Indices of Revealed Comparative Advantage and their Consistency with the Heckscher-Ohlin Theory

A Cross Sectional Analysis

Kaveri Deb *and* ***Partha Basu***

The issue of comparative advantage constitutes an important feature of the theory of international trade. Comparative advantage of countries has been measured in the literature by various alternative indices of 'revealed comparative advantage'. Balassa first coined the term 'revealed comparative advantage' and the index that he devised in the process, has been later modified by various authors in many ways to address one or more of the shortcomings of Balassa's index. However, the existing literature has not tried to determine empirically, the extent to which the different indices are consistent with the idea of comparative advantage as identified particularly by Eli Heckscher and Bertil Ohlin. The present paper makes an attempt in this regard. In the process, another index has been derived by considering the logarithmic transformation of the Balassa's index, and its consistency with the theory has been empirically tested in a similar manner to determine whether it performs better than the other indices. A theoretical review of the alternative indices and empirical findings in different situations exhibit the modified index certainly has an edge over other indices.

Keywords: *Revealed comparative advantage, Heckscher-Ohlin theory, labour intensive commodities, capital intensive commodities.*

JEL Classification: *F11, F14, C43, C12, C21, R12*

1. Introduction

COMPARATIVE advantage as a determinant of international trade, was developed and conceptualized by David Ricardo in 1817. In the Ricardian model, comparative advantage was the outcome of differences in technology or factor productivity between the countries. Later in 1930s, Eli Heckscher and Bertil Ohlin identified comparative advantage as the outcome of differences in relative factor endowments of two countries. With capital and labour as the two factors of production, the Heckscher-Ohlin theory propounded, a relatively labour abundant country will have comparative advantage in labour intensive goods and will export the same. A relatively capital abundant country on the other hand, will have comparative advantage in capital intensive goods, which it will export.

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 theories, the validity of the orthodox trade theories had been questioned. It had been argued that the recent splurge of intra-industry trade or trade in similar products can be explained by new trade theories only. However, this idea had been challenged by many theorists and they have demonstrated that even the orthodox trade theories can account for intra-industry trade (Falvey 1981; Davis 1995). Hence, the idea of comparative advantage and theories encompassing it, still remain an important strand of international trade theory.

While comparative advantage has been conceptualized theoretically to explain the pattern of international trade, the quantification of comparative advantage for empirical analysis is obviously not an easy task. This is because, economic theory is based upon certain restrictive assumptions which are difficult to quantify in the real world. The first problem relates to the fact that unlike as required by the theory, data available for measuring comparative advantage pertain to the post-trade situation. Moreover, the data are subjected to distortionary impacts of government intervention and imperfect information (Vollrath 1991, pp. 266-267). 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 same composite group (Vollrath 1991, p. 267).

Despite such shortcomings, several attempts have been made in the past to quantify comparative advantage and are still being actively pursued. Typically, comparative advantages of countries have been measured in the literature by various alternative indices of 'revealed comparative advantage' (RCA), using post-trade data. Balassa (1965) coined the term 'revealed comparative advantage'. However, over the years Balassa's RCA index has been criticized from various perspectives and alternative RCA indices have been suggested in the literature. Though each of these latter indices have tried to address one or more shortcomings of the Balassa's RCA index, the existing literature have not tried to determine empirically, how various alternative RCA indices are consistent with the idea of comparative advantage, as identified by Heckscher-Ohlin. Section 2 of this paper, reviews the RCA indices of Balassa and subsequent authors. Eventually, it introduces another index, by considering suitable modification of the index of Balassa, such that the new index has structural features superior to other existing indices. Section 3 attempts to examine empirically, using cross sectional data, how far the RCA indices considered in section 2, are consistent with the Heckscher-Ohlin theory and thereby permits a comparative study. Section 4 concludes that the suggested modified index could be used to determine the comparative advantage of countries as it has good structural features and in addition is empirically consistent with the theory.

2. Indices of RCA

Although the idea of 'revealed comparative advantage' is attributed to Balassa, prior to that Liesner (1958) actually 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. However, the index that he devised in the process, was limited in its coverage as it considered only a single commodity, and attempted to identify comparative advantage of a country, on the basis of its performance in exporting that commodity, relative to other countries (Vollrath 1991, P. 269).

Eventually, by adjusting Liesner's index, Balassa (1965) developed his index and named it as 'revealed comparative advantage index'. In

his view, comparative advantage can be 'revealed' through real life patterns of country or commodity trade, because actual exchange reflects differences in costs as well as non-price factors (Vollrath 1991, P. 266). The original RCA index of Balassa was defined in the following form:

$$(X^i_a / X^c_a) / (X^i_m / X^c_m)$$

Here, X stands for exports, m denotes all manufactured goods and a to any one of the manufactured goods. i and c denote any one of the 11 industrial countries that he considered and all the 11 countries together respectively. 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.

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.[1] Hence, the RCA index of Balassa takes the following form:

$$(X^i_a / X^i_t) / (X^w_a / X^w_t)$$

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. In the following analysis, while referring to Balassa's RCA index, this version of the original index would be considered.

2.1 Problems with the RCA index of Balassa

While Balassa's RCA index is useful in determining whether or not a country has comparative advantage in a commodity, its applicability in terms of compatibility with the theory is doubtful. 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 for a particular country, the values of the indices are independent of the comparative advantage as indicated by pre-trade relative prices. In addition, while comparing the RCA indices of a commodity for two countries, Hillman showed that the correspondence between RCA index values and the pre-trade relative prices would hold only under certain circumstances.[2]

While Hillman (1980) provided a theoretical evaluation of Balassa's index, several economists identified problems with the index in empirical estimation. Yeats (1985) pointed out that the traditional method of ranking industries in a country according to the value of 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 observed, this inconsistency in ranking arises out of the fact that different industries have different distribution of country index values.[3] He also observed Balassa's RCA index could give misleading results because, the index might signify stronger comparative advantage for countries with smaller market share in the world export market, i.e., smaller X^{i}_{t} / X^{w}_{t} and for commodities with smaller market share in the world export market, i.e., smaller X^{w}_{a} / X^{w}_{t}.

Laursen (1998) argued that the asymmetric distribution of the Balassa's index makes it unsuitable for econometric analysis. The lower bound of the distribution is zero. The upper bound can assume any value greater than unity and would generally tend to infinity.[4] Laursen commented, with such a skewed distribution, the error term in regression analysis would be non-normal and hence, t and F statistics cannot be used reliably.

Hoen and Oosterhaven (2006) recognized the fact that the distribution of the Balassa's index dependent upon the number of countries and the commodities in the analysis. Considering the original Balassa index, the distribution of the index will vary with the number of reference countries with respect to which the export performance of a particular country is compared. The distribution of the index will also 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. An unstable distribution will have unstable mean, making comparison of index values across country or commodity difficult. Hoen and Oosterhaven attributed this problem to the ratio form of the index. Yu *et al*. (2008) also recognized the problems associated

with the variability of mean for the index of Balassa. However, it is to be noted, foregoing arguments about the instability of the mean, hold for arithmetic mean of ratios. In case of ratios with both numerator and denominator varying, geometric mean is the appropriate concept. On averaging Balassa's index values across countries or commodities by using geometric mean (or ratios of arithmetic mean of numerator to the denominator), it would be evident that the average is not significantly different from unity.

2.2 Other Indices of RCA

With due recognition to the problems associated with the RCA index of Balassa for empirical estimation, some of the above-mentioned authors came up with a number of alternative RCA indices. Identifying the asymmetrical distribution of the index of Balassa, Laursen (1998) suggested a simple modification of the index in order to make its distribution symmetric. His index takes the following form:

$$[\{(X^i_a / X^i_t) / (X^w_a / X^w_t)\} - 1] / [\{(X^i_a / X^i_t) / (X^w_a / X^w_t)\} + 1]$$

The lower limit of the distribution of this index is -1 with upper limit tending to +1. The comparative advantage neutral point would be close to zero which also defines the mean of the distribution.[5] The mean value can be achieved by considering the geometric mean or ratio of the arithmetic mean of the numerator to the arithmetic mean of the denominator. Country *i* would 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. Benedictis and Tamberi (2001), however, point out that the economic interpretation of the index is not very clear. Moreover, this index, like the index of Balassa, might signify greater comparative advantage for countries or for commodities with smaller market share in the world export market.

Hoen and Oosterhaven (2006) suggested an alternative to the index of Balassa. They replaced the ratio form of the index with the deviation form. Their index takes the following form:

$$X^i_a / X^i_t - X^w_a / X^w_t$$

The authors term it as the Additive RCA index. They, however, 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 would be exactly equal to -1 and +1 respectively. Zero would be the comparative advantage neutral point. However, in that case, use of the index for inter country comparison is questionable. Hence, with entire world as the reference group, the lower and the upper limits of the distribution would tend to -1 and +1 respectively. If the index takes a value greater (less) than zero then country *i* would reveal comparative advantage (disadvantage) in product *a*. An approximate value of zero denotes comparative advantage neutral point.

If geometric mean or the ratio of arithmetic mean of the numerator to the arithmetic mean of the denominator, is considered for computing the mean of a commodity's (country's) index values across countries (commodities), the computed value would be close to zero, implying stability of the mean and therefore the distribution.

The index is, however, sensitive to the size of a sector. If commodity *a* has larger share in the world export market, X^w_a / X^w_t would be large and hence, it is possible for the index to have a low numerical value. In such cases Additive RCA index could also generate misleading results.

Yu *et al.* (2008) 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*, X^i_a, from its expected exports of the same product in a world of no relative advantages (X^i_t / X^w_t) X^w_a. 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 index of country *i* in product *a* is thus stated in the following form:

$$(X^i_a \;/\; X^w_t) - (X^i_t X^w_a) \;/\; (X^w_t X^w_t)$$

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

Yu *et al.* (2008) argue that their index satisfies most of the desirable properties of the RCA index for empirical analysis. *First*, the index is symmetric about zero with the lower and the upper limits of the distribution being -1/4 and +1/4 respectively. *Second*, 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. Given the stability of the mean and thereby the distribution of the index, it is possible to compare the indices across country, commodity and time (Yu *et al*. 2008, 274-275).

While Yu *et al.* (2008) themselves identified several advantages of their index, there are certain other distinctive features of their index that make it particularly attractive compared to other RCA indices, considered in this section. *First*, most of the other indices are sensitive to the size of the country or the sector or both. In comparison, the index of Yu *et al*. is not noticeably influenced by the size of the country or of the commodity, due to its deviation form and normalization of the index by total world exports. *Second*, there are no complications involved in computing the mean of the index values because unlike other indices, the denominator of the index is fixed. Therefore, contrary to other indices, simple arithmetic mean can be used for the purpose.

2.3 Modified Index of RCA

Incidentally, an attempt has been made to suggest another index of RCA with reasonably good structural features and which, if empirically analyzed, might be expected to generate more reliable results than the four existing indices. Hence, in this section a modification of the RCA index of Balassa on the line of Vollrath (1991) has been suggested.[6]

To derive the new index we consider the RCA index of Balassa:

$$(X^i_a / X^i_t) / (X^w_a / X^w_t)$$

With natural logarithmic transformation, we arrive at a new index:

$$\ln[(X^i_a / X^i_t) / (X^w_a / X^w_t)]$$

Thus, the new index measures on 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, the new index could be provided an economic interpretation. There are certain additional advantages of this index. *Firstly*, the index is symmetric about zero with the upper and lower limits of its distribution being $+\infty$ and $-\infty$. *Secondly*, due to the logarithmic form of the index, the mean of the index values, which is zero, could be computed by considering simple arithmetic mean, either across country or across commodity. Hence, the stability of mean ensures the comparison of

the index across commodity or across country. *Thirdly*, though this index, like the parent index of Balassa, remains sensitive to the size of a sector or country, this sensitivity is expected to be less compared to the indices of not only Balassa but also Hoen and Oosterhaven, and even Laursen, as it is measured on a log scale. *Lastly*, an added advantage of the modified index over all the four existing RCA indices considered is, being defined in logarithmic form, the estimated residuals from any regression with the modified index as the dependent variable, could be expected to be normally distributed as economic variables like exports are log-normally distributed. Its only limitation is that the index cannot be defined in case the export of a product by a country is zero.[7] However, this sort of boundary problem is not uncommon in economic theory.

3. Empirical Analyses of RCA Indices

Section 2 enumerates the problems with the RCA index of Balassa for empirical estimation, and puts forward various other indices of RCA, suggested in the literature to overcome the deficiencies of Balassa's index.[8] But the existing literature have not tried to analyze the extent to which the alternative indices of RCA, are consistent with the Heckscher-Ohlin theory on comparative advantage of countries. Hence, in addition to testing the consistency of the existing indices with the Heckscher-Ohlin theory, the modified index suggested in sub-section 2.3 is also empirically analyzed in this section.

3.1 Data and Methodology

For empirically analyzing the relevance of the Heckscher-Ohlin theory in the context of the RCA indices, necessary data are collected for each of 47 countries and for the world as a whole, for the year 2005, from United Nations COMTRADE database and *International Trade Statistics Yearbook 2009*.[9] Three labour intensive and three capital intensive sectors were selected for the analysis. The labour intensive sectors have SITC (revision 3) 3-digit codes of 652 (cotton fabrics, woven), 844 (women and girls clothing, knit) and 851 (footware). The capital intensive sectors have SITC (revision 3) 3-digit codes of 525 (radioactive and associated materials), 541 (medicinal and pharmaceutical products, other than medicaments of group 542) and 751 (office machines).[10]

In accordance with the Heckscher-Ohlin theorem, with two factors of production, labour and capital, labour abundant countries would produce and export relatively more of labour intensive products and could be considered to have comparative advantage in such products. Similarly, capital abundant countries could be considered to have comparative advantage in capital intensive products. Using the physical definition of factor abundance, if the ratio of labour force of a country relative to the world, to the gross capital formation of the same country relative to the world, exceeds unity, the country could be considered to be labour abundant. On the other hand, if the ratio falls short of unity, the country could be considered to be capital abundant.

Therefore, for each labour intensive and capital intensive manufactures on which data are collected, Spearman's rank correlation coefficients and Pearson's product moment correlation coefficients are computed between the alternative RCA indices of countries and the ratio of labour force of the corresponding countries relative to the world, to the gross capital formation of those countries relative to the world.

The parametric correlation results are further reaffirmed by reporting bivariate regression results with White's heteroscedasticity corrected robust estimate of standard error in each case.[11] Bivariate regressions of the following form are fitted to the data:

$$Y = \alpha + \beta X + u$$

Variable Y signifies the index values for the group of countries considered for each sector. Variable X refers to the ratio of labour force of each country relative to the world, to the gross capital formation of the same country relative to the world.

For each sector, initially the correlation and regression coefficients are computed by incorporating all countries, i.e., countries with comparative advantage and disadvantage in a product. At the next stage, to provide further support to the findings, correlation and regression coefficients are computed by considering only countries with comparative advantage in labour intensive products and subsequently by considering only those countries with comparative disadvantage in capital intensive products.

The parametric correlation and regression coefficients are reported both on linear and double log scales. However, it is to be noted, the modified index being a log transformation of the Balassa's index, double log form has not been tried, as the dependent variable would take the form of log of log. Instead, semi-log form for this index has been considered. The resulting correlation and regression coefficients would be similar to the corresponding double log form of the Balassa's index, except for the sign in the case of group of countries with comparative disadvantage in a product.[12]

Computation of parametric correlation and regression on double scale would be possible for the group of countries with only comparative advantage in labour intensive products and for the group of countries with only comparative disadvantage in capital intensive products. However, for the composite group of countries, parametric correlation and regression on double log scale could not be conducted. Because this group includes, among others, the countries with comparative disadvantage in a product. For these countries, the indices of Laursen, Hoen and Oosterhaven, and Yu *et al.* take negative values, log of which are undefined in real space.

Hence, considering the composite group of countries, i.e., countries with comparative advantage and disadvantage in selected product categories, more the labour force of each country relative to the world exceeds (falls short of) its gross capital formation relative to the world, more (less) will it produce and export labour intensive products, and can have higher revealed comparative advantage (disadvantage) in such products. Thus, higher the value of the ratio of relative labour force to the relative gross capital formation, greater would be the value of RCA indices of countries in labour intensive products. Following similar arguments, higher the ratio, smaller would be the value of RCA indices of countries in capital intensive products.

The hypotheses of this part can be stated as, Hypothesis 1: *RCA indices in labour intensive commodities would rise with a rise in ratio of relative labour force to relative gross capital formation. RCA indices in capital intensive commodities would fall with a rise in ratio of relative labour force to relative gross capital formation.*

Next, out of 47 selected countries, only the countries with comparative advantage in each labour intensive product are considered.

The hypothesis for this part can be stated as, Hypothesis 2: *RCA indices in labour intensive commodities would rise with a rise in ratio of relative labour force to relative gross capital formation.*

Subsequently, only countries with comparative disadvantage in each capital intensive product for the year 2005 are considered. For the computation of correlation and regression coefficients for this part of analysis the 'revealed comparative disadvantage' (RCD) indices in the selected products are considered. Here RCD is defined as the absolute value of the RCA index. Countries with comparative disadvantage in a product have negative values for the indices of Laursen, Hoen and Oosterhaven, Yu *et al.* and modified index since, the absolute values of the RCA indices are considered, higher revealed comparative disadvantage of countries would be associated with higher values of these four indices and higher ratio of relative labour to relative gross capital formation. Thus, a positive association between these four RCD indices and the ratio of relative labour to relative gross capital formation of countries for every capital intensive product is expected. But the index of Balassa is always positive even if a country has comparative disadvantage in a product. Thus, the association between the index of Balassa and the ratio of relative labour to relative gross capital formation would be negative for every product for the subset of countries considered.

The hypotheses for this part can be stated as, Hypothesis 3: *RCD indices of Laursen, Hoen and Oosterhaven, Yu* et al. *and modified index in capital intensive commodities would rise with a rise in ratio of relative labour force to relative gross capital formation. RCD index of Balassa in capital intensive commodities would fall with a rise in ratio of relative labour force to relative gross capital formation.*

The hypotheses 1, 2 and 3 are tested against the null hypothesis of no association between the variables.

3.2 Discussion of Results

The results corresponding to hypotheses 1, 2 and 3 are provided in Tables 1, 2 and 3 respectively. The results in each table are analyzed in two stages. In the first stage, the indices of Balassa, Laursen, Hoen and Oosterhaven and Yu *et al*., are studied to determine which, among the existing indices, are more consistent with the hypotheses. In the next stage, the results for the modified index are compared with the findings.

TABLE 1

CORRELATION AND REGRESSION ANALYSES FOR LABOUR AND CAPITAL INTENSIVE PRODUCTS

Index			Labour Intensive Products			Capital Intensive Products		
			652	844	851	525	541	751
Balassa	Correlation	Rank	0.378 *	0.731 *	0.332 **	-0.517 *	-0.637 *	-0.318 **
		Product Moment	0.181 (1.16)	0.224 (1.45)	0.249 (1.63)	-0.174 (-1.12)	-0.233 (-1.52)	-0.132 (-0.84)
	Regression	Relative Labour/ Relative Capital	0.326 (0.99)	0.514 ** (2.20)	0.158 (1.15)	-0.042 ** (-1.90)	-0.112 ** (-2.41)	-0.028 ** (-1.79)
		Constant	1.346 * (2.45)	2.875 ** (2.22)	0.911 * (3.48)	0.430 ** (2.23)	1.238 * (3.29)	0.612 * (4.34)
Laursen	Correlation	Rank	0.378 *	0.731 *	0.332 **	-0.517 *	-0.637 *	-0.318 **
		Product Moment	0.217 (1.41)	0.518 * (3.83)	0.328 ** (2.20)	-0.245 (-1.60)	-0.412 * (-2.86)	-0.276 ** (-1.82)
	Regression	Relative Labour/ Relative Capital	0.029 (1.27)	0.082 * (4.08)	0.045 * (2.84)	-0.026 * (-2.56)	-0.054 * (-3.75)	-0.030 * (-2.77)
		Constant	-0.309 * (-3.47)	-0.213 ** (-2.18)	-0.397 * (-4.59)	-0.712 * (-8.61)	-0.246 * (-2.72)	-0.419 * (-6.08)
Hoen & Oosterhaven	Correlation	Rank	0.378 *	0.731 *	0.332 **	-0.517 *	-0.637 *	-0.318 **
		Product Moment	0.181 (1.16)	0.224 (1.45)	0.249 (1.63)	-0.174 (-1.12)	-0.233 (-1.52)	-0.132 (-0.84)
	Regression	Relative Labour/ Relative Capital	0.001 (0.99)	0.001 ** (2.20)	0.001 (1.15)	-0.00003 ** (-1.90)	-0.001 ** (-2.41)	-0.00004 ** (-1.79)
		Constant	0.001 (0.63)	0.005 (1.45)	-0.0006 (-0.34)	-0.0004 * (-2.96)	0.002 (0.63)	-0.0006 * (-2.75)

Index			Labour Intensive Products			Capital Intensive Products		
			652	844	851	525	541	751
Yu *et al.*	Correlation	Rank	0.497 *	0.669 *	0.415 *	0.286 **	-0.322 **	0.444 *
		Product Moment	0.150 (0.96)	0.209 (1.35)	0.113 (0.72)	-0.043 (-0.27)	-0.166 (-1.06)	0.029 (0.18)
	Regression	Relative Labour/ Relative Capital	0.000004 (1.48)	0.000005 ** (2.14)	0.000008 (1.36)	-0.0000003 (-0.51)	-0.000008 ** (-1.72)	0.0000004 (0.64)
		Constant	-0.000004 (-0.28)	-0.00001 (-0.75)	-0.000008 (-0.18)	0.0000002 (0.004)	0.00005 (1.51)	0.000003 (0.36)
Modified Index	Correlation	Rank	0.378 *	0.793 *	0.332 **	-0.436 *	-0.637 *	-0.299 **
		Product Moment	0.169 (1.08)	0.454 * (3.18)	0.254 (1.66)	-0.367 ** (-2.30)	-0.442 * (-3.08)	-0.544 * (-4.05)
	Regression	Relative Labour/ Relative Capital	0.093 (1.57)	0.227 * (3.56)	0.131 * (2.78)	-0.234 ** (-2.41)	-0.229 * (-2.90)	-0.247 * (-2.61)
		Constant	-1.053 * (-3.02)	-0.596 ** (-1.76)	-1.370 * (-4.01)	-3.176 * (-6.52)	-0.775 ** (-2.24)	-1.018 * (-4.40)

Notes: * denotes significant at 1% level (1 tail).

** denotes significant at 5% level (1 tail).

The correlation and regression results are based on 42 observations as certain countries had to be dropped due to non-availability of relevant data. For the modified index, the number of observations are 42 each for the sectors 652 and 851; 41 each for the sectors 844, 541 and 751; and 36 for the sector 525. The figures in parentheses denote corresponding t values. Products are indicated by SITC codes as defined in the text.

TABLE 2

CORRELATION AND REGRESSION ANALYSES FOR LABOUR INTENSIVE PRODUCTS

Index			652		844		851	
			Linear	Double log / Semi log	Linear	Double log / Semi log	Linear	Double log / Semi log
Balassa	Correlation	Rank	0.376	-	0.334	-	0.121	-
		Product Moment	0.202 (0.71)	0.358 (1.33)	0.026 (0.11)	0.307 (1.41)	0.002 (0.007)	0.062 (0.22)
	Regression	Relative Labour/ Relative Capital	0.546 (0.70)	0.249 (1.36)	0.062 (0.24)	0.239 (1.67)	0.001 (0.01)	0.028 (0.22)
		Constant	3.861 ** (1.91)	0.920 * (4.83)	7.544 ** (2.29)	1.253 * (4.97)	3.271 * (3.98)	0.838 * (4.12)
Laursen	Correlation	Rank	0.376	-	0.334	-	0.121	-
		Product Moment	0.282 (1.02)	0.256 (0.92)	0.290 (1.32)	0.477 ** (2.37)	-0.178 (-0.63)	0.158 (0.55)
	Regression	Relative Labour/ Relative Capital	0.016 (1.35)	0.169 (1.45)	0.017 (1.64)	0.351 ** (2.30)	-0.008 (-0.83)	0.093 (0.49)
		Constant	0.358 * (4.71)	-1.235 * (-5.05)	0.457 * (5.05)	-1.185 * (-4.15)	0.397 * (4.24)	-1.433 * (-4.36)
Hoen & Oosterhaven	Correlation	Rank	0.376	-	0.334	-	0.121	-
		Product Moment	0.202 (0.71)	0.328 (1.20)	0.026 (0.11)	0.411 ** (1.97)	0.002 (0.007)	0.132 (0.46)
	Regression	Relative Labour/ Relative Capital	0.002 (0.70)	0.373 (1.49)	0.0002 (0.24)	0.526 ** (2.10)	0.000006 (0.01)	0.117 (0.43)
		Constant	0.008 (1.42)	-5.812 * (-15.72)	0.017 ** (1.99)	-5.570 * (-12.18)	0.014 * (2.76)	-5.246 * (-11.44)

Index			652		844		851	
			Linear	Double log / Semi log	Linear	Double log / Semi log	Linear	Double log / Semi log
Yu et al.	Correlation	Rank	0.042	-	0.175	-	-0.284	-
		Product Moment	-0.040 (-0.14)	0.022 (0.08)	-0.093 (-0.41)	0.200 (0.89)	-0.208 (-0.74)	-0.359 (-1.33)
	Regression	Relative Labour/ Relative Capital	-0.0000012 (-0.28)	0.034 (0.10)	-0.000002 (-0.98)	0.300 (1.11)	-0.00001 (-1.54)	-0.473 (-1.28)
		Constant	0.00008 ** (1.83)	-11.170 * (-19.21)	0.00005 ** (1.97)	-11.466 * (-21.01)	0.0003 ** (1.82)	-10.018 * (-17.56)
Modified Index	Correlation	Rank	0.376	-	0.334	-	0.121	-
		Product Moment	0.260 (0.93)	0.358 (1.33)	0.199 (0.89)	0.307 (1.41)	-0.127 (-0.44)	0.062 (0.22)
	Regression	Relative Labour/ Relative Capital	0.055 (1.01)	0.249 (1.36)	0.042 (1.16)	0.239 (1.67)	-0.017 (-0.60)	0.028 (0.22)
		Constant	0.849 * (3.95)	0.920 * (4.83)	1.248 * (3.99)	1.253 * (4.97)	0.932 * (3.90)	0.838 * (4.12)

Notes: * significant at 1% level (1 tail).

** significant at 5% level (1 tail).

Rank correlation coefficients for the linear and double log form (semi log form for the modified index) are equivalent and hence not reported for the latter case. Number of observations are 14 each for sectors 652 and 851. Number of observations for sector 844 is 21. The figures in parentheses denote corresponding t values. Products are indicated by SITC codes as defined in the text.

TABLE 3

CORRELATION AND REGRESSION ANALYSES FOR CAPITAL INTENSIVE PRODUCTS

Index			525		541		751	
			Linear	Double log / Semi log	Linear	Double log / Semi log	Linear	Double log / Semi log
Balassa	Correlation	Rank	-0.488 *	-	-0.393 **	-	-0.461 *	-
		Product Moment	-0.218 (-1.36)	-0.404 * (-2.69)	-0.269 (-1.50)	-0.349 ** (-2.01)	-0.342 ** (-2.12)	-0.614 * (-4.54)
	Regression	Relative Labour/ Relative Capital	-0.0094 ** (-2.40)	-0.60 ** (-2.39)	-0.018 * (-2.52)	-0.430 ** (-2.40)	-0.021 * (-3.83)	-0.674 * (-3.26)
		Constant	0.112 * (3.19)	-4.38 * (-11.52)	0.342 * (5.44)	-2.012 * (-6.28)	0.331 * (7.25)	-2.180 * (-7.91)
Laursen	Correlation	Rank	0.488 *	-	0.393 **	-	0.461 *	-
		Product Moment	0.236 (1.48)	0.272 ** (1.72)	0.307 ** (1.74)	0.282 (1.58)	0.404 * (2.58)	0.279 ** (1.69)
	Regression	Relative Labour/ Relative Capital	0.013 * (2.53)	0.071 ** (2.00)	0.022 * (2.68)	0.193 (1.51)	0.028 * (4.06)	0.125 ** (2.07)
		Constant	0.840 * (18.57)	-0.179 * (-3.10)	0.555 * (8.49)	-0.790 * (-4.16)	0.550 * (11.09)	-0.619 * (-5.46)
Hoen & Oosterhaven	Correlation	Rank	0.488 *	-	0.393 **	-	0.461 *	-
		Product Moment	0.218 (1.36)	0.259 (1.63)	0.269 (1.50)	0.264 (1.47)	0.342 ** (2.12)	0.231 (1.38)
	Regression	Relative Labour/ Relative Capital	0.000007 ** (2.40)	0.044 ** (1.89)	0.0001 * (2.52)	0.147 (1.38)	0.00003 * (3.83)	0.076 ** (1.75)
		Constant	0.0006 * (25.37)	-7.353 * (-195.90)	0.004 * (10.48)	-5.606 * (-35.69)	0.0010 * (14.66)	-6.933 * (-82.55)

Index			525		541		751	
			Linear	Double log / Semi log	Linear	Double log / Semi log	Linear	Double log / Semi log
Yu *et al.*	Correlation	Rank	-0.437 *	-	-0.221	-	-0.526 *	-
		Product Moment	-0.317 ** (-2.03)	-0.403 * (-2.68)	-0.213 (-1.17)	-0.210 (-1.16)	-0.347 ** (-2.16)	-0.469 * (-3.10)
	Regression	Relative Labour/ Relative Capital	-0.0000006 * (-3.77)	-0.490* (-3.69)	-0.000003 ** (-2.09)	-0.255 ** (-1.71)	-0.0000009 * (-3.49)	-0.548 * (-4.16)
		Constant	0.0000009 * (5.70)	-12.860 * (-45.37)	0.0000427 * (3.33)	-11.321 * (-34.80)	0.00001 * (5.60)	-12.608 * (-45.55)
Modified Index	Correlation	Rank	0.419 *	-	0.401 **	-	0.443 *	-
		Product Moment	0.345 ** (2.05)	0.404 * (2.46)	0.336 ** (1.89)	0.349 ** (1.97)	0.640 * (4.78)	0.614 * (4.47)
	Regression	Relative Labour/ Relative Capital	0.185 ** (2.13)	0.60 ** (2.39)	0.141 ** (1.82)	0.430 ** (2.40)	0.257 * (2.62)	0.674 * (3.26)
		Constant	3.701 * (8.38)	4.382 * (11.52)	1.638 * (4.03)	2.012 * (6.28)	1.360 * (6.31)	2.180 * (7.91)

Notes: * significant at 1% level (1tail).

** denotes significant at 5% level (1 tail).

Rank correlation coefficients for the linear and double log form (semi log form for the modified index) are equivalent and hence not reported for the latter case. Number of observations are 39 for the sector 525; 36 for the sector 751; and 31 for the sector 541. For the modified index, the number of observations are 33 for the sector 525; 35 for the sector 751; and 30 for the sector 541. The figures in parentheses denote corresponding t values. Products are indicated by SITC codes as defined in the text.

In Table 1, rank correlation coefficients have the expected signs and are significant for all cases except for the products with SITC codes 525 and 751 in the case of the index of Yu *et al*. The product moment correlation coefficients are however significant with expected signs only in the case of sectors 844, 851, 541 and 751 for the index of Laursen. Since, parametric tests based on actual values of the variables are more powerful than non-parametric tests, and also that the number of observations is reasonably large, the results generated by the product moment correlations could be attached more importance. This in effect implies the index of Laursen could be considered to be reasonably consistent with the hypothesis 1. Analysis of bivariate regression coefficients would reveal that the index of Laursen generates highest number of significant coefficients with expected signs. Hence, examining the correlation and regression results, the RCA index of Laursen seems to be largely consistent with the stated hypotheses among the first four indices.

The modified index however performs at par with the index of Laursen in terms of the number of significant correlation and regression coefficients with expected signs.

As per the results presented in Table 2, the rank correlation coefficients are insignificant in all cases. The corresponding product moment correlation coefficients and bivariate regression coefficients are insignificant in most cases, except for the indices of Laursen and Hoen and Oosterhaven, for the sector with SITC code 844, with log transformations of the variables.

The modified index also does not seem to generate significant coefficients for any of the labour intensive sectors.

Table 3 shows, for the indices of Balassa, Laursen and Hoen and Oosterhaven, rank correlation coefficients are significant with expected signs for all the sectors considered. The product moment correlation coefficients are significant in most cases for the indices of Balassa and Laursen. The index of Yu *et al*. however generates results which are significantly different from the stated hypothesis of positive association. Here again, as the number of observations for each considered sector is large, the results of parametric correlation test (which are more powerful than non-parametric correlation test) could be given preference. The regression coefficients are significant with expected sign for all cases for the index of Balassa, followed by the

indices of Laursen and Hoen and Oosterhaven. Hence, integrating the results of parametric correlation and regression analyses, the indices of Balassa and Laursen may be recommended for cross sectional analysis incorporating countries with comparative disadvantage in a product. But the fact that the distribution of the index of Balassa is asymmetric, cannot be ignored. Taking this point into consideration, the index of Laursen may be judged to be better than the index of Balassa.

The modified index however performs better than the index of Laursen in terms of the number of significant parametric correlation and regression coefficients with expected sign.

Incidentally, to test the reliability of the t statistics based on which testing of hypotheses have been performed in the above analyses, normal probability plots of the residuals estimated from the bivariate regressions, particularly in cases where the coefficients were significant, were attempted.[13] A graphical representation of the standardized normal probability plots showed, for the linear regressions, the index of Laursen generates, with a few exceptions, normally distributed residuals. The other three indices predominantly seem to generate non-normally distributed residuals. For the regressions in double log form, the index of Balassa generates normally distributed residuals in all applicable cases. The index of Yu *et al.* also seems to perform well in some cases.

The standardized normal probability plots for the modified index, supported the normality assumption in all the cases where the regression coefficients are significant. Better performance of the modified index compared to the other four indices from the standpoint of normality assumption is quite expected, since this index is in log form and economic variables like exports, are mostly log-normally distributed. As an illustration, the plots corresponding to each index for a few sectors is provided in the Appendix.

4. Conclusion

Empirical analyses, based on cross sectional data, of the existing and modified indices of RCA, were attempted in order to test their consistency with the Heckscher-Ohlin theory on comparative advantage of countries. It revealed that the index of Laursen (1998) performed empirically well for the composite group of countries as

also for the group of countries with comparative disadvantage in a product. However, its structural features are not entirely satisfactory. The index of Yu *et al.* although is structurally superior to the index of Laursen, it does not generate results consistent with the hypotheses in most cases. The modified RCA index has reasonably good structural features and at the same time generated results consistent with the Heckscher-Ohlin theory, particularly for the composite group of countries and for countries with comparative disadvantage in a product. For the group of countries with comparative advantage in a product, none of the indices generate encouraging results. Small number of observations could be a probable reason for this. Thus, taking into consideration the structural advantages and empirical findings in different situations the suggested modified RCA index seems to be a reasonable choice for examining the revealed comparative advantage of countries.

NOTES

[1] Refer to Vollrath (1991).

[2] Hillman (1980) showed, 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. He, however, takes note of the fact that at higher levels of commodity aggregation, the share of a commodity in the country's total exports would be closer to unity which could result in inconsistency between Balassa's RCA index and the comparative advantage as indicated by the pre-trade relative prices. Hence, finer the level of commodity disaggregation, greater the possibility that Balassa's RCA index would be consistent with the theory.

[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 could 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, p. 62).

[4] If the reference group of countries consists of the world less country i, the lower and the upper limits of the distribution would 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 would be as mentioned in the text. Use of entire world as the group, ensures comparability of the index values across country.

[5] If the reference group of countries consists of the world less country i, the lower and the upper limits of the distribution would be exactly equal to -1 and +1 respectively, with zero being the comparative advantage neutral point.

[6] Though Vollrath (1991) suggested an index similar to the one discussed in this section, the reference group of countries and commodities were rest of the world and rest of the commodities exported respectively. However, the modified index suggested in this section considers the entire world to be the reference group of countries and all commodities to the reference group of commodities. This consideration inevitably permits comparison of index values across countries and commodities.

[7] Laursen (1998) has in fact identified this problem with the logarithmic transformation of Balassa's RCA index.

[8] Several other indices have been suggested in the literature for measuring comparative advantage of countries, such as those by Proudman and Redding (1998) and Vollrath (1991). The RCA index of Proudman and Redding (1998) is not considered for analysis in this paper as the index is very similar to the

Balassa's index, except for the normalization term. It does not address most of the problems associated with the empirical examination of the Balassa's index, other than the issue of cross commodity comparability. The indices of Vollrath (1991) have also not been included because some of them depend upon import values and as pointed by Balassa, data on country imports could be highly distorted due to the incidence of subsidies, quotas and other import restrictions. Hence, the indices might not reflect the true comparative advantage of countries. Moreover, since Vollrath insisted on having rest of the world and rest of commodities as the reference group, the indices would not be suitable for cross sectional analysis as attempted in this section.

[9] The group of 47 countries include all the ASEAN countries, SAARC countries excluding Afghanistan; other Asian countries of China and Japan; the European countries of Austria, Belgium, Bulgaria, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland, and United Kingdom; Eurasian country of Turkey; Australia and New Zealand from Oceania; African countries of Kenya and Mauritius; Guatemala, Mexico, Peru and United States.

[10] SITC stands for Standard International Trade Classification. The classification of commodities into labour intensive or capital intensive sectors is based on UNCTAD *Trade and Development Report 2002*. The report classifies the sectors 652, 844 and 851 as 'labour intensive and resource-based manufactures.' The sectors 525, 541 and 751 are classified as 'manufactures with high skill and technology intensity' in the report and have been considered to be capital intensive sectors in this paper. This classification by UNCTAD has been made on the basis of mix of different skills, technology and capital intensities and scale characteristics.

[11] The parametric correlation and regression analyses would especially be relevant for sectors, where the sample size is large. In the case of two out of three labour intensive sectors, the sample sizes get reasonably reduced (about 14 observations) when countries with only comparative advantage are considered. In all other cases, since the sample sizes vary from 21 to 42, the parametric correlation and regression results can be relied upon.

[12] The modified index is correlated with and regressed upon log of the ratio of relative labour force to relative gross capital formation.

[13] Laursen (1998) tested the normality of estimated residuals from a fitted bivariate regression for his RCA index using Jarque Bera test statistics. He proved that for his index, the residuals are normally distributed for cases greater than that for the index of Balassa. But here to test the normality of residuals the normal probability plots have been used as the number of observations is not sufficiently large.

REFERENCES

Balassa, B. (1965), "Trade Liberalization and 'Revealed' Comparative Advantage", *Manchester School of Economics and Social Studies*, Vol. 33, pp. 99-123.

Benedictis, L.D. and M. Tamberi (2001), "A Note on Balassa Index of Revealed Comparative Advantage", Universitš di Macerata and Universitš di Ancona, Italy, Working Paper 158.

Davis, D.R. (1995), "Intra-Industry Trade: A Heckscher-Ohlin-Ricardo Approach", *Journal of International Economics*, Vol. 39, pp. 201-226.

Falvey, R.E. (1981), "Commercial Policy and Intra-Industry Trade", *Journal of International Economics*, Vol. 11, pp. 495-511.

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.

Hoen, A.R. and J. Oosterhaven (2006), "On the Measurement of Comparative Advantage", *The Annals of Regional Science*, Vol. 40, pp. 677-691.

Kunimoto, K. (1977), "Typology of Trade Intensity Indices", *Hitotsubashi Journal of Economics*, Vol. 17, pp. 15-32.

Laursen, K. (1998), "Revealed Comparative Advantage and the Alternatives as Measures of International Specialization", Danish Research Unit for Industrial Dynamics Working Paper 98-30.

Liesner, H.H. (1958), "The European Common Market and British Industry", *The Economic Journal*, Vol. 68, pp. 302-316.

Proudman, J. and S. Redding (1998), *Openness and Growth*, The Bank of England, UK.

Vollrath, T.L. (1991), "A Theoretical Evaluation of Alternative Trade Intensity Measures of Revealed Comparative Advantage", *Weltwirtschaftliches Archiv*, Vol. 127, pp. 265-280.

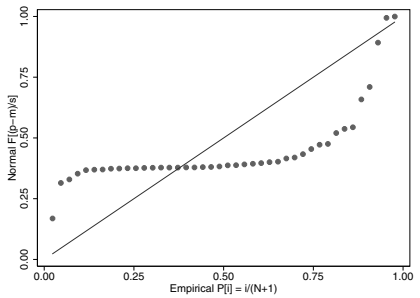
Yeats, A.J. (1985), "On the Appropriate Interpretation of the Revealed Comparative Advantage Index: Implications of a Methodology Based on Industry Sector Analysis", *Weltwirtschaftliches Archiv*, Vol. 121, pp. 61-73.

Yu, R., J. Cai and P. Leung (2008), "The Normalized Revealed Comparative Advantage Index", *The Annals of Regional Science*, Vol. 43, pp. 267-282.

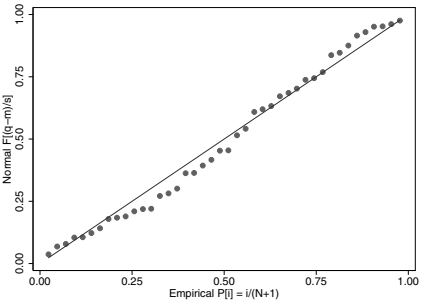
APPENDIX

STANDARDISED NORMAL PROBALITY PLOTS OF ESTIMATED RESIDUALS FROM THE REGRESSION OF RCA INDICES ON RELATIVE FACTOR INTENSITY FOR THE SECTOR 844 (CORRESPONDING TO THE RESULTS PRESENTED IN TABLE 1)

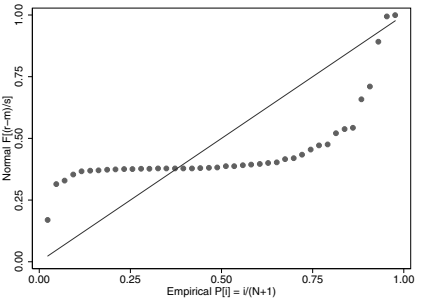
Balassa



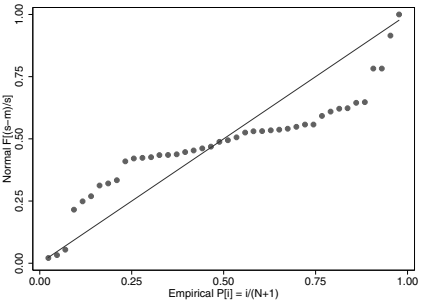
Laursen



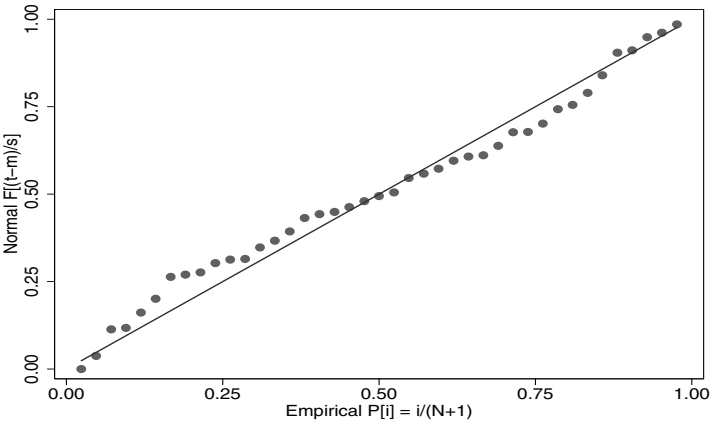
Hoen & Oosterhaven



Yu et al.



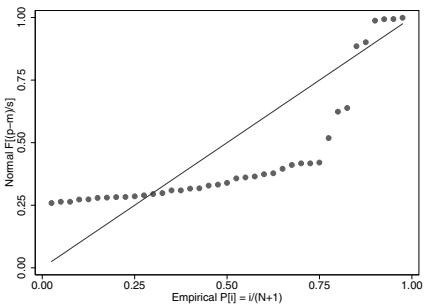
Modified Index



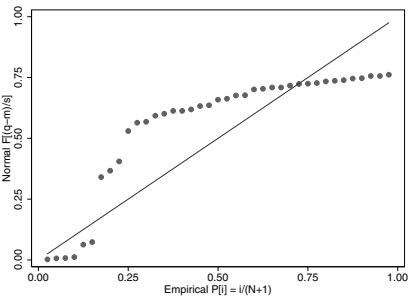
Note: p. q, r, s and t are estimated residuals from the corresponding bivariate regressions. Plots using STATA 10.

STANDARDISED NORMAL PROBALITY PLOTS OF ESTIMATED RESIDUALS FROM THE REGRESSION OF RCD INDICES ON RELATIVE FACTOR INTENSITY FOR THE SECTOR 525 (CORRESPONDING TO THE RESULTS PRESENTED IN TABLE 3)

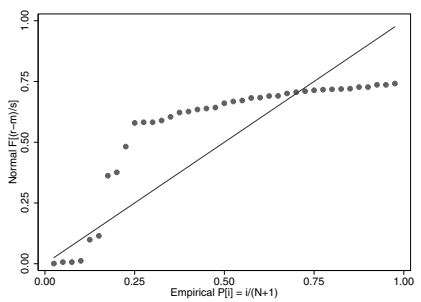
Balassa (Linear)



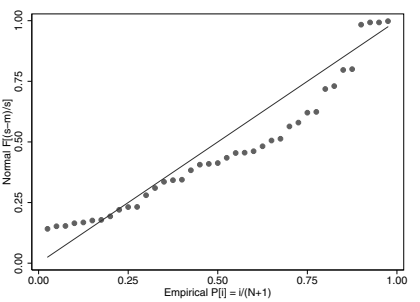
Laursen (Linear)



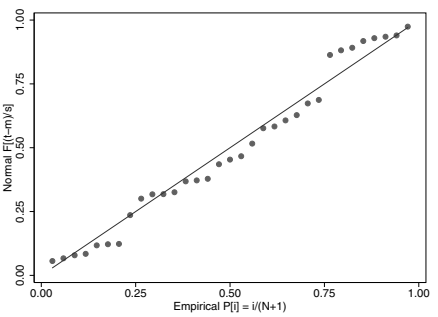
Hoen & Oosterhaven (Linear)



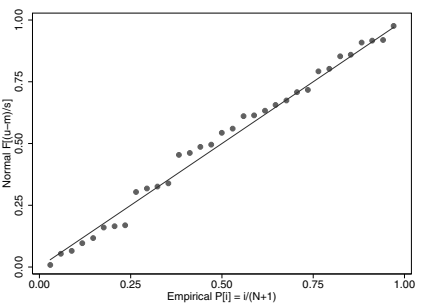
Yu et al. (Linear)



Modified Index (Linear)



Balassa (Double log)



Note: p. q, r, s, t and u are estimated residuals from the corresponding bivariate regressions. Plots using STATA 10.