).¹¹ For each sector there are

¹¹ Sector 355(manufacture of rubber products) and 383 (manufacture of machine apparatus, appliances and supplies) have not been considered in chapter 3. However, they are included in this chapter as sector 355 had to

observations on a set of six common countries. Wilcoxon's rank sum test is conducted for the year 1995 only. For analyzing the stability of index distributions over countries, the test is performed by considering the six countries in a group of two for each index. The six countries namely, Argentina, Japan, Korea, Portugal, Turkey and USA, have the largest shares in the exports of manufactures by all eight countries on which data are available for the year 1995.¹² For each country there are 17 sectoral observations. For analyzing the stability of index distributions over sectors, the selected five sectors are considered in a group of two. For each sector there are observations on eight countries.¹³

The summary of the results for signed rank test are presented in Tables 5.5 and 5.5.1. As in chapter 3, only the number of cases where the null hypothesis gets accepted for each index is reported for each country or sector, and as well as for all countries or all sectors together in the respective tables. In a similar manner the number of cases where the null hypotheses of similar distributions over sectors and over countries are accepted for each index for the rank sum tests are reported in Tables 5.6 and 5.7 respectively. Based on the combination of years, for each country and each sector, numbers of cases analyzed for Table 5.5 and 5.5.1 are 3. The reported numbers of cases where the null hypothesis gets accepted for each country and each sector in Tables 5.5 and 5.5.1 are out of those 3 cases. Based on the combination of years and number of countries, the total numbers of cases analyzed in Table 5.5 are 18. Based on the combination of years and number of sectors, the total numbers of cases analyzed in Table 5.5.1 are 15. In Table 5.6, based on the combination of two sectors out of five sectors, the total numbers of cases analyzed are 10 for each index. The reported numbers of cases for acceptance of null hypothesis in Table 5.6 are out of those 10 cases. In Table 5.7, based on a combination of two countries out of six, the total numbers of cases analyzed are 15 for each index. The reported numbers of cases for acceptance of null hypothesis in Table

be combined with the sectors 352 (manufacture of other chemical products) and 351 (manufacture of industrial chemicals), while sector 383 had to be combined with sector 385 (manufacture of professional, scientific and measuring equipments) for analysis in this chapter.

¹² The shares are calculated on the basis of TPP database. As noted in chapter 3 (p. 72), the measure could be influenced by the size of country.

¹³ Apart from observations on Argentina, Japan, Korea, Portugal, Turkey and USA, there are also sectoral observations on Egypt and Uruguay in 1995.

5.7 are out of those 15 cases. The detailed results for the tests corresponding to Tables 5.5, 5.5.1, 5.6 and 5.7 are presented in Tables D.8, D.9, D.10 and D.11 respectively in the Appendix D.

Table 5.5
Results Summary for Wilcoxon's Signed Rank Test for Countries

Country/ Index	Balassa	RSCA	ARCA	NRCA	Log of Balassa
Argentina	3	3	3	3	3
Japan	3	3	3	3	3
Korea	3	3	3	3	3
Portugal	3	3	3	3	3
Turkey	2	1	3	3	1
USA	3	3	3	3	3
Total number of accepted cases	17	16	18	18	16

Note: Reported numbers of cases for each index are out of 3 for each country, and out of 18 for all countries together.

As evident from Table 5.4, the ARCA and NRCA indices report the largest number of cases for acceptance of null hypothesis for equality of distributions. This implies, that the values generated by both of these indices can be compared over time. The difference between the reported number of cases for the Balassa index and that for ARCA or NRCA index is only marginal. The index of Balassa although marginally inferior to the ARCA and NRCA indices, is however found to be marginally superior to RSCA and Log of Balassa indices. The results thus reported in Table 5.4 are to some extent different from those reported in Table 3.2 in chapter 3, where Balassa index reported the least number of cases for acceptance of null hypothesis.

Table 5.5.1
Results Summary for Wilcoxon's Signed Rank Test for Sectors

ISIC Codes/ Index	Balassa	RSCA	ARCA	NRCA	Log of Balassa
321	2	2	3	3	2
322	3	3	3	3	3
323	3	3	3	3	3
351	3	3	3	3	3
383	3	3	3	3	3
Total number of accepted cases	14	14	15	15	14

Note: Reported numbers of cases for each index are out of 3 for each country, and out of 15 for all countries together.

The inferences that can be drawn from Table 5.5.1 are very similar to Table 5.5 as the two indices – ARCA and NRCA are still the most stable over time. The other three indices are comparatively less stable over time, although the difference is only marginal. The results presented in Table 5.4.1 differ to some extent from Table 3.2.1 in chapter 3. In chapter 3, only the distribution of NRCA index was noted to be most stable over years for each sector, with the other indices reporting fewer cases for stable distributions. In this chapter, after adjustment of the index values for domestic value-added in exports, the distribution of ARCA index is found to be as stable as that of NRCA index, with other indices only marginally falling short of the them. Thus, the stability of ARCA and NRCA indices over time is confirmed not only by the analysis of samples including all countries and all sectors together, but also by the analysis of sector specific and country specific samples.

Table 5.6
Results Summary for Wilcoxon's Rank Sum Test for Sectors

Summary/ Index	Balassa	RSCA	ARCA	NRCA	Log of Balassa
No. of Accepted Cases for Null Hypothesis	7	7	5	10	7

Note: Reported numbers of cases are out of a total of 10 cases for each index.

Table 5.6 shows that the NRCA index reports the largest number of cases for acceptance of null hypothesis for similarity of distributions over sectors. Thus, among all, the NRCA index will be most reliable as an ordinal measure with respect to a country and also as a cardinal measure with respect to a sector. It is not only possible to reliably determine the extent of comparative advantage of one country over another with respect to a sector, it is also possible to rank different sectors with respect to a country reliably using NRCA index. The results thus reported in Table 5.6 are very much consistent with those reported in the Table 3.3 in chapter 3. In chapter 3, the NRCA index based on gross export value is characterized by most stable distributions over sectors. In the current chapter, after adjusting the RCA indices for domestic value-added in exports, NRCA index is still the best, with other indices considerably falling short of it.

Table 5.7
Results Summary for Wilcoxon's Rank Sum Test for Countries

Summary/ Index	Balassa	RSCA	ARCA	NRCA	Log of Balassa
No. of Accepted Cases for Null Hypothesis	15	15	15	15	15

Note: Reported numbers of cases are out of a total of 15 cases for each index.

Table 5.7 shows that all the five indices perform equally well in generating stable distributions over countries. All of them therefore meet the ordinal ranking properties in the sense of ranking countries with respect to a sector. At the same time all of them can serve as cardinal measures for comparing the sectoral comparative advantage with respect to a country. The reported results in Table 5.7 are again different from the Tables 3.4 in chapter 3. Among the RCA indices based on gross trade figures, the ARCA index and NRCA index were identified to be most stable over countries, with other indices performing relatively poorly. But after adjusting the same RCA indices for domestic value-added in exports, all the indices are found to be characterized by stable distributions over countries.

Thus, based on the results reported in Tables 5.5, 5.5.1, 5.6 and 5.7, NRCA is the only index which is not only stable over time but also features stable distributions over countries and over sectors simultaneously. Thus the index is reliable as an ordinal or cardinal measure with respect to a sector or country, and with a stable distribution over years which aid in usage of the index in time series analysis. These findings are consistent with that of chapter 3. The ARCA index values can be compared across time. The index can also be used for ranking countries with respect to a sector and for comparing sectors with respect to a country. The index however is not so reliable for ranking sectors in a country or for comparing countries with respect to sector. The Balassa index is found to be almost as time stable as the ARCA index or NRCA index with respect to a country or sector, and as efficient as the ARCA index or NRCA index for ranking countries with respect to a sector or comparing sectors with respect to a country. As the ARCA index, Balassa index is however not so efficient for ranking sectors with respect to a country or for comparing countries with respect to a sector. The remaining two indices – RSCA index and Log of Balassa index are as stable over countries as any other index. But they are not so stable over sectors as the NRCA index. They are however almost as stable as the ARCA index or the NRCA index over time with respect to a sector. With respect to a country, the indices are unstable over time not only compared to ARCA and NRCA indices but also compared to the Balassa index. However, both RSCA and Log of Balassa indices only marginally fall short of the Balassa index in this respect. Thus, their time stability is not consistent across samples.

5.5 Conclusion

The analysis in this chapter mainly focused on adjusting the country wise gross export values to account for the amount of domestic value-added from each sector embedded in gross exports. Given the dataset and alternative estimation techniques involved, it is important to realize the difficulty in settling on any one index. If the purpose is to select an “ideal” index which is characterized by stable distributions over years, sectors and countries, and at the same time is empirically consistent with the theories on comparative advantage, the analyses fails to provide conclusive results. Based on the analyses in sections 5.3 Log of Balassa index can be selected to be the index which is empirically consistent with the models on comparative advantage. But on the basis of analyses in section 5.4, the index can only reliably serve as a cardinal measure with respect to a country or as an ordinal measure with respect to a sector. Thus the index has two problems. First, the index is not reliable as a cardinal measure with respect to a sector or as an ordinal measure with respect to a country. Second, the index is not so reliable for over time comparison, particularly with respect to a country. In the context of both problems, the NRCA index can be used. In fact the NRCA index is suitable as a cardinal or as an ordinal measure with respect to a country or sector over time. But its application for cross sector or cross country comparability is subject to the fact that the index is not so consistent with the theories on comparative advantage. The distribution of Log of Balassa index may not be as stable over time, sectors or countries as the NRCA index. However, the advantages associated with the index are twofold. First, the index is characterized by normal distribution, which facilitates the usage of the index in many estimation techniques that necessitate normally distributed variables. Second, as noted it has the theory ingrained in its structure.

Considering the results of analyses in this chapter, the Log of Balassa index and NRCA index are applied in chapter 6 to present a comparative study of China and India while exporting to United States. This will provide a representation of the performance of two indices in identifying the comparative advantage of the two countries.



Chapter 6

Application of Value-Added in Export based RCA indices in the context of China and India

6.1 Introduction

Research has been undertaken in this thesis to identify an “ideal” RCA index whose usage will facilitate theoretically consistent comparative analysis of countries with respect to a sector, or of sectors with respect to a country, over years. In this chapter the objective is to apply such an RCA index to present the relevance of RCA indices for determining potential areas for export expansion by China and India to United States. But before proceeding it is necessary to take note of two factors. First, as discussions in previous chapters have revealed that it is indeed difficult to establish any one index as the “ideal” index. Second, given the growing importance of production fragmentation, RCA indices based on domestic value-added in trade are more relevant for policy formulation aimed at country specific export expansion, than their gross trade based counterparts. Hence following the conclusions drawn in chapter 5, the domestic value-added in export based Normalized Revealed Comparative Advantage (NRCA) index and Log of Balassa index can be used in this chapter, with due recognition to the fact that none of the index on its own fits into the category of an “ideal” index.

Given greater compatibility of the Log of Balassa index with the comparative advantage theories, it seems plausible to rely more on the index for comparing sectors with respect to a country or ranking countries with respect to a sector over years. However as noted in chapter 5 (p.155), its application for comparing countries with respect to a sector or ranking sectors with respect to a country over time could be questionable due to its unstable distributions in certain cases. The NRCA index in this respect can be used not only for comparing countries with respect to a sector, or ranking sectors with respect to a country, but also for comparing sectors with respect to a country, or for ranking countries with respect to a sector. Moreover such comparison or ranking can be expected to be more stable over the years, than that of any

other index. Nevertheless, the results corresponding to NRCA index could probably be inconsistent with the underlying comparative advantage of countries. Despite such differences between the two indices, it is to be noted, NRCA index and Log of Balassa index will not differ in identifying sectors in which either country will reveal comparative advantage or disadvantage, since both the indices are derived from Balassa index.

The selection of two countries – China and India in this chapter, are largely driven by the recognition of China as a global economic giant, particularly in the area of manufacturing, and its implications for the other imminent Asian superpower - India. As observed by Bardhan (2010), the two countries China and India together with about two-fifth of the world's population, have experienced incredible high growth rates in their incomes over past quarter century. China and India have made significant contributions in the world trade. Economic reforms and integration of these large countries with the world economy have been considered extensively in the literature (Lau *et al.*, 2001; Chaudhuri and Ravallion, 2006; Maddison, 2007; Bosworth and Collins, 2008; Branstetter and Lardy, 2008; Mazumdar and Sarkar, 2008; Bardhan, 2010; Swamy 2010). With 1.2 billion and 1.3 billion of population size in India and China respectively, the two countries are very similar in terms of factor endowments.¹ In this context, the two countries are expected to be competitors while exporting labor intensive products to the world.

RCA indices have been used in the existing literature to study the export patterns of India and China. For instance, Yue and Hua (2002) apply the RCA index of Balassa to demonstrate that the policies of gradual trade liberalization and relaxing of exchange controls adopted by China have resulted in an export pattern consistent with the country's comparative advantage in labor intensive products. Based on 1 digit and 3 digit Standard International Trade Classification (SITC), they observe that over 1980 to 1997 China has moved from comparative advantage in resource and labor intensive sectors to more of labor intensive sectors such as clothing, footwear, travel goods and toys. In another study, Batra and Khan (2005) use Balassa index for analyzing the comparative advantages of India and China while exporting to the world. Their analysis of comparative advantages of China and India in manufactures suggests that both the countries have comparative advantage predominantly in

¹ Population estimates for 2012-2013 by United Nations Statistics Division.

resource and labor intensive products. Employing Balassa's RCA index Wignaraja (2011) note that in the year 2009, the top 10 sectors in which China reveals comparative advantage comprise primarily of low technology manufactures; while the top 10 sectors in which India reveals comparative advantage comprise of a mix of primary products, resource based manufactures and low technology manufactures. Nevertheless, both the countries have diversified their export structure considerably with China specializing more in exports of medium and high technology manufactures relative to India.

The existing literatures thus make use of only Balassa index based on gross exports. The present chapter will propose an alternative methodology. As already mentioned, in this chapter, the domestic value-added in exports based NRCA and Log of Balassa indices will be used for analyzing the comparative advantages of China and India. Moreover, although the existing literatures examine the export potentials of India and or China while exporting to the world market, the current study will be restricting itself to China and India's exports to United States. United States being the largest export market for commodities from both the countries as per the International Trade Statistics Yearbook 2013 published by United Nations, the selection seems to be quite justified and will help in understanding the dynamics of comparative advantage of China and India while trading with their principal export partner.²

The rest of the chapter is divided into three more sections. The data and methods of analysis are described in section 6.2. The results and policy recommendations are presented in section 6.3. The conclusions are drawn in section 6.4.

6.2 Methods and Data

As argued, RCA indices based on sectoral domestic value-added in total exports will be better able to reflect the export potential of a country in any product. With this justification in mind, the NRCA index and Log of Balassa index must first be adjusted for sectoral domestic value-added in exports. The index values must then be used to analyze and compare the

² United States accounted for 17% of the total merchandise exports of China and 12.1% of the total merchandise exports of India. Both China and India have positive trade balance in merchandise with United States, with the balance being larger in case of China (United Nations, 2014).

export performance of China and India while exporting to United States. The data on domestic value-added, gross output and input-output coefficient matrices on 16 industrial sectors are obtained from the World Input Output Tables (WIOT) 2008 and 2011 of the World Input-Output Database (WIOD). These two particular years are selected in order to evaluate the performance of indices during the onset of the financial crisis and years following the onset. Although more recent data after 2008 might have been preferable, the WIOT are available only till the year 2011. The descriptions on each of the 16 industrial sectors are provided in Appendix E.1. Since, the database does not provide any data on bilateral export figures, the final use of sectoral imports by United States from either country from the World Input Output Tables 2008 and 2011 are assumed to mirror final goods exports by China or India to USA.³ The summary statistics for the data used are provided in Table E.2 in the Appendix E. As already defined in chapter 5, following Hummels *et al.* (2001) and Koopman *et al.* (2012), the underlying expression for domestic value-added in exports will be used to adjust the gross bilateral export figures for domestic value-added in bilateral export figures:

$$VAX = \hat{A}_V(I - A^D)^{-1}X \quad (6.1)$$

With 16 industrial sectors, A_V is a (1×16) row vector of domestic value-added coefficients with its each element a_j^V representing sector j 's ratio of value-added to gross output. $\hat{A}_V$ is the (16×16) diagonal matrix corresponding to the row vector A_V . I is the (16×16) identity matrix. A^D is a (16×16) domestic input-output coefficient matrix. Its each element a_{ij}^D represents the amount of domestically produced input from sector i used to produce one unit of sector j 's output. X is the (16×1) vector of gross bilateral exports. $(I - A^D)^{-1}$ captures the total domestic output requirement from each sector i to produce exports from sector j . Each element in the column vector VAX measures, for a country the amount of domestic content from a sector embedded in its exports of 16 sectors.

After adjusting the gross export figures for domestic value-added in export figures using expression (6.1), the NRCA index values and the Log of Balassa index values for each sector

³ The database treats China and the special administrative regions of China – Hong Kong and Macao as one economic entity. This has to be kept in mind while analyzing the results corresponding to China.

are calculated for each country and for each year. It is to be noted, while calculating the indices, the reference group of commodities comprises of all 16 sectors and the reference group of countries consists of all countries exporting to United States, on which data are available from WIOD. A list of the countries included in the sample is provided as a note towards the end of Table E.2 in the Appendix E.

It is to be noted that the analysis in this chapter heavily draws from the inferences presented in chapter 5 which are based on years different from those to be considered in this chapter. It is assumed in this chapter that the inferences drawn about the stability of empirical distributions of RCA indices and about their empirical consistency with the comparative advantage theories, will hold even for more recent years.

6.3 Discussion of Results and Policy Recommendations

Table 6.1 presents the sectors in which China and India exhibit revealed comparative advantage as per the NRCA index for the years 2008 and 2011, arranged in descending order of magnitude. Thus, a sector with rank 1 implies that the sector has highest comparative advantage among all considered sectors on the basis of the calculated index values. Table 6.2 reports ranks corresponding to the sectors in which China and India reveal comparative advantage in 2008 and 2011 using the Log of Balassa index. The index values are reported in parenthesis corresponding to every sector in both the tables.

Table 6.1
NRCA Index Ranks for Sectors with Revealed Comparative Advantage

Rank	China		India	
	2008	2011	2008	2011
1	Electrical and Optical Equipment (0.01426)	Electrical and Optical Equipment (0.02219)	Manufacturing, Nec; Recycling (0.00442)	Manufacturing, Nec; Recycling (0.00656)
2	Basic Metals and Fabricated Metal (0.00889)	Basic Metals and Fabricated Metal (0.00913)	Textiles and Textile Products (0.00207)	Agriculture, Hunting, Forestry and Fishing (0.00252)
3	Agriculture, Hunting, Forestry and Fishing (0.00742)	Agriculture, Hunting, Forestry and Fishing (0.00565)	Agriculture, Hunting, Forestry and Fishing (0.00201)	Textiles and Textile Products (0.00129)
4	Leather, Leather and Footwear (0.00464)	Leather, Leather and Footwear (0.00501)	Wood and Products of Wood and Cork (0.00059)	Wood and Products of Wood and Cork (0.00069)
5	Rubber and Plastics (0.00251)	Textiles and Textile Products (0.00440)	Chemicals and Chemical Products (0.00004)	Chemicals and Chemical Products (0.00050)
6	Textiles and Textile Products (0.00216)	Rubber and Plastics (0.00327)		Basic Metals and Fabricated Metal (0.00037)
7	Other Non-Metallic Mineral (0.00177)	Other Non-Metallic Mineral (0.00176)		
8	Wood and Products of Wood and Cork (0.00168)	Wood and Products of Wood and Cork (0.00145)		
9	Manufacturing, Nec; Recycling (0.00022)			

Note: Nec implies not elsewhere classified. Figures in parenthesis denote calculated index values.

Table 6.2
Log of Balassa Index Ranks for Sectors with Revealed Comparative Advantage

Rank	China		India	
	2008	2011	2008	2011
1	Leather, Leather and Footwear (0.64434)	Leather, Leather and Footwear (0.71426)	Manufacturing, Nec; Recycling (1.45861)	Manufacturing, Nec; Recycling (1.59261)
2	Wood and Products of Wood and Cork (0.52887)	Other Non-Metallic Mineral (0.51908)	Wood and Products of Wood and Cork (1.35099)	Wood and Products of Wood and Cork (1.28564)
3	Other Non-Metallic Mineral (0.52683)	Wood and Products of Wood and Cork (0.47179)	Agriculture, Hunting, Forestry and Fishing (0.76083)	Agriculture, Hunting, Forestry and Fishing (0.68855)
4	Basic Metals and Fabricated Metal (0.41168)	Electrical and Optical Equipment (0.40954)	Textiles and Textile Products (0.61952)	Textiles and Textile Products (0.35581)
5	Agriculture, Hunting, Forestry and Fishing (0.31073)	Basic Metals and Fabricated Metal (0.37695)	Chemicals and Chemical Products (0.02114)	Chemicals and Chemical Products (0.17512)
6	Rubber and Plastics (0.30543)	Rubber and Plastics (0.35548)		Basic Metals and Fabricated Metal (0.15543)
7	Electrical and Optical Equipment (0.30061)	Agriculture, Hunting, Forestry and Fishing (0.21736)		
8	Textiles and Textile Products (0.07439)	Textiles and Textile Products (0.14821)		
9	Manufacturing, Nec; Recycling (0.01391)			

Note: Nec implies not elsewhere classified. Figures in parenthesis denote index values.

As evident from Tables 6.1 and 6.2, China exhibits comparative advantage in primary products, in some resource and labor intensive sectors requiring lower skill, technology and capital, such as textiles, footwear, wood, leather and other non-metallic minerals; in some medium to low skill and technology sectors such as basic and fabricated metals, rubber and plastics, and other manufacturing activities including production of jewellery, musical instruments, sports goods, toys etc. The country also reveals comparative advantage in one sector requiring higher intensity of skill, technology and capital, namely, electrical and optical equipments. India being an economy similar in resources with China, its comparative advantage in almost similar sectors comprising of primary products, other manufacturing activities, textiles and products of textiles, should not be surprising. Contrary to China, India however does not reveal comparative advantage in non-metallic minerals, leather, and products of leather and footwear that can be categorized as low skill labor intensive sectors.

Rubber and plastics requiring low to medium levels of skill, technology and capital, also do not figure into India's list of comparative advantage sectors. The country on the other hand reveals comparative advantage in one high skill and technology intensive sector namely, chemical and chemical products, over which China has no comparative advantage. Both the countries' comparative advantage in sectors apart from resource and labor intensive sectors, hints towards their changing comparative advantage. But still the relative importance of low skill and technology intensive sectors over high skill and technology intensive sectors, is an indicator of labor abundance in their economies.

It can be noted, India has fewer products to export to United States compared to China, based on their comparative advantage. The observation is quite justified given the kind of economic diversification and structural change followed by India and China. According to Ghosh and Chandrasekhar (2007), while China has followed a classic industrialization pattern by transiting from primary to manufacturing activities, India on the other hand has recorded a transition from agriculture to services with meager increases in the share of manufacturing in GDP. As a result, the potential of China to export higher number of products compared to India should not be considered unusual.

It is also relevant here to have a look at the sectors in which China and India reveal comparative disadvantage among the considered 16 sectors. Tables 6.3 and 6.4 report such sectors with associated ranks for the two countries, for the years 2008 and 2011. A sector with a rank of 1 implies the sector has the highest comparative disadvantage among all other sectors. The index values corresponding to each sector are reported in parenthesis in each table.

Table 6.3
NRCA Index Ranks for Sectors with Revealed Comparative Disadvantage

Rank	China		India	
	2008	2011	2008	2011
1	Transport Equipment (-0.02840)	Transport Equipment (-0.03180)	Transport Equipment (-0.00264)	Transport Equipment (-0.00356)
2	Food, Beverages and Tobacco (-0.00494)	Chemicals and Chemical Products (-0.00577)	Electrical and Optical Equipment (-0.00250)	Electrical and Optical Equipment (-0.00312)
3	Chemicals and Chemical Products (-0.00341)	Food, Beverages and Tobacco (-0.00561)	Machinery, Nec (-0.00119)	Machinery, Nec (-0.00153)
4	Coke, Refined Petroleum and Nuclear Fuel (-0.00254)	Mining and Quarrying (-0.00304)	Food, Beverages and Tobacco (-0.00096)	Food, Beverages and Tobacco (-0.00126)
5	Machinery, Nec (-0.001533)	Machinery, Nec (-0.00251)	Mining and Quarrying (-0.00078)	Mining and Quarrying (-0.00097)
6	Pulp, Paper, Paper , Printing and Publishing (-0.00146)	Coke, Refined Petroleum and Nuclear Fuel (-0.00231)	Rubber and Plastics (-0.00031)	Rubber and Plastics (-0.00050)
7	Mining and Quarrying (-0.00126)	Pulp, Paper, Paper , Printing and Publishing (-0.00140)	Pulp, Paper, Paper , Printing and Publishing (-0.00029)	Pulp, Paper, Paper , Printing and Publishing (-0.00040)
8		Manufacturing, Nec; Recycling (-0.00043)	Coke, Refined Petroleum and Nuclear Fuel (-0.00017)	Leather, Leather and Footwear (-0.00032)
9			Leather, Leather and Footwear (-0.00016)	Coke, Refined Petroleum and Nuclear Fuel (-0.00015)
10			Other Non-Metallic Mineral (-0.00009)	Other Non-Metallic Mineral (-0.00012)
11			Basic Metals and Fabricated Metal (-0.00003)	

Note: Nec implies not elsewhere classified. Figures in parenthesis denote index values.

Table 6.4
Log of Balassa Index Ranks for Sectors with Revealed Comparative Disadvantage

Rank	China		India	
	2008	2011	2008	2011
1	Transport Equipment (-2.07093)	Transport Equipment (-2.22077)	Transport Equipment (-2.83338)	Transport Equipment (-2.44422)
2	Coke, Refined Petroleum and Nuclear Fuel (-0.70360)	Coke, Refined Petroleum and Nuclear Fuel (-0.62039)	Food, Beverages and Tobacco (-1.63045)	Food, Beverages and Tobacco (-1.27625)
3	Food, Beverages and Tobacco (-0.43980)	Food, Beverages and Tobacco (-0.43131)	Electrical and Optical Equipment (-1.24638)	Electrical and Optical Equipment (-1.05163)
4	Pulp, Paper, Paper , Printing and Publishing (-0.31684)	Pulp, Paper, Paper , Printing and Publishing (-0.27974)	Pulp, Paper, Paper , Printing and Publishing (-0.99744)	Pulp, Paper, Paper , Printing and Publishing (-1.01516)
5	Chemicals and Chemical Products (-0.18815)	Chemicals and Chemical Products (-0.27587)	Machinery, Nec (-0.95448)	Leather, Leather and Footwear (-0.93107)
6	Mining and Quarrying (-0.08879)	Mining and Quarrying (-0.18803)	Mining and Quarrying (-0.93774)	Rubber and Plastics (-0.92279)
7	Machinery, Nec (-0.07077)	Machinery, Nec (-0.11259)	Rubber and Plastics (-0.69997)	Machinery, Nec (-0.90228)
8		Manufacturing, Nec; Recycling (-0.02817)	Other Non-Metallic Mineral (-0.55462)	Mining and Quarrying (-0.69142)
9			Coke, Refined Petroleum and Nuclear Fuel (-0.48590)	Other Non-Metallic Mineral (-0.52373)
10			Leather, Leather and Footwear (-0.43788)	Coke, Refined Petroleum and Nuclear Fuel (-0.33331)
11			Basic Metals and Fabricated Metal (-0.02286)	

Note: Nec implies not elsewhere classified. Figures in parenthesis denote index values.

From Tables 6.3 and 6.4 it can be noted that both China and India reveal the greatest comparative disadvantage in transport equipments followed by other sectors while exporting to United States. In fact the countries exhibit comparative disadvantage in mostly sectors requiring medium to high skill, technology and capital. However, their comparative disadvantage in some low skill and technology intensive sectors too will be apparent from Tables 6.3 and 6.4. The number of such sectors in which China reveals comparative disadvantage (e.g., food, beverages, tobacco, pulp, paper, printing and publishing) is less

compared to that of India (food, beverages, tobacco, pulp, paper, printing and publishing, leather and footwear, and other non-metallic mineral products).

Any direct comparison of the current findings with that of Wignaraja (2011), Batra and Khan (2005) and Yue and Hua (2002) are not feasible as the latter make use of detailed commodity categorization and different years for analyzing the exports of India and or China while exporting to the world market. Perhaps availability of commensurate data on input-output coefficients will permit calculation of index values for more disaggregated sectors and thereby enable more meaningful comparison of the generated results with that of Wignaraja (2011) or Batra and Khan (2005) or Yue and Hua (2002). Since the current study is based on more aggregated sectoral data than the existing studies, identifying the similarities rather than dissimilarities between the current and existing studies is perhaps more feasible. Hence, one cannot help noticing the fact that even with highly aggregated sectoral classifications, both India and China reveal comparative advantages primarily in labor intensive and resource intensive sectors. The existing studies reach similar conclusions on analyzing gross export based index values. Further, Wignaraja (2011) and Batra and Khan (2005) identify the relative importance of electrical and optical equipments in case of China and chemicals and chemical products in case of India. The present study based on sectoral domestic value-added in total exports could make similar predictions. This could imply that the factors inducing comparative advantage of countries in such sectors might be present.

In order to formulate country specific policies for export expansion, the products in which China and India reveal comparative advantage are more relevant than the products in which the respective countries reveal comparative disadvantage. For that reason, further detailed analysis in this chapter must be restricted to those sectors in which the countries reveal comparative advantage. Hence, the following discussions are based on sectors reported in Tables 6.1 and 6.2, instead of sectors with comparative disadvantage as reported in Tables 6.3 and 6.4.

6.3.1 Comparison of Country Sectoral Ranks over Years

Considering Tables 6.1 and 6.2, a comparison of the sectoral rankings for China between the years 2008 and 2011 reveals certain amount of discrepancy in case of both the NRCA index

and Log of Balassa index. For instance, on the basis of NRCA index, although rubber and plastics occupied fifth position in the year 2008, its position dropped to sixth in the year 2011. Textiles and textile products on the other hand gained its position from sixth to fifth over the years. The changes in sectoral rankings are however not substantial. On the basis of Log of Balassa index, the changes in sectoral rankings for China are much more pronounced. Changes in rankings are observed for most of the sectors excepting leather, leather products and footwear, rubber and plastics, and textiles and textile products. Notable changes in ranks are observed for electrical and optical equipments, and agriculture, hunting, forestry and fishing. Thus, the stability in sectoral rankings is more pronounced with the NRCA index. In manufacturing and recycling however, China lost comparative advantage in the year 2011, although it had comparative advantage in that sector in the year 2008, on the basis of both indices.

The ranking are relatively stable over the years for both the indices considered in case of India. On the basis of Log of Balassa index, no changes in sectoral rankings from 2008 to 2011 can be noted. On the basis of NRCA index however, some minor changes in sectoral rankings can be observed with textiles and textile products losing their position from second to third, and agriculture and allied primary sectors gaining their position from third to second over the years considered. However in basic and fabricated metals, India reported comparative disadvantage in the year 2008 but comparative advantage in the year 2011, according to the values reported by both the indices. Thus, the fact that the distribution of Log of Balassa index is almost as time stable as the NRCA index is evident from sectoral rankings for India in Table 6.2.

6.3.2 Comparison of Country Sectoral Ranks between Indices

The sectoral rankings are also observed to differ for the same country and year between the NRCA index and Log of Balassa index. In the year 2008, in China for instance, the NRCA index ranks electrical and optical equipments in the first place but the Log of Balassa index ranks the same sector in seventh place. NRCA ranks leather, leather goods and footwear in the fourth place while Log of Balassa index ranks it in the first place. In fact all sectors, other than manufacturing (not elsewhere classified) and recycling are ranked differently by the two indices. In 2011 also, discrepancies in rankings of most sectors excepting rubber and plastics

between the two indices are observed for China. But considering India, the sectoral rankings are relatively stable between the two indices. In the year 2008, both the NRCA index and Log of Balassa index place manufacturing (not elsewhere classified) and recycling in the first position, agriculture-hunting-forestry-and-fishing in the third position, and chemical and chemical products in the fifth position. An interchange in the positions of textiles sector and wood sector is observed. In the year 2011, manufacturing (not elsewhere classified) and recycling, chemicals and products of chemicals, basic and fabricated metals retain similar ranks for both the indices. Sectors like agriculture-hunting-forestry-and-fishing, textiles and textile products, wood and products of wood and cork, are ranked differently by the two indices, although the rankings do not differ to a significant extent.

6.3.3 Sectoral Exports Identification on the basis of NRCA Index Ranks

However, given the analysis in chapter 5, although Log of Balassa index is more representative of the theory of comparative advantage than the NRCA index, it is not as reliable as the NRCA index for ranking sectors with respect to a country. For that reason, results corresponding to NRCA index reported in Table 6.1 might be more preferable. Therefore, considering the NRCA index values for the more recent year 2011, China will be having the greatest comparative advantage in electrical and optical equipments, followed by basic and fabricated metals, agriculture-hunting-forestry-and-fishing, leather and leather products and footwear, textiles and textile products, rubber and plastics, other non-metallic minerals, and lastly wood and products of wood and cork, while exporting to United States. These findings are particularly interesting in the context of apparently growing importance of medium or high skill and technology intensive products in China's export basket. As noted by Hanson (2012), China's net exports of apparel and footwear as a percentage of GDP have declined considerably over the years 1994 to 2008. The exports from these sectors however constituted the largest share of country's GDP in 1994. Since 1994, the country has started exporting more sophisticated products like cell phones and computers. Currently electronics contributes most to the growth in net exports as a share of GDP, followed by metals and machinery. This phenomenon highlights the changing pattern of China's comparative advantage. The NRCA index rankings for the year 2011 for China are consistent with the findings of Hanson, as electrical and optical equipments are ranked above leather, leather

products, footwear, textiles and textile products. The index despite being based on sectoral domestic value-added in total exports, could not suppress the importance of electrical and optical equipments in the country's export basket. However, according to Rodrik (2006) China's export strength in electronics is more due to the country's industrial policies to increase high skill and technology intensive production rather than due to traditional forces of comparative advantage. The observations made by Rodrik are to some extent supported by the rankings offered by Log of Balassa index. The index ranks leather, leather goods and footwear over electrical and optical equipments. Interestingly however, in confirmation to Hanson's claim, textile and textile goods have lower contribution to China's export basket compared to the contribution of electrical and optical equipments. The observed phenomenon could be attributed to the fact that the Log of Balassa index is more consistent with the theory than NRCA index. Hence the sectoral rankings presented by the Log of Balassa index are more compatible with the comparative advantage of countries. Comparative advantage of China in electrical and optical equipments even after consideration of sector's domestic value added in exports, should not come as a surprise. Rodrik (2006) observes that most of these products could be categorized as low cost high volume products with limited technological sophistication. Rodrik compared China's unit values (US dollars per unit of output) in some electrical and electronic equipment with that of relatively rich countries like South Korea, Malaysia and Singapore. Compared to other countries, China was found to report lower unit values for most products, implying presence of lower quality products. Hence, the country's comparative advantage in mostly lower quality electrical and optical equipments is not entirely inconsistent with the country's labor abundance.

Consideration of NRCA index values for the year 2011 in case of India reveals different picture. The country has the highest comparative advantage in recycling and manufactures such as furniture, toys, sports goods, musical instruments, jewellery etc. The other sectors in succession from which India can export efficiently to United States are, agriculture-hunting-forestry-and-fishing, textiles and textile products, wood and products of wood and cork, chemicals and chemical products, and lastly, basic metals and fabricated metals. However, as already pointed out, these sectoral ranking do not differ substantially from the rankings corresponding to Log of Balassa index, although the latter index is more compatible with the

theory of comparative advantage out of the two indices. Hence, the importance of low skill resource and labor intensive sectors over more high skill and technology intensive sectors such as chemicals, is revealed by both the indices in case of India. India is thus one of the few cases where consistent rankings of NRCA index and Log of Balassa index implies that both the indices are compatible with the comparative advantage theories and both are reliable for ranking sectors with respect to a country.

For the purpose of determination of sectors on the basis of which inter country trade negotiations should take place, it is important to make a list of the sectors which either country can export to United States. In this respect one must realize that there are some sectors in which both the countries reveal comparative advantage and can be called overlapping sectors. There are also some other sectors in which one of them reveals comparative advantage and can be termed as non-overlapping sectors. In case of non-overlapping products in the year 2011 the exporting country can be distinctively defined. Hence, goods like electrical and optical equipments, leather, leather goods and footwear, rubber and plastics, and other non metallic mineral products can be exported by China. Chemicals and chemical products, manufactures such as jewellery, sports goods, toys, musical instruments etc. and recycled goods can be exported by India. However, in case of overlapping products, further analysis of the index values becomes necessary. Since, the NRCA index is characterized by most stable empirical distributions over sectors, comparison of countries with respect to a particular sector with NRCA index will be more reliable than with Log of Balassa index. Hence, for comparing India and China for overlapping sectors the NRCA index will be used. On comparing the NRCA index values for the year 2011, it will be evident that the extent of comparative advantage enjoyed by China is greater than India in all overlapping sectors such as basic and fabricated metals, agriculture-hunting-forestry-and-fishing, textiles and textile products, wood and products of wood and cork. In the year 2008 however, the extent of comparative advantage enjoyed by China exceeds that of India for agriculture-hunting-forestry-and-fishing, textiles and textile products, wood and products of wood and cork. Manufacturing and recycling is the only sector in which India has greater comparative advantage compared to China. China however loses comparative advantage in manufacturing and recycling in the year 2011. Thus considering the year 2011, China can export textiles, products of textiles, wood, products of wood and cork, products from

agriculture-hunting-forestry-and-fishing, and basic and fabricated metals, to United States. India however may not export any of the overlapping products to United States.

6.4 Conclusion

The analysis undertaken in this chapter focused on the export potential of two competing economies – China and India, exporting to their largest export market, United States. For the purpose, the NRCA index and the Log of Balassa index adjusted for domestic value-added in exports have been used. While the two indices do not differ in determination of sectors in which the respective countries have comparative advantage or disadvantage, they do represent differences in ranking of sectors in a country. Given the fact that the NRCA index is characterized by a stable distributions over time, sectors and countries, even though it is not as consistent with the theories on comparative advantage as the Log of Balassa index, the former is used to draw inferences about export potentials of India and China.

Sectoral rankings on the basis of NRCA index, place electrical and optical equipments in the first place in China, and manufacturing (not elsewhere specified) and recycling in the first place in India. Electrical and optical equipments comprise sectors requiring medium to higher intensity of skill, technology and capital. Manufacturing and recycling are based on medium to low skill, technology and capital intensity. Hence, the NRCA index rankings are compatible with the relative labor abundance of India but same phenomenon is not observed in case of China. The rankings with respect to Log of Balassa index however are more commensurate with the factor abundance of either country. On the basis of NRCA index values, the products which China can export to United States include electrical and optical equipments, basic and fabricated metals, rubber and plastics, other non-metallic minerals, leather-leather products and footwear, and wood and products of wood and cork, textiles and textile products, products from agriculture-hunting-forestry and fishing. India can export chemicals and chemical products, manufactures (not elsewhere classified) and products of recycling, to United States. Given the dependence of India on service oriented exports, the possibility of fewer product exports from India to United States is quite justified. China on the other hand being more dependent on commodity exports, have greater varieties of products to export to United States. Hence, India will lose out to China from the perspective of having a diversified product export basket while exporting to United States. However, the

study offers limited insights in the sense that the sectors considered are highly aggregated. Consideration of data at more disaggregated level will help in identifying the specific products within each sector that each country can export to United States. This will obviously be of more relevance to the policy makers as trade liberalization is usually undertaken at disaggregated product level rather than at high levels of commodity aggregation. The analysis in this chapter is restricted by the non-availability of data on input-output coefficients for more recent years at disaggregated product level, which are required for adjustment of gross export figures to reflect domestic value-added in exports. Nevertheless the present study offers a preliminary guide to sector based trade negotiations to be undertaken between the two Asian emerging economies and United States.





Chapter 7

Summary of Findings, Conclusions and Policy Suggestions

7.1 Introduction

Revealed Comparative Advantage (RCA) index has gained importance over the years because it facilitates identification of the products in which countries incur lower opportunity costs. Such facilitation therefore is an indication of the fact that they integrate the theories of comparative advantage into their structure. With this idea in mind, research is undertaken to determine the consistency of the RCA indices with the theories on comparative advantage. While Ricardian theory considered factor productivity differences, Heckscher-Ohlin theory focused on the factor endowment differences to be the root cause behind comparative advantages of countries. Identifying the fact that factor productivity and factor endowment based approaches to comparative advantage are not necessarily inconsistent, the current study seeks to identify an index consistent with the factor productivity and factor endowment theories of comparative advantage. In addition to ensure reliable usage of such indices for cross country, cross sector and cross time comparisons, stability of index distributions is also analyzed. By integrating these two criteria of theoretical consistency and stability, an “ideal” RCA index has been defined. The thesis ultimately tries to identify such as “ideal” RCA index. The RCA indices selected for analyses in the present work are a mix of exports and import based indices. The originally proposed Balassa index, and the subsequently suggested Revealed Symmetric Comparative Advantage (RSCA) index, Additive Revealed Comparative (ARCA) index, Normalized Revealed Comparative Advantage (NRCA) index and Log of Balassa index are based on only exports. The Relative Trade Advantage (RTA) index and Revealed Competitiveness (RC) index are based on both exports and imports, and are expected to adjust for incidences like, re-exports and intra-industry trade. With respect to the indices based on both exports and imports, the impact of trade barriers on imports has been taken into account. Thus such indices have been adjusted to factor out the effect of tariff

barriers on imports and their consistency with the comparative advantage theories have been re-analyzed.

Identification of an “ideal” index brings to the forefront the deficiencies that these RCA indices might face due to growth of global supply chains. Global supply chains encompass the phenomenon in which a finished product must pass through various countries with value-added by each country during the successive production stages of the product. Hence the gross exports of a product by a country include the value-added by that country and by all other countries involved in manufacture of the product at its various stages. In this context, analyzing the export potential of a country on the basis of gross trade based RCA indices, may fail to present the true picture. In fact RCA indices based on domestic value-added by countries should be more commensurate with the actual comparative advantage or disadvantage of countries than their gross trade based counterparts, and would help to understand the contribution of a country in exporting a product. For this reason, the analyses of RCA indices in this thesis have not been restricted only to gross trade based RCA indices. Attempts have been made to move beyond that and incorporate the effects of global value chains on RCA indices.

Before drawing the final conclusion in this thesis it seems worthwhile to present a summary of the findings in the following section.

7.2 Summary of Findings

7.2.1 Empirical Distributions of Gross Trade based RCA Indices

- i) On analyzing the year wise empirical distributions of all RCA indices by incorporating all countries and all sectors in the sample, RSCA index seems to be most well behaved. Not only is its distribution stable over time, but also approximately normal. Thus the index can be depended upon for cross time comparison. It can also be reliably used in parametric estimation techniques that necessitate normally distributed variables. The distributions of other indices although stable over time, do not fit into the specifications of normal distribution as well as that of RSCA index.

- ii) Country specific graphical analysis of the empirical distributions of RCA indices project the ARCA index and NRCA index to be most stable over time. The distributions of two indices although symmetric, are still non-normal. The other indices represent varying degrees of instability in distributions over time. This implies that the ARCA index and NRCA index are more reliable for cross time comparison with respect to countries than any other index analyzed.
- iii) Tests of significance on the shifts of empirical distributions of RCA indices over time reveal ARCA index, NRCA index and RC index to be most stable with respect to countries. With respect to sectors however, only NRCA index is found to be most stable over time. Thus, with respect to both countries and sectors NRCA index may be more relevant for cross time comparison than any other index.
- iv) Tests of significance on the shifts of empirical distributions of RCA indices over sectors reveal the NRCA index to be most stable. This implies that the NRCA index is more dependable as a cardinal measure with respect to a sector and as an ordinal measure with respect to a country, than any other index.
- v) Tests of significance on the shifts of empirical distributions of RCA indices over countries reveal ARCA index to be most stable, and NRCA index to be almost as stable as the ARCA index. Thus both ARCA and NRCA indices may be suitable as cardinal measures with respect to a country or as ordinal measures with respect to a sector.

7.2.2 Consistency of Gross Trade based RCA Indices with the Comparative Advantage Theories

- i) By controlling for heteroskedasticity in the sector specific cross sectional regressions, among the exports-only indices, RSCA index is found to be most consistent with the Heckscher-Ohlin predictions. By controlling for influential outliers, among the exports-only indices, RSCA index is again found to be most consistent with the Heckscher-Ohlin theory. Thus the RSCA index can be considered to be most reliable index, as it has the theory ingrained in its structure.
- ii) Among the other export-only indices, Log of Balassa index is found to be the other index most consistent with the Heckscher-Ohlin theory. Thus next to RSCA

- index, Log of Balassa index can be considered to reliably identify the comparative advantage of countries as per Heckscher-Ohlin theory.
- iii) The other three export-only indices of Balassa, ARCA and NRCA are relatively less consistent with the Heckscher-Ohlin theory.
 - iv) Among the export-and-import based indices, RC index is found to be more consistent with the Heckscher-Ohlin theory than the RTA index. However, the RC index performs at par with the RSCA index. Thus, consideration of imports in addition to exports in the RCA index formulation does not substantially improve the compatibility of RCA indices with the Heckscher-Ohlin theory.
 - v) By controlling for heteroskedasticity in the sector specific cross sectional regressions and then controlling for influential outliers, RSCA index is found to be most consistent with the Ricardian predictions. Thus among the exports-only indices, RSCA index can again be considered to be most reliable index, as it has the theory ingrained in its structure.
 - vi) Among the other export-only indices, Log of Balassa index is found to be the next index most consistent with the Ricardian theory. Thus next to RSCA index, Log of Balassa index can be considered to reliably identify the comparative advantage of countries.
 - vii) The other three exports-only indices of Balassa, ARCA and NRCA are relatively less consistent with the Ricardian theory.
 - viii) Among the export-and-import based indices, RC index is found to be more consistent with the Ricardian theory than the RTA index. However, the RC index is less consistent with the Ricardian theory not only compared to RSCA index but also compared to Log of Balassa index. Thus, consideration of imports in addition to exports in the RCA index formulation does not substantially improve the compatibility of the export-and-import based RCA indices with the Ricardian theory, relative to the exports-only indices.
 - ix) Adjustment of import figures for tariff barriers do not alter the compatibility of exports-and-import based RCA indices with the comparative advantage theories, relative to their unadjusted counterparts. This implies that tariff barriers on

imports do not have any noticeable impact on the consistency of these RCA indices with the theories.

- x) On comparing the consistency of exports-only RCA indices, import tariff unadjusted export-and-import indices, and import tariff adjusted export-and-import indices, with the Heckscher-Ohlin theory, RSCA index emerges to be most consistent with the theory. Thus the exports-only RSCA index is still the most reliable among all.
- xi) On comparing the consistency of exports-only RCA indices, import tariff unadjusted export-and-import indices, and import tariff adjusted export-and-import indices, with the Ricardian theory, RC index is found to be most consistent with the theory.

7.2.3 Value-Added Trade and RCA Indices

- i) For panel data regressions incorporating all countries and all sectors, Balassa index is found to be most consistent with the Ricardian theory. In other words, Ricardian theory and the Balassa index in most cases provide similar predictions about the sector in which countries on average have comparative advantage. Log of Balassa index only marginally falls short of Balassa index in this context. On the basis of rank correlation estimates, all the indices are equally compatible with the theory. Thus considering the results corresponding to both regression and correlation analyses, both Balassa index and Log of Balassa index can be considered to be more consistent with the theory than any other index.
- ii) Considering the sector specific panel regressions and correlation analyses, Log of Balassa index is more consistent with the Ricardian theory than any other index. However, with respect to the sector specific panel regression and correlation analyses in case of Heckscher-Ohlin model, Balassa index, ARCA index and Log of Balassa index are more consistent with the theory than other indices. Thus Log of Balassa index can be selected to be more reliable among all, given its consistent performance for sector specific regressions and correlations across comparative advantage theories. Overall, given the performance of Log of Balassa index for both

sector specific samples and country-sector inclusive samples, the index may be selected to most reliable out of all indices analyzed.

- iii) An analysis of the year wise empirical distributions of domestic value-added in export based RCA indices, by incorporating all countries and sectors in the sample, reveal RSCA index and Log of Balassa index to be more normally distributed than any other index considered. This implies reliable usage of these two indices in parametric estimation techniques requiring normally distributed variables. The ARCA and NRCA indices report more stable empirical distributions over the years than any other index. Thus these two indices are more dependable for cross time comparison.
- iv) Tests of significance on the shifts of empirical distributions of RCA indices over time reveal ARCA index and NRCA index be most stable with respect to countries. With respect to sectors also the two indices are found to be most stable over time. Thus, with respect to both countries and sectors the ARCA index and NRCA index may be more relevant for cross time comparison, than any other index.
- v) Tests of significance on the shifts of empirical distributions of RCA indices over sectors reveal the NRCA index to be most stable. This implies that the NRCA index is more dependable as a cardinal measure with respect to a sector and as an ordinal measure with respect to a country, than any other index.
- vi) Tests of significance on the shifts of empirical distributions of RCA indices over countries reveal all indices to be substantially stable. This implies all the indices are suitable as cardinal measures with respect to a country or as ordinal measures with respect to a sector.

7.2.4 Evaluation of China and India's Exports to United States on the basis of Value-Added Adjusted Indices

- i) China reveals comparative advantage in mainly products requiring low to medium levels of skill, technology and capital. Electrical and optical equipment is the only high skill and technology intensive sector in which China specializes. India also reveals comparative advantage in low to medium skill and technology intensive sectors. Chemical and chemical product is the only high skill and technology

intensive sector in which India specializes. Predominance of low skill and technology intensive sectors in the export baskets of China and India to United States, is an indicator of labor abundance in the two economies. However, presence of some high skill and technology intensive sectors in the respective country's export baskets are a representation of the countries' changing comparative advantage.

- ii) For China, stability in sectoral rank over the years is more observed with NRCA index. Log of Balassa index generates some amount of variations in the sectoral ranks over the years. For India, sectoral ranks over the years are observed to be relatively stable with both NRCA index and Log of Balassa index. Thus India represents one of the few cases for which the distribution of Log of Balassa index is almost as time stable as the NRCA index.
- iii) The discrepancies in ranks by the two indices are more apparent for China than for India.
- iv) NRCA index rankings are compatible with the relative labor abundance of India but same phenomenon is not observed in case of China. Thus India represents one of those few cases for which the NRCA index is consistent with the comparative advantage theory.
- v) As expected, the rankings with respect to Log of Balassa index are more commensurate with the factor abundance of either country than that of NRCA index.
- vi) NRCA index being more reliable as a cardinal measure with respect to a sector and as an ordinal measure with respect to a country, is used for identifying the possible areas for export expansion from China and India to United States. China can thus export electrical and optical equipments, basic and fabricated metals, rubber and plastics, other non-metallic minerals, leather-leather products and footwear, and wood and products of wood and cork, textiles and textile products, products from agriculture-hunting-forestry and fishing, to United States. India can export chemicals and chemical products, manufactures (not elsewhere classified) and products of recycling to United States. Thus compared to China, India has fewer products to export, given its dependence on service oriented exports.

7.3 Conclusion

When gross trade based RCA indices are analyzed using secondary sources of data and alternative cross sectional estimation techniques, the exports-only RSCA index is found to be most consistent with the factor productivity and factor endowment theories of comparative advantage. Thus in case of Heckscher-Ohlin theory, the hypotheses of positive association between RCA index and factor endowment ratio (for labor intensive products) and negative association between RCA index and factor endowment ratio (for capital intensive products) are valid in most cases for RSCA index. Also in case of Ricardian theory, the hypothesis of positive association between RCA index and factor productivity ratio is valid in most cases for RSCA index. The consistencies of the RCA indices based on export-and-imports with the theories are either at par or less than the RSCA index for all cases analyzed, excepting one. However, as argued in chapter 4, just one case may not provide sufficient evidence in favor of the RCA indices based on exports-and-imports.

Consideration of domestic value-added in export based indices for panel data analyses reveals the exports-only Log of Balassa index to be consistently compatible with the factor endowment and factor productivity theories of comparative advantage. Thus the hypothesized associations between the RCA index and factor endowment ratio or factor productivity ratio are consistently satisfied in case of Log of Balassa index.

However, when the empirical distributions of RCA indices (based on both gross trade and value-added trade) are examined for stability across sectors, countries and time, RSCA index or Log of Balassa index are no longer the indices with most stable distributions. The NRCA index, whether based on gross exports or on domestic value-added in exports, consistently features stable empirical distributions over time, sectors and countries. Thus, compared to either RSCA index or Log of Balassa index, NRCA index is more suitable as a cardinal or an ordinal measure with respect to countries or sectors. Further, the NRCA index is more reliable for cross time comparison than any other index.

Thus, the analyses of both gross trade based and domestic value-added in trade based RCA indices fail to identify an “ideal” index. None of the RCA indices existing in the literature satisfy the two required criteria for an “ideal” index. Selecting an index, which not only has

stable distributions, but also is compatible with the comparative advantage theories, is indeed a difficult task, given the data at hand and applicable techniques.

7.4 Policy Suggestions

The analysis in this thesis indicates two important issues. First, given the growing intensification of global value chains, analysis of a country's potential to export or import products should be based on domestic value-added in trade. Second, not a single existing RCA index is perfect. This is because, the index which is most consistent with the comparative advantage theories is not the one which features most stable empirical distributions over time, sectors and countries. Given these two issues the policy makers may have to make important changes in the methods of analyzing a countries exports or imports. First, the RCA indices must be adjusted for domestic value-added in exports and hence, the practice of using gross trade based RCA indices should be limited. Second, if the policy makers have to make use of the existing RCA indices to identify a country's comparative advantage in various sectors or to identify the countries having comparative advantage in a sector, it is advisable to study simultaneously the NRCA index and Log of Balassa index. The Log of Balassa index may be expected to generate results more consistent with the comparative advantage of a country, while the NRCA index may be expected to generate comparable results over sectors, countries and time. It is possible that the two measures of comparative advantage might coincide, by generating theoretically consistent reliable comparison across sectors, countries and time (as we notice in case of India in chapter 6), but the chances are expected to be less. Thus policy decisions should always be based on the careful analysis of performance of two indices. Nevertheless, given the results of analysis (referring to Table 5.7 in chapter 5), it is possible to say with some confidence that Log of Balassa index can be used for comparing sectors with respect to a country or to rank countries with respect to a sector. The added advantage of Log of Balassa index is its normal distribution. For empirical analysis requiring normally distributed variables, Log of Balassa index may be preferred. Thus, the advantages of Log of Balassa index exceed that of NRCA index. In all other instances of comparative analysis, the policy makers must make judicious use of both the indices to analyze the comparative advantage of countries and formulate policies after careful examination of results.

7.5 Limitations of Study

The study undertaken has certain limitations.

First, the study is constrained by commensurate data. As recognized in chapter 2, the instances of inter-industry trade would be more apparent at highly disaggregated product classification than at lower levels of product disaggregation. At higher levels of product aggregation intra-industry trade might be observed more. Hence, access to more disaggregated data might have produced more robust results. But since the analysis involves data on a wide variety of indicators, access to comparable data is a major constraint. Trade, Production and Protection (TPP) database in this regard is the best known source for data on a large number of indicators at as high as 3 digit level of ISIC product disaggregation. While more finely disaggregated internationally comparable data on sectoral imports, exports, value-added, output and employees will be available from other sources such as UN COMTRADE, UNIDO (United Nations Industrial Development Organization) etc., it is indeed difficult to get similar data on input-output matrices. The input-output coefficient matrices required for the adjustment of gross exports figures to generate domestic value-added in export figures are provided by the TPP database only for 3 digit ISIC codes. Other sources of data for input-output matrices e.g., World Input Output Database, IDE-JETRO, and GTAP report data at highly aggregated sectoral classifications and matching their commodity classification into more disaggregated international sectoral classifications might not be feasible.

Second, in order to adjust import figures for trade barriers, only the ad-valorem tariff rates have been taken into account. But given the restructuring in the pattern of trade barriers, non-tariff barriers have gained importance at the expense of tariff barriers. Under such circumstances, in order to gain insights into the influence of trade barriers, non-tariff barriers should also be taken into account apart from tariff barriers. But as recognized in chapter 4, data on ad-valorem equivalents of non-tariff barriers (the method has sufficient support in the literature compared to any other methods of measuring non-tariff barriers), is not exhaustive. Their inclusion into the analysis will reduce the sample sizes below the ones considered in chapter 4.

Third, the analysis offered in chapter 6 is limited due to the non-availability of disaggregated data on input-output matrices. The data accessible from World Input Output Database are at high levels of sectoral aggregation and those data fail to provide an in depth assay of the goods exports from China or India to United States. Access to more detailed sectoral classification will enable further investigation on the competitive nature of the two economies while exporting merchandise to United States, and that will be more relevant for policy analysis.

Fourth, the relation between country wise RCA indices and factor endowment ratio may not be as straight forward as presented in this thesis in the presence of “lumpiness” in factors of production. As defined by Courant and Deardorff (1992) the word “lumpiness” is used to describe a situation in which factors of production within a country are unevenly distributed. In such a scenario, a country’s trade will be determined by the factor which is relatively more “lumpy” or more unevenly distributed, but not by the factor which is relatively abundant. Nevertheless such trade will still be in line with the law of comparative advantage. Courant and Deardorff illustrates that the country using intensively a “lumpy” factor to produce a product will have lower relative autarky price in that product compared to any other country, and therefore comparative advantage in the considered product.

Finally labor is considered to be homogenous while constructing the indicator for factor endowment ratio. If labor is categorized on the basis of skill, the factor endowment ratio would deviate from what has been considered in the thesis. Obviously, in that case, consideration of heterogeneous labor input will imply deviating from the conventional Heckscher-Ohlin framework.

7.6 Scope for Future Research

The thesis has tried to address many issues related to the RCA indices. Still certain areas could be identified which leaves scope for future research.

First, the extensive empirical analysis in this thesis suggests that it is indeed difficult to identify one index among the existing RCA indices, which is not only consistent with the theories on comparative advantage but also features a stable distribution, to facilitate reliable

usage of the index for comparative analysis. Hence, future research can focus on developing a RCA index which will meet both the criteria for dependable usage of the index.

Second, availability of exhaustive data on ad-valorem equivalents to non-tariff barriers will permit a more rigorous examination of the behavior of RCA indices based on both exports and imports. It is quite possible that adjustment of imports for both tariff and non-tariff barriers will uphold the importance of exports-and-import based RCA indices over the exports-only RCA indices.

Third, only the export based RCA indices have been analyzed to examine the effects of domestic value-added in exports. The reason being, the RCA indices based on both exports and imports are hardly an improvement over the RCA indices based only on exports. However, this should not deter future researchers from adjusting the exports-and-import based RCA indices for domestic value-added in trade and examine their performance relative to the exports-only RCA indices. The import figures in that case must be adjusted for foreign value-added in imports to factor out the domestic contribution in imports.

Fourth, the factor endowment ratio can be adjusted for differing categories of skill content in labor. The modified ratio can then be considered for examining the consistency of RCA indices with Heckscher-Ohlin theory.

Appendix A

A.1 Derivation of limits of the RCA index distributions

A.1.1 Balassa Index

The Balassa index is expressed as

$$\frac{X_a^i/X_t^i}{X_a^w/X_t^w}$$

The X_a^w and X_t^w can be disintegrated into separate expressions for country i and rest of the world (ROW) and into commodity a and all other commodities $t-a$. X_t^i can similarly be broken into separate expressions for commodity a and $t-a$. The Balassa index can then be expressed as

$$\frac{X_a^i/(X_a^i+X_{t-a}^i)}{(X_a^i+X_a^{ROW})/(X_a^i+X_{t-a}^i+X_a^{ROW}+X_{t-a}^{ROW})} \quad (1)$$

If country i completely specializes in product a , $X_{t-a}^i = 0$ and $X_a^{ROW} = 0$. Expression (1) reduces to,

$$\frac{X_a^i/X_a^i}{(X_a^i)/(X_a^i+X_{t-a}^{ROW})}$$

The above expression would tend to infinity if $\frac{X_a^i}{X_a^i+X_{t-a}^{ROW}} \rightarrow 0$.

If ROW completely specializes in product a , $X_a^i = 0$ and $X_{t-a}^{ROW} = 0$. Expression (1) reduces to,

$$\frac{0/X_{t-a}^i}{X_{ROW}^a/(X_{t-a}^i+X_a^{ROW})} = 0$$

If $X_a^i/X_t^i \approx X_a^w/X_t^w$, the index value tends to 1, which is defined to be the comparative advantage neutral point.

A.1.2 RSCA Index

The RSCA index is defined as

$$\frac{\frac{X_a^i/X_t^i}{X_a^w/X_t^w} - 1}{\frac{X_a^i/X_t^i}{X_a^w/X_t^w} + 1}$$

The X_a^w and X_t^w can be disintegrated into separate expressions for country i and rest of the world (ROW) and into commodity a and all other commodities $t-a$. X_t^i can similarly be broken into separate expressions for commodity a and $t-a$. After rearranging terms, the RSCA index can then be expressed as,

$$\frac{\{X_a^i/(X_a^i + X_{t-a}^i)\} - \{(X_a^i + X_a^{ROW})/(X_a^i + X_{t-a}^i + X_a^{ROW} + X_{t-a}^{ROW})\}}{\{X_a^i/(X_a^i + X_{t-a}^i)\} + \{(X_a^i + X_a^{ROW})/(X_a^i + X_{t-a}^i + X_a^{ROW} + X_{t-a}^{ROW})\}} \quad (1)$$

If country i completely specializes in product a , $X_{t-a}^i = 0$ and $X_a^{ROW} = 0$. Expression (1) reduces to,

$$\frac{1 - (X_a^i)/(X_a^i + X_{t-a}^{ROW})}{1 + (X_a^i)/(X_a^i + X_{t-a}^{ROW})}$$

The above expression would tend to 1 if $\frac{X_a^i}{X_a^i + X_{t-a}^{ROW}} \rightarrow 0$.

If ROW completely specializes in product a , $X_a^i = 0$ and $X_{t-a}^{ROW} = 0$. Expression (1) reduces to,

$$\frac{0 - X_{ROW}^a/(X_{t-a}^i + X_a^{ROW})}{0 + X_{ROW}^a/(X_{t-a}^i + X_a^{ROW})} = -1$$

If $X_a^i/X_t^i \approx X_a^w/X_t^w$, the index value tends to 0, which is defined to be the comparative advantage neutral point.

A.1.3 ARCA Index

The ARCA index expressed as

$$\frac{X_a^i}{X_t^i} - \frac{X_a^w}{X_t^w}$$

The X_a^w and X_t^w can be disintegrated into separate expressions for country i and rest of the world (ROW) and into commodity a and all other commodities $t-a$. X_t^i can similarly be broken into separate expressions for commodity a and $t-a$. The ARCA index can then be expressed as,

$$\frac{X_a^i}{(X_a^i + X_{t-a}^i)} - \frac{(X_a^i + X_a^{ROW})}{(X_a^i + X_{t-a}^i + X_a^{ROW} + X_{t-a}^{ROW})} \quad (1)$$

If country i completely specializes in product a , $X_{t-a}^i = 0$ and $X_a^{ROW} = 0$. Expression (1) reduces to,

$$\frac{X_a^i}{X_t^i} - \frac{X_a^i}{(X_a^i + X_{t-a}^{ROW})}$$

The above expression would tend to 1 if $\frac{X_a^i}{X_a^i + X_{t-a}^{ROW}} \rightarrow 0$.

If ROW completely specializes in product a , $X_a^i = 0$ and $X_{t-a}^{ROW} = 0$. Expression (1) reduces to,

$$\frac{0}{X_{t-a}^i} - \frac{X_a^{ROW}}{(X_{t-a}^i + X_a^{ROW})}$$

The above expression would tend to -1 if $\frac{X_a^{ROW}}{(X_{t-a}^i + X_a^{ROW})} \rightarrow 1$.

If $X_a^i/X_t^i \approx X_a^w/X_t^w$, the index value tends to 0, which is defined to be the comparative advantage neutral point.

A.1.4 NRCA Index

The NRCA index is expressed as

$$\frac{X_a^i}{X_t^w} - \frac{X_t^i}{X_t^w} \frac{X_a^w}{X_t^w}$$

The X_a^w and X_t^w can be disintegrated into separate expressions for country i and rest of the world (ROW) and into commodity a and all other commodities $t-a$. X_t^i can similarly be broken into separate expressions for commodity a and $t-a$. After rearranging and Collecting terms, the NRCA index can then be expressed as,

$$\frac{X_a^i X_{t-a}^{ROW} - X_{t-a}^i X_a^{ROW}}{(X_a^i + X_{t-a}^i + X_a^{ROW} + X_{t-a}^{ROW})^2} \quad (1)$$

If country i completely specializes in product a , $X_{t-a}^i = 0$ and $X_a^{ROW} = 0$. Assuming balanced trade, $X_a^i = X_{t-a}^{ROW}$. After collecting terms, expression (1) reduces to,

$$\frac{X_a^i X_a^i}{4X_a^i X_a^i} = \frac{1}{4}$$

If ROW completely specializes in product a , $X_a^i = 0$ and $X_{t-a}^{ROW} = 0$. Assuming balanced trade, $X_{t-a}^i = X_a^{ROW}$. Expression (1) reduces to

$$-\frac{X_{t-a}^i X_{t-a}^i}{4X_{t-a}^i X_{t-a}^i} = -\frac{1}{4}$$

If $X_a^i = \left(\frac{X_t^i}{X_t^w}\right) X_a^w$, the index value takes the value of 0, which is defined to be the comparative advantage neutral point.

A.1.5 Log of Balassa Index

The Log of Balassa index is defined as,

$$\ln \frac{X_a^i / X_t^i}{X_a^w / X_t^w}$$

The X_a^w and X_t^w can be disintegrated into separate expressions for country i and rest of the world (ROW) and into commodity a and all other commodities $t-a$. X_t^i can similarly be broken into separate expressions for commodity a and $t-a$. The Log of Balassa index can then be expressed as

$$\ln \frac{X_a^i / (X_a^i + X_{t-a}^i)}{(X_a^i + X_a^{ROW}) / (X_a^i + X_{t-a}^i + X_a^{ROW} + X_{t-a}^{ROW})} \quad (1)$$

If country i completely specializes in product a , $X_{t-a}^i = 0$ and $X_a^{ROW} = 0$. Expression (1) reduces to,

$$\ln \frac{X_a^i / X_a^i}{(X_a^i) / (X_a^i + X_{t-a}^{ROW})}$$

The above expression would tend to $+\infty$ if $\frac{X_a^i}{X_a^i + X_{t-a}^{ROW}} \rightarrow 0$.

If ROW completely specializes in product a , $X_{t-a}^i = 0$ and $X_{t-a}^{ROW} = 0$. Expression (1) reduces to,

$$\ln \frac{0 / X_{t-a}^i}{X_a^{ROW} / (X_{t-a}^i + X_a^{ROW})} = -\infty$$

If $X_a^i / X_t^i \approx X_a^w / X_t^w$, the index value tends to 0, which is defined to be the comparative advantage neutral point.

A.1.6 RTA Index

The RTA index is expressed as

$$\frac{X_a^i/X_t^i}{X_a^w/X_t^w} - \frac{M_a^i/M_t^i}{M_a^w/M_t^w}$$

X_a^w , X_t^w , M_a^w and M_t^w can be disintegrated into separate expressions for country i and rest of the world (ROW) and into commodity a and all other commodities $t-a$. X_t^i , M_t^i can similarly be broken into separate expressions for commodity a and $t-a$. The RTA index can then be expressed as

$$\frac{X_a^i/(X_a^i + X_{t-a}^i)}{(X_a^i + X_a^{ROW})/(X_a^i + X_{t-a}^i + X_a^{ROW} + X_{t-a}^{ROW})} - \frac{M_a^i/(M_a^i + M_{t-a}^i)}{(M_a^i + M_a^{ROW})/(M_a^i + M_{t-a}^i + M_a^{ROW} + M_{t-a}^{ROW})} \quad (1)$$

If country i completely specializes in product a , $X_{t-a}^i = X_a^{ROW} = 0$ and $M_a^i = M_{t-a}^{ROW} = 0$. Expression (1) reduces to,

$$\frac{X_a^i/X_a^i}{(X_a^i)/(X_a^i + X_{t-a}^{ROW})} - \frac{0/M_{t-a}^i}{(M_a^{ROW})/(M_{t-a}^i + M_a^{ROW})}$$

The above expression would tend to $+\infty$ if $\frac{X_a^i}{X_a^i + X_{t-a}^{ROW}} \rightarrow 0$.

If ROW completely specializes in product a , $X_a^i = X_{t-a}^{ROW} = 0$ and $M_{t-a}^i = M_a^{ROW} = 0$. Expression (1) reduces to,

$$\frac{0/X_{t-a}^i}{X_a^{ROW}/(X_{t-a}^i + X_a^{ROW})} - \frac{M_a^i/M_a^i}{M_a^i/(M_a^i + M_{t-a}^{ROW})}$$

The above expression would tend to $-\infty$ if $\frac{M_a^i}{M_a^i + M_{t-a}^{ROW}} \rightarrow 0$.

If $\frac{X_a^i/X_t^i}{X_a^w/X_t^w} = \frac{M_a^i/M_t^i}{M_a^w/M_t^w}$, the index takes the value of 0, which is defined to be the comparative advantage neutral point.

A.1.7 RC Index

The RC index is expressed as

$$\ln \frac{X_a^i / X_t^i}{X_a^w / X_t^w} - \ln \frac{M_a^i / M_t^i}{M_a^w / M_t^w}$$

X_a^w , X_t^w , M_a^w and M_t^w can be disintegrated into separate expressions for country i and rest of the world (ROW) and into commodity a and all other commodities $t-a$. X_t^i M_t^i can similarly be broken into separate expressions for commodity a and $t-a$. The RC index can then be expressed as

$$\ln \frac{X_a^i / (X_a^i + X_{t-a}^i)}{(X_a^i + X_a^{ROW}) / (X_a^i + X_{t-a}^i + X_a^{ROW} + X_{t-a}^{ROW})} - \ln \frac{M_a^i / (M_a^i + M_{t-a}^i)}{(M_a^i + M_a^{ROW}) / (M_a^i + M_{t-a}^i + M_a^{ROW} + M_{t-a}^{ROW})} \quad (1)$$

If country i completely specializes in product a , $X_{t-a}^i = X_a^{ROW} = 0$ and $M_a^i = M_{t-a}^{ROW} = 0$. Expression (1) reduces to,

$$\ln \frac{X_a^i / X_a^i}{(X_a^i) / (X_a^i + X_{t-a}^{ROW})} - \ln \frac{0 / M_{t-a}^i}{(M_a^{ROW}) / (M_{t-a}^i + M_a^{ROW})}$$

The above expression would tend to $+\infty$ if $\frac{X_a^i}{X_a^i + X_{t-a}^{ROW}} \rightarrow 0$.

If ROW completely specializes in product a , $X_a^i = X_{t-a}^{ROW} = 0$ and $M_{t-a}^i = M_a^{ROW} = 0$. Expression (1) reduces to,

$$\ln \frac{0 / X_{t-a}^i}{X_a^{ROW} / (X_{t-a}^i + X_a^{ROW})} - \ln \frac{M_a^i / M_a^i}{M_a^i / (M_a^i + M_{t-a}^{ROW})}$$

The above expression would tend to $-\infty$ if $\frac{M_a^i}{M_a^i + M_{t-a}^{ROW}} \rightarrow 0$.

If $\frac{X_a^i / X_t^i}{X_a^w / X_t^w} = \frac{M_a^i / M_t^i}{M_a^w / M_t^w}$, the index takes the value of 0, which is defined to be the comparative advantage neutral point.

A.2 Calculation of Arithmetic Means of the RCA Indices

A.2.1 Balassa Index

Arithmetic mean across sectors:

$$\frac{1}{t} \sum_{a=1}^t \left(\frac{X_a^i / X_t^i}{X_a^w / X_t^w} \right) \neq 1$$

Arithmetic mean across countries:

$$\frac{1}{w} \sum_{i=1}^w \left(\frac{X_a^i / X_t^i}{X_a^w / X_t^w} \right) \neq 1$$

A.2.2 RSCA Index

Arithmetic mean across sectors:

$$\frac{1}{t} \sum_{a=1}^t \left(\frac{\frac{X_a^i / X_t^i}{X_a^w / X_t^w} - 1}{\frac{X_a^i / X_t^i}{X_a^w / X_t^w} + 1} \right) \neq 0$$

Arithmetic mean across countries:

$$\frac{1}{w} \sum_{i=1}^w \left(\frac{\frac{X_a^i / X_t^i}{X_a^w / X_t^w} - 1}{\frac{X_a^i / X_t^i}{X_a^w / X_t^w} + 1} \right) \neq 0$$

A.2.3 ARCA Index

Arithmetic mean across sectors:

$$\frac{1}{t} \sum_{a=1}^t \frac{X_a^i}{X_t^i} - \frac{1}{t} \sum_{a=1}^t \frac{X_a^w}{X_t^w} = 0$$

Arithmetic mean across countries:

$$\frac{1}{w} \sum_{i=1}^w \frac{X_a^i}{X_t^i} - \frac{1}{w} \sum_{i=1}^w \frac{X_a^w}{X_t^w} \neq 0$$

A.2.4 NRCA Index

Arithmetic mean across sectors:

$$\frac{1}{t} \sum_{a=1}^t \frac{X_a^i}{X_t^w} - \frac{1}{t} \sum_{a=1}^t \frac{X_t^i}{X_t^w} \frac{X_a^w}{X_t^w} = 0$$

Arithmetic mean across countries:

$$\frac{1}{w} \sum_{i=1}^w \frac{X_a^i}{X_t^w} - \frac{1}{w} \sum_{i=1}^w \frac{X_t^i}{X_t^w} \frac{X_a^w}{X_t^w} = 0$$

A.2.5 Log of Balassa index

Arithmetic mean across sectors:

$$\frac{1}{t} \sum_{a=1}^t \ln \left(\frac{X_a^i / X_t^i}{X_a^w / X_t^w} \right) \neq 0$$

Arithmetic mean across countries:

$$\frac{1}{w} \sum_{i=1}^w \ln \left(\frac{X_a^i / X_t^i}{X_a^w / X_t^w} \right) \neq 0$$

A.2.6 RTA Index

Arithmetic mean across sectors:

$$\frac{1}{t} \sum_{a=1}^t \frac{X_a^i / X_t^i}{X_a^w / X_t^w} - \frac{1}{t} \sum_{a=1}^t \frac{M_a^i / M_t^i}{M_a^w / M_t^w} \neq 0$$

Arithmetic mean across countries:

$$\frac{1}{w} \sum_{i=1}^w \frac{X_a^i/X_t^i}{X_a^w/X_t^w} - \frac{1}{w} \sum_{i=1}^w \frac{M_a^i/M_t^i}{M_a^w/M_t^w} \neq 0$$

A.2.7 RC Index

Arithmetic mean across sectors:

$$\frac{1}{t} \sum_{a=1}^t \ln \frac{X_a^i/X_t^i}{X_a^w/X_t^w} - \frac{1}{t} \sum_{a=1}^t \ln \frac{M_a^i/M_t^i}{M_a^w/M_t^w} \neq 0$$

Arithmetic mean across countries:

$$\frac{1}{w} \sum_{i=1}^w \ln \frac{X_a^i/X_t^i}{X_a^w/X_t^w} - \frac{1}{w} \sum_{i=1}^w \ln \frac{M_a^i/M_t^i}{M_a^w/M_t^w} \neq 0$$

A.3 Derivation of Expression (2.27) from Expression (2.26) in Chapter 2

Expression (24) is given by:

$$V_r B_{rr} E_{rs} + V_s B_{sr} E_{rs} = u E_{rs}$$

As defined in Chapter 2, E_{rs} is a (2 x 1) vector of exports from country r to country s .

Since exports of country r must either be absorbed in final consumption in country s (Y_{rs}), or for producing gross output of country s ($A_{rs}X_s$), expression (24) can be written as:

$$V_r B_{rr} (Y_{rs} + A_{rs}X_s) + V_s B_{sr} E_{rs} = u E_{rs}$$

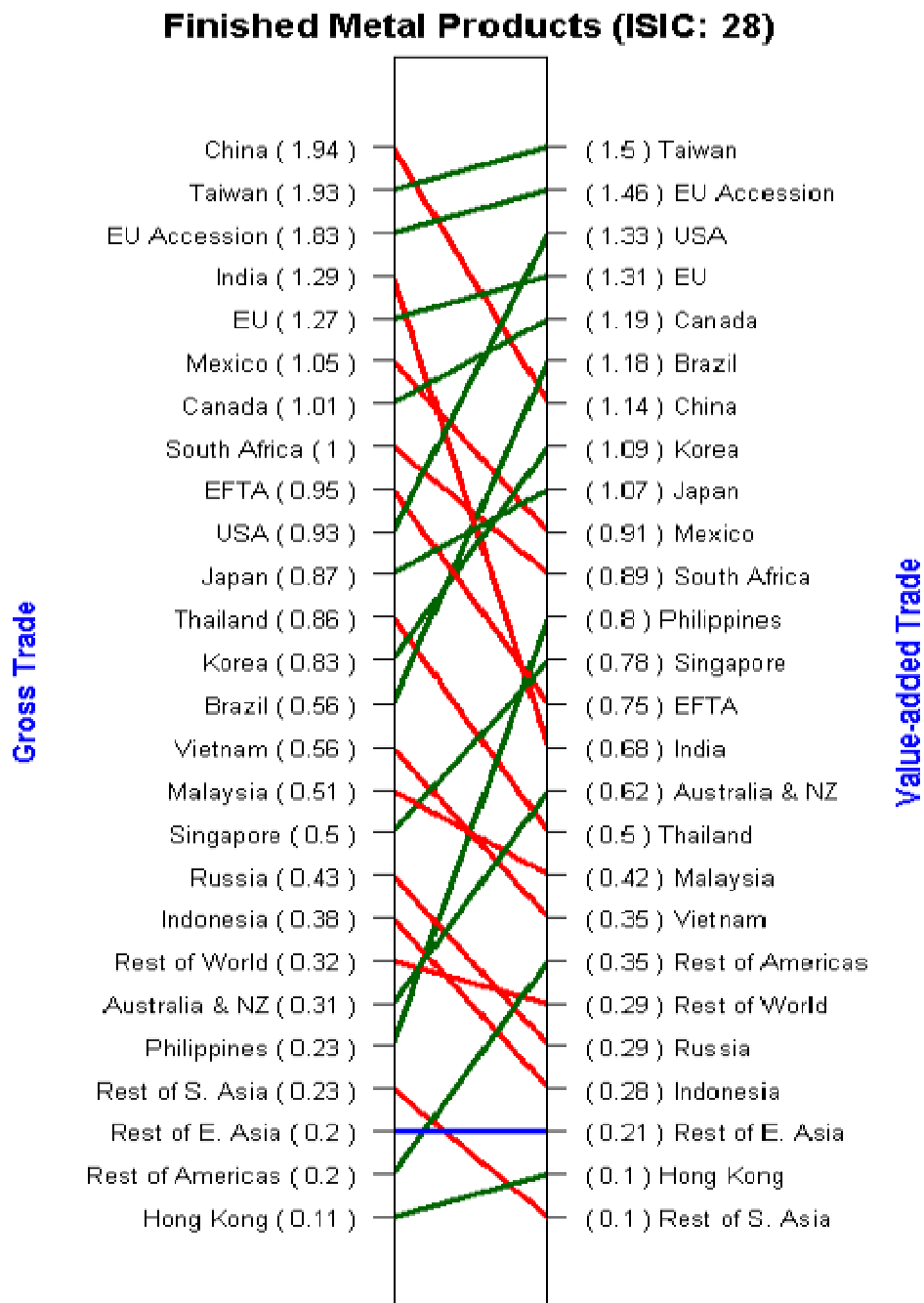
The gross output of s could be absorbed in s (X_{ss}), or in country t (X_{st}) or exported back to r (X_{sr}) after processing in s . Therefore the above expression can be written as:

$$V_r B_{rr} Y_{rs} + V_r B_{rr} A_{rs} (X_{ss} + X_{st} + X_{sr}) + V_s B_{sr} E_{rs} = u E_{rs}$$

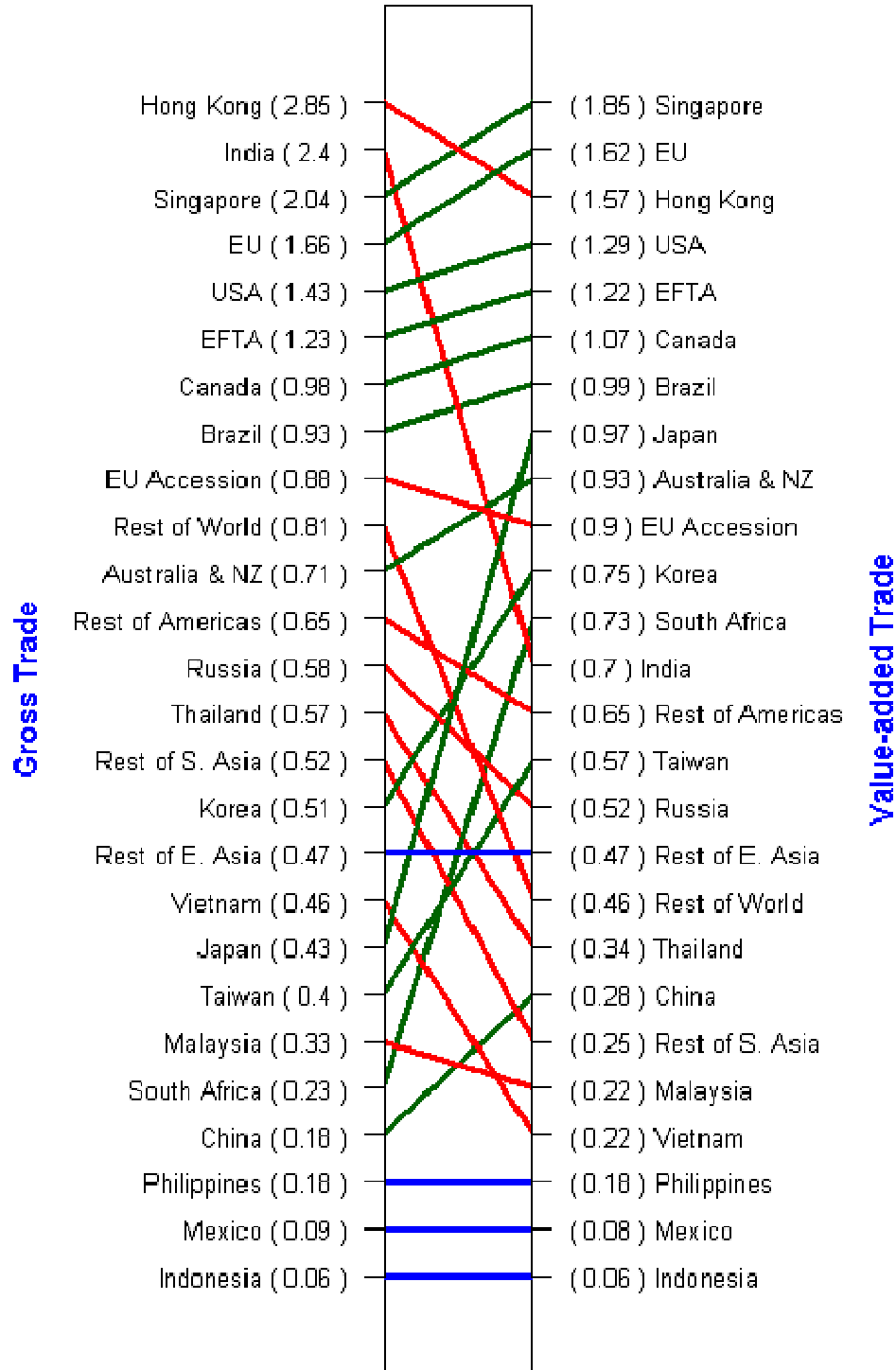
Rearranging and collecting terms in the above expression, we arrive at expression (25):

$$u E_{rs} = V_r B_{rr} Y_{rs} + V_r B_{rr} A_{rs} X_{ss} + V_r B_{rr} A_{rs} X_{st} + V_r B_{rr} A_{rs} X_{sr} + V_s B_{sr} E_{rs}$$

A.4: Comparison between Gross Trade based and Value-Added in Trade based Balassa Index Values for Selected Sectors



Real Estate, Renting and Business Activities (ISIC: K)



Source: Koopman *et al.* (2010, 2014)

Appendix B

Table B.1
Summary Statistics of the Data used in Chapter 3

Variable	No. of Observations	Mean	Median	Standard Deviation	Skewness	Kurtosis
Exports for 1998	1677	2472142	193107	9030313	8.583258	94.72834
Imports for 1998	1677	2422142	342177.5	8274205	10.83427	167.7656
Exports for 1999	1677	2547141	202829.7	9341102	8.467233	91.38276
Imports for 1999	1677	2547273	323409.5	9121551	11.54007	188.5304
Exports for 2000	1677	2788442	214887	10300000	8.499637	92.91693
Imports for 2000	1677	2809932	354286.9	10300000	11.80609	195.0212

Note: Export and Import figures are reported in thousands of US\$ in Trade, Production and Protection Database.

Figure B.1
Cumulative Distribution Plots for China

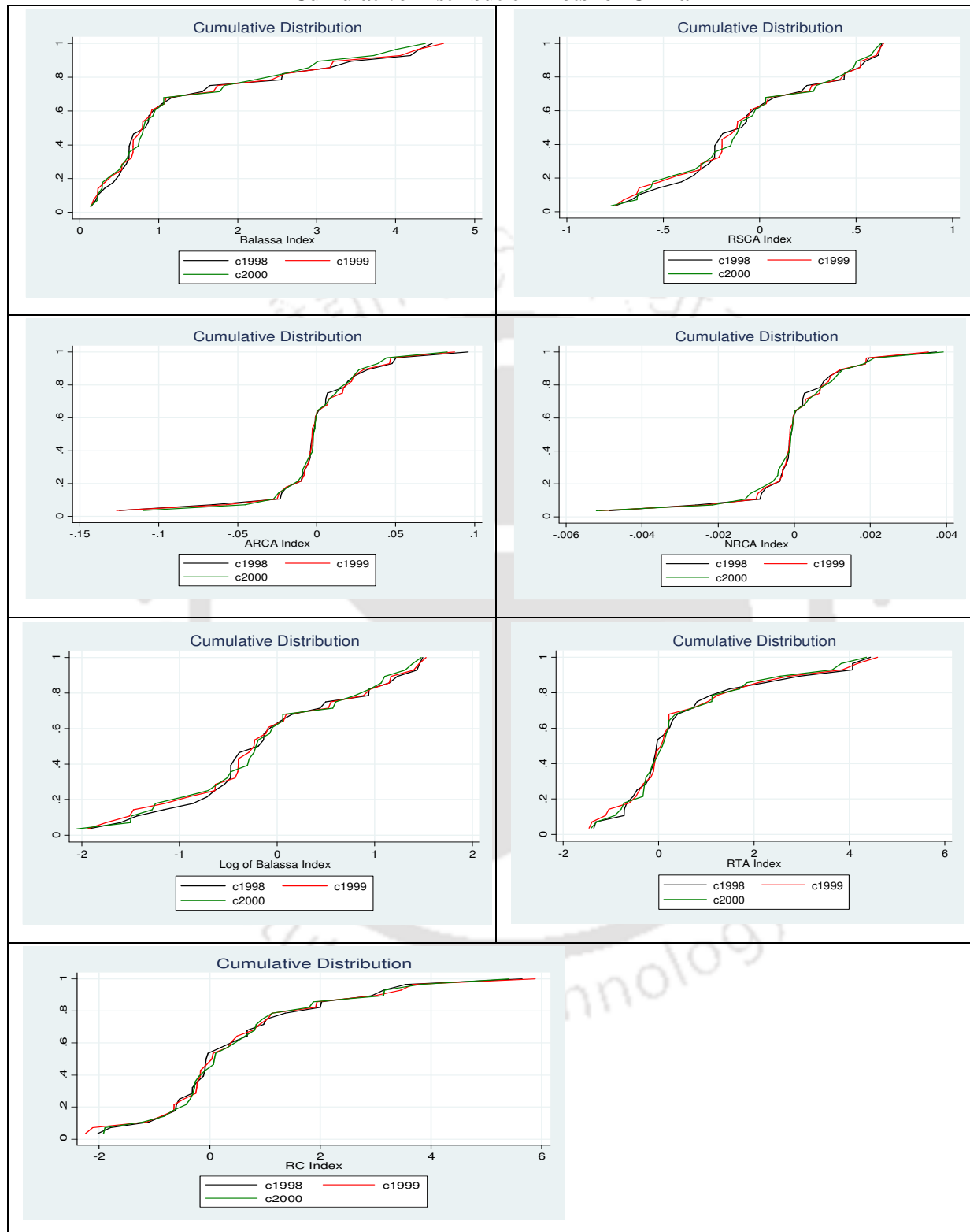


Figure B.2
Cumulative Distribution Plots for France

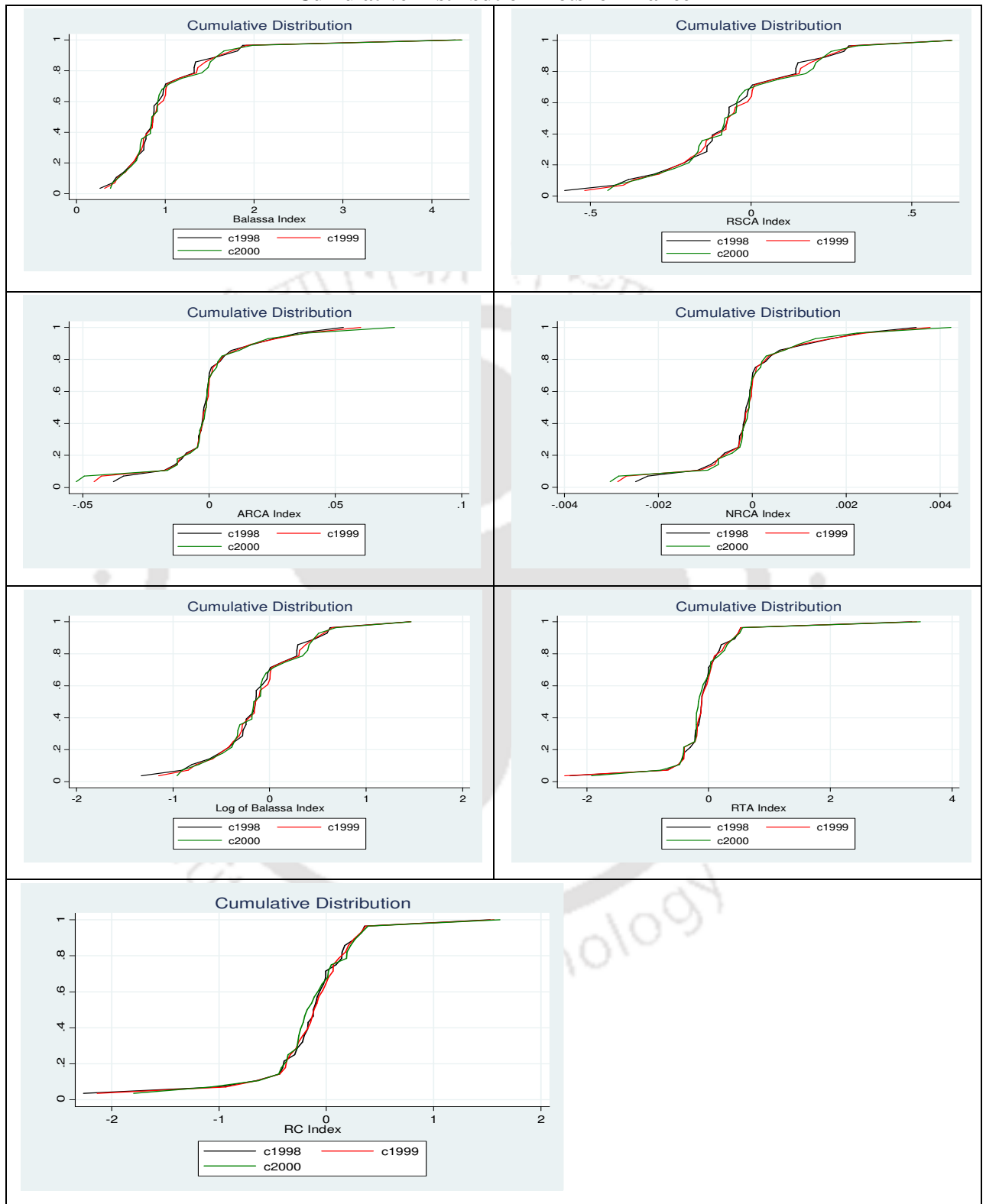


Figure B.3
Cumulative Distribution Plots for Germany

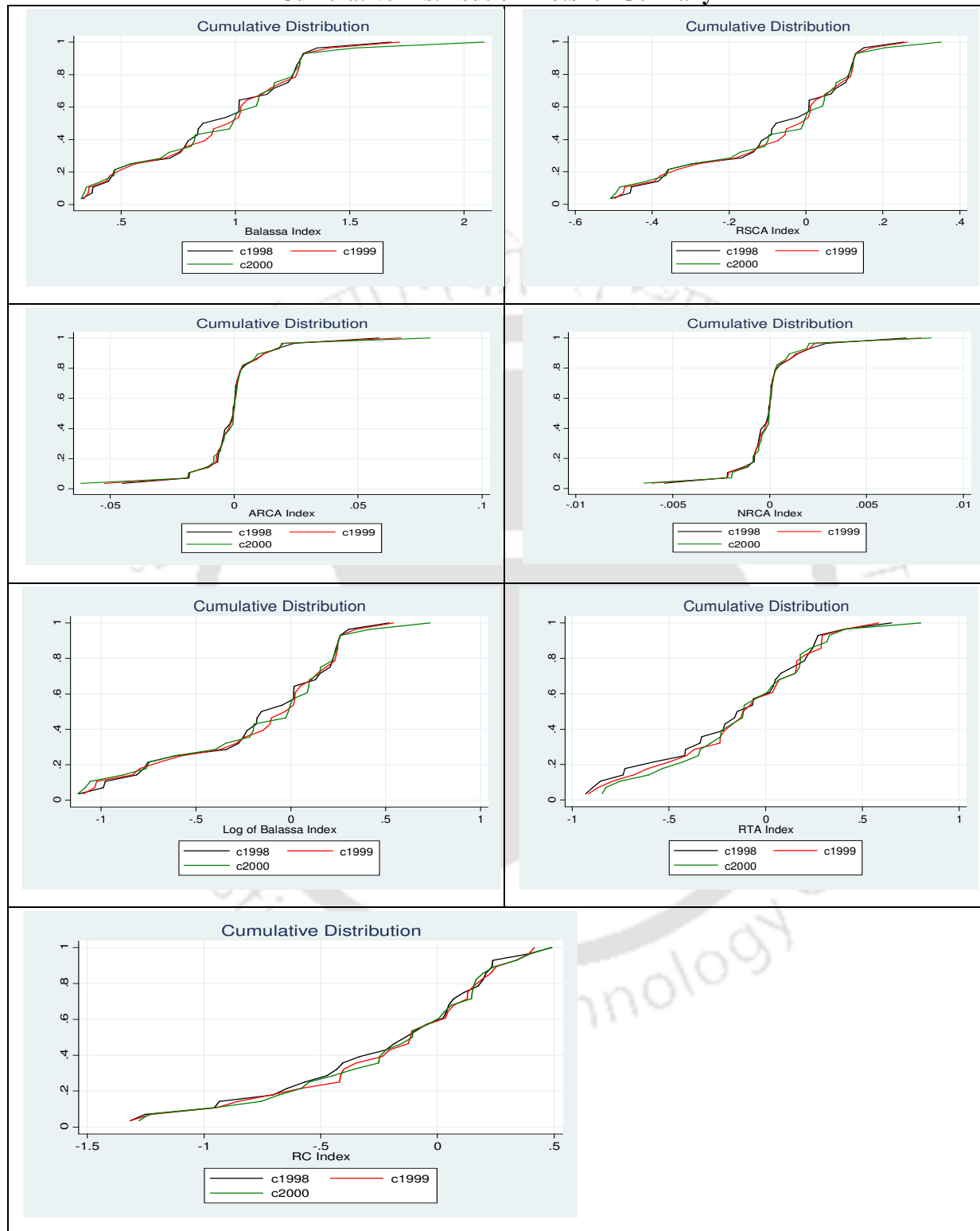


Figure B.4
Cumulative Distribution Plots for India

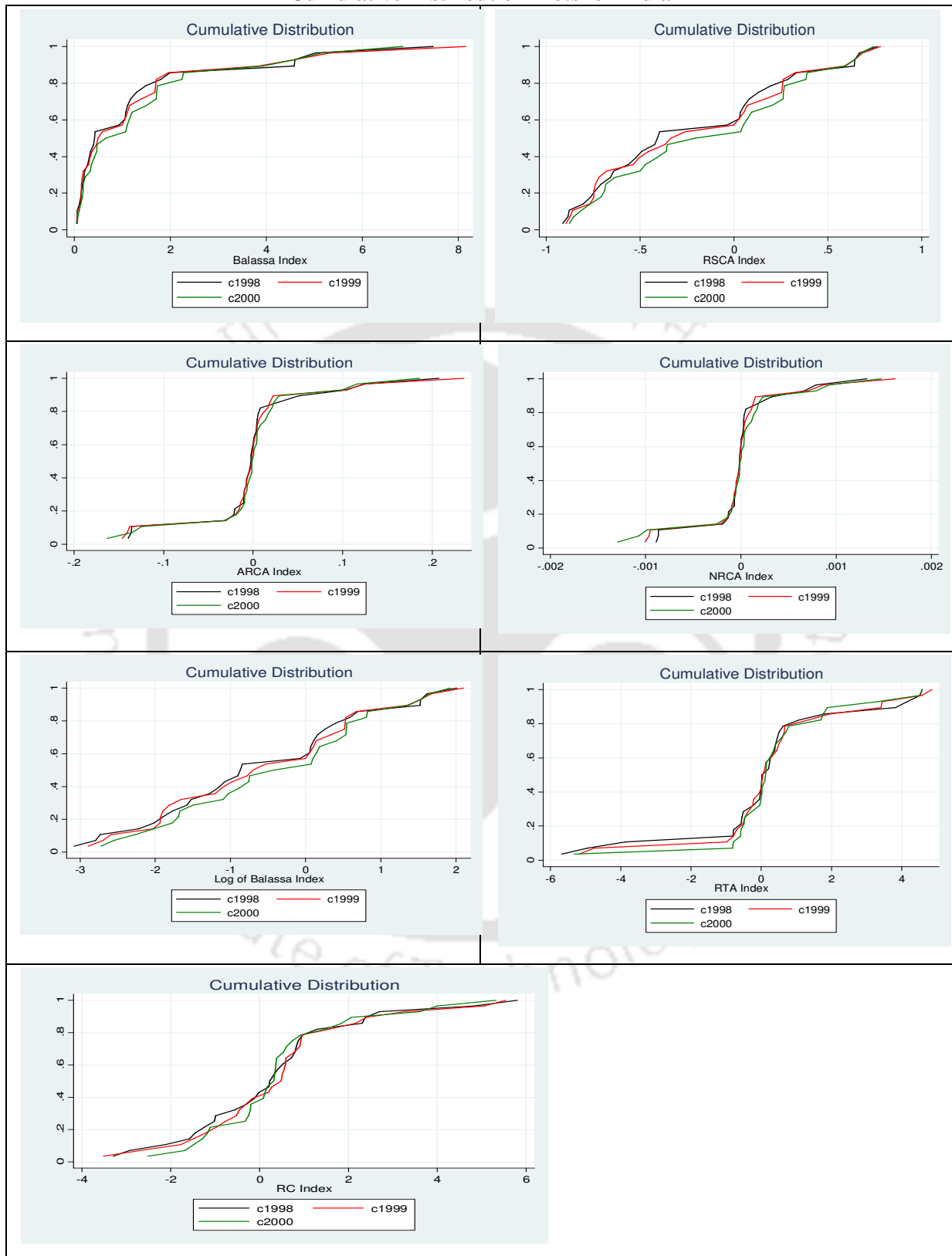


Figure B.5
Cumulative Distribution Plots for Italy

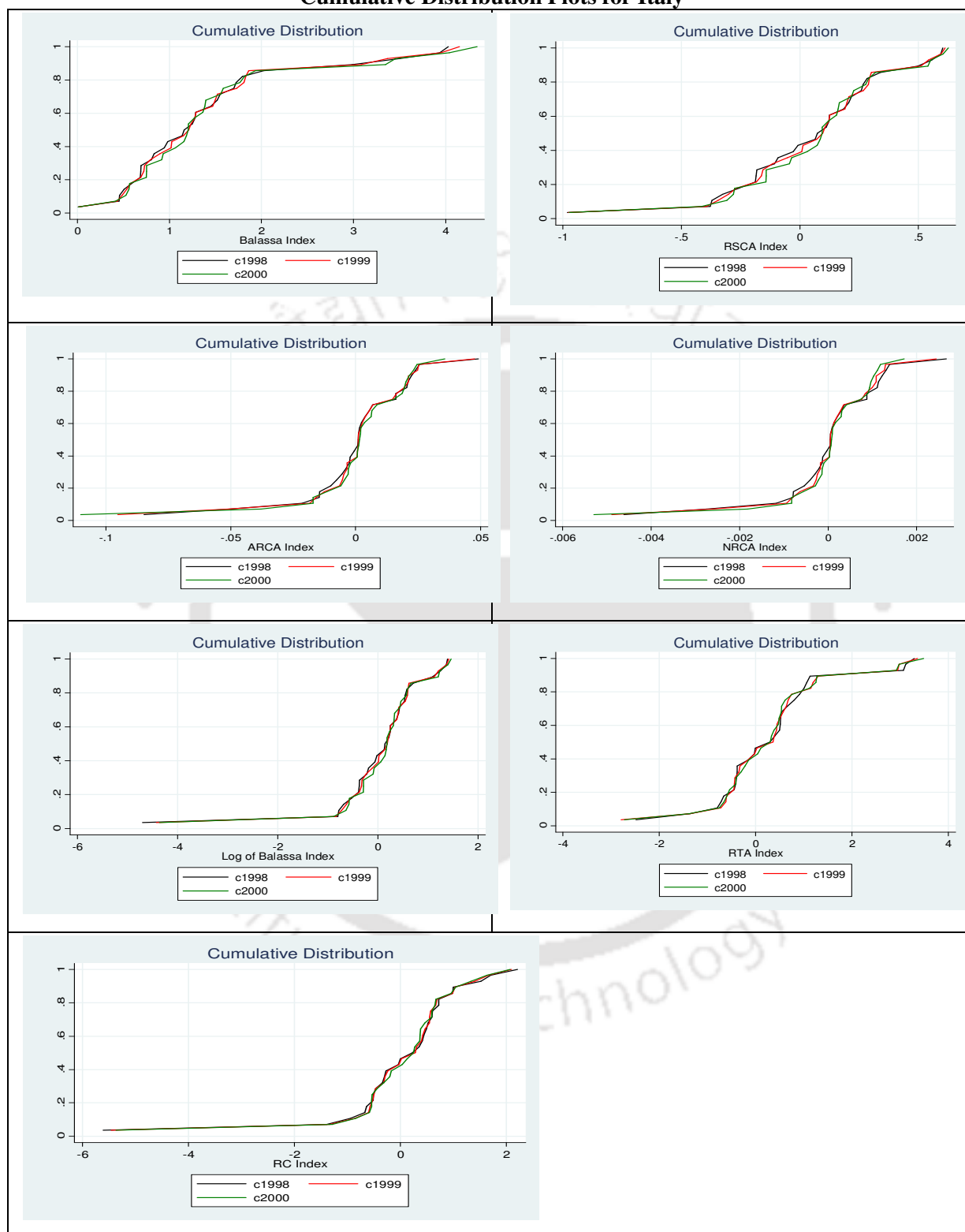


Figure B.6
Cumulative Distribution Plots for Japan

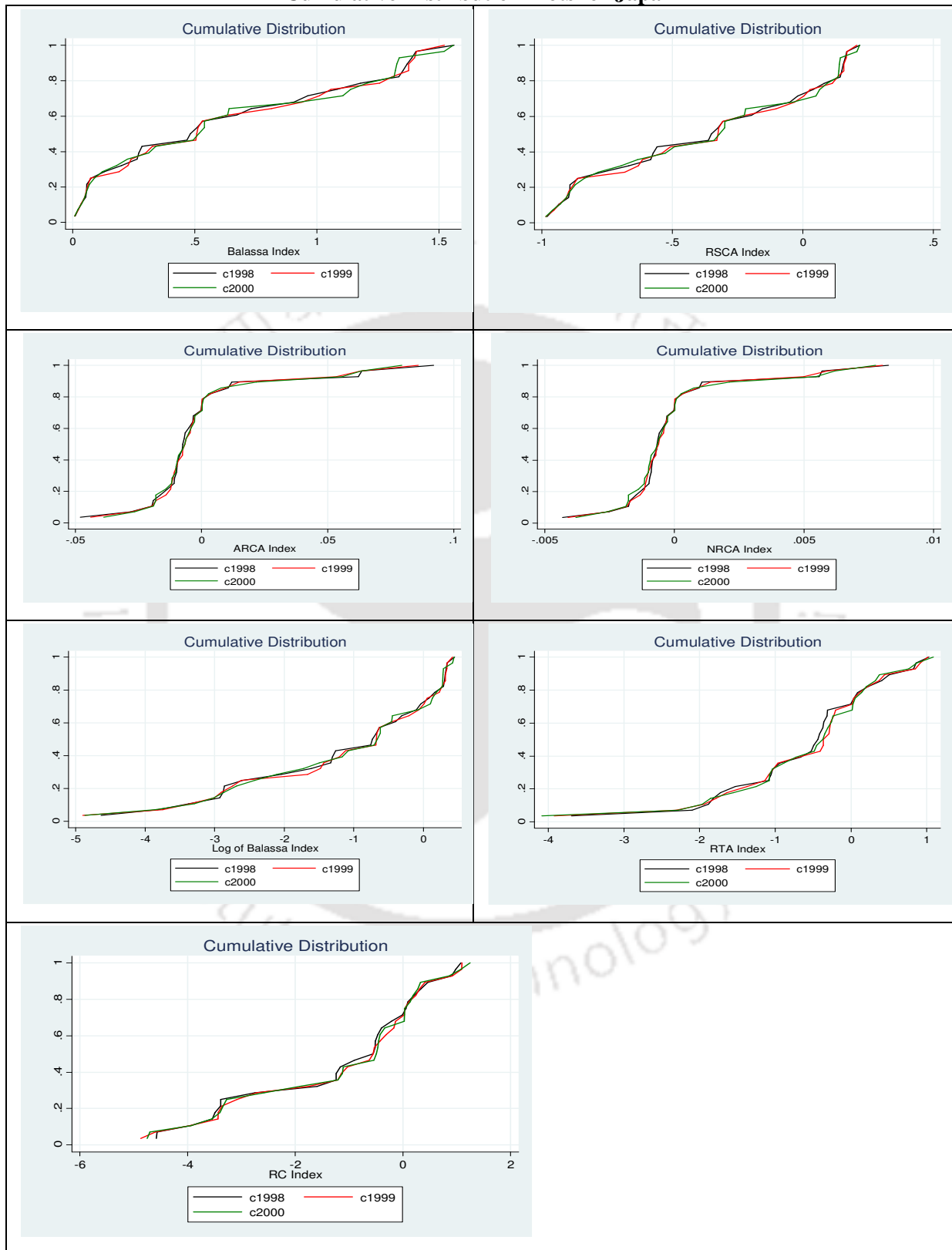


Figure B.7
Kernel Density Plots for China

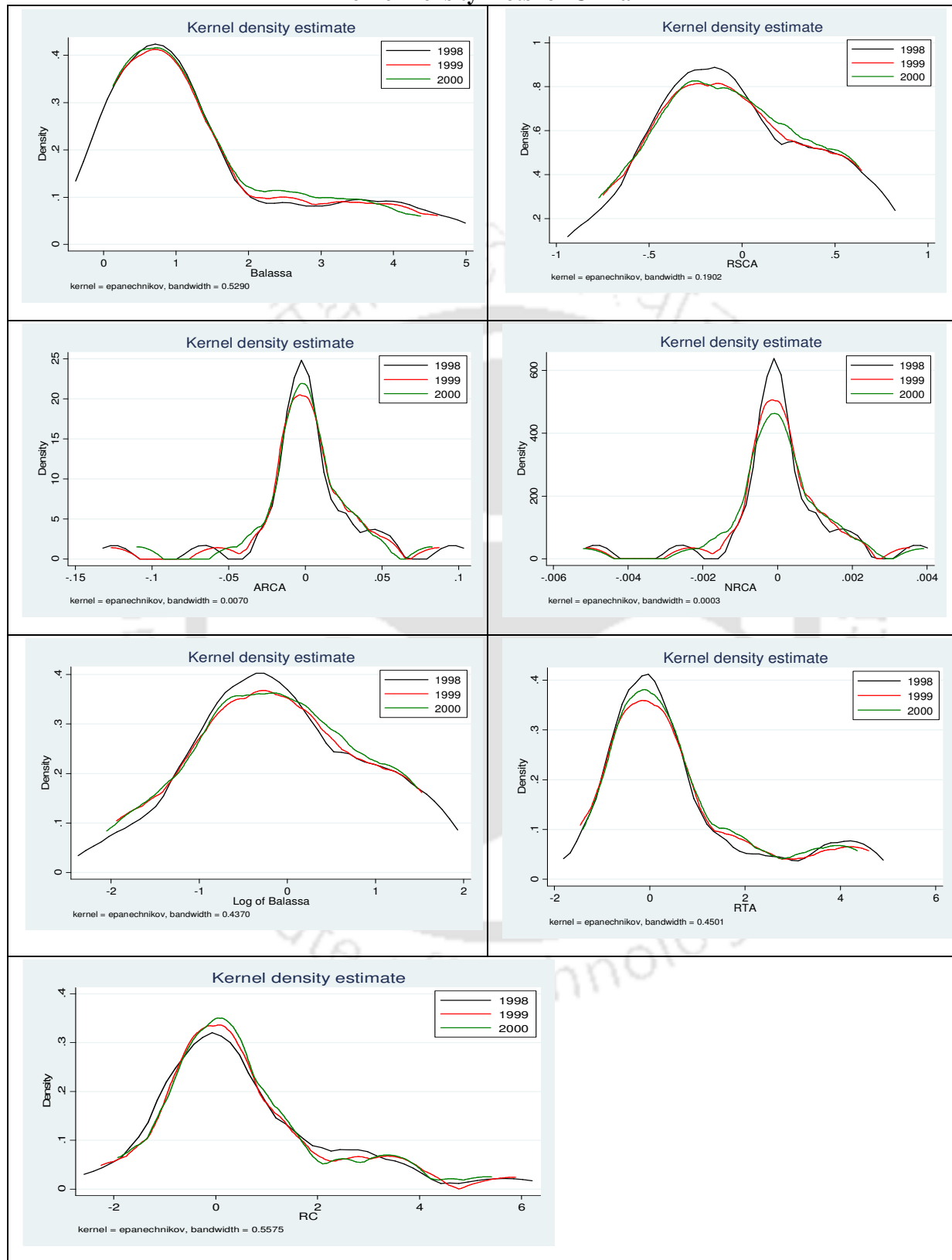


Figure B.8
Kernel Density Plots for France

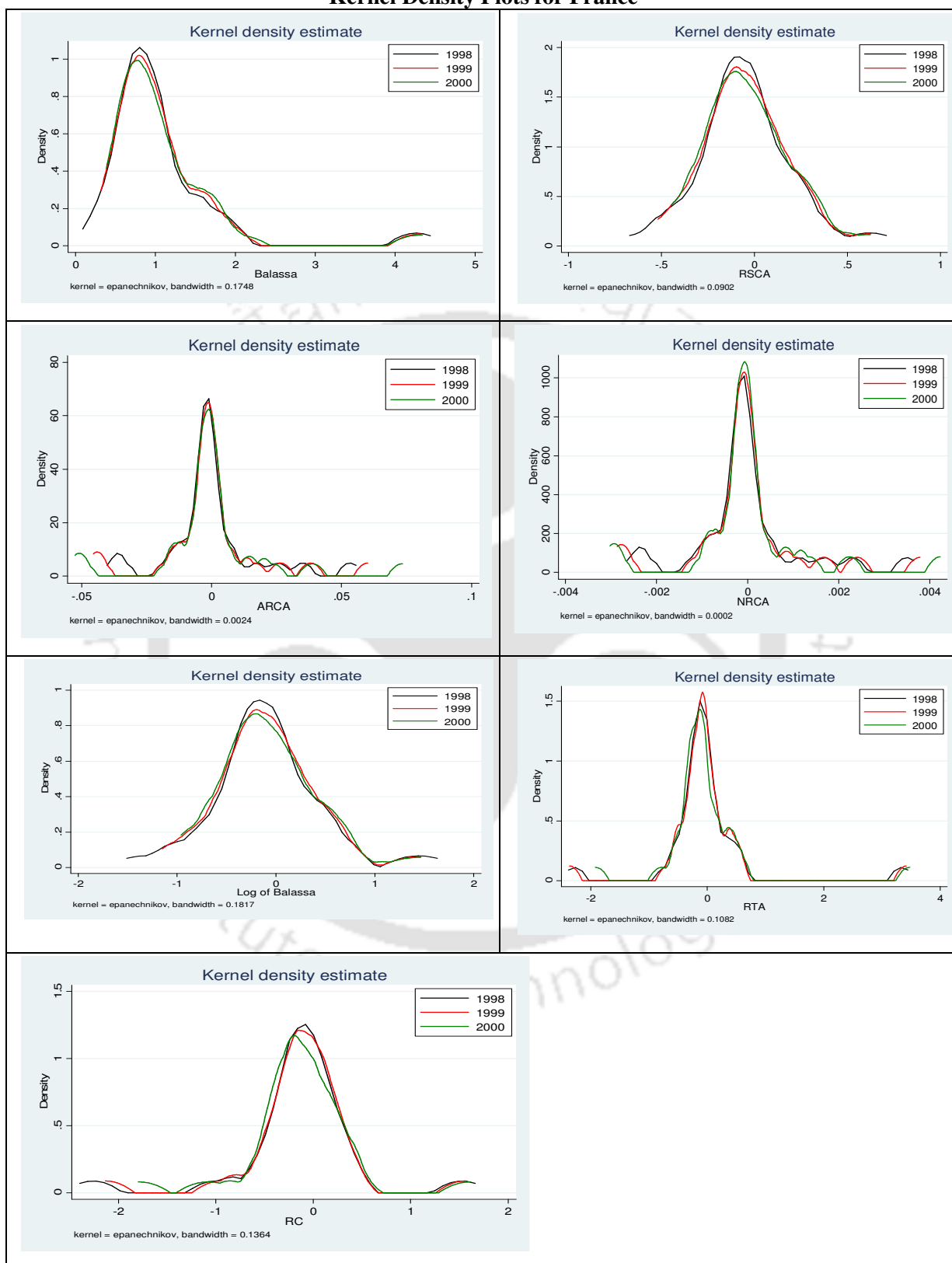


Figure B.9
Kernel Density Plots for Germany

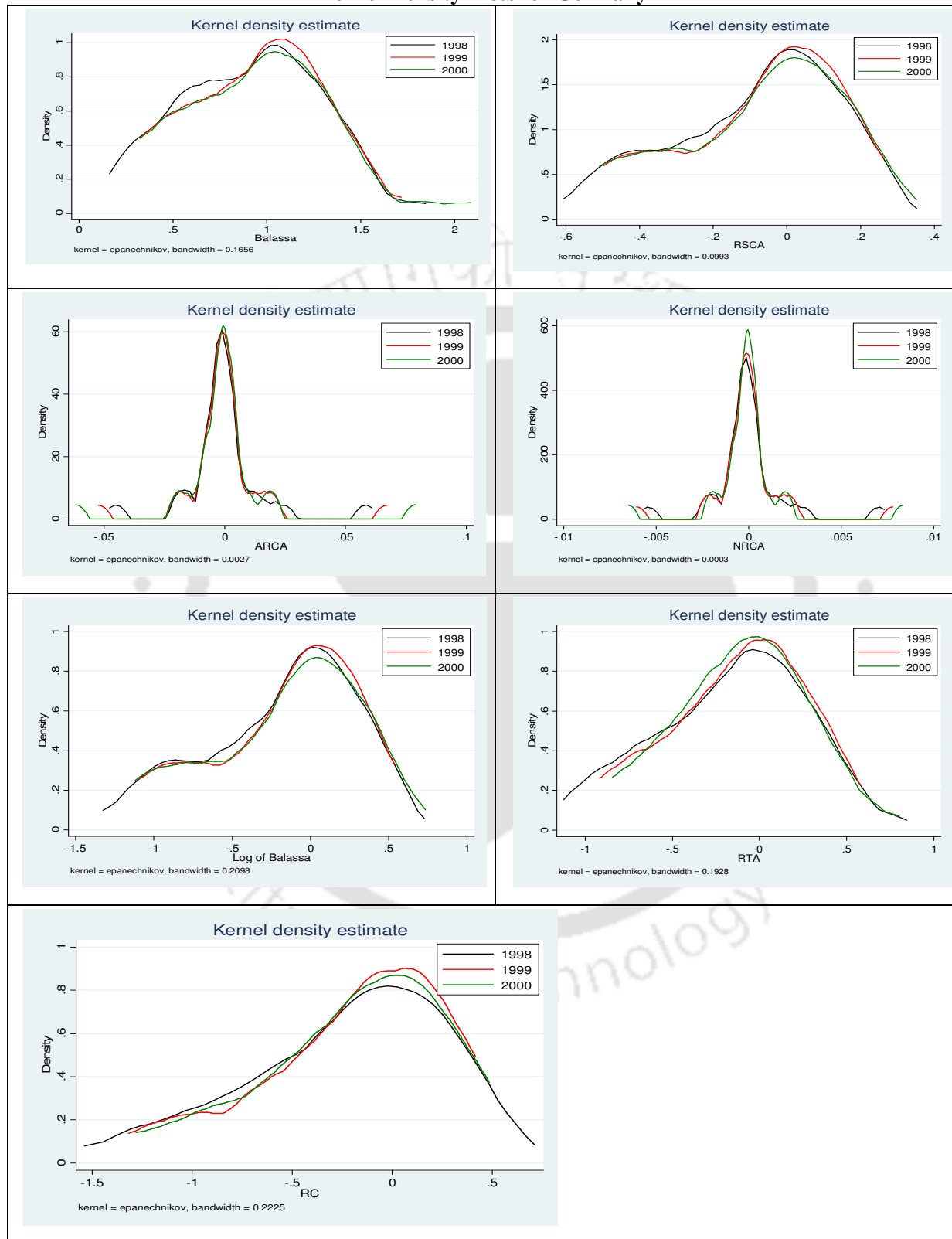


Figure B.10
Kernel Density Plots for India

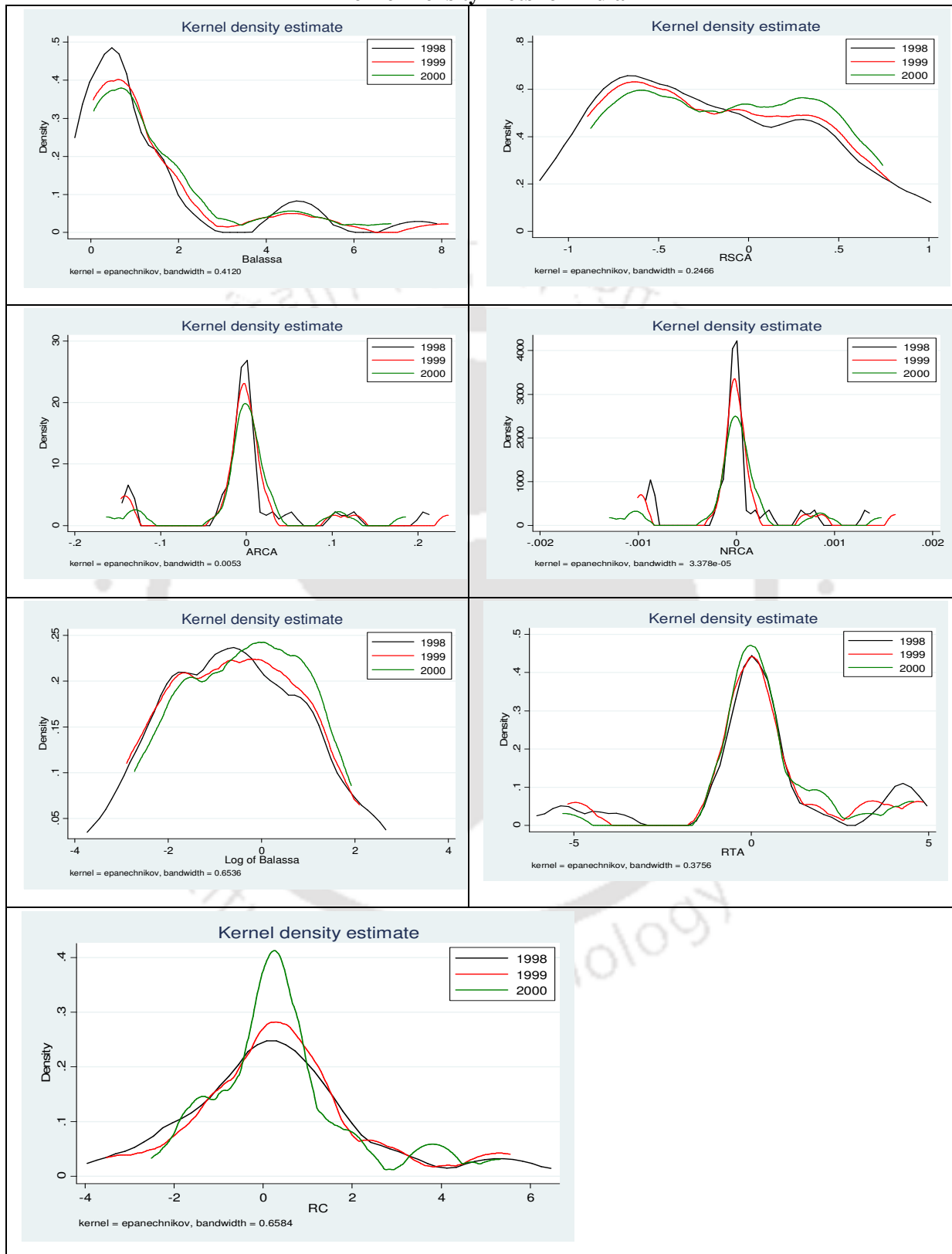


Figure B.11
Kernel Density Plots for Italy

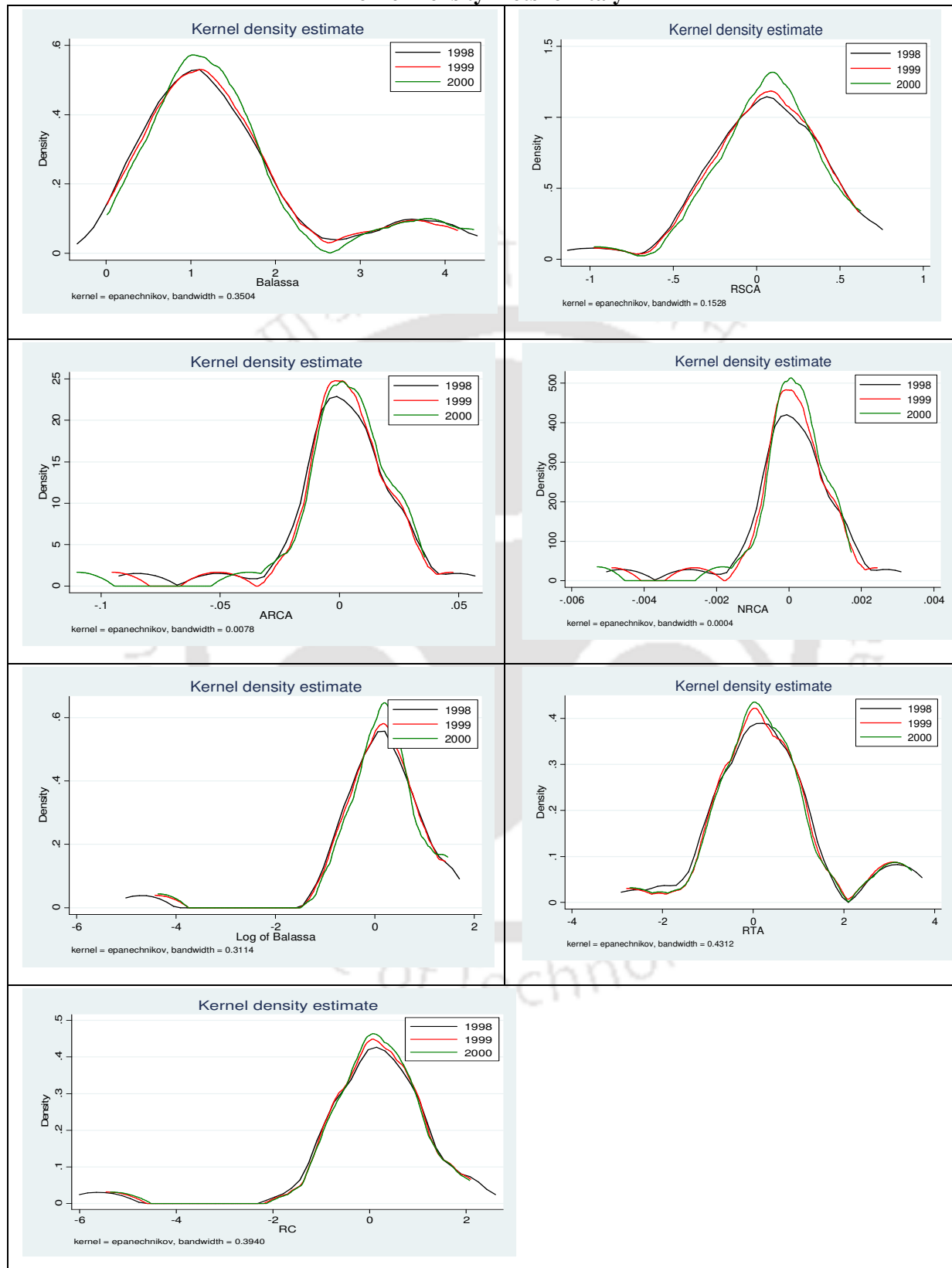


Figure B.12
Kernel Density Plots for Japan

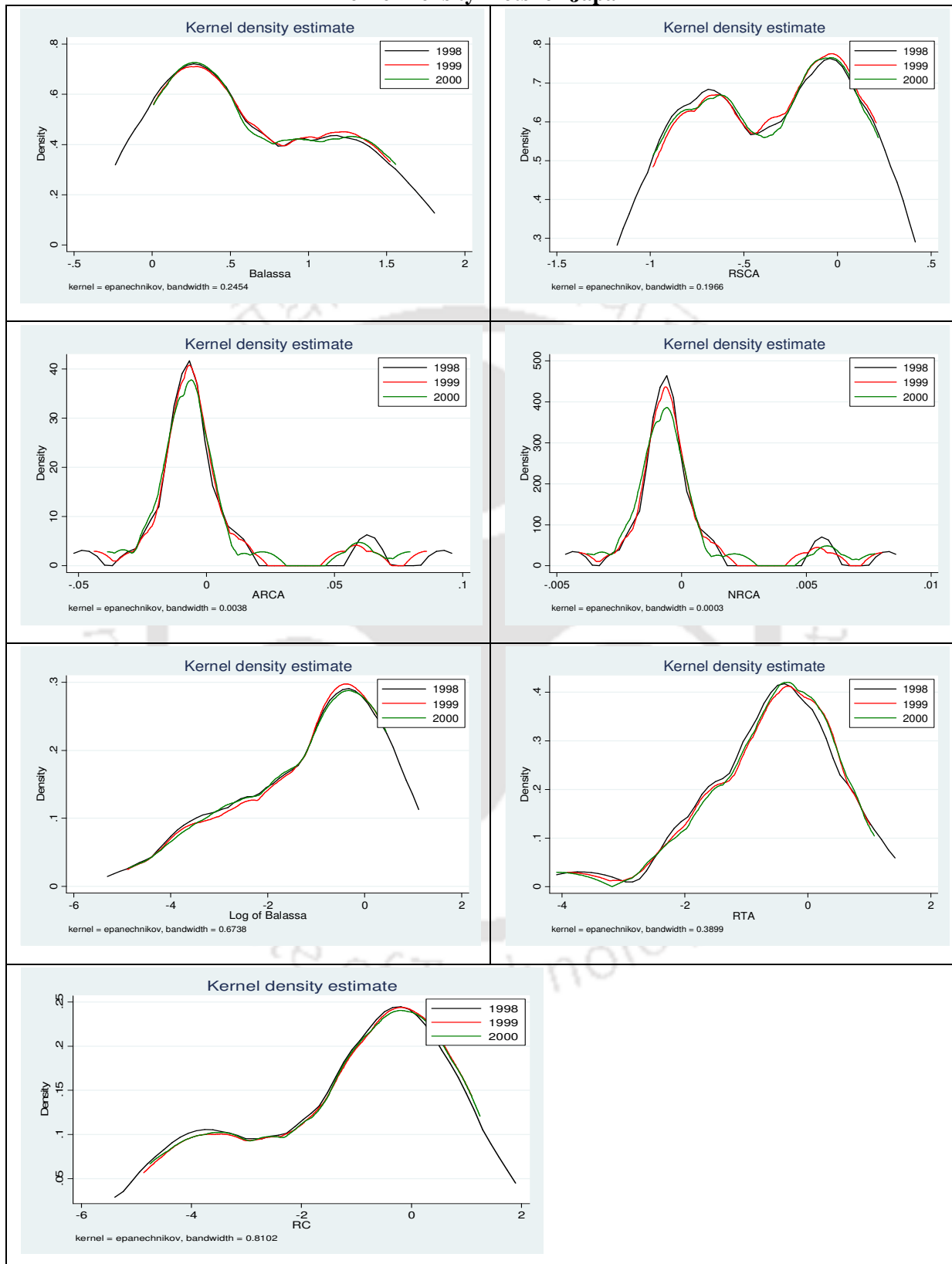


Table B.2
Sector Description on the basis of ISIC Rev.2

ISIC Code	Sector Description
311-312	Food manufacturing
313	Beverage industries
314	Tobacco manufactures
321	Manufacture of textiles
322	Manufacture of wearing apparel, except footwear
323	Manufacture of leather and products of leather, leather substitutes and fur, except footwear and wearing apparel
324	Manufacture of footwear, except vulcanized or moulded rubber or plastic footwear
331	Manufacture of wood and wood and cork products, except furniture
332	Manufacture of wooden furniture and fixtures, excluding furniture made primarily of metal
341	Manufacture of paper and paper products
342	Printing, publishing and allied industries
351	Manufacture of industrial chemicals
352	Manufacture of other chemical products
353	Petroleum refineries
354	Manufacture of miscellaneous products of petroleum and coal
355	Manufacture of rubber products
356	Manufacture of plastic products not elsewhere classified
361	Manufacture of pottery, china and earthenware
362	Manufacture of glass and glass products
369	Manufacture of other non-metallic mineral products
371	Iron and steel basic industries
372	Non-ferrous metal basic industries
381	Manufacture of fabricated metal products, except machinery and equipment
382	Manufacture of machinery except electrical
383	Manufacture of electrical machinery apparatus, appliances and supplies
384	Manufacture of transport equipment
385	Manufacture of professional and scientific, and measuring and controlling equipment not elsewhere classified, and of photographic and optical goods
390	Other Manufacturing Industries producing jewellery, sports goods, musical instruments, toys etc.

Note: ISIC denotes International Standard Industrial Classification of all Economic Activities.

Table B.3 (Corresponding to Table 3.2 in Chapter 3)
Results for Wilcoxon's Signed Rank Test for Countries

Country	Year combinations	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
China	1998-1999	-0.250 (0.802)	-0.137 (0.891)	1.093 (0.274)	0.729 (0.466)	0.046 (0.964)	0.410 (0.682)	-0.091 (0.927)
	1999-2000	0.911 (0.362)	0.296 (0.767)	0.888 (0.375)	0.182 (0.855)	0.524 (0.601)	0.091 (0.927)	-0.638 (0.524)
	1998-2000	0.911 (0.362)	0.159 (0.873)	0.478 (0.633)	0.091 (0.927)	0.25 (0.802)	0.797 (0.426)	-0.433 (0.665)
Accepted cases for null hyp.		3	3	3	3	3	3	3
Germany	1998-1999	-1.412 (0.158)	-1.002 (0.316)	-0.615 (0.539)	-0.729 (0.466)	-1.025 (0.306)	-1.435 (0.151)	-1.435 (0.151)
	1999-2000	0.25 (0.802)	0.683 (0.495)	0.182 (0.855)	-0.137 (0.891)	0.729 (0.466)	-1.753 (0.080)	-0.751 (0.452)
	1998-2000	0.046 (0.964)	0.273 (0.785)	0.319 (0.750)	-0.342 (0.733)	0.342 (0.733)	-1.89 (0.059)	-0.956 (0.339)
Accepted cases for null hyp.		3	3	3	3	3	3	3
France	1998-1999	-1.594 (0.111)	-1.594 (0.111)	-0.865 (0.387)	-1.184 (0.236)	-1.662 (0.096)	-0.455 (0.649)	-1.047 (0.295)
	1999-2000	-0.592 (0.554)	-0.455 (0.649)	-0.228 (0.820)	-0.228 (0.820)	-0.41 (0.682)	0.524 (0.601)	0.615 (0.539)
	1998-2000	-1.23 (0.219)	-1.002 (0.316)	-0.41 (0.682)	-0.455 (0.649)	-1.047 (0.295)	0.25 (0.802)	0.433 (0.665)
Accepted cases for null hyp.		3	3	3	3	3	3	3
India	1998-1999	-1.594 (0.111)	-1.594 (0.111)	-0.41 (0.682)	-0.068 (0.946)	-1.457 (0.145)	-2.095 (0.036)	-1.776 (0.076)
	1999-2000	-2.254* (0.024)	-2.71** (0.007)	-1.503 (0.133)	-0.25 (0.802)	-2.73** (0.006)	0.888 (0.375)	1.526 (0.127)
	1998-2000	-2.163* (0.031)	-2.85** (0.004)	-0.888 (0.375)	-0.296 (0.767)	-2.82** (0.005)	0.159 (0.873)	-0.023 (0.982)
Accepted cases for null hyp.		1	1	3	3	1	2	3
Italy	1998-1999	-1.571 (0.116)	-1.571 (0.116)	-0.387 (0.699)	0.114 (0.909)	-1.708 (0.088)	-0.364 (0.716)	-0.228 (0.820)
	1999-2000	-2.004* (0.045)	-1.708 (0.088)	-0.820 (0.412)	-0.182 (0.855)	-1.958 (0.050)	-0.114 (0.909)	0.114 (0.909)
	1998-2000	-2.049* (0.040)	-2.141* (0.032)	-1.025 (0.306)	-0.159 (0.873)	-2.300* (0.022)	-0.25 (0.802)	-0.387 (0.699)
Accepted cases for null hyp.		1	2	3	3	2	3	3
Japan	1998-1999	-1.366 (0.172)	-1.64 (0.101)	-0.911 (0.362)	-0.023 (0.982)	-1.344 (0.179)	-0.319 (0.750)	-1.776 (0.076)
	1999-2000	0.205 (0.838)	-0.068 (0.946)	-1.207 (0.228)	0.182 (0.855)	-0.273 (0.785)	0.228 (0.820)	<0.001 (1.000)
	1998-2000	-0.137 (0.891)	-0.843 (0.400)	-1.184 (0.236)	0.342 (0.733)	-0.865 (0.387)	<0.001 (1.000)	-1.298 (0.194)
Accepted cases for null hyp.		3	3	3	3	3	3	3
Total no. of accepted cases for null hypothesis		14	15	18	18	15	17	18

Note: For each index only the standardized normal approximation of the test statistic and the corresponding p value is reported. * denotes significant at 5% level. ** denotes significant at 1% level.

Table B.4 (Corresponding to Table 3.2.1 in Chapter 3)
Results for Wilcoxon's Signed Rank Test for Sectors

ISIC Codes	Year Combinations	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
321	1998-1999	0.994 (0.320)	1.458 (0.145)	1.163 (0.245)	1.244 (0.214)	1.428 (0.153)	1.494 (0.135)	2.319* (0.020)
	1999-2000	1.347 (0.178)	1.796 (0.073)	1.737 (0.082)	1.141 (0.254)	1.561 (0.119)	2.407* (0.016)	2.194* (0.028)
	1998-2000	1.759 (0.079)	2.746** (0.006)	2.047* (0.041)	1.708 (0.088)	2.466* (0.014)	1.936 (0.053)	3.489** (0.001)
Accepted cases for null hypothesis		3	2	2	3	2	2	0
322	1998-1999	0.493 (0.622)	1.038 (0.299)	0.486 (0.627)	0.037 (0.971)	0.913 (0.361)	0.530 (0.596)	2.15* (0.032)
	1999-2000	0.648 (0.517)	1.907 (0.057)	1.178 (0.239)	0.361 (0.718)	1.62 (0.105)	1.053 (0.293)	3.062** (0.002)
	1998-2000	0.670 (0.503)	1.708 (0.088)	1.112 (0.266)	0.773 (0.440)	1.340 (0.180)	0.751 (0.453)	3.254** (0.001)
Accepted cases for null hypothesis		3	3	3	3	3	3	0
323	1998-1999	1.310 (0.190)	2.002** (0.045)	1.259 (0.208)	1.185 (0.236)	1.737 (0.082)	1.259 (0.208)	1.575 (0.115)
	1999-2000	0.913 (0.361)	1.090 (0.276)	1.045 (0.296)	0.317 (0.752)	0.891 (0.373)	0.905 (0.365)	1.178 (0.239)
	1998-2000	1.053 (0.293)	1.929 (0.054)	1.148 (0.251)	1.296 (0.195)	1.656 (0.098)	0.751 (0.453)	2.238* (0.025)
Accepted cases for null hypothesis		3	2	3	3	3	3	2
324	1998-1999	-0.420 (0.675)	0.213 (0.831)	-0.802 (0.422)	-0.434 (0.664)	0.309 (0.757)	0.088 (0.930)	0.770 (0.441)
	1999-2000	2.746** (0.006)	3.467** (0.001)	0.604 (0.546)	-0.346 (0.729)	3.335** (0.001)	3.026** (0.003)	5.109** (<0.001)
	1998-2000	2.694** (0.007)	3.107** (0.002)	0.427 (0.669)	-0.604 (0.546)	2.830** (0.005)	2.967** (0.003)	4.061** (<0.001)
Accepted cases for null hypothesis		1	1	3	3	1	1	1
351	1998-1999	0.015 (0.988)	-0.346 (0.729)	-0.420 (0.675)	-0.928 (0.354)	-0.317 (0.752)	1.274 (0.203)	-0.353 (0.724)
	1999-2000	-0.059 (0.953)	-0.066 (0.947)	0.199 (0.842)	-0.839 (0.401)	<0.001 (1.000)	0.125 (0.900)	-0.037 (0.971)
	1998-2000	-0.515 (0.606)	-0.655 (0.512)	-0.707 (0.480)	-0.810 (0.418)	-0.508 (0.612)	1.119 (0.263)	-0.110 (0.912)
Accepted cases for null hypothesis		3	3	3	3	3	3	3
352	1998-1999	0.361 (0.718)	0.817 (0.414)	1.141 (0.254)	1.207 (0.227)	0.802 (0.422)	0.361 (0.718)	1.06 (0.289)
	1999-2000	-2.039* (0.041)	-2.128* (0.033)	-2.878** (0.004)	-1.031 (0.303)	-2.231** (0.026)	-0.081 (0.936)	-1.200 (0.230)
	1998-2000	-0.368 (0.713)	-0.236 (0.814)	-1.310 (0.190)	0.640 (0.522)	-0.169 (0.866)	0.633 (0.527)	1.097 (0.273)
Accepted cases for null hypothesis		2	2	2	3	2	3	3
385	1998-1999	0.670 (0.503)	0.449 (0.653)	3.291** (0.001)	1.561 (0.119)	0.536 (0.592)	-0.854 (0.393)	-0.762 (0.446)

	1999-2000	0.876 (0.381)	0.824 (0.410)	3.136** (0.002)	1.325 (0.185)	0.707 (0.480)	-1.789 (0.074)	-1.038 (0.299)
	1998-2000	1.583 (0.114)	1.458 (0.145)	3.401** (0.001)	3.320** (0.001)	1.298 (0.194)	-1.575 (0.115)	-1.034 (0.301)
Accepted cases for null hypothesis		3	3	0	2	3	3	3
Total no. of accepted cases for null hypothesis		18	16	16	20	17	18	12

Note: For each index only the standardized normal approximation of the test statistic and the corresponding p value is reported. * denotes significant at 5% level. ** denotes significant at 1% level.

Table B.5 (Corresponding to Table 3.3 in Chapter 3)
Results for Wilcoxon's Rank Sum Test for Sectors

Sector Combinations	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
321-322	-0.698 (0.485)	-0.698 (0.485)	-0.724 (0.469)	-0.835 (0.404)	-0.698 (0.485)	-3.506** (0.001)	-3.769** (0.0002)
321-323	0.388 (0.698)	0.388 (0.698)	-0.772 (0.440)	-0.304 (0.761)	0.388 (0.698)	-2.341* (0.019)	-2.131* (0.033)
321-324	2.813** (0.005)	2.813** (0.005)	-0.289 (0.773)	0.168 (0.867)	2.813** (0.005)	-0.966 (0.334)	-0.310 (0.757)
321-351	2.283* (0.022)	2.283* (0.022)	3.826** (0.0001)	0.630 (0.529)	2.283* (0.022)	1.386 (0.166)	1.554 (0.120)
321-352	2.141* (0.032)	2.141* (0.032)	2.771** (0.006)	0.446 (0.656)	2.141* (0.032)	1.622 (0.105)	1.832 (0.067)
321-385	5.144** (<0.001)	5.144** (<0.001)	5.983** (<0.001)	2.215* (0.027)	5.144** (<0.001)	1.275 (0.202)	2.981** (0.003)
322-323	1.107 (0.268)	1.107 (0.268)	0.525 (0.600)	0.877 (0.381)	1.107 (0.268)	1.538 (0.124)	2.173* (0.030)
322-324	3.181** (0.002)	3.181** (0.002)	0.924 (0.356)	1.081 (0.280)	3.181** (0.002)	2.519* (0.012)	2.462** (0.014)
322-351	2.677** (0.007)	2.677** (0.007)	4.351** (<0.001)	1.386 (0.166)	2.677** (0.007)	4.928** (<0.001)	4.745** (<0.001)
322-352	2.383* (0.017)	2.383* (0.017)	3.349** (0.001)	1.081 (0.280)	2.383* (0.017)	5.044** (<0.001)	4.970** (<0.001)
322-385	4.934** (<0.001)	4.934** (<0.001)	6.083** (<0.001)	2.467* (0.014)	4.934** (<0.001)	4.740** (<0.001)	5.543** (<0.001)
323-324	2.294* (0.022)	2.294* (0.022)	1.139 (0.255)	0.709 (0.479)	2.294* (0.022)	1.491 (0.136)	1.139 (0.255)
323-351	1.774 (0.076)	1.774 (0.076)	3.559** (0.0004)	1.391 (0.164)	1.774 (0.076)	4.225** (<0.001)	3.511** (0.0004)
323-352	1.359 (0.174)	1.359 (0.174)	2.467* (0.014)	0.919 (0.358)	1.359 (0.174)	4.314** (<0.001)	3.921** (0.0001)
323-385	4.267** (<0.001)	4.267** (<0.001)	6.876** (<0.001)	3.706** (0.0002)	4.267** (<0.001)	4.430** (<0.001)	4.792** (<0.001)
324-351	-1.197 (0.231)	-1.197 (0.231)	3.023** (0.003)	1.029 (0.304)	-1.197 (0.231)	2.414* (0.016)	1.097 (0.273)
324-352	-1.296 (0.195)	-1.296 (0.195)	1.979* (0.048)	0.504 (0.614)	-1.296 (0.195)	2.698** (0.007)	1.380 (0.168)
324-385	1.317 (0.188)	1.317 (0.188)	6.598** (<0.001)	3.338** (0.001)	1.317 (0.188)	1.806 (0.071)	2.194* (0.028)
351-352	-0.189 (0.850)	-0.189 (0.850)	-1.811 (0.070)	-0.226 (0.821)	-0.189 (0.850)	0.493 (0.622)	0.441 (0.659)

351-385	3.517** (0.0004)	3.517** (0.0004)	0.793 (0.428)	1.375 (0.170)	3.517** (0.0004)	-0.436 (0.663)	1.753 (0.080)
352-385	3.307** (0.001)	3.307** (0.001)	3.223** (0.001)	1.664 (0.096)	3.307** (0.001)	-1.485 (0.137)	1.244 (0.214)
No. of accepted cases for null hypothesis	9	9	8	17	9	10	9

Note: For each index only the standardized normal approximation of the test statistic and the corresponding p value is reported.

**Table B.6 (Corresponding to Table 3.4 in Chapter 3)
Results for Wilcoxon's Rank Sum Test for Countries**

Country Combinations	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
China-Germany	-0.115 (0.909)	-0.115 (0.909)	-0.164 (0.870)	0.033 (0.974)	-0.115 (0.909)	1.376 (0.169)	1.737 (0.082)
China-France	-0.164 (0.870)	-0.164 (0.870)	0.049 (0.961)	0.098 (0.922)	-0.164 (0.870)	1.295 (0.196)	1.524 (0.128)
India-China	0.688 (0.491)	0.688 (0.491)	-0.115 (0.909)	-0.606 (0.544)	0.688 (0.491)	-0.229 (0.819)	<0.001 (1.00)
China-Italy	-1.18 (0.238)	-1.18 (0.238)	-0.77 (0.441)	-0.754 (0.451)	-1.180 (0.238)	-0.016 (0.987)	0.557 (0.577)
China-Japan	2.392* (0.017)	2.392* (0.017)	1.180 (0.238)	1.737 (0.082)	2.392* (0.017)	2.704** (0.007)	3.097** (0.002)
Germany-France	-0.147 (0.883)	-0.147 (0.883)	0.311 (0.756)	0.213 (0.831)	-0.147 (0.883)	-0.098 (0.922)	-0.098 (0.922)
India-Germany	0.279 (0.781)	0.279 (0.781)	-0.262 (0.793)	-0.295 (0.768)	0.279 (0.781)	-1.737 (0.082)	-2.016* (0.044)
Germany-Italy	-2.114* (0.035)	-2.114* (0.035)	-1.311 (0.190)	-1.131 (0.258)	-2.114* (0.035)	-1.737 (0.082)	-1.999* (0.046)
Germany-Japan	2.016* (0.044)	2.016* (0.044)	1.671 (0.095)	1.671 (0.095)	2.016* (0.044)	1.835 (0.067)	1.852 (0.064)
India-France	0.541 (0.589)	0.541 (0.589)	-0.262 (0.793)	-1.032 (0.302)	0.541 (0.589)	-1.639 (0.101)	-1.999* (0.046)
France-Italy	-1.753 (0.080)	-1.753 (0.080)	-1.196 (0.232)	-1.180 (0.238)	-1.753 (0.080)	-1.442 (0.149)	-1.622 (0.105)
France-Japan	2.622** (0.009)	2.622** (0.009)	1.475 (0.140)	1.917 (0.055)	2.622** (0.009)	1.917 (0.055)	2.147* (0.032)
India-Italy	-1.376 (0.169)	-1.376 (0.169)	-0.541 (0.589)	-1.098 (0.272)	-1.376 (0.169)	0.049 (0.961)	0.328 (0.743)
India-Japan	1.721 (0.085)	1.721 (0.085)	1.344 (0.179)	2.376* (0.018)	1.721 (0.085)	2.72** (0.007)	3.179** (0.002)
Italy-Japan	3.490** (0.001)	3.490** (0.001)	2.130* (0.033)	2.524* (0.012)	3.490** (0.001)	2.720** (0.007)	2.819** (0.005)
No. of accepted cases for null hypothesis	10	10	14	13	10	12	8

Note: For each index only the standardized normal approximation of the test statistic and the corresponding p value is reported.

Table B.7 (Corresponding to Table 3.5 in Chapter 3)
Results for Kolmogorov Smirnov Test for Sectors

Sector Combinations	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
321-322	0.1500 (0.509)	0.1500 (0.509)	0.1500 (0.509)	0.2167 (0.120)	0.1500 (0.509)	0.3667** (0.001)	0.4500** (<0.001)
321-323	0.1167 (0.809)	0.1167 (0.809)	0.4000** (<0.001)	0.1500 (0.509)	0.1167 (0.809)	0.2333 (0.076)	0.2500* (0.047)
321-324	0.3500** (0.001)	0.3500** (0.001)	0.4167** (<0.001)	0.1833 (0.266)	0.3500** (0.001)	0.1667 (0.375)	0.2500* (0.047)
321-351	0.2500* (0.047)	0.2500* (0.047)	0.4333** (<0.001)	0.1500 (0.509)	0.2500* (0.047)	0.2000 (0.181)	0.1833 (0.266)
321-352	0.2333 (0.076)	0.2333 (0.076)	0.3500** (0.001)	0.100 (0.925)	0.2333 (0.076)	0.2500* (0.047)	0.1833 (0.266)
321-385	0.4000** (<0.001)	0.4000** (<0.001)	0.5167** (<0.001)	0.3667** (0.001)	0.4000** (<0.001)	0.2167 (0.120)	0.2500* (0.047)
322-323	0.1500 (0.509)	0.1500 (0.509)	0.3500** (0.001)	0.2500* (0.047)	0.1500 (0.509)	0.2000 (0.181)	0.3000** (0.009)
322-324	0.2833* (0.016)	0.2833* (0.016)	0.3667** (0.001)	0.2833* (0.016)	0.2833* (0.016)	0.2500* (0.047)	0.2333 (0.076)
322-351	0.3833** (<0.001)	0.3833** (<0.001)	0.4000** (<0.001)	0.2333 (0.076)	0.3833** (<0.001)	0.5000** (<0.001)	0.5667** (<0.001)
322-352	0.2667* (0.028)	0.2667* (0.028)	0.3000** (0.009)	0.2333 (0.076)	0.2667* (0.028)	0.4167** (<0.001)	0.5500** (<0.001)
322-385	0.4333** (<0.001)	0.4333** (<0.001)	0.5167** (<0.001)	0.4500** (<0.001)	0.4333** (<0.001)	0.5333** (<0.001)	0.5833** (<0.001)
323-324	0.2667* (0.028)	0.2667* (0.028)	0.2167 (0.120)	0.1667 (0.375)	0.2667* (0.028)	0.3167** (0.005)	0.3500** (0.001)
323-351	0.2833* (0.016)	0.2833* (0.016)	0.5833** (<0.001)	0.2333 (0.076)	0.2833* (0.016)	0.3667** (0.001)	0.3167** (0.005)
323-352	0.1667 (0.375)	0.1667 (0.375)	0.5500** (<0.001)	0.2333 (0.076)	0.1667 (0.375)	0.4500** (<0.001)	0.3333** (0.003)
323-385	0.4000** (<0.001)	0.4000** (<0.001)	0.7833** (<0.001)	0.4500** (<0.001)	0.4000** (<0.001)	0.3833** (<0.001)	0.3833** (<0.001)
324-351	0.2667* (0.028)	0.2667* (0.028)	0.5833** (<0.001)	0.1833 (0.266)	0.2667* (0.028)	0.2667* (0.028)	0.3667** (0.001)
324-352	0.2667* (0.028)	0.2667* (0.028)	0.5500** (<0.001)	0.1667 (0.375)	0.2667* (0.028)	0.2333 (0.076)	0.3500** (0.001)
324-385	0.2500* (0.047)	0.2500* (0.047)	0.8000** (<0.001)	0.3667** (0.001)	0.2500* (0.047)	0.3500** (0.001)	0.3833** (<0.001)
351-352	0.1833 (0.266)	0.1833 (0.266)	0.2667* (0.028)	0.0833 (0.985)	0.1833 (0.266)	0.1667 (0.375)	0.1167 (0.809)
351-385	0.3500** (0.001)	0.3500** (0.001)	0.2667* (0.028)	0.2333 (0.076)	0.3500** (0.001)	0.2167 (0.120)	0.1667 (0.375)
352-385	0.2833* (0.016)	0.2833* (0.016)	0.2833* (0.016)	0.2833* (0.016)	0.2833* (0.016)	0.3000** (0.009)	0.1333 (0.660)
No. of accepted cases for null hypothesis	6	6	2	14	6	8	4

Note: For each index the test statistics and the asymptotic p values are reported. Exact p values are also reported by Stata but they do not differ significantly from the asymptotic p values in the sense that the number of cases for acceptance of null hypothesis remains the same.

Table B.8 (Corresponding to Table 3.6 in Chapter 3)
Results for Kolmogorov Smirnov Test for Countries

Country Combinations	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
China-Germany	0.2857 (0.203)	0.2857 (0.203)	0.1786 (0.763)	0.1429 (0.938)	0.2857 (0.203)	0.2857 (0.203)	0.3929* (0.027)
China-France	0.2500 (0.346)	0.2500 (0.346)	0.1786 (0.763)	0.1429 (0.938)	0.2500 (0.346)	0.2857 (0.203)	0.3929* (0.027)
China-India	0.2500 (0.346)	0.2500 (0.346)	0.1429 (0.938)	0.2857 (0.203)	0.2500 (0.346)	0.1429 (0.938)	0.1429 (0.938)
China-Italy	0.2857 (0.203)	0.2857 (0.203)	0.2857 (0.203)	0.2857 (0.203)	0.2857 (0.203)	0.1786 (0.763)	0.1786 (0.763)
China-Japan	0.3214 (0.111)	0.3214 (0.111)	0.2857 (0.203)	0.3571 (0.056)	0.3214 (0.111)	0.3571 (0.056)	0.3929* (0.027)
Germany-France	0.2500 (0.346)	0.2500 (0.346)	0.2143 (0.541)	0.1786 (0.763)	0.2500 (0.346)	0.1786 (0.763)	0.1786 (0.763)
Germany-India	0.2857 (0.203)	0.2857 (0.203)	0.2500 (0.346)	0.2500 (0.346)	0.2857 (0.203)	0.3214 (0.111)	0.4286* (0.012)
Germany-Italy	0.3571 (0.056)	0.3571 (0.056)	0.2500 (0.346)	0.2143 (0.541)	0.3571 (0.056)	0.3929* (0.027)	0.4286* (0.012)
Germany-Japan	0.3929* (0.027)	0.3929* (0.027)	0.3214 (0.111)	0.3214 (0.111)	0.3929* (0.027)	0.3571 (0.056)	0.3571 (0.056)
France-India	0.3929* (0.027)	0.3929* (0.027)	0.2143 (0.541)	0.2500 (0.346)	0.3929* (0.027)	0.3214 (0.111)	0.4286* (0.012)
France-Italy	0.3571 (0.056)	0.3571 (0.056)	0.3214 (0.111)	0.3214 (0.111)	0.3571 (0.056)	0.3929* (0.027)	0.3929* (0.027)
France-Japan	0.4643** (0.005)	0.4643** (0.005)	0.3571 (0.056)	0.4286* (0.012)	0.4643** (0.005)	0.4286* (0.012)	0.4643** (0.005)
India-Italy	0.3929* (0.027)	0.3929* (0.027)	0.1786 (0.763)	0.2857 (0.203)	0.3929* (0.027)	0.1429 (0.938)	0.1429 (0.938)
India-Japan	0.3214 (0.111)	0.3214 (0.111)	0.2857 (0.203)	0.5714** (<0.001)	0.3214 (0.111)	0.3571 (0.056)	0.4286* (0.012)
Italy-Japan	0.4643** (0.005)	0.4643** (0.005)	0.3929* (0.027)	0.4643** (0.005)	0.4643** (0.005)	0.3571 (0.056)	0.3571 (0.056)
No. of accepted cases for null hypothesis	10	10	14	12	10	12	6

Note: For each index the test statistics and the corresponding asymptotic p values are reported.

Appendix C

Table C.1
Summary Statistics for Heckscher-Ohlin Sample used in Chapter 4

Variable	No. of Observations	Mean	Median	Standard Deviation	Skewness	Kurtosis
Export for 1998	420	2168740	307941.2	5151880	4.662197	31.28881
Export for 1999	420	2222453	295654.6	5272328	4.610586	30.33004
Export for 2000	420	2403529	324031.4	5806916	4.775795	32.42496
Import for 1998	420	2188623	455617.4	4681527	4.443676	28.68015
Import for 1999	420	2273205	437147.2	4945181	4.598262	30.63178
Import for 2000	420	2461783	485283.7	5559681	4.999566	36.13334
Employees for 1998	60	34900000	8523696	103000000	5.465145	34.34748
Employees for 1999	60	35500000	8617295	105000000	5.44125	34.08465
Employees for 2000	60	35900000	8850667	105000000	5.41038	33.72699
Capital stock (2005 PPP) for 1998	60	1100000000000	328000000000	2340000000000	3.922569	19.56326
Capital stock (2005 PPP) for 1999	60	1140000000000	335000000000	2460000000000	3.969915	20.06154
Capital stock (2005 PPP) for 2000	60	1190000000000	345000000000	2600000000000	4.024319	20.61705
GDP PPP (2005 constant prices) for 1998	60	639000000000	197000000000	1450000000000	5.134103	33.13573
GDP PPP (2005 constant prices) for 1999	60	662000000000	204000000000	1520000000000	5.184738	33.68062
GDP PPP (2005 constant prices) for 2000	60	692000000000	212000000000	1580000000000	5.174228	33.57956
GDP PPP (current	60	544000000000	163000000000	1240000000000	5.149003	33.26447

prices) for 1998						
GDP PPP (current prices) for 1999	60	572000000000	172000000000	1320000000000	5.203832	33.84646
GDP PPP (current prices) for 2000	60	613000000000	183000000000	1400000000000	5.185798	33.6696
PPP Conversion Factor to market exchange rate ratio for 1998	60	0.6764681	0.5738126	0.293964	0.4239061	2.058904
PPP Conversion Factor to market exchange rate ratio for 1999	60	0.6553715	0.5743979	0.2874585	0.5896894	2.460456
PPP Conversion Factor to market exchange rate ratio for 2000	60	0.6164193	0.549418	0.2528423	0.7298591	3.247708

Note: Exports and Import values are reported in thousand US dollars in Trade Production and Protection Database. PPP denotes Purchasing Power Parity. Countries included for each of the 7 sectors – Argentina, Australia, Austria, Bangladesh, Bulgaria, Brazil, Canada, Switzerland, Chile, China, Colombia, Costa Rica, Cyprus, Germany, Denmark, Egypt, Spain, Finland, France, United Kingdom, Greece, Guatemala, Hong Kong, Honduras, Hungary, Indonesia, India, Ireland, Israel, Italy, Jordan, Japan, Korea, Sri Lanka, Macau, Morocco, Mexico, Malta, Mongolia, Mauritius, Malaysia, Netherlands, Nepal, New Zealand, Pakistan, Panama, Peru, Philippines, Poland, Portugal, Romania, Singapore, El Salvador, Sweden, Thailand, Tunisia, Turkey, Uruguay, United States, South Africa.

Table C.2
Summary Statistics for Ricardian Sample used in Chapter 4

Variable	No. of Observations	Mean	Median	Standard Deviation	Skewness	Kurtosis
Export for 1998	1142	3043880	184530	10700000	7.405417	69.24347
Export for 1999	1142	3125745	187094.4	11000000	7.340852	67.35284
Export for 2000	1142	3406511	197424.4	12100000	7.415469	69.29656
Import for 1998	1142	2978923	359349.5	9790256	9.3959	123.5134
Import for 1999	1142	3130394	327357.5	10800000	10.00889	138.8296
Import for 2000	1142	3440400	360550.1	12200000	10.32882	145.7341
No. of Employees for 1998	1142	74216.33	11262.5	194863.7	5.608444	41.56837
No. of Employees for 1999	1142	72510.65	11153.5	185075.5	5.372644	38.46648
No. of Employees for 2000	1142	74434.73	11150	189137.7	5.263093	36.89808
Value added for 1998	1142	3990588	277403.2	16500000	8.674874	93.29306
Value added for 1999	1142	4180768	263132.8	17400000	8.737583	95.54843
Value added for 2000	1142	4257222	268640.5	17700000	8.495612	88.11419

Note: Export, Imports and Value added are reported in thousand US dollars in Trade Production and Protection Database.

Table C.2.1
Countries included for Each Sector in Table C.2

ISIC Code	Countries
311	Australia, Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, United Kingdom, Hong Kong, Hungary, Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Latvia, Macau, Morocco, Mexico, Malta, Netherlands, Norway, Oman, Panama, Portugal, Russian Federation, Senegal, Singapore, Sweden, Trinidad And Tobago, Tunisia, Turkey, Uruguay, United States
313	Australia, Bulgaria, Bolivia, Canada, Chile, Colombia, Cyprus, Germany, Spain, Ethiopia, Finland, Hungary, Indonesia, India, Ireland, Iran, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Latvia, Macau, Morocco, Mexico, Malta, Netherlands, Norway, Oman, Panama, Portugal, Russian Federation, Singapore, Sweden, Trinidad And Tobago, Tunisia, Turkey, Uruguay, United States
314	Australia, Bulgaria, Bolivia, Canada, Chile, Colombia, Cyprus, Germany, Ethiopia, Finland, Hungary, Indonesia, India, Ireland, Iran, Italy, Jordan, Japan, Korea, Sri Lanka, Macau, Morocco, Mexico, Malta, Myanmar, Netherlands, Poland, Portugal, Romania, Russian Federation, Senegal, Sweden, Trinidad And Tobago, Tunisia, Turkey, Uruguay, United States, Yemen
321	Australia, Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hong Kong, Hungary, Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Latvia, Macau, Morocco, Mexico, Malta, Netherlands, Norway, Oman, Panama, Poland, Portugal, Romania, Russian Federation, Senegal, Singapore, Sweden, Trinidad And Tobago, Tunisia, Uruguay, United States, Yemen
322	Australia, Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hong Kong, Hungary, Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Latvia, Macau, Morocco, Mexico, Netherlands, Norway, Oman, Panama, Portugal, Russian Federation, Singapore, Sweden, Trinidad And Tobago, Tunisia, Turkey, Uruguay, United States
323	Australia, Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hong Kong, Hungary, Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Macau, Morocco, Mexico, Myanmar, Netherlands, Norway, Panama, Portugal, Russian Federation, Singapore, Sweden, Turkey, Uruguay, United States

324	Australia, Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hungary, Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Macau, Morocco, Mexico, Malta, Myanmar, Netherlands, Norway, Oman, Panama, Portugal, Russian Federation, Singapore, Sweden, Trinidad And Tobago, Turkey, Uruguay, United States
331	Australia, Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hong Kong, Hungary, Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Latvia, Macau, Morocco, Mexico, Malta, Netherlands, Norway, Oman, Panama, Poland, Portugal, Romania, Russian Federation, Senegal, Singapore, Slovenia, Sweden, Trinidad And Tobago, Turkey, United States, Yemen
332	Australia, Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hong Kong, Hungary, Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Latvia, Morocco, Mexico, Malta, Netherlands, Norway, Oman, Portugal, Russian Federation, Singapore, Sweden, Turkey, United States
341	Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hong Kong, Hungary, Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Macau, Morocco, Mexico, Malta, Netherlands, Norway, Oman, Panama, Poland, Portugal, Romania, Russian Federation, Senegal, Singapore, Sweden, Trinidad And Tobago, Turkey, Uruguay, United States, Yemen
342	Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hong Kong, Hungary, Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Latvia, Macau, Morocco, Mexico, Malta, Netherlands, Norway, Oman, Panama, Poland, Portugal, Russian Federation, Senegal, Singapore, Sweden, Trinidad And Tobago, Turkey, Uruguay, Yemen
351	Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, Hungary, Indonesia, India, Ireland, Iran, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Morocco, Mexico, Netherlands, Oman, Panama, Portugal, Russian Federation, Senegal, Singapore, Sweden, Trinidad And Tobago, Tunisia, Turkey, Uruguay, United States
352	Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hungary, Indonesia, India, Ireland,

	Iran, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Morocco, Mexico, Netherlands, Norway, Oman, Panama, Portugal, Singapore, Sweden, Trinidad And Tobago, Tunisia, Turkey, Uruguay, United States
355	Australia, Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hong Kong, Hungary, Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Macau, Morocco, Mexico, Malta, Myanmar, Netherlands, Norway, Oman, Panama, Portugal, Russian Federation, Singapore, Sweden, Trinidad And Tobago, Tunisia, Turkey, Uruguay, United States
356	Australia, Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hong Kong, Hungary, Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Latvia, Macau, Morocco, Mexico, Malta, Myanmar, Netherlands, Norway, Oman, Panama, Portugal, Russian Federation, Singapore, Sweden, Trinidad And Tobago, Tunisia, Turkey, Uruguay, United States
361	Australia, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Finland, France, United Kingdom, Hungary, India, Ireland, Iran, Italy, Jordan, Japan, Kenya, Korea, Sri Lanka, Morocco, Mexico, Netherlands, Oman, Portugal, Sweden, Trinidad And Tobago, Tunisia, Turkey, United States
362	Australia, Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hong Kong, Hungary, India, Ireland, Iran, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Macau, Morocco, Mexico, Malta, Netherlands, Norway, Oman, Panama, Portugal, Russian Federation, Sweden, Trinidad And Tobago, Tunisia, Turkey, United States
369	Australia, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hong Kong, Hungary, India, Ireland, Iran, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Latvia, Morocco, Mexico, Netherlands, Norway, Oman, Panama, Portugal, Senegal, Singapore, Sweden, Trinidad And Tobago, Tunisia, Turkey, United States
371	Bolivia, Canada, Chile, Colombia, Costa Rica, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hong Kong, Hungary, Indonesia, India, Ireland, Iran, Italy, Jordan, Japan, Korea, Kuwait, Sri Lanka, Morocco, Mexico, Oman, Panama, Portugal, Russian Federation, Sweden, Trinidad And Tobago, Tunisia, Turkey, United States
381	Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hong Kong, Hungary,

	Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Latvia, Macau, Morocco, Mexico, Myanmar, Netherlands, Norway, Oman, Panama, Poland, Portugal, Romania, Russian Federation, Singapore, Sweden, Trinidad And Tobago, Turkey, Uruguay, United States, Yemen
382	Australia, Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hong Kong, Hungary, Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Macau, Morocco, Mexico, Malta, Netherlands, Oman, Poland, Portugal, Romania, Russian Federation, Senegal, Singapore, Sweden, Tunisia, Turkey, Uruguay, United States, Yemen
383	Australia, Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hong Kong, Hungary, Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Macau, Morocco, Mexico, Malta, Netherlands, Norway, Oman, Poland, Portugal, Russian Federation, Singapore, Slovenia, Sweden, Trinidad And Tobago, Tunisia, Turkey, Uruguay, United States, Yemen
384	Australia, Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Ethiopia, Finland, France, United Kingdom, Hong Kong, Hungary, Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Latvia, Morocco, Mexico, Malta, Netherlands, Norway, Oman, Poland, Portugal, Romania, Russian Federation, Singapore, Slovenia, Sweden, Tunisia, Turkey, Uruguay, United States, Yemen
385	Australia, Bulgaria, Bolivia, Canada, Chile, Costa Rica, Cyprus, Germany, Spain, Finland, France, United Kingdom, Hong Kong, Hungary, Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Sri Lanka, Morocco, Mexico, Malta, Netherlands, Norway, Oman, Poland, Portugal, Romania, Russian Federation, Singapore, Sweden, Turkey, Uruguay, United States
390	Australia, Bulgaria, Bolivia, Canada, Chile, Colombia, Costa Rica, Cyprus, Germany, Spain, Finland, France, United Kingdom, Hong Kong, Hungary, Indonesia, India, Ireland, Iran, Israel, Italy, Jordan, Japan, Kenya, Korea, Kuwait, Sri Lanka, Latvia, Morocco, Mexico, Malta, Myanmar, Netherlands, Norway, Oman, Portugal, Russian Federation, Singapore, Sweden, Trinidad And Tobago, Tunisia, Turkey, United States

Table C.3
Summary Statistics for Heckscher-Ohlin Sample with Tariff Adjusted Imports

Variable	No. of Observations	Mean	Median	Standard Deviation	Skewness	Kurtosis
Export for 1998	199	3623847	808521.2	6899198	3.41465	17.36041
Export for 1999	199	3693422	809714	7066720	3.370729	16.7465
Export for 2000	199	3969542	870469.5	7792905	3.516048	18.00841
Import for 1998	199	3794723	1510382	6198412	3.242147	15.94782
Import for 1999	199	3940904	1485546	6572907	3.359139	16.95667
Import for 2000	199	4227170	1542169	7431683	3.690812	20.09703
Import weighted average applied ad-valorem Tariff for 1998	199	8.13809	6.16	7.736911	1.652274	6.020483
Import weighted average applied ad-valorem Tariff for 1999	199	7.214322	7.19	6.026526	1.333414	5.42046
Import weighted average applied ad-valorem Tariff for 2000	199	6.489598	4.99	6.392812	1.713183	6.993911
Import demand elasticity	199	-1.033402	-1.00511	.2965989	-2.566142	18.15992
Employees for 1998	29	46000000	10600000	132000000	4.725063	24.40985
Employees for 1999	29	46500000	10900000	133000000	4.722777	24.39314
Employees for 2000	29	46900000	11100000	134000000	4.713938	24.3277
Capital stock (2005 PPP) for 1998	29	1850000000000	501000000000	3170000000000	2.65537	9.622941
Capital stock (2005 PPP) for 1999	29	1930000000000	535000000000	3330000000000	2.69285	9.893518
Capital	29	2020000000000	564000000000	3510000000000	2.735093	10.18762

stock (2005 PPP) for 2000						
GDP PPP (2005 constant prices) for 1998	29	1100000000000	396000000000	1970000000000	3.621434	16.9371
GDP PPP (2005 constant prices) for 1999	29	1140000000000	383000000000	2060000000000	3.666099	17.24923
GDP PPP (2005 constant prices) for 2000	29	1190000000000	380000000000	2150000000000	3.658464	17.19887
GDP PPP (current prices) for 1998	29	938000000000	339000000000	1690000000000	3.630717	16.98552
GDP PPP (current prices) for 1999	29	984000000000	332000000000	1790000000000	3.677939	17.31174
GDP PPP (current prices) for 2000	29	1190000000000	380000000000	2150000000000	3.658464	17.19887
PPP conversion factor to market exchange rate ratio for 1998	29	0.8559048	0.8817437	0.2760101	-0.3226586	2.223773
PPP conversion factor to market exchange rate ratio for 1999	29	0.8270371	0.8530416	0.2829532	-0.0781378	2.338533
PPP conversion factor to market exchange rate ratio for 2000	29	0.7603436	0.8223462	0.2515805	0.267823	3.276471

Note: Exports and Import values are reported in thousand US dollars in the Trade Production and Protection database. PPP denotes Purchasing Power Parity. Tariff data are in percentage terms. 29 countries included in each of the 6 sectors (excluding sector 324) – Argentina, Australia, Austria, Brazil, Canada, Switzerland, Chile, China, Germany, Denmark, Spain, Finland, France, United Kingdom, Greece, Ireland, Italy, Japan, Mexico, Netherlands, Nepal, New Zealand, Peru,

Philippines, Poland, Portugal, Sweden, Uruguay, United States. For sector 324 all of the above excepting Brazil, China, Mexico and Philippines are included.

Table C.4
Summary Statistics for Ricardian Sample with Tariff Adjusted Imports

Variable	No. of Observations	Mean	Median	Standard Deviation	Skewness	Kurtosis
Export for 1998	465	6242535	1261348	15700000	5.009582	31.73477
Export for 1999	465	6376739	1311753	16200000	4.979868	31.00913
Export for 2000	465	6808911	1364154	17600000	5.128245	32.99417
Import for 1998	465	6032043	1637918	14400000	6.527319	58.73833
Import for 1999	465	6390674	1628324	16000000	6.942288	65.58242
Import for 2000	465	6906351	1667724	17900000	7.288817	70.62663
Import weighted average applied ad-valorem Tariff for 1998	465	6.160796	3.49	12.87316	14.37651	263.2784
Import weighted average applied ad-valorem Tariff for 1999	465	5.98486	3.58	11.94135	13.91632	251.6849
Import weighted average applied ad-valorem Tariff for 2000	465	5.218022	2.5	11.88981	13.63716	242.6396
Import demand elasticity	465	-1.053571	-1.006865	.3064955	-2.32415	13.3713
Value added for 1998	465	8880433	1597362	25000000	5.569732	39.25388
Value added for 1999	465	9206471	1680745	26200000	5.641764	40.53344
Value added for 2000	465	9286139	1598413	26800000	5.486393	37.31455
No. of employees for 1998	465	115000.6	34558	243493.2	4.321177	24.89839
No. of	465	113542.3	34478	239406.1	4.312748	24.85255

employees for 1999						
No. of employees for 2000	465	116776.7	34851	244334.4	4.208196	23.69221

Note: Exports, Imports and Value added are reported in thousand US dollars in the Trade Production and Protection Database. PPP denotes Purchasing Power Parity. Tariff data are in percentage terms.

Table C.4.1
Countries included for Each Sector in Table C.4

ISIC code	Countries
311	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Portugal, Sweden, Uruguay, United States
313	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, Ireland, Italy, Japan, Mexico, Netherlands, Portugal, Sweden, Uruguay, United States
314	Australia, Bolivia, Canada, Chile, Germany, Finland, Ireland, Italy, Japan, Mexico, Netherlands, Poland, Portugal, Sweden, Uruguay, United States
321	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Poland, Portugal, Sweden, Uruguay, United States
322	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Portugal, Sweden, Uruguay, United States
323	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Portugal, Sweden, Uruguay, United States
324	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Netherlands, Portugal, Sweden, Uruguay, United States
331	Australia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Poland, Portugal, Sweden, United States
332	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Portugal, Sweden, United States
341	Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Poland, Portugal, Sweden, Uruguay, United States
342	Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Poland, Portugal, Sweden, Uruguay

351	Bolivia, Canada, Chile, Germany, Spain, Finland, France, Ireland, Italy, Japan, Mexico, Netherlands, Portugal, Sweden, Uruguay, United States
352	Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Portugal, Sweden, Uruguay, United States
353	Australia, Bolivia, Canada, Chile, France, Japan, Mexico, Netherlands, Portugal, Sweden, Uruguay
354	Chile, Japan, Mexico, Netherlands, Sweden
355	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Portugal, Sweden, Uruguay, United States
356	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Portugal, Sweden, Uruguay, United States
361	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Portugal, Sweden, United States
362	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Portugal, Sweden, United States
369	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Mexico, Netherlands, Portugal, Sweden, United States
371	Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Portugal, Sweden, United States
372	Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Portugal, Sweden, United States
381	Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Poland, Portugal, Sweden, Uruguay, United States
382	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Poland, Portugal, Sweden, Uruguay, United States
383	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Poland, Portugal, Sweden, Uruguay, United States
384	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Poland, Portugal, Sweden, Uruguay, United States
385	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Poland, Portugal, Sweden, Uruguay, United States

390	Australia, Bolivia, Canada, Chile, Germany, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Mexico, Netherlands, Portugal, Sweden, United States

Table C.5
Results for Breusch-Pagan/ Cook Weisberg Test for Heteroskedasticity from Regressions for
Heckscher-Ohlin Model

ISIC Codes	Regression Forms	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
321	Linear	61.52** (<0.001)	0.00 (0.966)	61.83** (<0.001)	0.63 (0.427)	0.16 (0.689)	43.06** (<0.001)	8.23** (0.004)
	Semi Log	49.53** (<0.001)	0.62 (0.432)	49.68** (<0.001)	9.57** (0.002)	3.09 (0.079)	32.90** (<0.001)	16.47** (<0.001)
322	Linear	4.90* (0.027)	0.79 (0.375)	4.90* (0.027)	0.12 (0.725)	0.88 (0.347)	5.62* (0.018)	3.89* (0.049)
	Semi Log	5.17* (0.023)	2.35 (0.125)	5.15* (0.023)	4.74* (0.030)	1.81 (0.178)	6.08* (0.014)	0.20 (0.657)
323	Linear	0.20 (0.654)	0.64 (0.422)	0.20 (0.651)	5.34* (0.021)	1.37 (0.241)	0.01 (0.911)	0.27 (0.606)
	Semi Log	0.78 (0.376)	0.51 (0.475)	0.78 (0.377)	29.65** (<0.001)	1.95 (0.163)	3.11 (0.078)	0.02 (0.880)
324	Linear	0.34 (0.561)	0.32 (0.574)	0.33 (0.563)	1.99 (0.159)	0.07 (0.787)	0.04 (0.841)	10.61** (0.001)
	Semi Log	0.09 (0.767)	0.35 (0.557)	0.09 (0.769)	11.15** (0.001)	0.02 (0.886)	0.39 (0.534)	8.60** (0.003)
351	Linear	1.64 (0.200)	0.01 (0.923)	1.65 (0.200)	6.44* (0.011)	0.02 (0.886)	0.08 (0.772)	0.34 (0.561)
	Semi Log	0.22 (0.639)	0.60 (0.438)	0.22 (0.640)	20.76** (<0.001)	0.38 (0.538)	0.51 (0.475)	1.12 (0.290)
352	Linear	0.01 (0.943)	5.16* (0.023)	0.00 (0.948)	7.54** (0.006)	28.10** (<0.001)	0.07 (0.793)	36.36** (<0.001)
	Semi Log	0.47 (0.494)	3.93* (0.048)	0.47 (0.495)	31.07** (<0.001)	15.13** (0.0001)	0.49 (0.486)	21.54** (<0.001)
385	Linear	4.69* (0.0304)	1.12 (0.2902)	4.69* (0.0304)	9.17** (0.0025)	3.21 (0.0730)	4.85* (0.0276)	6.03* (0.0140)
	Semi Log	22.37** (<0.001)	0.09 (0.767)	22.33** (<0.001)	38.01** (<0.001)	16.02** (0.0001)	19.06** (<0.001)	17.76** (<0.001)

Note: The reported values in the table are the chi-square test statistics and the corresponding p values. The default version of the test performed by Stata and presented in the above table, makes the assumption that the disturbances are normally distributed with a variance of σ^2 (Baum, 2006). There is another version of the test derived by Keonker (1981) which only assumes that the disturbance are independently and identically distributed (i.i.d) and as such drops the normality assumption. This version of the test may be considered more suitable for analysis given the fact that many of the indices are characterized by non-normal distributions. For that reason the test results presented in the above table has been cross checked with this second version of the test. Although some changes in p values were observed, consideration of such changes do alter the inferences drawn from the regression results, particularly in terms of the number of significant coefficients with expected signs. Therefore the results of the test derived by Koenker are not presented here.

Table C.6 (Corresponding to Table 4.1 in Chapter 4)
Regression Results for Heckscher-Ohlin Theory: With/without White's Correction

ISIC Codes	Regression Form		Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
321	Linear	$\hat{\beta}$	0.4546** (<0.001)	0.0478** (<0.001)	0.0133** (<0.001)	0.00003 (0.0605)	0.1293** (<0.001)	0.3117* (0.048)	0.0713* (0.024)
		R ²	0.4627	0.2669	0.4647	0.0409	0.2932	0.1540	0.0941
	Semi log	$\hat{\beta}$	1.2072** (0.002)	0.1511** (<0.001)	0.0354** (0.002)	0.0001* (0.026)	0.3817** (<0.001)	0.7272 (0.075)	0.1635 (0.078)
		R ²	0.2731	0.2234	0.2737	0.0900	0.2139	0.0701	0.0414
322	Linear	$\hat{\beta}$	0.6270** (0.004)	0.0604** (<0.001)	0.0178** (0.004)	0.00004* (0.039)	0.1690** (<0.001)	0.6671** (0.002)	0.2759** (<0.001)
		R ²	0.2529	0.2647	0.2523	0.0525	0.2789	0.2906	0.3804
	Semi log	$\hat{\beta}$	1.8395** (0.002)	0.2577** (<0.001)	0.0521** (0.002)	0.0002** (0.009)	0.6829** (<0.001)	2.0475** (0.001)	1.1977** (<0.001)
		R ²	0.1822	0.4031	0.1818	0.1432	0.3812	0.2291	0.5997
323	Linear	$\hat{\beta}$	0.1267* (0.017)	0.0387** (0.003)	0.0010* (0.017)	0.00001 (0.132)	0.0994** (0.002)	0.1579** (0.002)	0.1679** (<0.001)
		R ²	0.0756	0.1303	0.0760	0.0104	0.1341	0.1388	0.3723
	Semi log	$\hat{\beta}$	0.5355** (0.005)	0.1752** (<0.001)	0.0043** (0.005)	0.00003 (0.161)	0.4572** (<0.001)	0.5855** (0.001)	0.6334** (<0.001)
		R ²	0.1130	0.2241	0.1133	0.0273	0.2373	0.1598	0.4436
324	Linear	$\hat{\beta}$	0.0185 (0.373)	0.0126 (0.213)	0.0001 (0.371)	0.00001 (0.209)	0.0315 (0.239)	0.0621 (0.152)	0.1726* (0.043)
		R ²	0.0018	0.0110	0.0019	0.0114	0.0087	0.0182	0.1190
	Semi log	$\hat{\beta}$	0.2276 (0.123)	0.0955* (0.039)	0.0015 (0.122)	0.00004 (0.094)	0.2593* (0.044)	0.4545* (0.014)	0.8906** (<0.001)
		R ²	0.0232	0.0525	0.0234	0.0424	0.0494	0.0816	0.2651
351	Linear	$\hat{\beta}$	-0.0379* (0.025)	-0.027** (0.003)	-0.0026* (0.025)	-0.00001 (0.180)	-0.071** (0.005)	-0.060** (0.003)	-0.087** (<0.001)
		R ²	0.0649	0.1234	0.0649	0.0056	0.1106	0.1298	0.2113
	Semi log	$\hat{\beta}$	-0.1080 (0.054)	-0.0805* (0.010)	-0.0074 (0.054)	-0.00007 (0.121)	-0.2061* (0.016)	-0.259** (<0.001)	-0.320** (<0.001)
		R ²	0.0442	0.0904	0.0442	0.0263	0.0773	0.1994	0.2417
352	Linear	$\hat{\beta}$	-0.0143 (0.323)	-0.0165 (0.182)	-0.0006 (0.322)	-0.00001 (0.203)	-0.0764 (0.117)	-0.0023 (0.465)	-0.0667 (0.126)
		R ²	0.0037	0.0321	0.0037	0.0047	0.0927	0.0001	0.0928
	Semi log	$\hat{\beta}$	-0.0797 (0.229)	-0.0820* (0.035)	-0.0032 (0.229)	-0.0001 (0.166)	-0.2894* (0.021)	-0.0903 (0.154)	-0.2913* (0.012)
		R ²	0.0096	0.0669	0.0096	0.0229	0.1113	0.0180	0.1480
385	Linear	$\hat{\beta}$	-0.043** (0.001)	-0.032** (<0.001)	-0.002** (0.001)	-0.00001 (0.178)	-0.188** (<0.001)	-0.0103 (0.145)	-0.1359* (0.013)
		R ²	0.1144	0.1961	0.1145	0.0061	0.2619	0.0097	0.1801
	Semi log	$\hat{\beta}$	-0.256** (<0.001)	-0.189** (<0.001)	-0.010** (<0.001)	-0.0001 (0.068)	-0.713** (<0.001)	-0.1108* (0.028)	-0.514** (0.001)
		R ²	0.3465	0.4571	0.3466	0.0568	0.3148	0.0945	0.2153
Linear	No. of significant β coefficients		5	5	5	1	5	4	6
	Average R ²		0.1394	0.1464	0.1397	0.0188	0.1686	0.1059	0.2071
Semi log	No. of significant β coefficients		4	7	4	2	7	5	6
	Average R ²		0.1417	0.2168	0.1418	0.0584	0.1979	0.1218	0.2793

Note: The results are based on a sample size of 60 for each sector. Reported numbers of significant coefficients for each index are out of a total 7 cases.

Table C.7 (Corresponding to Table 4.2 in Chapter 4)
Robust Regression Results for Heckscher-Ohlin Theory

ISIC Codes	Regression Forms		Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
321	Linear	$\hat{\beta}$	0.2749** (<0.001)	0.0482** (<0.001)	0.0080** (<0.001)	0.000003 (0.078)	0.1267** (<0.001)	0.2872** (<0.001)	0.0603* (0.016)
		R ²	0.0778	0.2505	0.0778	0.0124	0.2508	-0.0057	0.0423
	Semi log	$\hat{\beta}$	0.3693** (0.001)	0.1642** (<0.001)	0.0108** (0.001)	0.00003** (0.002)	0.4099** (<0.001)	-0.0222 (0.622)	0.1081 (0.134)
		R ²	0.0804	0.2237	0.0802	0.0379	0.2180	0.0002	0.0132
322	Linear	$\hat{\beta}$	0.6450** (<0.001)	0.0612** (<0.001)	0.0183** (<0.001)	0.00003** (0.002)	0.1753** (<0.001)	0.5813** (<0.001)	0.4421** (<0.001)
		R ²	0.2464	0.2571	0.2489	0.0634	0.2695	0.3026	0.3804
	Semi log	$\hat{\beta}$	0.6793** (<0.001)	0.2998** (<0.001)	0.0193** (<0.001)	0.0001** (<0.001)	0.7368** (<0.001)	0.9291** (<0.001)	1.1369** (<0.001)
		R ²	0.1676	0.4343	0.1687	0.1489	0.3780	0.2417	0.4793
323	Linear	$\hat{\beta}$	0.1706** (<0.001)	0.0389** (0.004)	0.0014** (<0.001)	0.000002 (0.088)	0.0972** (0.004)	0.0950** (<0.001)	0.1619** (<0.001)
		R ²	0.1379	0.1267	0.1382	0.0128	0.1260	0.0941	0.3049
	Semi log	$\hat{\beta}$	0.5524** (<0.001)	0.1874** (<0.001)	0.0044** (<0.001)	0.00002** (<0.001)	0.4702** (<0.001)	0.4008** (<0.001)	0.6210** (<0.001)
		R ²	0.1564	0.2242	0.1566	0.0743	0.2222	0.1100	0.3801
324	Linear	$\hat{\beta}$	0.0377* (0.026)	0.0386* (0.038)	0.0002* (0.027)	0.000002 (0.085)	0.0359 (0.222)	0.0845** (0.002)	0.3423** (<0.001)
		R ²	0.0172	0.0598	0.0171	0.0173	0.0102	0.0536	0.2302
	Semi log	$\hat{\beta}$	0.1749** (0.009)	0.1086* (0.032)	0.0011** (0.008)	0.00002** (<0.001)	0.2901* (0.036)	0.3591** (<0.001)	0.9779** (<0.001)
		R ²	0.0330	0.0583	0.0329	0.0806	0.0513	0.1078	0.2344
351	Linear	$\hat{\beta}$	-0.0288* (0.019)	-0.028** (0.004)	-0.0020* (0.019)	0.000002 (0.696)	-0.070** (0.001)	-0.041** (0.005)	-0.0788** (<0.001)
		R ²	0.0559	0.1172	0.0560	0.0022	0.1058	0.0893	0.1780
	Semi log	$\hat{\beta}$	-0.1050* (0.013)	-0.087** (0.008)	-0.0072* (0.013)	0.00003 (0.973)	-0.218** (0.004)	-0.224** (<0.001)	-0.3181** (<0.001)
		R ²	0.0563	0.0886	0.0563	0.0121	0.0756	0.2021	0.2179
352	Linear	$\hat{\beta}$	-0.0408* (0.013)	-0.042** (0.005)	-0.0016* (0.013)	0.000001 (0.696)	-0.166** (<0.001)	0.0047 (0.589)	-0.1624** (<0.001)
		R ²	0.0218	0.1134	0.0217	0.0014	0.1448	0.0006	0.1491
	Semi log	$\hat{\beta}$	-0.147** (0.009)	-0.0897* (0.021)	-0.006** (0.009)	0.000004 (0.686)	-0.2414* (0.015)	-0.1183* (0.041)	-0.2548** (0.004)
		R ²	0.0456	0.0663	0.0457	0.0007	0.0539	0.0318	0.0786
385	Linear	$\hat{\beta}$	-0.032** (0.003)	-0.035** (0.001)	-0.001** (0.003)	0.000001 (0.632)	-0.120** (0.001)	-0.0029 (0.383)	-0.0697* (0.019)
		R ²	0.1154	0.1840	0.1151	0.0010	0.1438	0.0015	0.0650
	Semi log	$\hat{\beta}$	-0.193** (<0.001)	-0.194** (<0.001)	-0.007** (<0.001)	-0.000004 (0.375)	-0.531** (<0.001)	-0.0575* (0.0415)	-0.3507** (<0.001)
		R ²	0.3000	0.4290	0.2994	0.0007	0.3117	0.0424	0.1783
Linear	No. of significant coefficients	β	7	7	7	1	6	5	7
	Average R ²		0.0961	0.1584	0.0964	0.0158	0.1501	0.0765	0.1928
Semi log	No. of significant coefficients	β	7	7	7	4	7	6	6
	Average R ²		0.1199	0.2178	0.1200	0.0507	0.1872	0.1052	0.2260

Note: The results are based on a sample size of 60 for each sector. Reported numbers of significant coefficients for each index are out of a total 7 cases.

Table C.8
Results for Breusch-Pagan/ Cook Weisberg Test for Heteroskedasticity from Regressions for
Ricardian Model

ISIC Codes	Regression Forms	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
311	Linear	1.31 (0.2521)	1.37 (0.2411)	1.29 (0.2553)	0.24 (0.6255)	1.45 (0.2293)	0.68 (0.4083)	0.64 (0.4244)
	Semi Log	1.70 (0.1922)	1.14 (0.2863)	1.69 (0.1942)	0.05 (0.8220)	1.26 (0.2608)	1.12 (0.2901)	0.46 (0.4975)
313	Linear	0.20 (0.6584)	0.01 (0.9045)	0.19 (0.6614)	0.05 (0.8244)	0.03 (0.8616)	0.44 (0.5089)	0.00 (0.9679)
	Semi Log	2.82 (0.0932)	0.94 (0.3312)	2.80 (0.0944)	0.33 (0.5650)	1.32 (0.2511)	3.45 (0.0632)	0.97 (0.3235)
314	Linear	24.09** (<0.0001)	0.51 (0.4757)	24.12** (<0.0001)	7.23** (0.0072)	2.90 (0.0887)	2.24 (0.1343)	<0.01 (0.9492)
	Semi Log	27.31** (<0.0001)	1.64 (0.2000)	27.34** (0.0000)	13.97** (0.0002)	4.16* (0.0413)	6.64** (0.0099)	<0.01 (0.9575)
321	Linear	7.83** (0.0051)	0.33 (0.5636)	7.83** (0.0051)	1.33 (0.2488)	0.19 (0.6647)	3.51 (0.0610)	0.67 (0.4137)
	Semi Log	7.45** (0.0063)	0.49 (0.4850)	7.43** (0.0064)	2.39 (0.1224)	0.59 (0.4410)	2.15 (0.1422)	0.74 (0.3893)
322	Linear	23.10** (<0.0001)	0.14 (0.7064)	23.12** (<0.0001)	5.67 (0.0173)	0.51 (0.4732)	21.06** (<0.0001)	0.59 (0.4434)
	Semi Log	21.66** (<0.0001)	0.05 (0.8239)	21.67** (<0.0001)	6.06* (0.0138)	0.01 (0.9110)	20.18** (<0.0001)	0.03 (0.8599)
323	Linear	27.67** (<0.0001)	2.05 (0.1519)	27.66** (<0.0001)	2.25 (0.1340)	1.97 (0.1605)	28.64** (<0.0001)	0.02 (0.9025)
	Semi Log	16.53** (<0.0001)	2.03 (0.1545)	16.52** (<0.0001)	3.90 (0.0484)	3.10 (0.0783)	17.49** (<0.0001)	0.02 (0.8747)
324	Linear	29.64** (<0.0001)	0.76 (0.3839)	28.35** (<0.0001)	9.31** (0.0023)	3.24 (0.0720)	27.25** (<0.0001)	2.18 (0.1400)
	Semi Log	63.93** (<0.0001)	2.06 (0.1512)	61.06** (<0.0001)	9.90** (0.0017)	7.73** (0.0054)	58.36** (<0.0001)	1.85 (0.1742)
331	Linear	36.75** (<0.0001)	0.05 (0.8191)	36.62** (<0.0001)	8.74** (0.0031)	0.18 (0.6696)	36.28** (<0.0001)	0.24 (0.6225)
	Semi Log	27.28** (<0.0001)	0.03 (0.8548)	27.20** (<0.0001)	9.68** (0.0019)	<0.01 (0.9902)	27.47** (<0.0001)	0.03 (0.8585)
332	Linear	12.17** (0.0005)	0.77 (0.3790)	12.11** (0.0005)	15.77** (0.0001)	0.66 (0.4170)	7.05** (0.0079)	4.48* (0.0343)
	Semi Log	8.37** (0.0038)	0.54 (0.4628)	8.33** (0.0039)	12.03** (0.0005)	0.45 (0.5009)	5.02* (0.0251)	3.05 (0.0808)
341	Linear	24.04** (<0.0001)	0.00 (0.9717)	0.9717** (<0.0001)	5.79* (0.0161)	0.21 (0.6503)	37.73** (<0.0001)	1.88 (0.1699)
	Semi Log	15.30* (0.0001)	0.32 (0.5699)	15.31** (0.0001)	6.77** (0.0093)	0.86 (0.3529)	26.92** (<0.0001)	0.79 (0.3734)
342	Linear	8.23* (0.0041)	0.01 (0.9138)	7.85** (0.0051)	2.59 (0.1074)	1.69 (0.1940)	2.27 (0.1318)	2.67 (0.1023)
	Semi Log	9.35** (0.0022)	0.05 (0.8306)	9.01** (0.0027)	6.00* (0.0143)	1.74 (0.1876)	3.07 (0.0795)	2.17 (0.1406)
351	Linear	0.99 (0.3189)	0.10 (0.7568)	1.00 (0.3184)	0.07 (0.7894)	0.00 (0.9492)	1.43 (0.2315)	0.00 (0.9630)
	Semi Log	3.19 (0.0741)	0.17 (0.6766)	3.19 (0.0741)	0.26 (0.6132)	<0.01 (0.9880)	4.01* (0.0453)	0.03 (0.8637)
352	Linear	0.05	0.41	0.05	0.79	1.35	0.48	1.84

		(0.8185)	(0.5196)	(0.8195)	(0.3756)	(0.2454)	(0.4896)	(0.1749)
	Semi Log	0.04 (0.8433)	0.07 (0.7939)	0.04 (0.8427)	1.31 (0.2521)	0.87 (0.3519)	0.01 (0.9155)	0.95 (0.3307)
355	Linear	2.39 (0.1218)	2.58 (0.1080)	2.39 (0.1225)	0.00 (0.9599)	25.71** (<0.0001)	10.77** (0.001)	24.06** (<0.001)
	Semi Log	2.53 (0.1115)	0.76 (0.3840)	2.52 (0.1122)	0.48 (0.4873)	15.53** (0.0001)	19.81** (<0.0001)	16.47** (<0.0001)
356	Linear	5.34* (0.0209)	0.01 (0.9203)	5.33* (0.0210)	1.13 (0.2871)	1.66 (0.1976)	2.21 (0.1370)	0.08 (0.7786)
	Semi Log	4.24* (0.0396)	0.04 (0.8515)	4.23* (0.0398)	0.16 (0.6856)	2.20 (0.1377)	1.63 (0.2018)	0.14 (0.7065)
361	Linear	<0.01 (0.9746)	<0.01 (0.9792)	<0.01 (0.9726)	0.31 (0.5792)	0.02 (0.8933)	<0.01 (0.9843)	0.16 (0.6859)
	Semi Log	0.68 (0.4095)	0.30 (0.5848)	0.68 (0.4110)	0.05 (0.8184)	0.20 (0.6522)	0.84 (0.3585)	0.30 (0.5841)
362	Linear	1.70 (0.1923)	1.56 (0.2114)	1.69 (0.1933)	0.05 (0.8243)	0.53 (0.4648)	1.26 (0.2617)	0.12 (0.7300)
	Semi Log	2.73 (0.0985)	2.04 (0.1528)	2.72 (0.0992)	0.24 (0.6240)	1.03 (0.3097)	2.09 (0.1484)	0.39 (0.5332)
369	Linear	0.76 (0.3847)	0.90 (0.3431)	0.77 (0.3794)	1.25 (0.2644)	1.32 (0.2500)	1.64 (0.2007)	2.99 (0.0840)
	Semi Log	0.50 (0.4809)	0.81 (0.3672)	0.50 (0.4781)	0.52 (0.4717)	1.48 (0.2245)	1.54 (0.2150)	3.47 (0.0626)
371	Linear	0.10 (0.7572)	0.06 (0.8130)	0.09 (0.7662)	0.29 (0.5886)	0.30 (0.5834)	0.17 (0.6818)	0.82 (0.3653)
	Semi Log	1.80 (0.1794)	1.12 (0.2902)	1.76 (0.1842)	0.03 (0.8529)	0.06 (0.8036)	2.14 (0.1431)	0.01 (0.9347)
381	Linear	7.31** (0.0069)	<0.01 (0.9966)	7.37** (0.0066)	8.78** (0.0030)	3.71 (0.0542)	86.54** (<0.0001)	3.98* (0.0460)
	Semi Log	2.10 (0.1476)	1.07 (0.3021)	2.12 (0.1456)	4.68* (0.0305)	2.37 (0.1235)	68.99** (<0.0001)	0.52 (0.4707)
382	Linear	0.09 (0.7585)	1.38 (0.2408)	0.10 (0.7465)	5.42* (0.0199)	0.65 (0.4197)	3.43 (0.0639)	4.21* (0.0403)
	Semi Log	0.08 (0.7783)	0.59 (0.4432)	0.09 (0.7687)	5.94 (0.0148)	2.67 (0.1021)	2.54 (0.1106)	8.96** (0.0028)
383	Linear	31.26** (<0.0001)	10.59** (0.0011)	31.40** (<0.0001)	0.01 (0.9050)	2.37 (0.1239)	26.47** (<0.0001)	0.31 (0.5752)
	Semi Log	13.07** (0.0003)	4.62* (0.0317)	13.07** (0.0003)	0.79 (0.3727)	1.17 (0.2791)	9.91** (0.0016)	0.14 (0.7125)
384	Linear	0.16 (0.6897)	0.01 (0.9181)	0.17 (0.6843)	1.94 (0.1637)	0.01 (0.9413)	0.01 (0.9221)	0.41 (0.5239)
	Semi Log	0.11 (0.7372)	0.03 (0.8630)	0.12 (0.7296)	3.07 (0.0796)	0.05 (0.8191)	0.09 (0.7629)	0.55 (0.4567)
385	Linear	3.19 (0.0739)	0.10 (0.7462)	3.13 (0.0769)	3.46 (0.0628)	1.92 (0.1659)	0.59 (0.4422)	5.64* (0.0175)
	Semi Log	1.11 (0.2925)	0.24 (0.6274)	1.07 (0.3002)	4.75* (0.0293)	4.43* (0.0353)	0.11 (0.7433)	9.16** (0.0025)
390	Linear	20.81** (<0.0001)	1.82 (0.1768)	20.84** (<0.0001)	0.28 (0.5998)	5.51* (0.0189)	31.72** (<0.0001)	8.11** (0.0044)
	Semi Log	47.62** (<0.0001)	3.75 (0.0528)	47.69** (<0.0001)	1.81 (0.1787)	13.05** (0.0003)	81.86** (<0.0001)	17.53** (<0.0001)

Note: the reported values in the table are the chi-square test statistics and the corresponding p values. Note: The reported values in the table are the chi-square test statistics and the corresponding p values. The default version of the test performed by Stata and presented in the above table, makes the assumption that the disturbances are normally distributed with a variance of σ^2 (Baum, 2006). There is another version of the test derived by Keonker (1981) which only assumes that the

disturbance are independently and identically distributed (i.i.d) and as such drops the normality assumption. This version of the test may be considered more suitable for analysis given the fact that many of the indices are characterized by non-normal distributions. For that reason the test results presented in the above table has been cross checked with this second version of the test. Although changes in p values were observed, consideration of such changes do not alter the inferences drawn from the regression results, particularly in terms of the number of significant regression coefficients with expected signs. Therefore the results of the test derived by Koenker are not presented here.

Table C.9 (Corresponding to Table 4.3 in Chapter 4)
Regression Results for Ricardian Theory: With/without White's Correction

ISIC Codes	Regression Forms		Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
311	Linear	$\hat{\beta}$	1.1966 (0.205)	0.4055* (0.042)	0.0418 (0.205)	0.0004 (0.171)	0.9616 (0.056)	2.3169* (0.048)	1.3296* (0.012)
		R ²	0.0155	0.0669	0.0155	0.0206	0.0568	0.0616	0.1114
	Semi log	$\hat{\beta}$	1.1380 (0.247)	0.4600* (0.043)	0.0398 (0.247)	0.0005 (0.156)	1.0771 (0.060)	2.4234 (0.065)	1.4980* (0.013)
		R ²	0.01	0.0655	0.0107	0.0232	0.0543	0.0514	0.1077
313	Linear	$\hat{\beta}$	0.0045 (0.497)	-0.0636 (0.735)	0.00002 (0.498)	-0.00001 (0.631)	-0.1934 (0.769)	-0.1234 (0.573)	-0.1689 (0.735)
		R ²	<0.0001	0.0102	<0.0001	0.0029	0.0139	0.0009	0.0105
	Semi log	$\hat{\beta}$	0.8128 (0.220)	0.0605 (0.362)	0.0037 (0.221)	-0.00001 (0.614)	0.1386 (0.378)	0.4470 (0.347)	0.1103 (0.405)
		R ²	0.0154	0.0032	0.0153	0.0022	0.0025	0.004	0.0015
314	Linear	$\hat{\beta}$	0.8635 (0.203)	-0.0109 (0.682)	0.0024 (0.203)	0.000005 (0.156)	-0.0391 (0.709)	0.5875 (0.220)	-0.1051 (0.877)
		R ²	0.0304	0.0063	0.0303	0.0231	0.0087	0.0167	0.0383
	Semi log	$\hat{\beta}$	5.3329 (0.140)	0.0088 (0.469)	0.0146 (0.140)	0.00004 (0.107)	0.1212 (0.384)	2.4429 (0.208)	-0.4695 (0.857)
		R ²	0.0494	0.0002	0.0492	0.0423	0.0035	0.0123	0.0323
321	Linear	$\hat{\beta}$	0.1172 (0.418)	0.1795 (0.065)	0.0032 (0.416)	0.0003 (0.063)	0.3489 (0.127)	-2.0202 (0.989)	-0.1612 (0.709)
		R ²	0.0011	0.0481	0.0012	0.0491	0.0276	0.1073	0.0065
	Semi log	$\hat{\beta}$	0.0709 (0.464)	0.2545 (0.063)	0.0021 (0.461)	0.0003 (0.088)	0.5212 (0.112)	-2.2735 (0.966)	-0.0678 (0.566)
		R ²	0.0002	0.0491	0.0002	0.0389	0.0313	0.0689	0.0006
322	Linear	$\hat{\beta}$	5.9531* (0.014)	0.3577* (0.010)	0.1379* (0.014)	0.0003 (0.124)	1.2320** (0.003)	5.2369* (0.022)	0.3484 (0.274)
		R ²	0.2453	0.1204	0.2452	0.0332	0.1689	0.2041	0.0085
	Semi log	$\hat{\beta}$	5.7663* (0.025)	0.4234* (0.013)	0.1336* (0.025)	0.0003 (0.156)	1.4178** (0.003)	5.0330* (0.034)	0.5208 (0.233)
		R ²	0.1509	0.1106	0.1508	0.0202	0.1467	0.1236	0.0124
323	Linear	$\hat{\beta}$	7.9710 (0.125)	0.3963* (0.020)	0.0555 (0.125)	0.00007 (0.283)	1.2760* (0.011)	7.5656 (0.138)	0.8998 (0.081)
		R ²	0.0641	0.1022	0.064	0.0083	0.1255	0.0575	0.0483
	Semi log	$\hat{\beta}$	4.8195 (0.232)	0.2352 (0.110)	0.0336 (0.232)	0.00006 (0.232)	0.7842 (0.079)	4.5145 (0.244)	0.5506 (0.194)
		R ²	0.0244	0.0374	0.0243	0.005	0.0493	0.0213	0.0188
324	Linear	$\hat{\beta}$	-0.6158 (0.634)	0.2465* (0.018)	-0.0024 (0.626)	0.00004 (0.166)	0.6609* (0.028)	-0.5924 (0.629)	0.6537 (0.110)
		R ²	0.0069	0.1017	0.0059	0.0161	0.0849	0.0062	0.0356
	Semi log	$\hat{\beta}$	-2.7864 (0.749)	0.2966 (0.054)	-0.0115 (0.743)	0.00006 (0.176)	0.8002 (0.176)	-2.7964 (0.750)	0.7407 (0.187)

		R^2	0.0584	0.0607	0.0548	0.0134	0.0513	0.0572	0.0189
331	Linear	$\hat{\beta}$	7.7804* (0.015)	0.6296** (0.002)	0.0864* (0.015)	0.0002 (0.073)	2.0452** (0.001)	7.0631* (0.022)	1.1199 (0.052)
		R^2	0.2008	0.1692	0.2015	0.026	0.1961	0.1754	0.0552
	Semi log	$\hat{\beta}$	6.3900* (0.029)	0.5552** (0.007)	0.0710* (0.029)	0.0002 (0.107)	1.8175** (0.004)	5.7479* (0.040)	0.9742 (0.088)
		R^2	0.125	0.1214	0.1254	0.0192	0.1429	0.1072	0.0385
332	Linear	$\hat{\beta}$	1.1401 (0.124)	0.2857 (0.146)	0.0106 (0.124)	0.0001 (0.269)	0.5671 (0.205)	0.1890 (0.414)	-0.9874 (0.864)
		R^2	0.0466	0.0284	0.0465	0.012	0.0176	0.0016	0.0297
	Semi log	$\hat{\beta}$	0.6510 (0.104)	0.1980 (0.117)	0.0060 (0.104)	0.00006 (0.306)	0.4203 (0.159)	0.0972 (0.417)	-0.4738 (0.799)
		R^2	0.0402	0.0362	0.0401	0.0057	0.0255	0.0011	0.0181
341	Linear	$\hat{\beta}$	1.6496* (0.026)	0.6019** (0.001)	0.0422* (0.026)	0.0005* (0.019)	1.4781** (0.001)	1.5574* (0.046)	1.3693** (0.001)
		R^2	0.1568	0.2179	0.1568	0.0749	0.2015	0.134	0.2045
	Semi log	$\hat{\beta}$	1.4023* (0.017)	0.5272** (0.001)	0.0358* (0.017)	0.0004* (0.018)	1.2864** (0.001)	1.4007* (0.024)	1.2389** (0.001)
		R^2	0.1458	0.2152	0.1458	0.0662	0.1967	0.1395	0.2158
342	Linear	$\hat{\beta}$	0.5453 (0.228)	0.1446 (0.190)	0.0031 (0.230)	0.00002 (0.374)	0.6648 (0.090)	0.7406 (0.110)	0.8994* (0.032)
		R^2	0.0175	0.0176	0.0171	0.0024	0.0406	0.034	0.076
	Semi log	$\hat{\beta}$	0.8129 (0.107)	0.2706* (0.048)	0.0046 (0.108)	0.00003 (0.173)	1.0611* (0.015)	0.8675 (0.054)	1.1942** (0.006)
		R^2	0.0388	0.0617	0.0385	0.0078	0.1034	0.0467	0.1337
351	Linear	$\hat{\beta}$	0.9377** (0.005)	0.3142** (<0.001)	0.0612** (0.005)	0.0001 (0.266)	0.6965** (<0.001)	0.7227* (0.019)	0.5261** (0.003)
		R^2	0.1764	0.3343	0.1768	0.011	0.303	0.1136	0.1996
	Semi log	$\hat{\beta}$	1.3166** (0.004)	0.4766** (<0.001)	0.0860** (0.004)	0.0002 (0.246)	1.0392** (<0.001)	1.0296** (0.003)	0.8108** (0.001)
		R^2	0.183	0.4047	0.1833	0.0133	0.3548	0.1213	0.2495
352	Linear	$\hat{\beta}$	0.3008 (0.286)	0.2274 (0.149)	0.0120 (0.286)	-0.0001 (0.553)	0.6807 (0.110)	0.9520* (0.028)	1.1001* (0.019)
		R^2	0.009	0.0302	0.009	0.0005	0.0415	0.0982	0.1145
	Semi log	$\hat{\beta}$	0.3236 (0.210)	0.2266 (0.083)	0.0129 (0.210)	-0.00002 (0.520)	0.6565 (0.057)	0.8336* (0.013)	0.9895** (0.006)
		R^2	0.0182	0.0526	0.0182	0.0001	0.0679	0.1323	0.1626
355	Linear	$\hat{\beta}$	0.4158 (0.062)	0.2698* (0.019)	0.0042 (0.062)	0.00004 (0.145)	0.2138 (0.427)	1.0081 (0.196)	0.2135 (0.437)
		R^2	0.053	0.0951	0.0529	0.0255	0.0041	0.0305	0.0029
	Semi log	$\hat{\beta}$	0.5950* (0.019)	0.4106** (0.001)	0.0060* (0.019)	0.00005 (0.126)	0.9151 (0.175)	1.6028 (0.107)	1.0705 (0.173)
		R^2	0.0945	0.1917	0.0943	0.0297	0.066	0.0671	0.0644
356	Linear	$\hat{\beta}$	0.8376 (0.149)	0.3340 (0.066)	0.0103 (0.149)	0.00003 (0.430)	0.5985 (0.183)	1.0596 (0.099)	0.4719 (0.267)
		R^2	0.0402	0.0496	0.0402	0.0007	0.0183	0.0366	0.0087
	Semi log	$\hat{\beta}$	0.6093 (0.112)	0.2643 (0.057)	0.0075 (0.112)	0.00003 (0.410)	0.4657 (0.175)	0.7965 (0.099)	0.4419 (0.219)
		R^2	0.0376	0.0548	0.0375	0.0012	0.0195	0.0365	0.0134
361	Linear	$\hat{\beta}$	0.9874 (0.099)	0.3821** (0.006)	0.0016 (0.100)	0.00001 (0.292)	0.8946* (0.011)	1.0383 (0.0995)	0.8675* (0.030)
		R^2	0.0549	0.1914	0.0545	0.0101	0.1643	0.0543	0.1132
	Semi log	$\hat{\beta}$	0.9389 (0.078)	0.3595** (0.003)	0.0015 (0.078)	0.00001 (0.239)	0.8467** (0.006)	1.0202 (0.072)	0.8897* (0.012)

		R^2	0.0663	0.2264	0.0659	0.017	0.1966	0.0701	0.1591
362	Linear	$\hat{\beta}$	0.0678 (0.417)	-0.0514 (0.626)	0.0004 (0.417)	-0.00001 (0.577)	-0.2263 (0.682)	0.3030 (0.168)	-0.0333 (0.529)
		R^2	0.0011	0.0027	0.0011	0.001	0.0058	0.0238	0.0001
	Semi log	$\hat{\beta}$	0.1856 (0.282)	0.0122 (0.470)	0.0010 (0.282)	0.000002 (0.475)	-0.0909 (0.575)	0.4008 (0.100)	0.0847 (0.428)
		R^2	0.0086	0.0002	0.0086	0.0001	0.0009	0.0417	0.0009
369	Linear	$\hat{\beta}$	0.1228 (0.399)	0.0989 (0.228)	0.0007 (0.410)	<-0.0001 (0.524)	0.3252 (0.157)	-0.5897 (0.806)	0.0742 (0.425)
		R^2	0.0018	0.0148	0.0014	0.0001	0.0268	0.0196	0.001
	Semi log	$\hat{\beta}$	0.2923 (0.346)	0.2392 (0.119)	0.0018 (0.352)	0.00002 (0.447)	0.7175 (0.073)	-0.1861 (0.570)	0.5693 (0.170)
		R^2	0.0042	0.0365	0.0039	0.0005	0.055	0.0008	0.024
371	Linear	$\hat{\beta}$	0.0492 (0.431)	0.0514 (0.315)	0.0012 (0.432)	0.0001 (0.309)	0.3924 (0.108)	-0.0631 (0.598)	0.3034 (0.160)
		R^2	0.001	0.0074	0.0009	0.0078	0.0489	0.0019	0.0319
	Semi log	$\hat{\beta}$	0.4305 (0.129)	0.2732* (0.026)	0.0105 (0.130)	0.0001 (0.228)	1.1781** (0.002)	0.3277 (0.172)	1.1005** (0.003)
		R^2	0.0399	0.1127	0.0397	0.0176	0.2357	0.0281	0.2245
381	Linear	$\hat{\beta}$	1.3159** (0.003)	0.7076** (<0.001)	0.0394** (0.003)	0.0003 (0.059)	1.8365** (0.002)	3.6510 (0.061)	2.1614** (<0.001)
		R^2	0.1948	0.2766	0.1949	0.0448	0.1738	0.2472	0.2027
	Semi log	$\hat{\beta}$	0.5574** (0.007)	0.3135** (0.001)	0.0167** (0.007)	0.0001 (0.121)	0.7644* (0.011)	2.9151** (0.005)	1.2172** (<0.001)
		R^2	0.1308	0.2031	0.1307	0.0223	0.1127	0.5895	0.2405
382	Linear	$\hat{\beta}$	0.4764* (0.038)	0.4605** (0.007)	0.0873* (0.038)	0.0008 (0.086)	1.7493** (0.002)	0.5907** (0.007)	1.9477** (0.006)
		R^2	0.07	0.1311	0.0704	0.0193	0.1724	0.1307	0.2394
	Semi log	$\hat{\beta}$	0.3819* (0.016)	0.3645** (0.002)	0.0699* (0.015)	0.0005 (0.106)	1.3961** (<0.001)	0.4120** (0.005)	1.4760** (0.001)
		R^2	0.1017	0.1857	0.1021	0.0142	0.2483	0.1437	0.3109
383	Linear	$\hat{\beta}$	0.7883 (0.097)	0.4007 (0.125)	0.1365 (0.096)	0.0006 (0.275)	1.5649** (0.009)	0.4600 (0.067)	1.4009** (0.006)
		R^2	0.1192	0.0996	0.1202	0.008	0.1195	0.1805	0.1314
	Semi log	$\hat{\beta}$	0.6075** (0.010)	0.3892** (0.003)	0.1051** (0.010)	0.0003 (0.345)	1.8016** (<0.001)	0.3837** (0.001)	1.6771** (<0.001)
		R^2	0.1499	0.1988	0.1508	0.0036	0.3353	0.2657	0.3986
384	Linear	$\hat{\beta}$	-0.0527 (0.602)	-0.0384 (0.610)	-0.0099 (0.605)	0.0009 (0.150)	-0.1928 (0.684)	0.0999 (0.282)	-0.0371 (0.540)
		R^2	0.0015	0.0018	0.0016	0.0239	0.0051	0.0075	0.0002
	Semi log	$\hat{\beta}$	0.0024 (0.495)	0.0185 (0.437)	0.0001 (0.499)	0.0010 (0.072)	-0.0121 (0.514)	0.1384 (0.169)	0.1449 (0.319)
		R^2	<0.0001	0.0006	<0.0001	0.0469	<0.0001	0.0204	0.0049
385	Linear	$\hat{\beta}$	0.6042* (0.013)	0.4913** (0.007)	0.0230* (0.014)	0.0003 (0.224)	1.5057** (0.005)	0.1242 (0.273)	0.9623* (0.041)
		R^2	0.1238	0.1498	0.1221	0.0153	0.165	0.0097	0.0973
	Semi log	$\hat{\beta}$	0.3196* (0.013)	0.2588** (0.007)	0.0122* (0.014)	0.0002 (0.135)	0.7670* (0.017)	0.0541 (0.310)	0.4582 (0.085)
		R^2	0.1236	0.1483	0.18	0.0124	0.1526	0.0065	0.0787
390	Linear	$\hat{\beta}$	-1.8791 (0.829)	-0.0606 (0.630)	-0.0591 (0.829)	0.0001 (0.432)	-0.0691 (0.551)	-1.6206 (0.821)	0.0867 (0.419)
		R^2	0.0414	0.0027	0.0414	0.0007	0.0005	0.0501	0.0011
	Semi log	$\hat{\beta}$	-3.2401 (0.865)	-0.0852 (0.683)	-0.1020 (0.865)	0.00004 (0.452)	-0.1273 (0.565)	-2.9173 (0.870)	0.0445 (0.472)

		R ²	0.1277	0.0055	0.1278	0.0004	0.0016	0.1687	0.0003
Linear	No. of significant β coefficients		7	12	7	1	11	7	10
	Average R ²		0.0669	0.0910	0.0699	0.0175	0.0876	0.0721	0.0707
Semi log	No. of significant β coefficients		9	13	9	1	11	8	10
	Average R ²		0.0698	0.1033	0.1519	0.0169	0.1061	0.0930	0.1012

Note: The sample size for each sector varies between 46 and 32. Reported numbers of significant coefficients for each index are out of a total 25 cases.

Table C.10 (Corresponding to Table 4.4 in Chapter 4)
Robust Regression Results for Ricardian Theory

ISIC Codes	Regression Forms		Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
311	Linear	$\hat{\beta}$	1.853** (0.009)	0.4240* (0.048)	0.0642* (0.011)	<0.0001 (0.519)	1.0383* (0.048)	2.566** (0.002)	1.3196** (0.008)
		R ²	0.0569	0.0637	0.0550	0.00003	0.0582	0.1071	0.0968
	Semi log	$\hat{\beta}$	2.0024* (0.014)	0.5058* (0.040)	0.0692* (0.016)	<0.0001 (0.499)	1.2657* (0.036)	2.8672** (0.003)	1.6247** (0.005)
		R ²	0.0500	0.0659	0.0487	<0.0001	0.0614	0.1008	0.1003
313	Linear	$\hat{\beta}$	0.3365 (0.192)	0.2104 (0.132)	0.0016 (0.191)	<0.0001 (0.325)	0.5329 (0.135)	-0.0803 (0.587)	0.4318 (0.196)
		R ²	0.0120	0.0355	0.0118	0.0014	0.0340	0.0009	0.0181
	Semi log	$\hat{\beta}$	0.0706 (0.427)	0.0737 (0.346)	0.0003 (0.423)	0.00001 (0.247)	0.1979 (0.341)	0.2051 (0.366)	0.1763 (0.359)
		R ²	0.0006	0.0039	0.0006	0.0031	0.0037	0.0018	0.0032
314	Linear	$\hat{\beta}$	-0.0310 (0.798)	-0.0138 (0.704)	-0.0001 (0.802)	<0.0001 (0.495)	0.0631 (0.249)	-0.0829 (0.898)	0.0239 (0.419)
		R ²	0.0111	0.0087	0.0097	0.000002	0.0125	0.0175	0.0014
	Semi log	$\hat{\beta}$	-0.1511 (0.824)	-0.0006 (0.502)	-0.0004 (0.817)	0.000001 (0.312)	0.0490 (0.445)	-0.3192 (0.835)	-0.4676 (0.848)
		R ²	0.0094	<0.0001	0.0086	0.0017	0.0005	0.0100	0.0262
321	Linear	$\hat{\beta}$	0.7651* (0.014)	0.2139* (0.044)	0.0207* (0.014)	0.00002 (0.161)	0.5269* (0.036)	0.2054 (0.192)	0.1772 (0.235)
		R ²	0.0645	0.0574	0.0651	0.0058	0.0512	0.0059	0.0065
	Semi log	$\hat{\beta}$	0.9549* (0.026)	0.2986* (0.045)	0.0257* (0.026)	0.00003 (0.110)	0.7650* (0.032)	0.3309 (0.147)	0.3541 (0.151)
		R ²	0.0498	0.0565	0.0499	0.0076	0.0524	0.0075	0.0130
322	Linear	$\hat{\beta}$	0.1518 (0.356)	0.3682* (0.014)	0.0035 (0.356)	0.0001 (0.148)	1.1758** (0.006)	-0.0997 (0.589)	0.3265 (0.284)
		R ²	0.0013	0.1124	0.0013	0.0137	0.1306	0.0004	0.0076
	Semi log	$\hat{\beta}$	0.1924 (0.343)	0.4317* (0.017)	0.0044 (0.343)	0.0001 (0.202)	1.2975* (0.012)	0.0458 (0.465)	0.4218 (0.276)
		R ²	0.0021	0.1040	0.0021	0.0086	0.1072	<0.0001	0.0084
323	Linear	$\hat{\beta}$	0.8923* (0.033)	0.4377* (0.017)	0.0063* (0.032)	0.00003* (0.026)	1.3744** (0.007)	0.9817* (0.023)	0.5882 (0.128)
		R ²	0.0272	0.1051	0.0272	0.0297	0.1116	0.0341	0.0192
	Semi log	$\hat{\beta}$	0.4870 (0.148)	0.535** (0.009)	0.0034 (0.149)	0.00003* (0.045)	1.6205** (0.005)	0.6084 (0.088)	0.1683 (0.373)
		R ²	0.0144	0.1380	0.0146	0.0224	0.1485	0.0227	0.0012
324	Linear	$\hat{\beta}$	0.2070* (0.047)	0.2877* (0.012)	0.001** (0.050)	0.00001 (0.264)	0.8534** (0.007)	0.4118 (0.061)	0.8278 (0.060)
		R ²	0.0168	0.1163	0.0163	0.0062	0.1167	0.0184	0.0506

	Semi log	$\hat{\beta}$	0.3185 (0.070)	0.568** (0.0035)	0.0014 (0.072)	0.00001 (0.277)	1.7312** (0.002)	0.4251 (0.162)	1.0667 (0.103)
		R ²	0.0149	0.1843	0.0147	0.0045	0.2033	0.0091	0.0290
331	Linear	$\hat{\beta}$	0.2212 (0.164)	0.682** (0.002)	0.0025 (0.165)	0.0001* (0.014)	2.1289** (0.001)	-0.2511 (0.747)	1.1974 (0.050)
		R ²	0.0037	0.1664	0.0037	0.0450	0.1768	0.0022	0.0554
	Semi log	$\hat{\beta}$	0.2214 (0.191)	0.588** (0.008)	0.0024 (0.193)	0.00009* (0.046)	1.8827** (0.005)	-0.1064 (0.606)	1.0694 (0.080)
		R ²	0.0036	0.1186	0.0034	0.0273	0.1297	0.0005	0.0398
332	Linear	$\hat{\beta}$	0.1935 (0.364)	0.3067 (0.147)	0.0018 (0.363)	-0.00002 (0.635)	0.5967 (0.213)	-0.7437 (0.937)	-0.4931 (0.757)
		R ²	0.0029	0.0279	0.0029	0.0013	0.0152	0.0375	0.0086
	Semi log	$\hat{\beta}$	0.1938 (0.285)	0.2129 (0.119)	0.0018 (0.284)	<-0.0001 (0.545)	0.4585 (0.162)	-0.3679 (0.886)	-0.2485 (0.715)
		R ²	0.0082	0.0363	0.0082	0.0001	0.0243	0.0234	0.0061
341	Linear	$\hat{\beta}$	0.761** (0.004)	0.616** (0.001)	0.019** (0.004)	0.00009* (0.011)	1.4732** (0.002)	0.7603** (<0.001)	1.3047** (0.002)
		R ²	0.0849	0.2058	0.0845	0.0332	0.1811	0.0482	0.1501
	Semi log	$\hat{\beta}$	0.674** (0.009)	0.540** (0.0005)	0.017** (0.009)	0.0001* (0.011)	1.3016** (0.002)	0.4771** (0.009)	1.1496** (0.002)
		R ²	0.0688	0.2069	0.0686	0.0330	0.1836	0.0293	0.1665
342	Linear	$\hat{\beta}$	0.1295 (0.341)	0.1142 (0.263)	0.0007 (0.343)	-0.00001 (0.718)	-0.0051 (0.505)	0.4180 (0.159)	0.5046 (0.097)
		R ²	0.0023	0.0089	0.0022	0.0041	0.000004	0.0170	0.0287
	Semi log	$\hat{\beta}$	0.3276 (0.150)	0.2650 (0.067)	0.0019 (0.152)	<-0.0001 (0.637)	0.3076 (0.225)	0.5382 (0.099)	0.8788* (0.013)
		R ²	0.0141	0.0476	0.0138	0.0016	0.0064	0.0274	0.0680
351	Linear	$\hat{\beta}$	0.825** (<0.001)	0.319** (<0.001)	0.054** (<0.001)	0.0001 (0.087)	0.6662** (<0.001)	0.6706** (0.002)	0.5262** (0.005)
		R ²	0.1335	0.3059	0.1336	0.0193	0.2647	0.1148	0.1785
	Semi log	$\hat{\beta}$	0.498** (0.0025)	0.490** (<0.001)	0.033** (0.003)	0.0001 (0.151)	1.0092** (<0.001)	0.8508** (0.005)	0.8614** (0.001)
		R ²	0.0809	0.3727	0.0808	0.0116	0.3111	0.1133	0.2237
352	Linear	$\hat{\beta}$	0.3751 (0.181)	0.2342 (0.163)	0.0149 (0.181)	-0.0001 (0.938)	0.5810 (0.161)	1.0455** (0.008)	0.9250* (0.038)
		R ²	0.0148	0.0276	0.0148	0.0142	0.0250	0.1223	0.0686
	Semi log	$\hat{\beta}$	0.3238 (0.147)	0.2293 (0.098)	0.0129 (0.147)	-0.00003 (0.789)	0.59001 (0.090)	0.8537** (0.003)	0.9069* (0.010)
		R ²	0.0240	0.0486	0.0239	0.0056	0.0463	0.1545	0.1169
355	Linear	$\hat{\beta}$	0.699** (<0.001)	0.546** (<0.001)	0.007** (<0.001)	0.000003 (0.289)	1.2072** (0.001)	0.6059* (0.012)	1.4966** (0.001)
		R ²	0.1225	0.2603	0.1233	0.0037	0.1420	0.0530	0.1384
	Semi log	$\hat{\beta}$	0.703** (<0.001)	0.667** (<0.001)	0.007** (<0.001)	0.000003 (0.305)	1.8617** (<0.001)	0.7177** (0.006)	2.0561** (<0.001)
		R ²	0.1644	0.3042	0.1627	0.0029	0.1906	0.0817	0.1716
356	Linear	$\hat{\beta}$	0.4801* (0.040)	0.3930* (0.034)	0.0059* (0.040)	<-0.0001 (0.605)	0.7861* (0.035)	0.4631 (0.161)	1.0906* (0.029)
		R ²	0.0389	0.0590	0.0387	0.0005	0.0437	0.0109	0.0526
	Semi log	$\hat{\beta}$	0.3757* (0.035)	0.3136* (0.024)	0.0046* (0.035)	<-0.0001 (0.659)	0.6441* (0.022)	0.5489 (0.0595)	0.9205* (0.015)
		R ²	0.0412	0.0700	0.0411	0.0011	0.0558	0.0279	0.0707
361	Linear	$\hat{\beta}$	0.981** (0.001)	0.386** (0.009)	0.002** (0.001)	0.000002 (0.371)	0.8996* (0.011)	1.0752* (0.010)	0.8996* (0.036)
		R ²	0.1920	0.1895	0.1929	0.0021	0.1687	0.1422	0.1169
	Semi log	$\hat{\beta}$	0.734** (0.005)	0.364** (0.004)	0.001** (0.005)	0.000003 (0.291)	0.8564** (0.004)	0.8470* (0.016)	0.9254* (0.012)
		R ²	0.1289	0.2271	0.1290	0.0056	0.2076	0.1141	0.1662

362	Linear	$\hat{\beta}$	0.0011 (0.499)	-0.0355 (0.583)	0.00001 (0.497)	0.00001 (0.104)	0.1549 (0.320)	0.1939 (0.228)	-0.0079 (0.510)
		R ²	<0.0001	0.0009	<0.0001	0.0154	0.0019	0.0107	<0.0001
	Semi log	$\hat{\beta}$	0.0917 (0.381)	0.0474 (0.391)	0.0005 (0.381)	0.00001 (0.058)	0.2384 (0.234)	0.2623 (0.155)	0.2130 (0.248)
		R ²	0.0021	0.0018	0.0021	0.0234	0.0068	0.0207	0.0063
369	Linear	$\hat{\beta}$	0.4932* (0.013)	0.1015 (0.243)	0.0031* (0.017)	0.000001 (0.458)	0.3120 (0.188)	-0.6239 (0.877)	-0.0980 (0.596)
		R ²	0.0623	0.0146	0.0590	0.0001	0.0225	0.0219	0.0017
	Semi log	$\hat{\beta}$	0.7903* (0.015)	0.2572 (0.122)	0.0050* (0.017)	<0.0001 (0.482)	0.7485 (0.084)	-0.9227 (0.826)	0.1454 (0.409)
		R ²	0.0554	0.0358	0.0549	0.00002	0.0454	0.0205	0.0011
371	Linear	$\hat{\beta}$	0.3598 (0.111)	0.3122* (0.038)	0.0089 (0.108)	<-0.0001 (0.542)	1.2620** (0.005)	-0.1938 (0.827)	0.7867* (0.031)
		R ²	0.0344	0.0956	0.0350	0.0002	0.1560	0.0212	0.0413
	Semi log	$\hat{\beta}$	0.3219 (0.093)	0.3004* (0.024)	0.0080 (0.092)	0.000006 (0.445)	1.2228** (0.002)	0.2019 (0.249)	1.1201** (0.002)
		R ²	0.0444	0.1124	0.0441	0.0003	0.1808	0.0088	0.1457
381	Linear	$\hat{\beta}$	0.952** (<0.001)	0.747** (<0.001)	0.029** (<0.001)	0.00002 (0.227)	1.5994** (<0.001)	0.9226** (0.008)	1.7442** (<0.001)
		R ²	0.1753	0.2455	0.1774	0.0050	0.2145	0.0826	0.1518
	Semi log	$\hat{\beta}$	0.698** (0.0005)	0.574** (<0.001)	0.021** (0.001)	0.00001 (0.328)	1.2546** (<0.001)	0.7241** (0.004)	1.1752** (<0.001)
		R ²	0.1824	0.2446	0.1844	0.0022	0.2156	0.1058	0.2687
382	Linear	$\hat{\beta}$	0.809** (0.001)	0.620** (0.001)	0.148** (0.001)	-0.00002 (0.531)	1.9584** (<0.001)	0.7476** (<0.001)	1.6733** (0.001)
		R ²	0.1504	0.1580	0.1506	0.0001	0.1114	0.2425	0.1334
	Semi log	$\hat{\beta}$	0.456** (0.003)	0.412** (0.001)	0.083** (0.003)	-0.0001 (0.629)	1.3176** (<0.001)	0.4878** (<0.001)	1.3351** (<0.001)
		R ²	0.1426	0.1992	0.1424	0.0010	0.1496	0.2508	0.1877
383	Linear	$\hat{\beta}$	1.078** (<0.001)	0.956** (<0.001)	0.187** (<0.001)	0.0001 (0.193)	2.7703** (<0.001)	0.7882** (<0.001)	1.5725** (<0.001)
		R ²	0.1914	0.3130	0.1932	0.0055	0.2085	0.3036	0.0869
	Semi log	$\hat{\beta}$	0.445** (0.0055)	0.484** (<0.001)	0.077** (0.006)	0.0001 (0.311)	2.1480** (<0.001)	0.3549** (<0.001)	1.7560** (<0.001)
		R ²	0.1089	0.2199	0.1098	0.0018	0.2460	0.2194	0.1634
384	Linear	$\hat{\beta}$	-0.0167 (0.532)	-0.0382 (0.601)	-0.0032 (0.534)	0.00004 (0.422)	-0.2050 (0.682)	0.0793 (0.336)	-0.0156 (0.516)
		R ²	0.0002	0.0015	0.0002	0.0005	0.0051	0.0041	0.00004
	Semi log	$\hat{\beta}$	0.0395 (0.410)	0.0245 (0.423)	0.0069 (0.415)	-0.0001 (0.784)	-0.0025 (0.503)	0.1302 (0.201)	0.1957 (0.278)
		R ²	0.0011	0.0009	0.0010	0.0046	0.000001	0.0142	0.0070
385	Linear	$\hat{\beta}$	0.5667* (0.024)	0.4984* (0.011)	0.0216* (0.025)	-0.00001 (0.551)	1.3215* (0.012)	0.1560 (0.228)	0.6780 (0.067)
		R ²	0.0959	0.1344	0.0945	0.0002	0.1177	0.0130	0.0428
	Semi log	$\hat{\beta}$	0.2961* (0.026)	0.2665* (0.010)	0.0113* (0.027)	-0.00002 (0.660)	0.6762* (0.013)	0.0680 (0.270)	0.3297 (0.081)
		R ²	0.0942	0.1360	0.0928	0.0018	0.1065	0.0086	0.0324
390	Linear	$\hat{\beta}$	0.0238 (0.449)	-0.0345 (0.573)	0.0008 (0.449)	0.00001 (0.431)	0.0025 (0.498)	0.0905 (0.337)	-0.1265 (0.630)
		R ²	0.0003	0.0007	0.0003	0.0027	0.000002	0.0027	0.0022
	Semi log	$\hat{\beta}$	0.0998 (0.306)	-0.0386 (0.582)	0.0031 (0.306)	-0.00001 (0.551)	0.3570 (0.243)	0.17530 (0.212)	-0.5269 (0.926)
		R ²	0.0026	0.0006	0.0026	0.0001	0.0071	0.0069	0.0254
Linear	No. of significant β coefficients		14	16	14	3	16	10	11
	Average R ²		0.0598	0.1086	0.0597	0.0084	0.0948	0.0574	0.0583

Semi log	No. of significant β coefficients	12	16	12	3	16	9	12
	Average R^2	0.0523	0.1174	0.0522	0.0069	0.1076	0.0552	0.0819

Note: The sample size for each sector varies between 46 and 32. Reported numbers of significant coefficients for each index are out of a total 25 cases.

C.11 Testing Normality of Distribution of OLS Regression Residuals in Chapter 4

An attempt is made to test the reliability of the regression results reported in Tables C.6 and C.9 through normality of the regression residuals. Normally distributed residuals are not required to estimate the parameters of the regression model. But if the objective is estimation as well as inference, then the residuals are assumed to be normally distributed. Since in the present analysis the estimated β coefficients are tested for acceptance or rejection of null hypothesis of no association between the RCA indices and the factor endowment ratio or factor productivity ratio, through t statistics and associated p-values, the assumption of normality of residuals need to be invoked. Test of null hypothesis against alternative hypothesis through t test statistic, is based on the assumption that the error term is normally distributed. Obviously it is not possible to observe the residuals directly, so the practice is to consider the estimated errors from the estimated regression equations.

Although several tests exist to determine the normality of distribution of any variable, the present analysis would restrict itself to only one such test known as the Shapiro-Wilk (swilk) test. The choice for swilk is based in its power, which many studies have recognized to be better than other recommended tests of normality.¹ As observed by Royston (1991) comparison of powers may be the best way to select one among alternative tests of normality.² In fact the swilk test is considered to be an omnibus test of normality in the literature (D'Augostino *et al.*, 1990). D'Augostino *et al.* (1990) define an omnibus test to be

¹ Mendes and Pala (2003), Keskin (2006), Razali and Wah (2011) compared the power of Shapiro-Wilk test with D'Agostino-Pearson, Lilliefors, Kolmogorov-Smirnov, Anderson-Darling tests of normality and came to the conclusion that swilk is more powerful than any of the tests of normality considered in their papers.

² Swilk may not be the recommended test for normality in case of grouped or tied data (Gould and Rogers, 1991). A tie is when identical values occur more than once in the dataset. However for the dataset used in this report no such tie is observed, so there seems to be no harm in resorting to the swilk test.

the one which is able to determine deviations from normality either due to skewness or kurtosis.

For the swilk test, the test statistic is derived from the regression of ordered sample values on the corresponding expected standard normal ordered statistics, and the plot tends to be linear in case the sample is from a normally distributed population (Shapiro and Wilk, 1965). The test statistic W is defined to be the square of a linear combination of an ordered sample $x_1 \leq x_2 \leq \dots \leq x_n$ divided by the usual symmetric unbiased estimate of the variance of the sample (Shapiro and Wilk, 1965). It could be represented in the following manner:

$$W = \frac{(\sum_{i=1}^n a_i x_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

In the above formulation, x_i is the i th order statistics, a_i is the constant generated from means, variances and covariances of the order statistics of independent and identically distributed random variables of size n , sampled from a normal distribution (Keskin, 2006). The numerator is actually deduced from the square of the normalized best linear unbiased estimate of the slope of the linear regression of ordered sample values, on the expected standard normal ordered statistics. If indeed the sample is from a normal population, the slope coefficient would estimate the sample variance. Thus, both the numerator and the denominator of the above expression would both be estimating the same thing (Shapiro and Wilk 1965). So, closer is the value of the test statistic to 1, greater is the extent of normality in the data. Non normality would be represented by deviations from 1.³ The statistic takes a positive value always. The null hypothesis of normality distributed residuals is tested against the alternative hypothesis of non normally distributed residuals.

The results for the residual normality test corresponding to Tables C.6 and C.9 are reported in Tables C.11.1 and C.11.2 respectively. For each sector and for each index the value of the test statistic and the corresponding p values are reported.

³ Shapiro and Wilk's original test statistics was valid for sample sizes between 3 and 50. But later Royston (1982) by incorporating certain changes to the null distribution of the test statistics, made it valid for sample sizes between 7 and 2000.

Table C.11.1
Residual Normality Test corresponding to Table C.6

ISIC Codes	Regression Form	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
321	linear	0.7204** (<0.001)	0.9858 (0.714)	0.7218** (<0.001)	0.8083** (<0.001)	0.9580* (0.038)	0.7719** (<0.001)	0.9671 (0.106)
	semi log	0.7582** (<0.001)	0.9831 (0.571)	0.7581** (<0.001)	0.8046** (<0.001)	0.9481* (0.013)	0.8230** (<0.001)	0.9715 (0.173)
322	linear	0.6409** (<0.001)	0.9740 (0.227)	0.6405** (<0.001)	0.7455** (<0.001)	0.9855 (0.696)	0.6472** (<0.001)	0.993 (0.985)
	semi log	0.6883** (<0.001)	0.9745 (0.240)	0.6878** (<0.001)	0.7552** (<0.001)	0.9609 (0.052)	0.6869** (<0.001)	0.9762 (0.289)
323	linear	0.7218** (<0.001)	0.9640 (0.074)	0.7221** (<0.001)	0.6829** (<0.001)	0.9936 (0.989)	0.7062** (<0.001)	0.9522* (0.020)
	semi log	0.7525** (<0.001)	0.9807 (0.460)	0.7529** (<0.001)	0.6634** (<0.001)	0.9875 (0.799)	0.7583** (<0.001)	0.9904 (0.919)
324	linear	0.7073** (<0.001)	0.9221** (0.001)	0.7082** (<0.001)	0.7262** (<0.001)	0.9783 (0.362)	0.7667** (<0.001)	0.9890 (0.865)
	semi log	0.7295** (<0.001)	0.9513* (0.018)	0.7302** (<0.001)	0.7188** (<0.001)	0.9854 (0.693)	0.7678** (<0.001)	0.98874 (0.855)
351	linear	0.8524** (<0.001)	0.9842 (0.627)	0.8523** (<0.001)	0.7727** (<0.001)	0.9123** (0.0003)	0.9243** (0.001)	0.9698 (0.142)
	semi log	0.8384** (<0.001)	0.9887 (0.854)	0.8384** (<0.001)	0.80997** (<0.001)	0.9261** (0.001)	0.8702** (<0.001)	0.9680 (0.117)
352	linear	0.7730** (<0.001)	0.9579* (0.037)	0.7728** (<0.001)	0.7767** (<0.001)	0.9927 (0.976)	0.8841** (<0.001)	0.9913 (0.947)
	semi log	0.7627** (<0.001)	0.9562* (0.031)	0.7624** (<0.001)	0.8050** (<0.001)	0.9804 (0.447)	0.8689** (<0.001)	0.9827 (0.550)
385	linear	0.7584** (<0.001)	0.9805 (0.451)	0.7589** (<0.001)	0.6440** (<0.001)	0.8302** (<0.001)	0.7624** (<0.001)	0.8031** (<0.001)
	semi log	0.7785** (<0.001)	0.9906 (0.927)	0.7793** (<0.001)	0.7367** (<0.001)	0.7922** (<0.001)	0.7907** (<0.001)	0.7805** (<0.001)
No. of accepted cases for null hypothesis	Linear	0	5	0	0	4	0	5
	Semi log	0	5	0	0	4	0	6

Table C.11.2
Residual Normality Test corresponding to Table C.9

ISIC Codes	Regression Form	Balassa	RSCA	ARCA	NRCA	Log of Balassa	RTA	RC
311	linear	0.7674** (<0.001)	0.9673 (0.220)	0.7707** (<0.001)	0.6996** (<0.001)	0.9826 (0.714)	0.8341** (<0.001)	0.9511 (0.052)
	semi log	0.7702** (<0.001)	0.9681 (0.235)	0.7735** (<0.001)	0.7014** (<0.001)	0.9814 (0.664)	0.8352** (<0.001)	0.9532 (0.063)
313	linear	0.7108** (<0.001)	0.9513 (0.078)	0.7113** (<0.001)	0.8030** (<0.001)	0.9786 (0.623)	0.8372** (<0.001)	0.9800 (0.690)
	semi log	0.7545** (<0.001)	0.9522 (0.084)	0.7549** (<0.001)	0.8025** (<0.001)	0.9725 (0.415)	0.8425** (<0.001)	0.9855 (0.880)
314	linear	0.3489** (<0.001)	0.9139** (0.006)	0.3481** (<0.001)	0.5776** (<0.001)	0.9818 (0.794)	0.5465** (<0.001)	0.9678 (0.351)
	semi log	0.3888** (<0.001)	0.9110** (0.005)	0.3878** (<0.001)	0.6284** (<0.001)	0.9851 (0.891)	0.5389** (<0.001)	0.9641 (0.272)
321	linear	0.7294** (<0.001)	0.9773 (0.458)	0.7294** (<0.001)	0.8330** (<0.001)	0.9645 (0.146)	0.8698** (0.0001)	0.9681 (0.202)
	semi log	0.7317** (<0.001)	0.9767 (0.436)	0.7315** (<0.001)	0.8189** (<0.001)	0.9680 (0.201)	0.8393** (<0.001)	0.9663 (0.171)
322	linear	0.8129** (<0.001)	0.9429* (0.028)	0.8127** (<0.001)	0.8258** (<0.001)	0.9832 (0.749)	0.7967** (<0.001)	0.9711 (0.318)
	semi log	0.7596** (<0.001)	0.9416* (0.025)	0.7594** (<0.001)	0.8178** (<0.001)	0.9847 (0.809)	0.7550** (<0.001)	0.9733 (0.379)
323	linear	0.3771** (<0.001)	0.9606 (0.155)	0.3771** (<0.001)	0.6610** (<0.001)	0.9752 (0.486)	0.3665** (<0.001)	0.9137** (0.004)
	semi log	0.3349** (<0.001)	0.9530 (0.083)	0.3349** (<0.001)	0.6554** (<0.001)	0.9683 (0.290)	0.3221** (<0.001)	0.9186** (0.006)
324	linear	0.5161** (<0.001)	0.9231** (0.006)	0.5206** (<0.001)	0.6496** (<0.001)	0.9748 (0.440)	0.5568** (<0.001)	0.9492 (0.051)
	semi log	0.6586** (<0.001)	0.9307* (0.011)	0.6593** (<0.001)	0.6438** (<0.001)	0.9741 (0.418)	0.6781** (<0.001)	0.9535 (0.074)
331	linear	0.6393** (<0.001)	0.9535 (0.051)	0.6409** (<0.001)	0.639** (<0.001)	0.9924 (0.987)	0.6466** (<0.001)	0.9898 (0.945)
	semi log	0.5935** (<0.001)	0.9431* (0.020)	0.5951** (<0.001)	0.6377** (<0.001)	0.9884 (0.906)	0.6058** (<0.001)	0.9894 (0.936)
332	linear	0.7738** (<0.001)	0.9563 (0.117)	0.7735** (<0.001)	0.6979** (<0.001)	0.9638 (0.213)	0.8122** (<0.001)	0.9021** (0.002)
	semi log	0.7579** (<0.001)	0.9586 (0.140)	0.7575** (<0.001)	0.6795** (<0.001)	0.9660 (0.253)	0.8100** (<0.001)	0.8983 (0.002)
341	linear	0.7577** (<0.001)	0.9602 (0.109)	0.7577** (<0.001)	0.7322** (<0.001)	0.9787 (0.553)	0.7323** (<0.001)	0.9802 (0.614)
	semi log	0.7376** (<0.001)	0.9638 (0.152)	0.7375** (<0.001)	0.7394** (<0.001)	0.9801 (0.609)	0.7256** (<0.001)	0.9824 (0.706)
342	linear	0.7296** (<0.001)	0.9737 (0.379)	0.7329** (<0.001)	0.6694** (<0.001)	0.9215** (0.004)	0.8803** (<0.001)	0.8906** (0.0004)
	semi log	0.7420** (<0.001)	0.9718 (0.322)	0.7455** (<0.001)	0.6869** (<0.001)	0.9384* (0.017)	0.8801** (<0.001)	0.8897** (0.0004)
351	linear	0.5841** (<0.001)	0.9854 (0.891)	0.5847** (<0.001)	0.7062** (<0.001)	0.9724 (0.461)	0.7412** (<0.001)	0.9856 (0.896)
	semi log	0.6196** (<0.001)	0.9901 (0.979)	0.6202** (<0.001)	0.7177** (<0.001)	0.9709 (0.417)	0.7626** (<0.001)	0.9792 (0.691)
352	linear	0.8151** (0.00002)	0.9753 (0.553)	0.8148** (<0.001)	0.7553** (<0.001)	0.9706 (0.407)	0.8784** (0.0007)	0.9546 (0.126)

	semi log	0.8155** (0.00002)	0.9768 (0.606)	0.8152** (<0.001)	0.7564** (<0.001)	0.9769 (0.606)	0.8801** (0.0007)	0.9652 (0.278)
355	linear	0.8061** (<0.001)	0.9826 (0.712)	0.8073** (<0.001)	0.8338** (<0.001)	0.7796** (<0.001)	0.5155** (<0.001)	0.7968** (<0.001)
	semi log	0.7928** (<0.001)	0.9711 (0.303)	0.7942** (<0.001)	0.8407** (<0.001)	0.7341** (<0.001)	0.5360** (<0.001)	0.7530** (<0.001)
356	linear	0.7456** (<0.001)	0.9677 (0.215)	0.7456** (<0.001)	0.5096** (<0.001)	0.8145** (<0.001)	0.8843** (0.0002)	0.9131** (0.002)
	semi log	0.7423** (<0.001)	0.9638 (0.153)	0.7423** (<0.001)	0.5107** (<0.001)	0.8111** (<0.001)	0.8712** (<0.001)	0.9076** (0.001)
361	linear	0.5931** (<0.001)	0.9811 (0.830)	0.5917** (<0.001)	0.8856** (0.003)	0.9692 (0.476)	0.7281** (<0.001)	0.9814 (0.839)
	semi log	0.6250** (<0.001)	0.9728 (0.580)	0.6233** (<0.001)	0.8869** (0.003)	0.9540 (0.187)	0.7484** (<0.001)	0.9742 (0.621)
362	linear	0.9176** (0.006)	0.9716 (0.387)	0.9176** (0.006)	0.7302** (<0.001)	0.8627** (0.0002)	0.9252* (0.010)	0.8337** (<0.001)
	semi log	0.9294* (0.014)	0.9667 (0.267)	0.9294* (0.014)	0.7260** (<0.001)	0.8587** (0.0001)	0.9299* (0.014)	0.8287** (<0.001)
369	linear	0.7933** (<0.001)	0.9440* (0.047)	0.7926** (<0.001)	0.5597** (<0.001)	0.9594 (0.159)	0.9366* (0.027)	0.9491 (0.071)
	semi log	0.7901** (<0.001)	0.9509 (0.081)	0.7892** (<0.001)	0.5604** (<0.001)	0.9630 (0.211)	0.9339* (0.022)	0.9506 (0.079)
371	linear	0.8032** (0.00003)	0.9536 (0.157)	0.8029** (<0.001)	0.7740** (<0.001)	0.9215* (0.020)	0.9284* (0.028)	0.8651** (0.001)
	semi log	0.8234** (0.0001)	0.9712 (0.495)	0.8223** (0.0001)	0.7790** (<0.001)	0.9436 (0.086)	0.9459 (0.093)	0.8895** (0.003)
381	linear	0.8077** (<0.001)	0.9673 (0.208)	0.8069** (<0.001)	0.7385** (<0.001)	0.7263** (<0.001)	0.6547** (<0.001)	0.7431** (<0.001)
	semi log	0.7889** (<0.001)	0.9693 (0.249)	0.7880** (<0.001)	0.7064** (<0.001)	0.7259** (<0.001)	0.9254** (0.005)	0.6819** (<0.001)
382	linear	0.9226** (0.005)	0.9877 (0.903)	0.9230** (0.005)	0.8877** (0.0004)	0.9295** (0.008)	0.8747** (0.0001)	0.9336* (0.011)
	semi log	0.9232** (0.005)	0.9861 (0.852)	0.9237** (0.005)	0.8835** (0.0003)	0.9440* (0.028)	0.8806** (0.0002)	0.952 (0.056)
383	linear	0.8815** (0.0002)	0.9570 (0.082)	0.8814** (0.0002)	0.8305** (<0.001)	0.8649** (0.0001)	0.9480* (0.036)	0.7837** (<0.001)
	semi log	0.8775** (0.0002)	0.9737 (0.363)	0.8776** (0.0002)	0.8281** (<0.001)	0.9266** (0.006)	0.9504* (0.045)	0.8809** (0.0002)
384	linear	0.8737** (0.0001)	0.9105** (0.002)	0.8740** (0.0001)	0.8484** (<0.001)	0.9406* (0.019)	0.9819 (0.672)	0.9691 (0.244)
	semi log	0.8723** (0.0001)	0.9123** (0.002)	0.8724** (0.0001)	0.8762** (<0.001)	0.9393* (0.017)	0.9793 (0.562)	0.9672 (0.206)
385	linear	0.9445* (0.049)	0.9418* (0.040)	0.9441* (0.048)	0.7283** (<0.001)	0.9518 (0.087)	0.9876 (0.933)	0.9394* (0.033)
	semi log	0.92151** (0.009)	0.9583 (0.147)	0.9211** (0.008)	0.7286** (<0.001)	0.9595 (0.161)	0.9858 (0.889)	0.9424* (0.042)
390	linear	0.5241** (<0.001)	0.9473* (0.048)	0.5239** (<0.001)	0.7138** (<0.001)	0.9651 (0.212)	0.5056** (<0.001)	0.9675 (0.258)
	semi log	0.6735** (<0.001)	0.9515 (0.068)	0.6735** (<0.001)	0.7137** (<0.001)	0.9650 (0.210)	0.6794** (<0.001)	0.9672 (0.252)
No. of accepted cases for null hypothesis	Linear	0	18	0	0	16	2	14
	Semi Log	0	20	0	0	17	3	16

The normality tests reveal that the regression results corresponding to RSCA index reported in Tables C.9 are most reliable as the index generates the largest number of normally distributed residuals. Log of Balassa index is the second most reliable index to report normally distributed residuals. In Table C.11.1, the RC index performs as good as the RSCA index in terms of the number of normally distributed residuals corresponding to the linear regressions, but marginally better than RSCA index corresponding to the semi log form of regressions. It however falls short of the performance of not only the RSCA index, but also the Log of Balassa index, on the basis of similar criterion corresponding to results reported in Table C.11.2. Thus, for the RC index, while the results corresponding to Heckscher-Ohlin theory reported in Table C.6 are as reliable as that for RSCA index, the results corresponding to the Ricardian theory reported in Table C.9 are not as reliable as the RSCA index or the Log of Balassa index. The rest of the indices analysed – Balassa index, ARCA index, NRCA index and RTA index mostly report non-normally distributed residuals thus generating concerns about the reliability of results corresponding to these indices reported in Tables C.6 and C.9. These results are in fact at par with the non normality in distributions noted in cases of Balassa index, ARCA index, NRCA index and RTA index.

C.12 Deriving the Adjusted Import Formula used in Chapter 4⁴

By definition, the import demand elasticity for sector a in country i is

$$e_{i,a} = \frac{\text{p. c. change in quantity of imports}}{\text{p. c. change in price of imports with imposition of tariffs}}$$

$$e_{i,a} = \frac{d \log m_{i,a}}{d \log p_{i,a}}$$

Where $e_{i,a}$ is the import demand elasticity for the sector, $m_{i,a}$ is the quantity of imports in the sector, and $p_{i,a}$ is the price paid by domestic consumers for the good produced in the sector.

Rearranging terms gives us

$$e_{i,a} d \log p_{i,a} = d \log m_{i,a}$$

⁴ In personal communication with William Hauk Jr., University of South Carolina, USA.

The above expression can be written as,

$$e_{i,a}[\log P_{i,a} - \log P_{i,a}^*] = \log \bar{m}_{i,a} - \log \bar{m}_{i,a}^* \quad (1)$$

$P_{i,a}^*$ is the domestic price before the imposition of tariff. $\bar{m}_{i,a}^*$ is the quantity of imports before the imposition of tariff. $P_{i,a}$ is the domestic price after the imposition of tariff. $\bar{m}_{i,a}$ is the quantity of imports after the imposition of tariff.

It is assumed that country i is small so that the world price of good a is unaffected by country i 's imposition of tariff on that good.⁵ Let P_a^w be the world price of good at a , at which imports are valued. Expression (1) could then be represented as,

$$e_{i,a}[\log P_{i,a} - \log P_{i,a}^*] = \log \frac{P_a^w \bar{m}_{i,a}}{P_a^w \bar{m}_{i,a}^*} \quad (2)$$

It is further assumed that the domestic price of good a changes in proportion to an ad-valorem tariff rate $\tau_{i,a}$, which implies the domestic price after the imposition of tariff could be expressed as,

$$P_{i,a} = (1 + \tau_{i,a})P_{i,a}^* \quad (3)$$

Substituting expression (3) in expression (2) and rearranging terms gives the final expression for import value M adjusted for tariff at time t ,

$$M_{i,a,t}^* = \exp[\log M_{i,a,t} - e_{i,a} \log(1 + \tau_{i,a,t})]$$

⁵ The concern here is that in a large country, a tariff might lower the world price of the good, and hence the full effect of the tariff may not be passed along to domestic consumers. In other words, the amount of imports demanded may not fall to the extent desired in order to capture the change in the level of imports due to imposition of tariffs. The import demand elasticity would be much lower than in the case where tariff has no effect on the world price of the product. In making this assumption, we are following the same partial equilibrium methodology found in Kee *et al.* (2009) and Hauk (2012). However, to the extent that protection in a large country may have an effect on world prices that are similar across countries, this effect will be mitigated when imports are divided by the total level of world imports while calculating the RCA indices.

Table C.12.1 (Corresponding to Table 4.5 in Chapter 4)
Rank Correlation Results for Heckscher-Ohlin Theory

ISIC Codes	Balassa	RSCA	ARCA	NRCA	Log of Balassa	Unadjusted RTA	Unadjusted RC	Adjusted RTA	Adjusted RC
321	0.3064 (0.053)	0.3167* (0.047)	0.3064 (0.053)	0.3044 (0.054)	0.3064 (0.053)	0.0793 (0.341)	0.1128 (0.280)	0.033 (0.433)	0.07 (0.359)
322	0.441** (0.008)	0.441** (0.008)	0.439** (0.009)	0.532** (0.002)	0.441** (0.008)	0.6695** (0.0001)	0.664** (0.0001)	0.67** (0.0001)	0.6468** (0.0001)
323	0.446** (0.008)	0.441** (0.008)	0.446** (0.008)	0.4163* (0.012)	0.446** (0.008)	0.53** (0.002)	0.5374** (0.001)	0.5389** (0.001)	0.5502** (0.001)
324	0.0254 (0.452)	-0.0138 (0.474)	0.0254 (0.452)	0.3362 (0.050)	0.0254 (0.452)	0.28 (0.088)	0.2108 (0.156)	0.27 (0.096)	0.2154 (0.151)
351	-0.4103* (0.014)	-0.4103* (0.014)	-0.4044* (0.015)	-0.0202 (0.459)	-0.4103* (0.014)	-0.667** (0.0001)	-0.6443** (0.0001)	-0.661** (0.0001)	-0.613** (0.0002)
352	-0.457** (0.007)	-0.469** (0.005)	-0.457** (0.006)	-0.2522 (0.093)	-0.457** (0.006)	-0.4714** (0.005)	-0.4754** (0.005)	-0.441** (0.008)	-0.451** (0.007)
385	-0.722** (<0.001)	-0.722** (<0.001)	-0.722** (<0.001)	-0.2424 (0.103)	-0.722** (<0.001)	-0.4478** (0.007)	-0.6537** (0.0001)	-0.4123* (0.013)	-0.629** (0.0002)
No. of significant coefficients	5	6	5	2	5	5	5	5	5

Note: The sample size is 29 each for sectors 321, 322, 323, 351, 352 and 385, and 25 for sector 324. Reported numbers of significant coefficients for each index are out of a total 7 cases.

Table C.12.2 (Corresponding to Table 4.6 in Chapter 4)
Rank Correlation Results for Ricardian Theory

ISIC Codes	Balassa	RSCA	ARCA	NRCA	Log of Balassa	Unadjusted RTA	Unadjusted RC	Adjusted RTA	Adjusted RC
311	0.2083 (0.211)	0.2083 (0.211)	0.2083 (0.211)	0.3235 (0.103)	0.2083 (0.211)	0.1299 (0.310)	0.2941 (0.126)	0.1422 (0.293)	0.2721 (0.145)
313	0.2824 (0.145)	0.2882 (0.130)	0.2882 (0.130)	0.1000 (0.356)	0.2824 (0.145)	0.2500 (0.175)	0.1059 (0.348)	0.2647 (0.161)	0.1059 (0.348)
314	0.3324 (0.104)	0.3324 (0.104)	0.3324 (0.104)	0.2265 (0.200)	0.3324 (0.104)	0.4794* (0.030)	0.4294* (0.048)	0.5029* (0.024)	0.4706* (0.033)
321	0.4193* (0.037)	0.4193* (0.037)	0.4193* (0.037)	0.5263* (0.010)	0.4193* (0.040)	0.6018** (0.003)	0.5807** (0.005)	0.6018** (0.003)	0.5825** (0.004)
322	0.1084 (0.334)	0.1084 (0.334)	0.1084 (0.334)	0.1476 (0.280)	0.1084 (0.334)	-0.0712 (0.611)	-0.0898 (0.638)	-0.0733 (0.614)	-0.0898 (0.638)
323	0.4778* (0.022)	0.4778* (0.022)	0.4778* (0.022)	0.3849 (0.057)	0.4778* (0.022)	0.2508 (0.158)	0.2178 (0.193)	0.2766 (0.132)	0.2446 (0.164)
324	0.5319* (0.014)	0.5319* (0.014)	0.5319* (0.014)	-0.0294 (0.5446)	0.5319* (0.014)	0.2672 (0.150)	0.4632* (0.031)	0.2672 (0.150)	0.4755* (0.027)
331	0.3652 (0.075)	0.3652 (0.075)	0.3652 (0.075)	0.2157 (0.203)	0.3652 (0.075)	0.1961 (0.225)	0.1348 (0.303)	0.1961 (0.225)	0.1078 (0.340)
332	0.1373 (0.300)	0.1373 (0.300)	0.1373 (0.300)	-0.0735 (0.610)	0.1373 (0.300)	-0.0833 (0.625)	-0.027 (0.541)	-0.0637 (0.596)	-0.0417 (0.563)
341	0.728** (0.0003)	0.728** (0.0003)	0.728** (0.0003)	0.548** (0.010)	0.728** (0.0003)	0.7792** (0.0001)	0.8122** (<0.001)	0.7957** (0.0001)	0.8287** (<0.001)
342	0.3775 (0.068)	0.3775 (0.068)	0.3848 (0.064)	0.2917 (0.128)	0.3775 (0.068)	0.3395 (0.056)	0.3701 (0.072)	0.3995 (0.056)	0.3701 (0.072)
351	0.3706 (0.079)	0.3706 (0.079)	0.3706 (0.079)	0.1059 (0.348)	0.3706 (0.079)	0.5353* (0.016)	0.5441* (0.015)	0.5794** (0.010)	0.5441* (0.015)
352	0.5411* (0.012)	0.5411* (0.012)	0.5411* (0.012)	-0.0294 (0.545)	0.5411* (0.012)	0.6765** (0.001)	0.6691** (0.002)	0.6765** (0.001)	0.6691** (0.002)
353	0.1091	0.1091	0.1091	-0.2182	0.1091	-0.4818	-0.2455	-0.4273	-0.3364

	(0.375)	(0.375)	(0.375)	(0.740)	(0.375)	(0.933)	(0.767)	(0.905)	(0.844)
354	0.2 (0.374)	0.2 (0.374)	0.2 (0.374)	0.2 (0.374)	0.2 (0.374)	-0.1 (0.564)	-0.1 (0.564)	-0.1 (0.564)	-0.1 (0.564)
355	0.4448* (0.032)	0.4241* (0.040)	0.4448* (0.032)	0.4654* (0.026)	0.4448* (0.032)	0.3106 (0.105)	0.3932 (0.053)	0.3106 (0.105)	0.3767 (0.062)
356	0.3911 (0.054)	0.3911 (0.054)	0.3911 (0.054)	0.2362 (0.173)	0.3911 (0.054)	0.645** (0.002)	0.6182** (0.003)	0.645** (0.002)	0.6223** (0.003)
361	0.2696 (0.148)	0.2696 (0.148)	0.2843 (0.134)	0.0392 (0.441)	0.2696 (0.148)	0.2181 (0.200)	0.2279 (0.189)	0.2181 (0.200)	0.2279 (0.189)
362	0.3088 (0.114)	0.3015 (0.120)	0.3088 (0.114)	0.1176 (0.326)	0.3088 (0.114)	0.3382 (0.092)	0.3456 (0.087)	0.3578 (0.079)	0.3775 (0.068)
369	0.5235* (0.019)	0.5235* (0.019)	0.5235* (0.019)	0.4471* (0.041)	0.5235* (0.019)	0.4529* (0.039)	0.5294* (0.018)	0.4529* (0.039)	0.5294* (0.018)
371	0.2321 (0.203)	0.2321 (0.203)	0.2321 (0.203)	0.0964 (0.366)	0.2321 (0.203)	0.1107 (0.347)	0.15 (0.297)	0.1143 (0.343)	0.15 (0.297)
372	0.2615 (0.183)	0.2615 (0.183)	0.2615 (0.183)	0.2352 (0.209)	0.2615 (0.183)	0.2835 (0.163)	0.367 (0.098)	0.3187 (0.133)	0.3846 (0.087)
381	0.847** (<0.001)	0.847** (<0.001)	0.847** (<0.001)	0.645** (<0.001)	0.847** (<0.001)	0.839** (<0.001)	0.8762** (<0.001)	0.839** (<0.001)	0.8824** (<0.001)
382	0.553** (0.007)	0.553** (0.007)	0.560** (0.006)	0.2544 (0.147)	0.553** (0.007)	0.5491** (0.007)	0.5526** (0.007)	0.5491** (0.007)	0.5509** (0.007)
383	0.3772 (0.057)	0.3772 (0.057)	0.3772 (0.057)	-0.0772 (0.623)	0.3772 (0.057)	0.4193* (0.037)	0.4509* (0.026)	0.4123* (0.040)	0.4456* (0.028)
384	0.586** (0.004)	0.591** (0.004)	0.591** (0.004)	0.588** (0.004)	0.586** (0.004)	0.4982* (0.015)	0.5474** (0.008)	0.4842* (0.018)	0.5404** (0.008)
385	0.1509 (0.269)	0.1439 (0.278)	0.1439 (0.278)	-0.0807 (0.629)	0.1509 (0.2688)	-0.0702 (0.612)	<0.0001 (0.500)	-0.0702 (0.612)	<0.0001 (0.500)
390	0.1422 (0.293)	0.1422 (0.293)	0.1422 (0.293)	0.1348 (0.303)	0.1422 (0.293)	0.2843 (0.134)	0.174 (0.252)	0.2843 (0.134)	0.174 (0.252)
No. of significant coefficients	10	10	10	6	10	11	12	11	12

Note: The sample size for each sector varies between 19 and 15. Reported numbers of significant coefficients for each index are out of a total 28 cases.

Appendix D

D.1 Derivation of the Expression for Domestic Value-Added in Exports used in Chapter 5

Vertical specialization in Hummels *et al.*'s framework is simply foreign value-added in exports. While deriving the expression for foreign value-added in exports, Hummels *et al.* (2001) allow for the fact that apart from imported inputs being directly used to produce exports, they could also be used to produce domestic intermediate goods which in turn could be used to produce exports of any country. It follows many such possibilities could be identified where imported inputs are embodied in exports of a country at the third, fourth or fifth stage and so on. These possibilities could thus be expressed mathematically as:

$$\begin{aligned} & u(A^M X + A^M A^D X + A^M A^D A^D X + A^M A^D A^D A^D X + \dots) \\ & = uA^M(I - A^D)^{-1}X \end{aligned} \quad (1)$$

where u is the $(1 \times N)$ vector of 1s, A^M is the $(N \times N)$ imported input-output coefficient matrix (its each element a_{ij}^M represent the amount of imported input from sector i used to produce one unit of sector j 's output), I is the identity matrix, A^D is $(N \times N)$ domestic input-output coefficient matrix (its each element a_{ij}^D represent the amount of domestically produced input from sector i used to produce one unit of sector j 's output), and X is the $(N \times 1)$ vector of exports. In this scenario, $(I - A^D)^{-1}$ captures the total domestic output requirement from each sector i to produce exports from sector j .

Koopman *et al.* (2012) used Hummels *et al.*'s (2001) conceptual framework to show that the total output in any sector equals the sum of direct value-added in that particular sector and the cost of intermediate inputs from all domestically produced and imported products. Using matrix notations, the statement can be expressed as:

$$uA^D + uA^M + A_V = u \quad (2)$$

u , A^D , A^M have already been defined. A_V is the $(1 \times N)$ row vector of domestic value-added coefficients and its each element a_V^j represents each sector j 's ratio of value-added to gross output.

Defining DVS as the share of domestic value-added and FVS as the share of foreign value-added, it is possible to arrive at the expression for domestic content or domestic value added share as:

$$\begin{aligned} DVS &= u - FVS \\ &= u - uA^M(I - A^D)^{-1} \end{aligned} \quad (3)$$

Replacing uA^M by the equivalent expression from equation (2), equation (3) reduces to:

$$DVS = A_V(I - A^D)^{-1} \quad (4)$$

Koopman *et al.* (2012) define expression (4) as the domestic content share in exports of a country. It provides an estimate of how much of the domestic content from various sectors could be found in a country's final exports of a product. But in this thesis a minor modification is introduced in expression (4) to obtain domestic value-added from each of the N sectors in a country's exports of all sectors. Using the usual matrix notations, the expression for domestic value-added in exports is now:

$$VAX = \hat{A}_V(I - A^D)^{-1}X$$

where $\hat{A}_V$ is a $(N \times N)$ diagonal matrix corresponding to the row vector A_V .

Table D.1
Summary Statistics for Ricardian Sample used in Chapter 5

Variable	No. of observations	Mean	Median	Standard deviation	Skewness	Kurtosis
Export	1156	6426474	479611.3	19200000	4.246556	20.96003
Value added	1156	17000000	1761501	39000000	3.179896	13.17147
Output	1156	41300000	4471034	92100000	3.043361	11.83816
No. of employees	1156	202670	49801	389606.7	3.104764	13.16145
Share of output of 17 sectors sold to sector 311	340	0.0178138	0.0047251	0.0319014	3.12091	15.00576
Share of output of 17 sectors sold to sector 313	340	0.0088731	0.0019031	0.0182713	5.410059	48.80178
Share of output of 17 sectors sold to sector 321	340	0.0138841	0.0008637	0.042819	4.287157	22.83307
Share of output of 17 sectors sold to sector 322	340	0.0162931	0.0005627	0.0484442	4.019947	20.0834
Share of output of 17 sectors sold to sector 323	340	0.0104645	0.0005502	0.0389388	6.052706	50.61205
Share of output of 17 sectors sold to sector 331	340	0.0135945	0.0025616	0.033777	3.791842	17.1844
Share of output of 17 sectors sold to sector 341	340	0.0149214	0.0020087	0.0455417	4.390849	22.76865
Share of output of 17 sectors sold to sector 351	340	0.0232039	0.0075509	0.0417599	3.308142	14.93669
Share of output of 17 sectors sold to sector 353	340	0.0155168	0.0011381	0.0520189	4.838135	27.76943
Share of output of 17 sectors sold to sector 361	340	0.0094786	0.0021271	0.023737	4.646387	29.60059
Share of output of 17 sectors sold to sector 371	340	0.0159055	0.0010788	0.0483994	5.625078	44.62193
Share of output of 17	340	0.0096056	0.0005797	0.0392994	7.325031	68.00406

sectors sold to sector 372						
Share of output of 17 sectors sold to sector 381	340	0.0202862	0.0034542	0.0464705	3.479288	16.27048
Share of output of 17 sectors sold to sector 382	340	0.0244333	0.0064755	0.0387157	2.498279	10.5266
Share of output of 17 sectors sold to sector 383	340	0.014627	0.0028016	0.0298088	3.372058	15.90232
Share of output of 17 sectors sold to sector 384	340	0.0220643	0.0058659	0.0377021	2.82873	12.61518
Share of output of 17 sectors sold to sector 390	340	0.0066553	0.0017612	0.014032	4.149295	22.65924

Note: Exports, Value added and Output are reported in thousand US dollars in the Trade Production and Protection Database. Countries included in the sample – Argentina, Australia, Chile, Colombia, Ecuador, Egypt, Spain, Finland, United Kingdom, Greece, Iran, Japan, Korea, Peru, Poland, Portugal, Turkey, Uruguay, United States, South Africa.

Table D.2
Summary Statistics for Heckscher-Ohlin Sample used in Chapter 5

Variable	No. of Observations	Mean	Median	Standard Deviation	Skewness	Kurtosis
Export	1156	3681406	290849.4	9906413	4.111272	21.55679
Employees	68	30700000	19300000	34800000	1.835263	5.545782
Capital stock (2005 PPP)	68	2600000000000	518000000000	3840000000000	1.367491	3.09886
GDP PPP (2005 constant prices)	68	1440000000000	460000000000	2310000000000	2.16643	6.796003
GDP PPP (current prices)	68	1150000000000	303000000000	1840000000000	2.08015	6.41666
PPP conversion factor to market exchange rate ratio	68	0.7967342	0.7505341	0.4018329	0.648734	2.819858

Note: Exports are reported in thousand US dollars in Trade Production and Protection Database. PPP denotes Purchasing Power Parity. Countries included in the sample - Argentina, Australia, Chile, Colombia, Ecuador, Egypt, Spain, Finland, United Kingdom, Greece, Iran, Japan, Korea, Peru, Poland, Portugal, Turkey, Uruguay, United States, South Africa.

Table D.3 (Corresponding to Table 5.2 in Chapter 5)
Regression and Correlation Results for Ricardian Theory: Sector Specific Samples

ISIC Codes	Regression / Correlation Results	Balassa	RSCA	ARCA	NRCA	Log of Balassa
311	β	-1.8438 (0.749)	0.5428** (<0.001)	0.0492* (0.033)	0.0027 (0.205)	0.9653** (<0.001)
	R^2	0.1550	0.0009	0.0181	0.0078	0.0108
	ρ	-0.0581 (0.681)	-0.0581 (0.681)	-0.1540 (0.895)	0.2151* (0.039)	-0.0581 (0.681)
313	β	0.5232 (0.201)	0.0102 (0.437)	0.0023 (0.219)	0.0002 (0.126)	0.1208 (0.283)
	R^2	0.1425	0.0270	0.0259	<0.0001	0.0452
	ρ	0.0527 (0.335)	0.0527 (0.335)	0.0853 (0.245)	0.0644 (0.301)	0.0527 (0.335)
321	β	0.8640 (0.145)	0.0654 (0.221)	0.0218 (0.318)	0.0015 (0.245)	0.2161 (0.145)
	R^2	0.0016	<0.0001	0.0046	0.0012	<0.0001
	ρ	0.0558 (0.326)	0.0558 (0.326)	0.0508 (0.340)	0.1486 (0.113)	0.0558 (0.326)
322	β	5.9516** (0.001)	-0.0137 (0.577)	0.0343** (<0.001)	0.0004 (0.198)	0.4375** (0.003)
	R^2	0.1458	0.0164	0.0303	0.0039	0.0369
	ρ	0.2085* (0.044)	0.2085* (0.044)	0.1449 (0.119)	0.2737* (0.012)	0.2085* (0.044)
323	β	3.7501** (<0.001)	0.4825** (0.001)	0.0151** (0.001)	0.0002 (0.192)	1.1040** (0.002)
	R^2	0.1566	0.2257	0.2215	0.0055	0.2138
	ρ	0.4581** (0.0001)	0.4581** (0.0001)	0.5351** (<0.001)	0.4020** (0.0004)	0.4581** (0.0001)
331	β	2.2909 (0.061)	0.5613* (0.023)	0.0204* (0.018)	0.0007* (0.016)	1.3104* (0.023)
	R^2	0.0209	0.0551	0.0001	0.0088	0.0585
	ρ	-0.1640 (0.909)	-0.1640 (0.909)	-0.1280 (0.851)	-0.3354 (0.997)	-0.1640 (0.909)
341	β	0.0705 (0.405)	0.2373* (0.014)	0.0140 (0.108)	0.0004 (0.365)	0.6814* (0.010)
	R^2	0.0332	0.1287	0.0684	<0.0001	0.1585
	ρ	0.5887** (<0.001)	0.5887** (<0.001)	0.5523** (<0.001)	0.1928 (0.058)	0.5887** (<0.001)
351	β	0.7680** (<0.001)	0.4795** (<0.001)	0.1197** (<0.001)	0.0045 (0.167)	1.0068** (<0.001)
	R^2	0.3886	0.4134	0.3895	0.0774	0.4149
	ρ	0.6048** (<0.001)	0.6048** (<0.001)	0.5993** (<0.001)	0.2553* (0.018)	0.6048** (<0.001)
353	β	1.3997 (0.067)	0.0718** (0.007)	0.0135* (0.033)	0.0002* (0.027)	0.1988** (0.008)
	R^2	0.2907	0.3553	0.3050	0.1124	0.3577
	ρ	0.7222** (<0.001)	0.7222** (<0.001)	0.7227** (<0.001)	0.5578** (<0.001)	0.7222** (<0.001)
361	β	0.7590** (<0.001)	0.3801** (<0.001)	0.0158** (<0.001)	0.0003 (0.213)	0.8148** (<0.001)
	R^2	0.0400	0.0293	0.0333	0.0006	0.0352

	ρ	0.1273 (0.151)	0.1273 (0.151)	0.1300 (0.145)	0.1284 (0.148)	0.1273 (0.151)
371	β	1.0943** (<0.001)	0.3084** (<0.001)	0.0516** (<0.001)	0.0025** (<0.001)	0.8262** (<0.001)
	R^2	0.4214	0.2203	0.4153	0.1006	0.2549
	ρ	0.5625** (<0.001)	0.5625** (<0.001)	0.5624** (<0.001)	0.6094** (<0.001)	0.5625** (<0.001)
372	β	2.2461** (<0.001)	0.0863** (0.003)	0.0511** (<0.001)	0.0012** (<0.001)	0.2682** (<0.001)
	R^2	0.7299	0.2094	0.7499	0.2115	0.3026
	ρ	0.5277** (<0.001)	0.5277** (<0.001)	0.5327** (<0.001)	0.4655** (0.0001)	0.5277** (<0.001)
381	β	0.4489** (<0.001)	0.2618** (<0.001)	0.0204** (<0.001)	0.0009 (0.165)	0.5375** (<0.001)
	R^2	0.2855	0.2713	0.2927	0.0341	0.2579
	ρ	0.6578** (<0.001)	0.6578** (<0.001)	0.6664** (<0.001)	0.5248** (<0.001)	0.6578** (<0.001)
382	β	0.3520** (0.005)	0.2919** (0.009)	0.0452* (0.030)	0.0028 (0.092)	1.1913** (0.007)
	R^2	0.4173	0.4511	0.3298	0.0098	0.4772
	ρ	0.8121** (<0.001)	0.8121** (<0.001)	0.8005** (<0.001)	0.2859** (0.009)	0.8121** (<0.001)
383	β	0.0591 (0.339)	0.0006 (0.497)	0.0915* (0.036)	0.0036 (0.197)	-0.0263 (0.523)
	R^2	0.0012	0.0025	0.0427	0.0014	0.0001
	ρ	0.3755** (0.001)	0.3755** (0.001)	0.3090** (0.005)	0.0782 (0.263)	0.3755** (0.001)
384	β	-0.1030 (0.693)	-0.0543 (0.605)	-0.0118 (0.656)	-0.0003 (0.630)	0.3700 (0.281)
	R^2	0.0001	0.0093	0.0131	0.0005	0.0694
	ρ	0.3929** (0.001)	0.3929** (0.001)	0.3653** (0.001)	0.2232* (0.034)	0.3929** (0.001)
390	β	0.1255 (0.233)	0.1379 (0.158)	0.0018 (0.324)	-0.0002 (0.831)	0.3351 (0.183)
	R^2	0.0772	0.0610	0.0835	0.0019	0.0483
	ρ	0.3278** (0.003)	0.3278** (0.003)	0.2989** (0.007)	-0.0741 (0.726)	0.3278** (0.003)
No. of significant β coefficients		8	11	12	4	12
Average R^2		0.1946	0.1457	0.1779	0.0340	0.1613
No. of significant ρ coefficients		12	12	11	10	12

Note: values in parenthesis are p values for 1 tail when the alternative hypothesis is $p > t$. ** represents significant at 1%; * represents significant at 5%. Each regression and correlation results are based on sample size of 68.

Table D.4 (Corresponding to Table 5.3 in Chapter 5)
Regression and Correlation Results for Heckscher-Ohlin Theory: Sector Specific Samples

ISIC Codes	Regression / Correlation Results	Balassa	RSCA	ARCA	NRCA	Log of Balassa
321	β	0.2356** (<0.001)	0.0020 (0.259)	0.0059** (0.001)	0.0001 (0.259)	0.0153* (0.021)
	R^2	0.4434	0.0106	0.3827	0.0189	0.1000
	ρ	0.741** (<0.001)	0.741** (<0.001)	0.7449** (<0.001)	0.5839** (<0.001)	0.741** (<0.001)
322	β	0.6328** (<0.001)	0.0053 (0.265)	0.0014* (0.014)	0.00004 (0.239)	0.0526** (<0.001)
	R^2	0.3363	0.0521	0.0752	0.0168	0.1669
	ρ	0.6919** (<0.001)	0.6919** (<0.001)	0.6419** (<0.001)	0.5541** (<0.001)	0.6919** (<0.001)
323	β	-0.0395 (0.607)	-0.0185 (0.985)	-0.00004 (0.570)	-0.000002 (0.539)	-0.0445 (0.973)
	R^2	0.0206	0.0049	0.0034	0.0001	0.0073
	ρ	0.3811** (0.001)	0.3811** (0.001)	0.2969** (0.007)	0.3289** (0.003)	0.3811** (0.001)
351	β	0.0051 (0.803)	0.0046 (0.861)	0.0009 (0.828)	-0.00004 (0.318)	0.0101 (0.870)
	R^2	0.0641	0.0207	0.0623	0.0018	0.0174
	ρ	-0.3566** (0.001)	-0.3566** (0.001)	-0.3641** (0.001)	0.0291 (0.593)	-0.3566** (0.001)
383	β	-0.0034 (0.122)	-0.0055** (<0.001)	-0.0042** (0.006)	-0.0001 (0.272)	-0.0719** (<0.001)
	R^2	0.0528	0.2885	0.3341	0.0246	0.4149
	ρ	-0.7805** (<0.001)	-0.7805** (<0.001)	-0.8036** (<0.001)	-0.2705* (0.013)	-0.7805** (<0.001)
No. of significant β coefficients		2	1	3	0	3
Average R^2		0.1834	0.0754	0.1715	0.0124	0.1413
No. of significant ρ coefficients		5	5	5	4	5

Note: values in parenthesis are p values for 1 tail when the alternative hypothesis is $p > t$. ** represents significant at 1%; * represents significant at 5%. Each regression and correlation results are based on sample size of 68.

Table D.5
Regression and Correlation Results for Ricardian Theory including Implausible Observations:
Combined Sample

RCA Index	Regression results with country FE and year FE		Regression results with country- year FE		Rank correlation results
	β	R ²	β	R ²	
Balassa	0.6142** (<0.001)	0.1399	0.6253** (<0.001)	0.1453	0.3681** (<0.001)
RSCA	0.1015** (<0.001)	0.1391	0.1021** (<0.001)	0.1421	0.3681** (<0.001)
ARCA	0.0123** (<0.001)	0.1110	0.0128** (<0.001)	0.1116	0.3930** (<0.001)
NRCA	0.0002** (0.002)	0.0070	0.0002** (0.002)	0.0071	0.2656** (<0.001)
Log of Balassa	0.2970** (<0.001)	0.1615	0.2997** (<0.001)	0.1656	0.3681** (<0.001)

Note: values in parenthesis are p values for 1 tail when the alternative hypothesis is $p > t$. ** represents significant at 1% level. Each regression and correlation results are based on a sample size of 3094. Apart from the countries notified below Tables D.1 and D.2, the sample includes additionally observations on Austria, Bangladesh, Canada, Costa Rica, Denmark, Indonesia, India, Italy, Mexico, Malaysia, Norway, Pakistan, Philippines, Sweden, Thailand, Taiwan and Venezuela.

Table D.6
Regression and Correlation Results for Ricardian Theory including Implausible Observations:
Sector Specific Samples

ISIC Codes	Regression/Correlation Results	Balassa	RSCA	ARCA	NRCA	Log of Balassa
311	β	0.7504 (0.190)	0.1663* (0.032)	0.0153 (0.166)	0.0006* (0.021)	0.3943* (0.027)
	R ²	0.0285	0.0121	0.0004	0.0011	0.0154
	ρ	-0.0549 (0.769)	-0.0549 (0.769)	-0.0832 (0.868)	0.0347 (0.321)	-0.0549 (0.769)
313	β	0.2207 (0.156)	0.0434* (0.031)	0.0010 (0.067)	<0.0001 (0.290)	0.1348* (0.027)
	R ²	0.0724	0.0264	0.0335	0.0001	0.0391
	ρ	-0.0307 (0.660)	-0.0307 (0.660)	0.0141 (0.425)	-0.0905 (0.888)	-0.0307 (0.660)
321	β	0.3566 (0.103)	0.0238 (0.305)	0.0131 (0.130)	0.0003 (0.240)	0.1202 (0.145)
	R ²	0.0083	0.0077	0.0039	0.0017	0.0125
	ρ	0.0348 (0.320)	0.0348 (0.320)	0.0359 (0.315)	0.0887 (0.117)	0.0348 (0.320)
322	β	0.8420 (0.074)	0.0374 (0.197)	0.0177** (0.007)	-0.0003 (0.900)	0.1725 (0.122)
	R ²	0.0240	0.0028	0.0214	0.0084	0.0127
	ρ	0.0388 (0.302)	0.0388 (0.302)	0.0288 (0.350)	0.0757 (0.155)	0.0388 (0.302)
323	β	1.0559** (<0.001)	0.1553* (0.017)	0.0072** (<0.001)	<-0.0001 (0.505)	0.5003** (0.008)
	R ²	0.2050	0.1180	0.2390	<0.0001	0.1462
	ρ	0.3323** (<0.001)	0.3323** (<0.001)	0.3523** (<0.001)	0.1939** (0.004)	0.3323** (<0.001)
331	β	1.5465**	0.4124**	0.0305**	0.0008	1.1634**

		(0.002)	(<0.001)	(<0.001)	(0.120)	(<0.001)
	R ²	0.0726	0.0568	0.0759	0.0045	0.0864
	ρ	0.2583** (0.0002)	0.2583** (0.0002)	0.2392** (0.001)	0.1370* (0.033)	0.2583** (0.0002)
341	β	0.3260** (<0.001)	0.1298** (<0.001)	0.0193** (<0.001)	0.0003** (<0.001)	0.2855** (<0.001)
	R ²	0.0466	0.0691	0.0439	0.0011	0.0667
	ρ	0.5190** (<0.001)	0.5190** (<0.001)	0.4616** (<0.001)	0.3385** (<0.001)	0.5190** (<0.001)
351	β	0.4914** (<0.001)	0.3198** (<0.001)	0.0693** (<0.001)	0.0010 (0.059)	0.6873** (<0.001)
	R ²	0.2532	0.2649	0.2557	0.0727	0.2648
	ρ	0.6028** (<0.001)	0.6028** (<0.001)	0.6035** (<0.001)	0.4306** (<0.001)	0.6028** (<0.001)
353	β	0.6835** (0.003)	0.0787* (0.014)	0.0125** (0.006)	0.0003* (0.047)	0.2240* (0.010)
	R ²	0.0809	0.2007	0.0730	0.0459	0.1946
	ρ	0.5742** (<0.001)	0.5742** (<0.001)	0.5476** (<0.001)	0.5393** (<0.001)	0.5742** (<0.001)
361	β	0.2959** (0.002)	0.1742** (0.001)	0.0069** (0.001)	0.0004* (0.044)	0.3705** (<0.001)
	R ²	0.0584	0.0722	0.0555	0.0250	0.0729
	ρ	0.2808** (0.0001)	0.2808** (0.0001)	0.2788** (0.0001)	0.2921** (0.0001)	0.2808** (0.0001)
371	β	0.2680* (0.014)	0.1387** (0.007)	0.0149* (0.021)	0.0003 (0.135)	0.3384** (<0.008)
	R ²	0.0030	0.0043	0.0028	0.0027	0.0070
	ρ	0.3146** (<0.001)	0.3146** (<0.001)	0.3025** (<0.001)	0.2437** (0.001)	0.3146** (<0.001)
372	β	0.7478** (<0.001)	0.1462** (<0.001)	0.0231** (<0.001)	0.0002 (0.091)	0.4462** (<0.001)
	R ²	0.4375	0.3455	0.4314	0.0781	0.3405
	ρ	0.5420** (<0.001)	0.5420** (<0.001)	0.5407** (<0.001)	0.5384** (<0.001)	0.5420** (<0.001)
381	β	0.4380** (<0.001)	0.2681** (<0.001)	0.0204** (<0.001)	0.0006 (0.115)	0.5829** (<0.001)
	R ²	0.0871	0.1413	0.0705	0.0159	0.1381
	ρ	0.4390** (<0.001)	0.4390** (<0.001)	0.4078** (<0.001)	0.2618** (0.0002)	0.4390** (<0.001)
382	β	0.0889** (<0.001)	0.0372** (<0.001)	0.0138** (<0.001)	0.0003** (<0.001)	0.0894** (0.002)
	R ²	0.0798	0.0313	0.0750	0.0008	0.0236
	ρ	0.6151** (<0.001)	0.6151** (<0.001)	0.5788** (<0.001)	0.0795 (0.143)	0.6151** (<0.001)
383	β	0.0081 (0.470)	0.0286 (0.309)	0.0530* (0.010)	0.0020 (0.209)	0.1497 (0.205)
	R ²	0.0000	0.0003	0.0136	0.0001	0.0071
	ρ	0.1468* (0.024)	0.1468* (0.024)	0.1469* (0.024)	-0.0507 (0.752)	0.1468* (0.024)
384	β	0.0820** (<0.001)	0.0944** (<0.001)	0.0074* (0.016)	0.0002 (0.238)	0.8304** (<0.001)
	R ²	0.0061	0.0157	0.0153	0.0001	0.0337
	ρ	0.2702** (0.0001)	0.2702** (0.0001)	0.2222** (0.001)	-0.0212 (0.612)	0.2702** (0.0001)

390	β	0.6054* (0.018)	0.2052** (<0.001)	0.0180* (0.027)	0.00003 (0.371)	0.6480** (<0.001)
	R^2	0.0859	0.0573	0.0862	0.0036	0.0761
	ρ	0.1989** (0.004)	0.1989** (0.004)	0.1772** (0.008)	-0.0133 (0.571)	0.1989** (0.004)
No. of significant β coefficients		12	14	14	5	14
Average R^2		0.0911	0.0839	0.0881	0.0154	0.0904
No. of significant ρ coefficients		13	13	13	9	13

Note: values in parenthesis are p values for 1 tail when the alternative hypothesis is $p > t$. ** represents significant at 1%; * represents significant at 5%. Each regression and correlation results are based on sample size of 182. Apart from the countries notified below Tables D.1 and D.2, the sample includes additionally observations on Austria, Bangladesh, Canada, Costa Rica, Denmark, Indonesia, India, Italy, Mexico, Malaysia, Norway, Pakistan, Philippines, Sweden, Thailand, Taiwan and Venezuela.

Table D.7
Regression and Correlation Results for Heckscher-Ohlin Theory including Implausible
Observations: Sector Specific Samples

ISIC Codes	Regression/ Correlation Results	Balassa	RSCA	ARCA	NRCA	Log of Balassa
321	β	0.0808 (0.093)	0.0029 (0.296)	0.0016 (0.219)	-0.0001 (0.823)	0.0083 (0.295)
	R^2	0.1659	0.0383	0.0975	0.0118	0.0455
	ρ	0.5801** (<0.001)	0.5801** (<0.001)	0.5834** (<0.001)	0.4877** (<0.001)	0.5801** (<0.001)
322	β	0.3128* (0.023)	0.0123* (0.037)	0.0045* (0.013)	-0.00004 (0.906)	0.0310 (0.059)
	R^2	0.413	0.1713	0.3905	0.0381	0.1576
	ρ	0.6437** (<0.001)	0.6437** (<0.001)	0.6434** (<0.001)	0.5192** (<0.001)	0.6437** (<0.001)
323	β	-0.0534 (0.806)	0.0019 (0.390)	0.00002 (0.468)	0.1800 (0.067)	0.0015 (0.467)
	R^2	0.0146	0.0235	0.0057	0.0108	0.0088
	ρ	0.3996** (<0.001)	0.3996** (<0.001)	0.3915** (<0.001)	0.4593** (<0.001)	0.3996** (<0.001)
351	β	-0.0064 (0.110)	-0.0020 (0.230)	-0.0010 (0.103)	<-0.0001 (0.454)	-0.0041 (0.240)
	R^2	0.0525	0.0601	0.0555	<0.0001	0.0574
	ρ	-0.1246* (0.049)	-0.1246* (0.049)	-0.1209 (0.054)	0.1751 (0.990)	-0.1246* (0.049)
383	β	-0.0072 (0.067)	-0.0020 (0.228)	-0.0001 (0.456)	0.00003 (0.665)	-0.0078 (0.190)
	R^2	0.0464	0.0288	0.0143	0.0021	0.0362
	ρ	-0.5035** (<0.001)	-0.5035** (<0.001)	-0.4933** (<0.001)	-0.1151 (0.064)	-0.5035** (<0.001)
No. of significant β coefficients		1	1	1	0	0
Average R^2		0.13848	0.0644	0.1127	0.01256	0.0611
No. of significant ρ coefficients		5	5	4	3	5

Note: values in parenthesis are p values for 1 tail when the alternative hypothesis is $p > t$. ** represents significant at 1%; * represents significant at 5%. Each regression and correlation results are based on sample size of 177. Apart from the countries notified below Tables D.1 and D.2, the sample includes additionally observations on Austria, Bangladesh, Canada, Costa Rica, Denmark, Indonesia, India, Italy, Mexico, Malaysia, Norway, Pakistan, Philippines, Sweden, Thailand, and Venezuela.

Table D.8 (Corresponding to Table 5.5 in Chapter 5)
Results for Wilcoxon's Signed Rank Test for Countries

Country	Year combinations	Balassa	RSCA	ARCA	NRCA	Log of Balassa
Argentina	1993-1994	-0.450 (0.653)	-1.349 (0.177)	-0.876 (0.381)	-0.686 (0.493)	-1.018 (0.309)
	1993-1995	-0.781 (0.435)	-1.681 (0.093)	-0.402 (0.687)	-0.686 (0.493)	-1.349 (0.177)
	1994-1995	-1.065 (0.287)	-1.254 (0.210)	-0.402 (0.687)	-1.112 (0.266)	-1.16 (0.246)
Accepted cases for null hyp.		3	3	3	3	3
Japan	1993-1994	-0.497 (0.619)	-0.402 (0.687)	0.213 (0.831)	-0.97 (0.332)	-0.26 (0.795)
	1993-1995	0.071 (0.943)	0.828 (0.407)	-0.118 (0.906)	-1.065 (0.287)	0.923 (0.356)
	1994-1995	0.26 (0.795)	0.544 (0.586)	-0.876 (0.381)	-0.923 (0.356)	0.592 (0.554)
Accepted cases for null hyp.		3	3	3	3	3
Korea	1993-1994	1.396 (0.163)	1.065 (0.287)	1.349 (0.177)	1.207 (0.227)	1.065 (0.287)
	1993-1995	1.775 (0.076)	1.207 (0.227)	1.018 (0.309)	1.349 (0.177)	0.876 (0.381)
	1994-1995	1.112 (0.266)	0.781 (0.435)	0.402 (0.687)	1.112 (0.266)	0.639 (0.523)
Accepted cases for null hyp.		3	3	3	3	3
Portugal	1993-1994	0.402 (0.687)	0.26 (0.795)	0.26 (0.795)	0.213 (0.831)	0.402 (0.687)
	1993-1995	1.302 (0.193)	0.923 (0.356)	0.923 (0.356)	0.308 (0.758)	1.065 (0.287)
	1994-1995	1.207 (0.227)	0.497 (0.619)	1.065 (0.287)	1.112 (0.266)	0.639 (0.523)
Accepted cases for null hyp.		3	3	3	3	3
Turkey	1993-1994	0.213 (0.831)	-1.16 (0.246)	-0.402 (0.687)	0.118 (0.906)	-0.876 (0.381)
	1993-1995	-1.775 (0.076)	-2.864** (0.004)	-0.639 (0.523)	-0.592 (0.554)	-2.864** (0.004)
	1994-1995	-2.627** (0.009)	-2.343* (0.019)	-1.065 (0.287)	-1.681 (0.093)	-2.249* (0.025)
Accepted cases for null hyp.		2	1	3	3	1
United States	1993-1994	1.254 (0.210)	1.16 (0.246)	0.828 (0.407)	0.686 (0.493)	1.16 (0.246)
	1993-1995	0.26 (0.795)	0.118 (0.906)	-0.118 (0.906)	-0.166 (0.868)	0.118 (0.906)
	1994-1995	-0.97 (0.332)	-1.207 (0.227)	-1.065 (0.287)	-0.923 (0.356)	-1.207 (0.227)
Accepted cases for null hyp.		3	3	3	3	3
Total number of accepted cases		17	16	18	18	16

Note: For each index only the standardized normal approximation of the test statistic and the corresponding p value is reported. . ** represents significant at 1%; * represents significant at 5%. For each country, 17 sectoral observations are available in each year.

Table D.9 (Corresponding to Table 5.5.1 in Chapter 5)
Results for Wilcoxon's Signed Rank Test for Sectors

ISIC codes	Year Combinations	Balassa	RSCA	ARCA	NRCA	Log of Balassa
321	1993-1994	1.572 (0.116)	0.943 (0.345)	1.572 (0.116)	-0.943 (0.345)	0.943 (0.345)
	1993-1995	1.992* (0.046)	1.992* (0.046)	1.782 (0.075)	0.314 (0.753)	1.992* (0.046)
	1994-1995	1.153 (0.249)	1.572 (0.116)	1.572 (0.116)	0.524 (0.600)	1.572 (0.116)
Accepted cases for null hyp.		2	2	3	3	2
322	1993-1994	0.943 (0.345)	0.105 (0.917)	0.943 (0.345)	-0.314 (0.753)	0.524 (0.600)
	1993-1995	0.943 (0.345)	0.105 (0.917)	0.943 (0.345)	0.105 (0.917)	0.314 (0.753)
	1994-1995	-0.105 (0.917)	-0.314 (0.753)	-0.105 (0.917)	-0.314 (0.753)	-0.105 (0.917)
Accepted cases for null hyp.		3	3	3	3	3
323	1993-1994	0.734 (0.463)	1.572 (0.116)	0.734 (0.463)	-0.314 (0.753)	1.572 (0.116)
	1993-1995	1.572 (0.116)	1.153 (0.249)	1.153 (0.249)	0.105 (0.917)	1.782 (0.075)
	1994-1995	0.314 (0.753)	0.524 (0.600)	-0.105 (0.917)	-0.524 (0.600)	0.524 (0.600)
Accepted cases for null hyp.		3	3	3	3	3
351	1993-1994	-1.363 (0.173)	-1.572 (0.116)	-1.153 (0.249)	0.314 (0.753)	-1.572 (0.116)
	1993-1995	-0.734 (0.463)	-0.943 (0.345)	-0.314 (0.753)	0.314 (0.753)	-0.943 (0.345)
	1994-1995	-0.314 (0.753)	-0.314 (0.753)	0.105 (0.917)	-0.314 (0.753)	-0.314 (0.753)
Accepted cases for null hyp.		3	3	3	3	3
383	1993-1994	-0.105 (0.917)	-0.105 (0.917)	0.524 (0.600)	0.943 (0.345)	-0.105 (0.917)
	1993-1995	0.105 (0.917)	-0.314 (0.753)	0.943 (0.345)	0.943 (0.345)	-0.524 (0.600)
	1994-1995	0.524 (0.600)	0.314 (0.753)	0.943 (0.345)	0.734 (0.463)	0.314 (0.753)
Accepted cases for null hyp.		3	3	3	3	3
Total no. of accepted cases for null hypothesis		14	14	15	15	14

Note: For each index only the standardized normal approximation of the test statistic and the corresponding p value is reported. . ** represents significant at 1%; * represents significant at 5%. For each sector 6 country observations are available. Since Stata provides z statistic, the question might arise whether normal approximation of the test statistic could be used for small sample sizes. Wilcoxon (1947) tabulated the z statistic starting with as small as 6 pairs of observations in case of signed rank test and 5 observations per sample in case of rank sum test, after analyzing the asymptotic properties of the test statistic. Further, Mann and Whitney (1947) with independent samples of unequal sizes found that from 8 observations per sample, the distribution of the test statistic differs only by a negligible amount from that of a normal distribution. Bellera (2010) has in fact found support for their claim by graphically superimposing normal distribution curves over the exact distributions of Wilcoxon's signed rank and rank sum test statistics. Hence, usage of normal approximation of the Wilcoxon's test statistics for small sample sizes is justified in the literature.

Table D.10 (Corresponding to Table 5.6 in Chapter 5)
Results for Wilcoxon's Rank Sum Test for Sectors

Sector combinations	Balassa	RSCA	ARCA	NRCA	Log of Balassa
321-322	-0.525 (0.600)	-0.525 (0.600)	0.525 (0.600)	0.21 (0.834)	-0.525 (0.600)
321-323	0.315 (0.753)	0.315 (0.753)	1.26 (0.208)	0.63 (0.529)	0.315 (0.753)
321-351	1.89 (0.059)	1.89 (0.059)	2.626** (0.009)	1.155 (0.248)	1.89 (0.059)
321-383	2.626** (0.009)	2.626** (0.009)	2.310* (0.021)	0.630 (0.527)	2.626** (0.009)
322-323	0.21 (0.834)	0.21 (0.834)	0.84 (0.401)	0.525 (0.600)	0.21 (0.834)
322-351	1.68 (0.093)	1.68 (0.093)	2.521* (0.012)	1.155 (0.248)	1.68 (0.093)
322-383	2.310* (0.021)	2.310* (0.021)	2.415* (0.016)	0.840 (0.401)	2.310* (0.021)
323-351	1.155 (0.248)	1.155 (0.248)	2.31* (0.021)	1.575 (0.115)	1.155 (0.248)
323-383	2.100* (0.036)	2.100* (0.036)	1.890 (0.059)	1.260 (0.208)	2.100* (0.036)
351-383	1.155 (0.248)	1.155 (0.248)	1.260 (0.208)	0.630 (0.529)	1.155 (0.248)
No. of Accepted Cases for Null Hypothesis	7	7	5	10	7

Note: For each index only the standardized normal approximation of the test statistic and the corresponding p value is reported. For each sector, observations on 8 countries are available.

Table D.11 (Corresponding to Table 5.7 in Chapter 5)
Results for Wilcoxon's Rank Sum Test for Countries

Country combinations	Balassa	RSCA	ARCA	NRCA	Log of Balassa
Argentina-Japan	1.671 (0.095)	1.671 (0.095)	0.637 (0.524)	1.292 (0.197)	1.671 (0.095)
Argentina-Korea	0.603 (0.547)	0.603 (0.547)	0.534 (0.593)	0.741 (0.459)	0.603 (0.547)
Argentina-Portugal	-0.017 (0.986)	-0.017 (0.986)	-0.327 (0.744)	-0.362 (0.718)	-0.017 (0.986)
Argentina-Turkey	0.121 (0.904)	0.121 (0.904)	0.121 (0.904)	0.086 (0.931)	0.121 (0.904)
Argentina-United States	0.465 (0.642)	0.465 (0.642)	0.327 (0.744)	0.155 (0.877)	0.465 (0.642)
Japan-Korea	-1.257 (0.209)	-1.257 (0.209)	0.155 (0.877)	-0.603 (0.547)	-1.257 (0.209)
Japan-Portugal	-1.843 (0.065)	-1.843 (0.065)	-0.465 (0.642)	-1.223 (0.221)	-1.843 (0.065)
Japan-Turkey	-1.567 (0.117)	-1.567 (0.117)	-0.086 (0.931)	-1.119 (0.263)	-1.567 (0.117)
Japan-United States	-1.533 (0.125)	-1.533 (0.125)	-0.775 (0.438)	-0.224 (0.823)	-1.533 (0.125)
Korea-Portugal	-0.741 (0.459)	-0.741 (0.459)	-0.293 (0.770)	-0.603 (0.547)	-0.741 (0.459)
Korea-Turkey	-0.293 (0.770)	-0.293 (0.770)	0.052 (0.959)	-0.362 (0.718)	-0.293 (0.770)
Korea-United States	-0.121 (0.904)	-0.121 (0.904)	-0.534 (0.593)	-0.086 (0.931)	-0.121 (0.904)
Portugal-Turkey	0.603 (0.547)	0.603 (0.547)	0.465 (0.642)	0.431 (0.667)	0.603 (0.547)
Portugal-United States	0.741 (0.459)	0.741 (0.459)	0.534 (0.593)	0.052 (0.959)	0.741 (0.459)
Turkey-United States	0.396 (0.692)	0.396 (0.692)	0.121 (0.904)	0.121 (0.904)	0.396 (0.692)
No. of Accepted Cases for Null Hypothesis	15	15	15	15	15

Note: For each index only the standardized normal approximation of the test statistic and the corresponding p value is reported. For each country, observations on 17 sectors are available.

Appendix E

Table E.1
Sector Description for Chapter 6

Sector Code	Sector Description
1	Agriculture, Hunting, Forestry and Fishing
2	Mining and Quarrying
3	Food, Beverages and Tobacco
4	Textiles and Textile Products
5	Leather, Leather and Footwear
6	Wood and products of Wood and Cork (excluding furniture)
7	Pulp, Paper, Printing and Publishing
8	Coke, Refined Petroleum and Nuclear Fuel
9	Chemicals and Chemical Products
10	Rubber and Plastics
11	Other Non-Metallic Minerals such as production of glass and glass products, ceramic ware, cement, lime and plaster, articles of asbestos, asphalt etc.
12	Basic Metals and Fabricated Metals (excludes machinery and equipment). Basic metal includes manufacture of basic iron and steel and their castings; manufacture of non-ferrous metals such as copper, lead, aluminium, zinc etc. and their castings. Fabricated metal includes manufacture of structural metal products such as bridges, parts of bridge, towers, columns, metal doors, windows and their frames made of iron, steel or aluminium; manufacture of tanks, reservoirs, metal containers; treatment and coating of metals such as plating, polishing, anodizing, colouring etc.; manufacture of cutlery, hand tools and general hardware.
13	Machinery (not elsewhere classified)
14	Electrical and Optical Equipments (includes office machinery and computers, radio, television, communication equipments, watches, clocks, medical instruments, precision and optical equipments)
15	Transport equipments
16	Manufacturing (not elsewhere classified); Recycling. Manufactures under this category includes furniture, jewellery, musical instruments, sports goods, toys etc.

Table E.2
Summary Statistics of Variables used in Chapter 6

Variable	Year	No. of Observations	Mean	Median	Standard deviation	Skewness	Kurtosis
Export	2008	576	1483.088	63.66238	5825.27	8.563892	98.00806
Value added	2008	576	22057.89	5040.204	81427.3	14.33511	261.5927
Output	2008	576	68026.79	17062.27	177737.5	8.021941	95.70271
Domestic output from 16 sectors used by sector 1 within country	2008	576	1444.16	54.88936	8334.16	11.17049	140.2153
Domestic output from 16 sectors used by sector 2 within country	2008	576	1572.58	12.78163	25285.83	23.62093	563.7266
Domestic output from 16 sectors used by sector 3 within country	2008	576	3202.223	121.3481	17458.98	11.34952	151.3615
Domestic output from 16 sectors used by sector 4 within country	2008	576	1117.055	17.62849	11495.65	20.44105	454.8102
Domestic output from 16 sectors used by sector 5 within country	2008	576	242.6112	3.319728	1796.612	14.89544	265.7709
Domestic output from 16 sectors used by sector 6 within country	2008	576	427.9082	18.05634	2975.479	17.74393	365.0117

country							
Domestic output from 16 sectors used by sector 7 within country	2008	576	778.4333	42.19066	4380.334	12.43192	178.9175
Domestic output from 16 sectors used by sector 8 within country	2008	576	892.045	12.22385	6583.626	13.46173	210.8734
Domestic output from 16 sectors used by sector 9 within country	2008	576	1674.646	69.79675	10067.7	14.41114	256.7745
Domestic output from 16 sectors used by sector 10 within country	2008	576	769.5028	38.85087	5305.241	14.29617	227.9021
Domestic output from 16 sectors used by sector 11 within country	2008	576	632.8663	35.55074	3542.874	12.23107	173.8606
Domestic output from 16 sectors used by sector 12 within country	2008	576	3051.991	101.7074	22128.48	14.67576	249.5597
Domestic output from 16 sectors used by sector 13	2008	576	1500.437	44.99556	8751.975	12.32666	178.9595

within country							
Domestic output from 16 sectors used by sector 14 within country	2008	576	2335.031	60.14509	18551.39	16.68922	318.9209
Domestic output from 16 sectors used by sector 15 within country	2008	576	2258.907	68.96558	13262.22	13.00188	202.5341
Domestic output from 16 sectors used by sector 16 within country	2008	576	356.1887	42.00872	928.8883	5.441041	42.58666
Export	2011	576	1533.226	67.20093	6305.126	9.978387	137.1184
Value added	2011	576	27087.4	5188.843	103595	12.93996	217.5205
Output	2011	576	83400.4	17522.33	241348.1	7.613154	80.30954
Domestic output from 16 sectors used by sector 1 within country	2011	576	1944.247	55.41109	11988.67	11.19395	138.7232
Domestic output from 16 sectors used by sector 2 within country	2011	576	2269.837	12.2451	35749.2	23.55349	561.528
Domestic output from 16 sectors used by sector 3 within country	2011	576	4126.167	113.9459	25229.85	11.99178	165.3778
Domestic	2011	576	1570.977	18.97417	18264.14	21.34506	485.5525

output from 16 sectors used by sector 4 within country							
Domestic output from 16 sectors used by sector 5 within country	2011	576	336.3285	3.03777	2841.271	16.08681	302.4333
Domestic output from 16 sectors used by sector 6 within country	2011	576	558.9527	16.24801	4478.053	18.40414	384.94
Domestic output from 16 sectors used by sector 7 within country	2011	576	943.1197	40.9554	6050.936	14.20356	229.497
Domestic output from 16 sectors used by sector 8 within country	2011	576	1192.22	11.68527	9858.13	14.92417	252.958
Domestic output from 16 sectors used by sector 9 within country	2011	576	2219.707	69.95955	15119.62	15.69592	297.8057
Domestic output from 16 sectors used by sector 10 within country	2011	576	1072.302	36.96245	8191.388	14.68392	238.2313

Domestic output from 16 sectors used by sector 11 within country	2011	576	839.5312	30.19623	5399.432	11.91343	159.2766
Domestic output from 16 sectors used by sector 12 within country	2011	576	4044.592	97.74553	31210.89	16.18742	310.0653
Domestic output from 16 sectors used by sector 13 within country	2011	576	1896.383	45.12793	12514.38	13.93057	221.0932
Domestic output from 16 sectors used by sector 14 within country	2011	576	3432.868	58.36692	31047.44	17.88202	360.1817
Domestic output from 16 sectors used by sector 15 within country	2011	576	2780.711	69.84143	15863.33	11.54641	157.6318
Domestic output from 16 sectors used by sector 16 within country	2011	576	416.4806	39.89597	1194.018	6.66546	65.3245

Note: Database provides data on indicators for millions of US dollar. 36 countries included for each year – Australia, Austria, Belgium, Bulgaria, Brazil, Canada, China, Czech Republic, Germany, Denmark, Spain, Estonia, Finland, France, United Kingdom, Greece, Hungary, Indonesia, India, Ireland, Italy, Japan, Korea, Lithuania, Mexico, Malta, Netherlands, Poland, Portugal, Romania, Russian Federation, Slovakia, Slovenia, Turkey, Taiwan and Rest of the World (excluding Cyprus, Latvia, Luxembourg, Sweden and USA which had to be deleted at various stages of calculation).

References

- Alma O.G. (2011), “Comparison of Robust Regression Methods in Linear Regression”, *International Journal of Contemporary Math. Sciences*, Vol. 6, No. 9, pp. 409-421.
- Anderson, J. and P. Neary (1994), “Measuring the restrictiveness of trade policy”, *World Bank Economic Review*, Vol. 8, No. 2, pp. 151–169.
- Anderson, J. and P. Neary (1996), “A new approach to evaluating trade policy”, *Review of Economic Studies*, Vol. 63, No. 1, pp. 107–125.
- Anderson, J. and P. Neary (2003), “The Mercantilist index of trade policy”, *International Economic Review*, Vol. 44, No. 2, pp. 627–649.
- Anderson, J. and P. Neary (2005), *Measuring the Restrictiveness of Trade Policy*, MIT Press, Boston.
- Ando, M. (2006), “Fragmentation and vertical intra-industry trade in East Asia”, *North American Journal of Economics and Finance*, Vol. 17, pp. 257-281.
- Andriamananjara, S., J.M. Dean, M.J. Ferrantino, R.M. Feinberg, R.T. Ludema and M.E. Tsigas (2004), “The Effects of Non-Tariff Measures on Prices, Trade, and Welfare: CGE Implementation of Policy-Based Price Comparisons”, Working Paper No. 2004-04-A, U.S. International Trade Commission, Office of Economics.
- Arestoff, F. and C. Hurlin (2008), “Estimates of Government Net Capital Stocks for 26 Developing Countries, 1970-2002: Are Public Investment Efficient in Creating Capital Stocks?”, Document de Recherche n° 2008-07, Laboratoire d'Economie d'Orléans,.
- Arndt, S.W. (1997), “Globalization and the Open Economy”, *North American Journal of Economics and Finance*, Vol. 81, No. 1, pp. 71-79.
- Balassa, B. (1963), “An Empirical Demonstration of Classical Comparative Cost Theory”, *The Review of Economics and Statistics*, Vol. 45, No. 3, pp. 231-238.

- Balassa, B. (1965), "Trade Liberalization and 'Revealed' Comparative Advantage", *Manchester School of Economics and Social Studies*, Vol. 33, No. 2, pp. 99-123.
- Balassa, B. (1977), "Revealed Comparative Advantage Revisited: An Analysis of Relative Export Shares of the Industrial Countries 1953-1971", *Manchester School of Economics and Social Studies*, Vol. 45, No. 4, pp. 327-344.
- Ballance, R.H., H. Forstner and T. Murray (1987), "Consistency Tests of Alternative Measures of Comparative Advantage", *The Review of Economics and Statistics*, Vol. 69, No. 1, pp. 157-161.
- Bardhan, P. (2010), *Awakening Giants: Feet of Clay*, Oxford University Press, New Delhi.
- Batra, A. and Z. Khan (2005), "Revealed Comparative Advantage: An Analysis for India and China", Working Paper No. 168, Indian Council for Research on International Economic Relations (ICRIER), New Delhi.
- Baum, C.F. (2006), *An Introduction to Modern Econometrics Using Stata*, Stata Press, Texas.
- Bellera, C.A. (2010), "Normal Approximations to the Distributions of the Wilcoxon Statistic: Accurate to What N? Graphical Insights", *Journal of Statistics Education*, Vol. 18, No. 2, pp. 1-17.
- Bender, S. and K. Li (2002), "The Changing Trade and Revealed Comparative Advantages of Asian and Latin American Manufacture Exports", Discussion Paper Series No. 843, Economic Growth Center, Yale University.
- Benedictis, L.D. and M. Tamberi, (2001), "A Note on Balassa Index of Revealed Comparative Advantage", Working Paper No. 158, Università de Macerata and Università de Ancona, Italy. Accessed from <http://docs.dises.univpm.it/web/quaderni/pdf/158.pdf> on 20.8.2012.
- Benedictis, L.D. and M. Tamberi (2004), "Overall Specialization Empirics: Techniques and Applications", *Open Economies Review*, Vol. 15, No. 1, pp. 323-346.

- Blanes, J.V. and C. Martin (2000), “The Nature and Causes of Intra-Industry Trade: Back to the Comparative Advantage Explanation? The Case of Spain”, *Weltwirtschaftliches Archiv*, Vol. 136, No. 3, pp. 423-441.
- Bosworth, B. and S.M. Collins (2007), “Accounting for Growth: Comparing China and India”, *Journal of Economic Perspectives*, Vol. 22, No. 1, pp. 45-66.
- Bowen, H. P. (1983), “On the Theoretical Interpretation of Indices of Trade Intensity and Revealed Comparative Advantage”, *Weltwirtschaftliches Archiv*, Vol. 119, pp. 464-472.
- Bowen, H., E. Leamer and L. Sveikauskas (1987), “Multicountry, Multifactor Tests of the Factor Abundance Theory”, *American Economic Review*, Vol. 77, No. 5, pp. 791-809.
- Bradford, S. (2003), “Paying the Price: Final Goods Protection in OECD Countries”, *Review of Economic and Statistics*, Vol. 85, No. 1, pp. 24–37.
- Branstetter, L. and N. Lardy (2008), “China’s Embrace of Globalization”, in Brandt, L. and T.G. Rawski (ed.), *China’s Great Economic Transformation*, Cambridge University Press, New York.
- Brasili, A., P. Epifani and R. Helg (2000), “On the Dynamics of Trade Patterns”, *De Economist*, Vol. 148, No. 2, pp. 233-257.
- Breusch, T.S. and A.R. Pagan (1979), “A Simple Test for Heteroscedasticity and Random Coefficient Variation”, *Econometrica*, Vol. 47, No. 5, pp. 1287-1294.
- Burange, L. and S. Chaddha (2008), “India’s Revealed Comparative Advantage in Merchandise Trade”, Working Paper no. UDE28/6/2008, Department of Economics, University of Mumbai.
- Campbell, M.J. (2006), “Teaching Non-parametric Statistics to Students in Health Sciences”, ICOTS 7, IASE, ISI, Salvador, Bahia, Brazil. Accessed from www.stat.auckland.ac.nz/~iase/publications/17/3F3_CAMP.pdf on 19.12.2014.

- Cantwell, J. (1989), *Technological Innovation and Multinational Corporations*, Basil Blackwell, Oxford.
- Chatterjee, S., A. Hadi and B. Price (2000), *Regression Analysis by Example*, Wiley, New York.
- Chaudhuri, S. and M. Ravallion (2006), “Partially Awakened Giants: Uneven Growth in China and India”, in Winters, L.A. and S. Yusuf (ed.) *Dancing with Giants: China, India and Global Economy*, World Bank and Institute of Policy Studies, Washington DC and Singapore.
- Chen, K. and Y. Xie (2007), “Performances of Tests for Two Sample Location Problem”, Accessed from www.stat.uiowa.edu/~kcowles/sl66_2007/chenxieproject.pdf on 19.12.2014.
- Cobbold, T. (2003), “A Comparison of Gross Output and Value-Added Methods of Productivity Estimation”, Productivity Commission Research Memorandum, Canberra.
- Conover, W.J. (1999), *Practical Nonparametric Statistics*, Wiley India, New Delhi.
- Cordella, T. (1998), “Patterns of Trade and Oligopoly Equilibria: An Example”, *Review of International Economics*, Vol. 6, No. 4, pp. 554-563.
- Costinot, A., D. Donaldson and I. Komunjer (2012), “What Goods do Countries Trade. A Quantitative Exploration of Ricardo’s Ideas”, *Review of Economic Studies*, Vol. 79, pp. 581-608.
- Courant, P.N. and A.V. Deardorff (1992), “International Trade with Lumpy Countries”, *Journal of Political Economy*, Vol. 100, No. 1, pp. 198-210.
- Dalum, B., K. Laursen and G. Villumsen (1998), “Structural Change in OECD Export Specialization Patterns: Despecialization and ‘Stickiness’”, *International Review of Applied Economics*, Vol. 12, No. 3, pp. 423-443.

- D'Agostino, R.B., A. Belanger and R.B. D'Agostino Jr. (1990), "A Suggestion for Using Powerful and Informative Tests of Normality", *The American Statistician*, Vol. 44, No. 4, pp. 316-321.
- Daudin, G., C. Riffart, and D. Schweisguth (2011), "Who Produces for Whom in the World Economy?", *Canadian Journal of Economics*, Vol. 44, No. 4, pp. 1403-1437.
- Davis, D. R. (1995), "Intra-industry trade: A Heckscher-Ohlin-Ricardo approach" *Journal of International Economics*, Vol. 39, pp. 201-226.
- Davis, D.R. and D.E. Weinstein (2001), "An Account of Global Factor Trade", *American Economic Review*, Vol. 91, No. 5, pp. 1423-1453.
- Deardorff, A.V. (1980), "The General Validity of the Law of Comparative Advantage", *Journal of Political Economy*, Vol. 88, No. 5, pp. 941-957.
- Deardorff, A. (1984), "Testing Trade Theories and Predicting Trade Flows", in Jones, R. and P.B. Kenen (ed.) *Handbook of International Economics*, Vol. 1, North Holland, Amsterdam.
- Deardorff, A. (2001), "Fragmentation in Simple Trade Models", *North American Journal of Economics and Finance*, Vol. 12, No. 2, pp. 121-137.
- Deardorff, A.V. and R.M. Stern (1997), "Measurement of Non-Tariff Barriers", *Working Paper No. 179, OECD Economics Department*, OECD Publishing, Paris. Accessed on 25.7.2014. doi: 10.1787/568705648470.
- Deb, K. and P. Basu (2011), "Indices of Revealed Comparative Advantage and their Consistency with the Heckscher-Ohlin Theory: A Cross Sectional Analysis", *Foreign Trade Review*, Vol. 46, No. 3, pp. 3-28.
- Deb, K. and P. Basu (2012), "Consistency of Revealed Comparative Advantage Indices with the Heckscher-Ohlin Theory", *Conference Proceedings of Knowledge Globalization Conference, Pune, India, January 5-7*, Vol. 6, No. 1, pp. 122-131.

- Deb, K. and W.R. Hauk (2015), “RCA Indices, Multination Production and the Ricardian Trade Model”, *International Economics and Economic Policy*. doi: 10.1007/s10368-015-0317-z.
- Dimaranan, B., E. Ianchovichina and W.J. Martin (2007), “China, India, and the Future of the World Economy: Fierce Competition or Shared Growth?”, Policy Research Working Paper No. 4304, World Bank, Washington, DC.
- Dixit, A.K. and J.E. Stiglitz (1977), “Monopolistic Competition and Optimum Product Diversity”, *American Economic Review*, Vol. 67, No. 3, pp. 297-308.
- Drabicki, J.Z. and A. Takayama (1979), “An Antinomy in the Theory of Comparative Advantage”, *Journal of International Economics*, Vol. 9, No. 2, pp. 211-223.
- Eaton, J. and H. Kierzkowski (1984), “Oligopolistic Competition, Product Variety and International Trade” in Kierzkowski, H. (ed.) *Monopolistic Competition and International Trade*, Oxford University Press, Oxford.
- Edgeworth, F.Y. (1887), “On Observations relating to several quantities”, *Hermathena*, Vol. 6, No. 13, pp. 279-285.
- Ekram, H. M., D. Huang and Q. Tran (2013), “Competitiveness of Textile and Apparel Industry of Bangladesh with China in International Trade”, Paper presented at International Conference on Information, Business and Education Technology 2013. Accessed on 10.8.2013. doi: 10.2991/icibet.2013.250.
- Epanechnikov, V.A. (1969), “Non-Parametric Estimation of a Multivariate Probability Density”, *Theory of Probability and Applications*, Vol. 14, No. 1, pp. 153-158.
- Falvey, R.E. (1981), “Commercial Policy and Intra-Industry Trade”, *Journal of International Economics*, Vol. 11, No. 4, pp. 495-511.
- Falvey, R. E. and H. Kierzkowski (1987), “Product Quality, Intra-Industry Trade and (Im)Perfect Competition”, in Kierzkowski, H. (ed.) *Protection and Competition in International Trade*, Basil Blackwell, New York.

- Faustino, H. (2008), “Intra-Industry Trade and Revealed Comparative Advantage: An Inverted-U Relationship”, Working Paper No. 03/2008/DE/SOCIUS, Department of Economics, School of Economics and Management, Technical University of Lisbon.
- Feenstra, R. (1998), “Integration of Trade and Disintegration of Production in Global Economy”, *Journal of Economic Perspectives*, Vol. 12, No. 4, pp. 31-50.
- Feenstra, R.C. (2004), *Advanced International Trade: Theory and Evidence*, Princeton University Press, USA.
- Feenstra, R.C. and G.H. Hanson (1997), “Foreign direct investment and relative wages: Evidence from Mexico’s maquiladoras”, *Journal of International Economics*, Vol. 42, pp. 371-393.
- Flam, H. and E. Helpman (1987), “Vertical Product Differentiation and North-South Trade”, *American Economic Review*, Vol. 77, No. 5, pp. 810-822.
- Frantzen, D. (2008), “Technology, Competitiveness and Specialization in OECD Manufacturing”, *Journal of Economic Studies*, Vol. 35, No. 1, pp. 44-68.
- Fraser, D.A.S. (1957), *Nonparametric Methods in Statistics*, John Wiley & Sons, New York.
- Galor, O. (1992), “A Two Sector Overlapping Generation Model: A Global Characterization of the Dynamical System”, *Econometrica*, Vol. 60, No. 6, pp. 351-386.
- Galton, F. (1889), *Natural Inheritance*, Macmillan, London.
- Ghosh, J. and C.P. Chandrasekhar (2007), “Recent Employment Trends in India and China: An Unfortunate Convergence”, *The Indian Journal of Labour Economics*, Vol. 50, No. 3, pp. 383-406.
- Golub, S.S. and C. Hsieh (2000), “Classical Ricardian Theory of Comparative Advantage Revisited”, *Review of International Economics*, Vol. 8, No. 2, pp. 221-234.
- Gould W.W and W.H. Rogers (1991), “sg3.4: Summary of tests of normality”, *Stata Technical Bulletin* 3: 20-23, Reprinted in *Stata Technical Bulletin* Reprints 1: 106-110. Stata Press, College Station, TX.

Greenaway, D. and C. Milner (1986), *The Economics of Intra-Industry Trade*, Basil Blackwell, New York.

Greenaway, D., R. Hine, and C. Milner (1994), “Country-Specific Factors and the Pattern of Horizontal and Vertical Intra-Industry Trade in the UK”, *Weltwirtschaftliches Archiv* Vol. 130, No. 1, pp. 77-100.

_____ (1995), “Vertical and Horizontal Intra-Industry Trade: A Cross Industry Analysis for the United Kingdom”, *The Economic Journal*, Vol. 105, No. 433, pp. 1505-1518.

Grubel, H. and P.J. Lloyd (1975), *Intra-Industry Trade: The Theory and Measurement of International Trade in Differentiated Product*, Macmillan, London.

Gujarati, D.N. and Sangeetha (2007), *Basic Econometrics*, Tata McGraw-Hill, New Delhi.

Hamilton, L.C. (2009), *Statistics with STATA, Updated for Version 10*, Brooks/Cole, Cengage Learning, Belmont, CA.

Hanson, G.H. (2012), “The Rise of Middle Kingdoms: Emerging Economies in Global Trade”, *Journal of Economic Perspectives*, Vol. 26, No. 2, pp. 41-64.

Harrigan, J. (1993), “OECD imports and trade barriers in 1983”, *Journal of International Economics*, Vol. 35, No. 1, pp. 91-111.

Hart, A. (2001), “Mann-Whitney test is not just a test of medians: differences in spread can be important”, *British Medical Journal*, Vol. 323, No. 7309, pp. 391-393.

Hart, P.E. and S.J. Prais (1956), “The Analysis of Business Concentration: A Statistical Approach”, *Journal of the Royal Statistical Society, Series A (General)*, Vol. 119, No. 2, pp. 150-191.

Hart, P.E. (1995), “The Convergence of Labour Productivity in British and German Industries”, *International Journal of the Economics of Business*, Vol. 2, No. 3, pp. 453-464.

- Hassan, M. (2013), “An Analysis of Competitiveness of Pakistan’s Agricultural Export Commodities”, *Journal of Business and Management*, Vol. 11, No. 5, pp. 29-34.
- Hatzichronoglou, T. (1997), “Revision of the High Technology Sector and Product Classification”, OECD Science Technology and Industry Working Papers 1997/02, OECD Publishing. Accessed on 19.3.2015. doi: 10.1787/134337307632.
- Hauk, W. (2012), “Trade Restriction Indices and U.S. Trade Policy”, *Applied Economics Letters*, Vol. 19, pp. 795-799. Accessed on 20.8.2012. doi: 10.1080/13504851.2011.605347
- Hayes, A.F. and L. Cai (2007), “Using Heteroskedasticity-Consistent Standard Error Estimators in OLS Regression: An Introduction and Software Implementation”, *Behavior Research Methods*, Vol. 39, No. 4, pp. 709-722.
- Helpman, E. (1981), “International Trade in the Presence of Product Differentiation, Economies of Scale and Monopolistic Competition: A Chamberlin-Heckscher-Ohlin Approach” *Journal of International Economics*, Vol. 11, pp. 305-340.
- Helpman, E. (1985), “Multinational Corporations and Trade Structure”, *The Review of Economic Studies*, Vol. 52, No. 3, pp. 443-457.
- Helpman, E. and P. Krugman (1985), *Market Structure and Foreign Trade: Increasing Returns Imperfect Competition and the International Economy*, MIT Press, USA.
- Higgins, J.J. (2004), *Introduction to Modern Nonparametric Statistics*, Duxbury Press, CA.
- Hiley, M. (1999), “The dynamics of changing comparative advantage in the Asia-Pacific region”, *Journal of the Asia Pacific Economy*, Vol. 4, No. 3, pp. 446-467. Accessed on 19.9.2012. doi: 10.1080/13547869908724693.
- Hillman, A. L. (1980), “Observations on the Relation between “Revealed Comparative Advantage” and Comparative Advantage as Indicated by Pre-Trade Relative Prices”, *Weltwirtschaftliches Archiv*, Vol. 116, pp. 315-321.

- Hinloopen, J. and C. Marrewijk (2001), “On Empirical Distribution of the Balassa Index”, *Weltwirtschaftliches Archiv*, Vol. 137, No. 1, pp. 1-35.
- Hinloopen, J. and C. Marrewijk (2004), “Dynamics of Chinese Comparative Advantage”, Tinbergen Institute Discussion Paper No. 2004-034/2.
- Hoen, A. R., and J. Oosterhaven (2006), “On the Measurement of Comparative Advantage”, *The Annals of Regional Science*, Vol. 40, pp. 677-691. Accessed on 13.8.2010. doi: 10.1007/s00168-006-0076-4.
- Huber, P.J. (1964), “Robust Estimation of a Location Parameter”, *Annals of Mathematical Statistics*, Vol. 35, No. 1, pp. 73-101.
- Huber, P. J. (1967), “The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions”, *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, pp. 221-233.
- Hummels, D., D. Rapoport and K. Yi (1998), “Vertical Specialization and the Changing Nature of the World Trade”, *Federal Reserve Bank of New York Economic Policy Review*, Vol. June, pp. 79-99.
- Hummels, D., J. Ishii and K. Yi (2001), “The nature and growth of vertical specialization in world trade”, *Journal of International Economics*, Vol. 54, pp. 75–96.
- Johnson. R. (2014), “Five Facts about Value-Added Exports and Implications for Macroeconomics and Trade Research”, *Journal of Economic Perspectives*, Vol. 28, No. 2, pp. 119-142.
- Johnson, R., and G. Noguera (2012), “Accounting for Intermediates: Production Sharing and Trade in Value-added”, *Journal of International Economics*, Vol. 86, pp. 224-236.
- Johnson, R., and G. Noguera (2014), “A Portrait of Trade in Value Added over Four Decades”, Unpublished paper, Dartmouth College.
- Johnston, J. and J. DiNardo (1997), *Econometric Methods*, McGraw-Hill, London.

- Jones, R.W. and H. Kierzkowski (2001), "A Framework for Fragmentation." In Arndt, S.W. and H. Kierzkowski (ed.) *Fragmentation. New Production Patterns in the World Economy*, Oxford University Press, Oxford.
- Jones, R.W., H. Kierzkowski, and G. Leonard (2002), "Fragmentation and Intra-Industry Trade." in Lloyd, P.J. and H.H. Lee (ed.) *Frontiers of Research in Intra-Industry Trade*, Palgrave MacMillan, New York.
- Karaalp, H. S. (2011), "Competitiveness of Turkey in Eurasia: A Comparison with CIS Countries", *China-USA Business Review*, Vol. 10, No. 9, pp. 727-744.
- Kee, H.L., A. Nicita and M. Olarreaga (2008), "Import Demand Elasticities and Trade Distortions", *Review of Economics and Statistics*, Vol. 90, No. 4, pp. 666-682.
- Kee, H.L., A. Nicita and M. Olarreaga (2009), "Estimating Trade Restriction Indices", *Economic Journal*, Vol. 119, No. 1, pp. 172-199.
- Keskin, S. (2006), "Comparison of Several Univariate Normality Tests Regarding Type I Error Rate and Power of the Test in Simulation based Small Samples", *Journal of Applied Science Research*, Vol. 2, No. 5, pp. 296-300.
- Kijboonchoo, T. and K. Kalayanakupt (2003), "Comparative Advantage and Competitive Strength of Thai Canned Tuna Export in the World Market: 1982-1998", *ABAC Journal*, Vol. 23, No. 1, pp. 19-33.
- Kinoshita, Y. and N.F. Campos (2006), "Determinants and effects of foreign direct investment in transition economies", in Nakagawa, J. (ed.) *Managing Development: Globalization, Economic Restructuring and Social Policy*, Routledge, England.
- Koenker, R. (1981), "A note on Studentizing a test for heteroscedasticity", *Journal of Econometrics*, Vol. 17, No. 1, pp. 107-112.
- Koopman, R., W. Powers, Z. Wang and S. Wei (2010), "Give credit where credit is due: tracing value added in global production chains", NBER Working Paper No. 16426.

- Koopman, R., Z. Wang and S. Wei (2012), “Estimating Domestic Content in Exports when Processing Trade is Pervasive”, *Journal of Development Economics*, Vol. 99, pp. 178-189.
- Koopman, R., Z. Wang and S. Wei (2014), “Tracing Value-Added and Double Counting in Gross Exports”, *American Economic Review*, Vol. 104, No. 2, pp. 459-494.
- Krugman, P. (1979), “Increasing Returns, Monopolistic Competition and International Trade”, *Journal of International Economics*, Vol. 9, No. 4, pp. 469-79.
- Krugman, P. (1980), “Scale Economies, Product Differentiation and the Pattern of Trade”, *America Economic Review*, Vol. 70, No. 5, pp. 950-959.
- Krugman, P. (1981), “Intra-industry Specialization and the Gains from Trade”, *Journal of Political Economy*, Vol. 89, No. 5, pp. 959-73.
- Krugman, P. (1995), “Growing World Trade: Causes and Consequences”, *Brooking Paper on Economic Activity*, Vol. 26, No. 1, pp. 327-362.
- Kunimoto, K. (1977), “Typology of Trade Intensity Indices”, *Hitotsubashi Journal of Economics*, Vol. 17, No. February, pp. 15-32.
- Lafay, G. (1992), “The measurement of revealed comparative advantages”, in Dagenais, M.G. and P.A. Muet (ed.) *International Trade Modeling*, Chapman & Hall, London.
- Lancaster, K. (1975), “Socially Optimal Product Differentiation”, *American Economic Review*, Vol. 65, No. 4, pp. 567-85.
- Lancaster, K. (1979), *Variety, Equity, and Efficiency*, Columbia University Press, New York.
- Lancaster, K. (1980), “Intra-Industry Trade under Perfect Monopolistic Competition”, *Journal of International Economics*, Vol. 10, No. 2, pp. 151-170.
- Lary, H.B. (1968), *Imports of Manufactures from Less Developed Countries*, National Bureau of Economic Research, New York.

- Lau, L.J., Y. Qian and G. Roland (2001), "Reform without Losers: An Interpretation of China's Dual Track Approach to Transition", *Journal of Political Economy*, Vol. 108, No. 1, pp. 120-143.
- Laursen, K. (1998), "Revealed Comparative Advantage and the Alternatives as Measures of International Specialization", Working Paper No. 98-30, Danish Research Unit for Industrial Dynamics.
- Lawrence, R. (1994), "Trade, Multinationals, and Labor", NBER Working Paper No. 4836.
- Leamer, E.E. (1980), "The Leontief Paradox, Reconsidered", *Journal of Political Economy*, Vol. 88, pp. 495-503.
- Leamer, E. (1990), "Latin America as a Target of Trade Barriers Erected by the Major Developed Countries in 1983", *Journal of Development Economics*, Vol. 32, No. 2, pp. 337-368.
- Lehmann, E.L. (2006), *Nonparametrics: Statistical Methods Based on Ranks*, Springer, USA.
- Leontief, W.W. (1953), "Domestic Production and Foreign Trade: The American Capital Position Re-examined." *Proceedings of the American Philosophical Society*, Vol. 97, pp. 332-349. Reprinted in Caves, R. and H.G. Johnson (ed.) *Readings in International Economics* (Homewood, Ill.: Irwin, 1968).
- Leromain, E. and G. Orefice (2014), "New Revealed Comparative Advantage Index: Dataset and Empirical Distribution", *International Economics*, Vol. 139, pp. 48-70.
- Liesner, H. H. (1958), "The European Common Market and British Industry", *The Economic Journal*, Vol. 68, No. 270, pp. 302-316.
- MacDougall, G.D.A. (1951), "British and American Exports: A Study Suggested by the Theory of Comparative Costs. Part I", *The Economic Journal*, Vol. 61, No. 244, pp. 697-724.
- Maddala, G.S. (1988), *Introduction to Econometrics*, MacMillan, New York.

- Maddison, A. (2007), *Chinese Economic Performance in the Log Run*, OECD Development Centre, Paris.
- Mann, H.B. and D.R. Whitney (1947), “On a Test of Whether One Of Two Random Variables Is Stochastically Larger Than The Other”, *The Annals of Mathematical Statistics*, Vol. 18, No. 1, pp. 50-60.
- Marquetti, A. and D. Foley (2011), Extended Penn World Tables, Version 4. Accessed from <http://sites.google.com/a/newschool.edu/duncan-foley-homepage/home/EPWT> on 15.3.2013.
- Mazumdar, D. and S. Sarkar (2008), *Globalization, Labor Markets and Inequality in India*, Routledge, New York.
- McDonald, B. (2002), “ A Teaching Note on Cook’s Distance – A Guideline”, *Res. Lett. Inf. Math. Sci*, Vol. 3, No. April, pp. 127-188. Accessed from <http://www.massey.ac.nz/~wwiims/research/letters> on 11.12.2014.
- Mendes, M. and A. Pala (2003), “Type I Error Rate and Power of Three Normality Tests”, *Pakistan Journal of Information and Technology*, Vol. 2, No. 2, pp. 135-139.
- Meng, B., Y. Fang and N. Yamano (2012), “Measuring Global Value Chains and Regional Economic Integration: An International Input-Output Approach”, IDE Discussion Paper No. 362.
- Miroudot, S., R. Lanz and A. Ragoussis (2009), “Trade in Intermediate Goods and Services”, OECD Trade Policy Working Paper No. 93.
- Mountford, A. (1998), “Trade, Convergence and Overtaking”, *Journal of International Economics*, Vol. 46, No. 1, pp. 167-182.
- Nicita, A. and M. Olarreaga (2007), “Trade, Production and Protection Database 1976-2004”, *World Bank Economic Review*, Vol. 21, No. 1, pp. 165-171.

- Nogués, J.J., A. Olechowski and L.A. Winters (1986), “The Extent of Nontariff Barriers to Industrial Countries’ Imports”, *The World Bank Economic Review*, Vol. 1, No. 1, pp. 181-199.
- NRC. (2006), “Analyzing the U.S. content of imports and the foreign content of exports”, National Research Council, The National Academies Press, Washington DC.
- OECD (2015), “Purchasing Power Parities: Frequently Asked Questions.” Accessed from <http://www.oecd.org/std/ppp/faq> on 11.12.2014.
- Palley, T.I. (2008), “Institutionalism and New Trade Theory: Rethinking Comparative Advantage and Trade Policy”, *Journal of Economic Issues (Association for Evolutionary Economics)*, Vol. 42, No. 1, pp. 195-208.
- Pittiglio, R. (2014), “An Essay on Intra-Industry Trade in Intermediate Goods”, *Modern Economy*, Vol. 5, pp. 468-488.
- Pritchett, L. (1996), “Mind your P.s and Q.s. The cost of Public Investment is Not the Value of Public Capital”, Policy Research Working Paper No. 1660, The World Bank.
- Proudman, J. and S. Rdding (1998), *Openness and Growth*, Bank of England, UK.
- PTI, (2015, August 18), “India’s turn to become factory of the world: China state media”, *The Times of India*. Accessed from <http://timesofindia.indiatimes.com/business/india-business> on 3.9.2015.
- Razali, N.M. and Y.B. Wah (2011), “Power Comparison of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests”, *Journal of Statistical Modeling and Analytics*, Vol. 2, No. 1, pp. 21-33.
- Richardson, D.J. and C. Zhang (2001), “Revealing Comparative Advantage: Chaotic or Coherent Patterns across Time and Sector and U.S. Trading Partner?” in Blomström, M. and L.S. Goldberg (ed.) *Topics in Empirical Economics: A Festschrift in Honor of Robert E. Lipsey*, University of Chicago Press, USA.

- Rodriguez, G. (2015), “Generalised Linear Models”, Princeton University. Accessed from <http://data.princeton.edu/wws509/notes/c2s9.html> on 12.3.2015.
- Rodrik, D. (2006), “What’s So Special about China’s Exports?” *China and the World Economy*, Vol. 14, No. 5, pp. 1–19.
- Rosenblatt, M. (1956), “Remarks on Some Nonparametric Estimates of Density Function”, *Annals of Mathematical Statistics*, Vol. 27, No. 3, pp. 832-837.
- Rousseeuw, P.J. and B.C. van Zomeren (1990), “Unmasking Multivariate Outliers and Leverage points”, *Journal of the American Statistical Association*, Vol. 85, No. 411, pp. 633-639.
- Royston, J.P. (1982) “An Extension of Shapiro and Wilk’s W Test for Normality to Large Samples”, *Applied Statistics* Vol. 31, No. 2, pp. 115-124.
- Sanidas, E. and Y. Shin (2010), “Comparison of Revealed Comparative Advantage Indices with Application to Trade Tendencies of East Asian Countries”, Paper Presented at the 9th Korea and the World Economy Conference, Incheon, 2010.
- Sanidas, E. and Y. Shin (2011), “Convergence towards the Revealed Comparative Advantage Neutral Point for East Asia: Similarities and Differences between the Three Countries”, *Seoul Journal of Economics*, Vol. 24, No. 1, pp. 21-50.
- Shaked, A. and J. Sutton (1984), “Natural oligopolies and international trade”, in Kierzkowski, H. (ed.) *Monopolistic Competition and International Trade*, Oxford University Press, Oxford.
- Shapiro, S.S. and M.B. Wilk (1965), “An Analysis of Variance Test for Normality (Complete Samples)”, *Biometrika*, Vol. 52, No. (3/4), pp. 591-611.
- Sharma, A. and M. Dietrich (2007), “The Structure and Composition of India's Exports and Industrial Transformation (1980-2000)”, *International Economic Journal*, Vol. 21, No. 2, pp. 207-231.

- Siegel, S. (1956), *Nonparametric Statistics for the Behavioural Sciences*, McGraw-Hill, New York.
- Smirnov, N.V. (1939), “Estimate of Deviation between Empirical Distribution Functions in Two Independent Samples”, (in Russian) *Bulletin Moscow University*, Vol. 2, No. 2, pp. 3-16.
- Smyth, D. A. (2005), “Ireland’s Revealed Comparative Advantage”, *Quarterly Bulletin*, Central Bank of Ireland, Dublin, Vol. 1, pp. 101-114.
- Spence, A.M. (1976), “Product Selection, Fixed Costs, and Monopolistic Competition”, *Review of Economic Studies*, Vol. 43, No. 2, pp. 217-235.
- StataCorp. (2013), *Stata 13 Base Reference Manual*, College Station, Stata Press, TX.
- Suranovic, S.M. (2015), “The Theory of Comparative Advantage – Overview”, Accessed from <http://internationalecon.com/Trade/Tch40/T40-0.php> on 25.4.2015.
- Swamy, S. (2010), *Economic Development and Reforms in India and China: A Comparative Perspective*, Har-Anand Publications, New Delhi.
- Tharakan, P.K. and B. Kerstens (1995), “Does North-South Horizontal Intra-Industry Trade Really Exist? An Analysis of the Toy Industry”, *Weltwirtschaftliches Archiv*, Vol. 131, No. 1, pp. 86-105.
- The World Bank (2013), *World Development Indicators*, The World Bank, Washington, DC. Accessed from <http://data.worldbank.org/data-catalog/world-development-indicators> on 12.3.2013.
- Trefler, D. (1993), “International Factor Price Differences: Leontief Was Right!”, *Journal of Political Economy*, Vol. 101, No. 6, pp. 961-987.
- Trefler, D. (1995), “The Case of Missing Trade and Other Mysteries”, *American Economic Review*, Vol. 85, No. 5, pp. 1029-1046.

- Tuong, H.D. and A.J. Yeats (1980), “On Factor Proportions as a Guide to the Future Composition of Developing Country Exports”, *Journal of Development Economics* Vol. 7, No. 4, pp. 521-539.
- Türkcan, K. (2011), “Vertical Intra-Industry Trade and Product Fragmentation in the Auto-Parts Industry”, *Journal of Industry, Competition and Trade*, Vol. 11, No. 2, pp. 149-186.
- Türkcan, K. and A. Ates (2011), “Vertical Intra-industry Trade and Fragmentation: An Empirical Examination of the U.S. Auto-Parts Industry”, *The World Economy*, Vol. 34, No. 1, pp. 154-172.
- Ullah, A. (1988), “Non-Parametric Estimations of Econometric Functionals”, *The Canadian Journal of Economics*, Vol. 21, No. 3, pp. 625-658.
- UNCTAD (1996), *Trade and Development Report*, UNCTAD, Geneva.
- United Nations (2014), *International Trade Statistics Yearbook 2013*, United Nations, New York.
- Vanek, J. (1968), “The Factor Proportions Theory: The N-Factor Case”, *Kyklos*, Vol. 21, No. 4, pp. 749-754.
- Verardi, V. and C. Croux (2009), “Robust Regression in Stata”, *The Stata Journal*, Vol. 9, No. 3, pp. 439-453.
- Vollrath, T.L. and D. Vo. (1990), “Agricultural Competitiveness in an Interdependent World”, International Association of Agricultural Economists Occasional Paper No. 5, Hants, England.
- Vollrath, T. L. (1991), “A Theoretical Evaluation of Alternative Trade Intensity Measures of Revealed Comparative Advantage”, *Weltwirtschaftliches Archiv*, Vol. 127, pp. 265-280.
- White, H. (1980), “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity”, *Econometrica*, Vol. 48, No. 4, pp. 817–838.

- Wignaraja, G. (2011), “Economic Reforms, Regionalism and Exports: Comparing China and India”, East-West Centre Policy Studies 60.
- Wilcoxon, F. (1945), “Individual Comparisons by Ranking Methods”, *Biometric Bulletin*, Vol. 1, No. 6, pp. 80-83.
- Wilcoxon, F. (1947), “Probability Tables for Individual Comparisons by Ranking Methods”, *Biometrics*, Vol. 3, No. 3, pp. 119-122.
- Yeats, A. J. (1985), “On the Appropriate Interpretation of the Revealed Comparative Advantage Index: Implications of a Methodology based on Industry Sector Analysis”, *Weltwirtschaftliches Archiv*, Vol. 121, pp. 61-73.
- Yi, K. (2003), “Can Vertical Specialization Explain the Growth of World Trade?”, *Journal of Political Economy*, Vol. 111, No. 1, pp. 52–102.
- Yu, R., J. Cai, and P. Leung (2009), “The Normalized Revealed Comparative Advantage Index”, *The Annals of Regional Science*, Vol. 43, pp. 267-282. Accessed on 13.8.2010. doi: 10.1007/s00168-008-0213-3.
- Yue, C. and P. Hua (2002), “Does comparative advantage explains export patterns in China”, *China Economic Review*, Vol. 13, pp. 276-296.