

# A note on the Balassa Index of Revealed Comparative Advantage

Luca De Benedictis\* and Massimo Tamberi†

15 December 2001

## Abstract

It is customary in the empirical trade literature to analyze specialization patterns of countries using Revealed Comparative Advantage (RCA) measures. This paper explores the pros and cons of the most commonly used RCA index: the Balassa (1965) Index of RCA.

After describing the properties of the Balassa Index, we analyze two different normalizations of the original Index recently proposed by Laursen (2000) and Proudman and Redding (1998) to remedy some of the shortcomings of the original index, and we propose an alternative and more traditional strategy in order to analyze the dynamics of specialization.

**KEYWORDS:** International Trade, Specialization, Revealed Comparative Advantage.

**JEL CLASSIFICATION:** C10, F14

---

\*European University Institute and DIFE - Università di Macerata - Italy. debene@unimc.it

†DE - Università di Ancona - Italy. tamberi@deanovell.unian.it . We wish to thank (without implicating) Riccardo "Jack" Lucchetti, Francesco Palumbo, Søren Johansen, the Rerum group and seminar participants at the University of Ancona, University of LIUC - Castellanza, the Ente Einaudi for their helpful comments. Luca De Benedictis wishes to thank the EUI R.Schuman Centre for Advanced Studies for supporting his research with a 2000 J.Monnet Fellowship. Both authors wish to thank the MURST for financial support (1999 - 9913163117-005).

# Contents

<b>1</b>	<b>Introduction: Comparative advantage and revelation.</b>	<b>3</b>
<b>2</b>	<b>The Balassa Index of Revealed Comparative Advantage.</b>	<b>4</b>
2.1	Two major points and a note of caution . . . . .	8
2.1.1	Measuring what ? . . . . .	8
2.1.2	Point 1: Upper bound variability and asymmetry. . . . .	8
2.1.3	Point 2: Compositional effect and Mean variability. . . . .	12
2.2	Two alternative normalizations. . . . .	14
2.2.1	Fixing the mean: Proudman and Redding (1998). . . . .	14
2.2.2	Symmetry: Laursen (2000). . . . .	18
<b>3</b>	<b>Back to the origin ... with some caution.</b>	<b>22</b>
3.1	The dynamics of RCA . . . . .	22
3.2	Stationarity . . . . .	25
3.3	Concentration . . . . .	27
<b>4</b>	<b>Summary and conclusion.</b>	<b>29</b>
	<b>Appendix 1: The <i>TradeCan</i> Database.</b>	<b>31</b>
	<b>Appendix 2: The three dimensions of <i>BI</i>: sectors, countries and time.</b>	<b>32</b>
	<b>References.</b>	<b>37</b>

# 1 Introduction: Comparative advantage and revelation.

How to predict of international trade flows is one of the main issues in economic theory. Although the emergence of the *New Trade Theory* in the 1980s attributed comparative advantage and economies of scale equal importance as explanations of why countries trade, as far as trade flows prediction between similar and different (in factor endowments or technology) countries is concerned, comparative advantage is still the main theoretical explanation (Davies, 1997).

Under free trade, countries will specialize in and be net exporters of goods in which they have a comparative advantage. Theoretically, taking account of both static demand and supply effects, the following proposition is clear and sufficiently general (Deardorff, 1980): to identify in which good or industry a country has a comparative advantage one needs only to observe the sign of the difference between autarkic and free trade relative prices; if the sign is positive the country has a comparative advantage in the production and export of that particular good; if the sign is negative the country has a comparative disadvantage. When we move from theory to measurement, however, a major problem arises.

Relative autarkic prices are unobservable variables, and this unobservability hampers the identification of true or shadow comparative advantages. To overcome this obstacle, it has become customary practice in the empirical trade literature to analyze specialization patterns of countries using Revealed Comparative Advantage (RCA) measures. Although there have been attempts to link RCA to the theory of comparative advantage (Hillman, 1980; Bowen, 1983), the usual approach is to compare national sectoral shares with their international analogs and to infer the existence of comparative advantage through the examination of actual output and/or trade flows, as done by Balassa (1965), and Hoover (1936) and Liesner (1958) before him. The implication of that approach is that there can be as many RCA indexes as there are the combinations and transformations of the variables (production, imports and export sectoral data, more or less aggregated) used to infer comparative advantage (Vollrath, 1991).

The greater independence of measurement issues from theory<sup>1</sup> implies more

---

<sup>1</sup>A partial contribution to the disentangling of measurement from theory has been made by the theory itself. The recent literature on overlapping-generations (Galor, 1992; Mountford, 1998) and on oligopoly theory (Cordella, 1998) indicates that the autarkic equilibrium is not necessarily a good predictor of long run equilibrium when the economy opens up to trade. This finding implies that even if autarkic relative prices were observable they might not predict true comparative advantage.

freedom in the selection of the specific index of RCA to use in applied research on international specialization patterns, but it also requires greater awareness of the implications of that selection.

The issue is all the more important given the substantial amount of research that has followed the advent of trade models of agglomeration and diffusion of economic activity (Fujita et al., 1999) and their influence on the analysis of the stability of specialization patterns, especially in Europe. Amiti (1998, 1999); Brulhart and Torstensson (1996); Brulhart (2001); Laursen (2000); Proudman and Redding (1998); Sapir (1996) have studied if and how the process of economic integration among the European Union member states confirms the prediction of Economic Geography models: a decrease in transport costs will generate a non-monotonic tendency toward agglomeration and specialization.

In spite of the different answers given to the initial question, many of these studies use an approach based on a RCA index. This paper concentrates on their shared starting point. It explores the pros and cons of the most commonly used (also by the studies cited above) of the RCA indexes: the Balassa (1965) Index of RCA.

After describing the statistical properties of the Balassa Index, we analyze (and criticize) two different normalizations of the original Index, recently proposed by Laursen (2000) and Proudman and Redding (1998) to remedy some of its shortcomings. We propose that a step back should be taken to the original *BI*, and to a more traditional and general strategy of analysis when dealing with RCA and the dynamics of specialization.

## 2 The Balassa Index of Revealed Comparative Advantage.

Restricting the analysis to global measures of RCA, and therefore excluding bilateral indexes, the various measures that have been proposed to infer the existence of comparative advantage from actual data can be roughly classified according to the variable used<sup>2</sup>: exports,  $X$ ; imports,  $M$ ; and net trade,  $(X - M)$ . We will focus our analysis only on the first variable<sup>3</sup>.

---

<sup>2</sup>Ballance, Forstner, and Murray (1987) proposed two classes of indices: the *trade-only* class and the *trade-cum-production* class (with four indices,  $X/Y$ , the ratio of exports to production,  $M/C$ , the ratio of imports to consumption,  $(X - M)/Y$ , the ratio of net trade to production and some other indices based on Bowen's definition of comparative advantage neutrality (Bowen, 1983)). We will consider only the former class of indices.

<sup>3</sup>For a comparison between indexes based on exports and on net trade see Ballance, Forstner, and Murray (1987), Bowen (1983), and Laursen (1998). For an application to

The first and still most widely used RCA measure built on exports as the only information variable is the Balassa (1965) Index. If we use  $c$  to denote a specific country,  $w$  the world economy or the entire set of countries considered in the analysis,  $s$  a specific sector (and  $S$  the total number of sectors considered), and  $t$  the time period considered, the Balassa Index is:

$$BI_{[cst|w]} = \frac{\frac{X_{cs}}{X_c}}{\frac{X_{ws}}{X_w}} \equiv \frac{\frac{X_{cs}}{X_{ws}}}{\sum_{s=1}^S \frac{X_{cs}}{X_{ws}} \cdot \frac{X_{ws}}{X_w}} \equiv \frac{\frac{X_{cs}}{X_{ws}}}{\frac{X_c}{X_w}}. \quad (1)$$

The  $BI_{[cst|w]}$  (henceforth  $BI$ ) is a sectoral relative export measure in terms of share of world exports. It can also be interpreted<sup>4</sup> as the country's share of sectoral export,  $\frac{X_{cs}}{X_{ws}}$ , normalized by a weighted sum of exports shares in all sectors which is equivalent to  $\frac{X_c}{X_w}$ <sup>5</sup>. Since the numerator ranges from 0 (sectoral exports are nought, according to the first interpretation) to 1 (international monopoly), and the denominator - depicting the economic dimension of  $c$  - also ranges as well from 0 to 1, the  $BI$  ranges between 0 and  $\frac{X_w}{X_c}$ <sup>6</sup> and is equal to 1 when  $\frac{X_{cs}}{X_{ws}} = \frac{X_c}{X_w}$ . Therefore, the  $BI$  reveals that country  $c$  has a comparative disadvantage in sector  $s$  if  $0 < BI < 1$ , while it has a comparative advantage in sector  $s$  if  $1 < BI < \frac{X_w}{X_c}$ . The demarcation is given by each specific comparison with the relative value in the chosen benchmark. In other words, when  $BI \lesseqgtr 1$ , then  $\frac{X_{ws}}{X_{cs}} \lesseqgtr \frac{X_w}{X_c}$ ; so for a  $BI > 1$  the country displays a sectoral share greater than the same sectoral share for the totality of the set of countries considered,  $w$ .

To put it briefly, the  $BI$  follows an asymmetric distribution with a lower bound and with a variable (across countries and across time) upper bound, it has a fixed demarcation value in 1, a variable (across countries and across time) mean value,  $\mu(BI)$ , and standard deviation,  $\sigma(BI)$ .

---

the Italian case see ?. For alternative approaches see Theil (1967) and Lafay (1992).

<sup>4</sup>The Balassa index has been expressed in terms of expected export by Ballance, Forstner, and Murray (1987):

$$BI = \frac{X_{cs}}{X_{ws}} \cdot \frac{X_w}{X_c} \equiv \frac{X_{cs}}{\frac{X_c}{X_w} \cdot X_{ws}} \equiv \frac{X_{cs}}{E[X_{cs}]}$$

where  $E[X_{cs}]$  is the expected level of exports of sector  $s$  from country  $c$  assuming that the country's exports of  $s$  are in proportion to the country's share of world exports.

<sup>5</sup>The country's share of sectoral exports,  $\frac{X_{cs}}{X_{ws}}$ , is normalized by  $\frac{X_c}{X_w}$ , which is equivalent to  $\sum_{s=1}^S \frac{X_{cs}}{X_w} \cdot \frac{X_{ws}}{X_{ws}}$ ; that is, a weighted sum of export shares in all sectors, where the weights are the total sectoral shares,  $\frac{X_{ws}}{X_w}$ .

<sup>6</sup>The general definition of the  $BI$  states that the index ranges from 0 to  $+\infty$ , but the effective upper bound is  $\frac{X_w}{X_c}$ , which tends to  $\infty$  when  $X_c$  tends to 0, or in other words, when the economic weight of country  $c$  in terms of exports is irrelevant.

Following Ballance, Forstner, and Murray (1987), it is possible to give a threefold interpretation to the Balassa Index. The most common interpretation of it is that each index provides a demarcation between countries that reveal a comparative advantage in a particular sector and those countries that do not. Secondly, it quantifies the sector-specific degree of comparative advantage enjoyed by one country with respect to any other country or set of countries. Thirdly, the index generates possible cross-country (with respect to a sector) or cross-sector (inside a country) rankings, ordered according to the specific value of the *BI*.

Let us take as an example the data in Table 1. These represent an exemplary subset of the dataset used throughout the paper and drawn from the ECLAC(UN))-World Bank *TradeCan* Database (1999). The characteristics of the database are discussed in Appendix 1.

Table 1 gives the 1986 and 1996 *BI* values for three OECD countries (Italy, France and Germany). Of the 538 manufacturing sectors in the *TradeCan* Database (1999), we have chosen three sectors identified by their four-digit Sitc (rev. 2) code: Fabrics, woven, of sheep's or lamb's wool or of fine animal hair (6543), Perfumery, cosmetics and toilet preparations (excluding soaps) (5530), Articles of iron or steel, n.e.s. (6997).

	Italy		France		Germany	
<b>Sitc (rev.2)</b>	1986	1996	1986	1996	1986	1996
<b>6543</b>	10.8	11.4	0.88	1.07	0.05	0.06
<b>5530</b>	0.71	1.01	5.86	5.63	0.01	0.02
<b>6997</b>	1.21	1.27	1.04	0.84	6.95	4.59
$\frac{X_w}{X_c}$	17.92	21.17	15.82	18.10	12.12	12.76
$\frac{X_c}{X_w}$	0.056	0.047	0.063	0.055	0.083	0.078
$\mu(BI)$	1.15	1.23	1.09	1.10	1.42	1.17
$\sigma(BI)$	1.28	1.40	0.87	1.08	1.23	1.02

Table 1: Balassa Index

In addition to the *BI* values (rows 1 to 3), the table comprises four more rows containing the upper bound of the *BI* ( $\frac{X_w}{X_c}$ ) and its reciprocal, the national share of world export flows, the sectoral arithmetic mean of the entire distribution,  $\mu(BI)$ , and the standard deviation,  $\sigma(BI)$ .

For the moment, we shall use Table 1 to elaborate on the three possible interpretations of the *BI* outlined above: quantification, ranking, and demarcation.

Starting with demarcation, the *BI* reveals that in 1986 Italy had a com-

parative advantage in the 6543 textile sector, while France and Germany did not (although France reversed its position in 1996); France had a comparative advantage in the 5530 cosmetics sector, while Germany and the Italy did not (but Italy was slightly specialized in 1996); and all the three countries explicitly considered in Table 1 revealed, in 1986, a comparative advantage in the 6997 iron and steel sector, although they had *different* degrees of specialization. In 1996 only Italy and Germany maintained an advantage in this sector, with Germany displaying a lower degree of specialization than in 1986.

As regards ranking, as said, the *BI* has been used for both cross-sector comparisons and cross-country comparisons (for an analytical treatment see Appendix 2). On the one hand, it reveals that in both years Italy had a greater comparative advantage in the textiles sector than in the iron and steel sector<sup>7</sup>. On the other hand, in 1986, Germany had a greater comparative advantage in the iron and steel sector than Italy, and that the Italian comparative advantage was greater than the French one<sup>8</sup>.

One can go even further if quantification is taken into account. What the *BI* shows is that, e.g., in 1986, the share of Italian exports in the textiles sector was proportional to the world share to an order of magnitude of 10.8; and that in 1986 the share of the Italian textiles sector with respect to the Italian cosmetics sector (the relative contribution to total export) was  $\frac{10.8}{0.71} = 15.2$  times greater than for the sum of the countries considered in the sample set; and, finally, that in 1986 the share of the Italian textiles sector was  $\frac{10.8}{0.05} = 216$  times greater than the same share for Germany. Therefore, using Table 1 in a cardinal way allows the information content of raw export data to be preserved, adding the possibility of demarcation and of sectoral and country ranking. But when the *BI* is used as a cardinal measure, some problems arise. We shall explicitly consider two of these problems: asymmetry (variability of the upper bound) and across-time ranking (variability of the mean value). Before doing so, however, we shall start the next section with a note on the measurement of the *BI* and on its meaning. The content of this note is self evident and only needs a brief mention. But since it is one of the main obstacles against comparison among studies we have decided to give some emphasis to it.

---

<sup>7</sup>If only two hypothetical sectors - sector 1 and sector 2 of country *c* - are considered, it is always true that  $BI_{c2} \equiv BI_{c1} \cdot \frac{X_{cs2}}{X_{ws2}} \cdot \frac{X_{ws1}}{X_{cs1}}$ ; therefore,  $BI_{c2} > BI_{c1} \iff \frac{X_{cs2}}{X_{ws2}} > \frac{X_{cs1}}{X_{ws1}}$ .

<sup>8</sup>If only two hypothetical countries - country 1 and country 2 - and one sector *s* are considered, it is always true that  $BI_{2s} \equiv BI_{1s} \cdot \frac{X_{c2s}}{X_{c2}} \cdot \frac{X_{c1}}{X_{c1s}}$ ; therefore,  $BI_{2s} > BI_{1s} \iff \frac{X_{c2s}}{X_{c2}} > \frac{X_{c1s}}{X_{c1}}$ .

## 2.1 Two major points and a note of caution

### 2.1.1 Measuring what ?

As said, when discussing equation 1, the *BI* is a measure of country *c*'s export share in sector *s* with respect to some *particular* aggregation of sectors ( $X_c$ ), relative to some *particular* benchmark of countries ( $X_w$ ), at a particular instant of time *t*, and the choice of sectoral aggregation and of the benchmark and the way they are calculated influence the value of the *BI* obtained.

Revealing comparative advantage depends on how the *BI* has been measured<sup>9</sup>. Comparison among different analyses is possible (and makes sense) only if they share the same focus in terms of sectoral aggregation and benchmark. Analysis of Italian RCA will (of course) yield different results if the sectors considered are manufactures (as in our analysis), goods, goods and services, or any other possible sectoral aggregation, and if the benchmark chosen is the EU, the OECD countries, the world, or any other possible aggregation of countries<sup>10</sup>.

### 2.1.2 Point 1: Upper bound variability and asymmetry.

As pointed out in section 2, the upper bound of the *BI* is not fixed, and it inversely depends on the share of country exports on world exports. This dependence on *size* may cause confusion in the interpretation of cross-country comparisons<sup>11</sup>. Let us use again Table 1 as an example. It is evident that the upper bound of the *BI* ( $\frac{X_w}{X_c}$ : row 4) varies across countries and, for each country, across time. In 1986, the Italian upper bound was a slightly less than 1.5 times the German upper bound (17.92 vs 12.12). Now note that the *BI* associated with the Italian clothing sector was, in 1996, almost as high as the German upper bound: a value that could be reached by Germany only in the case of a sectoral international monopoly. But if different cardinal values of the *BI* imply very different economic conditions for two countries like Germany and Italy, this raises questions about the international comparison

---

<sup>9</sup>Another important issue is the effect of the particular exchange rate chosen on revealing comparative advantages. Since the *BI* is a ratio, the choice of the numeraire is irrelevant, but when the individual components of the *BI* are examined, and  $X_{ws}$  and  $X_w$  are calculated, the choice of the numeraire becomes non neutral.

<sup>10</sup>See Richardson and Zhang (2001) for a similar view.

<sup>11</sup>Consider the following illustrative example. Suppose that two countries share the total world export of a good *s*; their world export shares, in the considered sector, be equal: 50% for both; if the two countries are very different in *size* ( $\frac{X_c}{X_w}$ ), the *BI* will make one of them appear much more specialized than the other. From many perspectives, this can be misleading.



- the real informational economic content - of  $BI$ <sup>12</sup>.

The explanation for the lack of clarity<sup>13</sup> associated with the numerical value *per se* of the  $BI$  is that this same value - as described in equation 1 - can be determined by the influence of two different factors. The first one is the level (and the change in the same level) of the economy's overall export share,  $\frac{X_c}{X_w}$ , which is the denominator of the  $BI$ . The second factor is the level and change of the sectoral export share,  $\frac{X_{cs}}{X_{ws}}$ , which is the numerator. A high value of the  $BI$  may result from a high numerator, a low denominator, or a combination of the two; and one should always be careful in determining the relative weight of both factors in the case of dynamic international comparisons<sup>14</sup>. Although it should be clearly stated that the dynamics of specialization is the result of the combination of sectoral and overall changes, from an analytical viewpoint it is useful to control *both* the total effect (the change in  $BI$ ) and the sectoral and overall effect separately (the change in the components of the  $BI$ ).

A second issue relates to the asymmetry of the  $BI$ . Observe the characteristics of the probability density function of the  $BI$  for Italy, France, Japan and Germany, in 1986 and 1996.

In the cases described in Table 1, and shown in Figure 1 with the inclusion of Japan, the mean,  $\mu(BI)$ , is above the median,  $m(BI)$ , so that the distribution is skewed to the right; in all cases except the German 1986,  $m(BI)$  is below 1; and in most cases  $\mu(BI)$  is above 1, but cases of  $\mu(BI)$  below 1 are as well possible<sup>15</sup>. Note that the vertical lines in the four quadrants in figure 1 depict the fixed demarcation value. Finally, the distribution *is not* monotonically declining, as stated by Hinloopen and Marrewijk (2001)<sup>16</sup>, and the

---

<sup>12</sup>The  $BI$  has been criticized by Yeats (1985) from a different perspective. He observed that the  $BI$  does not reflect the comparative advantage for each country, across sectors. The correlation between the ranks across sectors and the ranks across countries (based on 1976-1978 trade data for 47 countries) was positive and significant only in 60% of the cases examined. Laursen (1998) asserts (using 1970-1993 trade data for 19 OECD countries) that Yeats' results overestimate the problem, at least when the sample consists of similar countries. He found a less than 2% insignificant correlation, and when all the data were pooled, the rank correlation coefficient was 0.80.

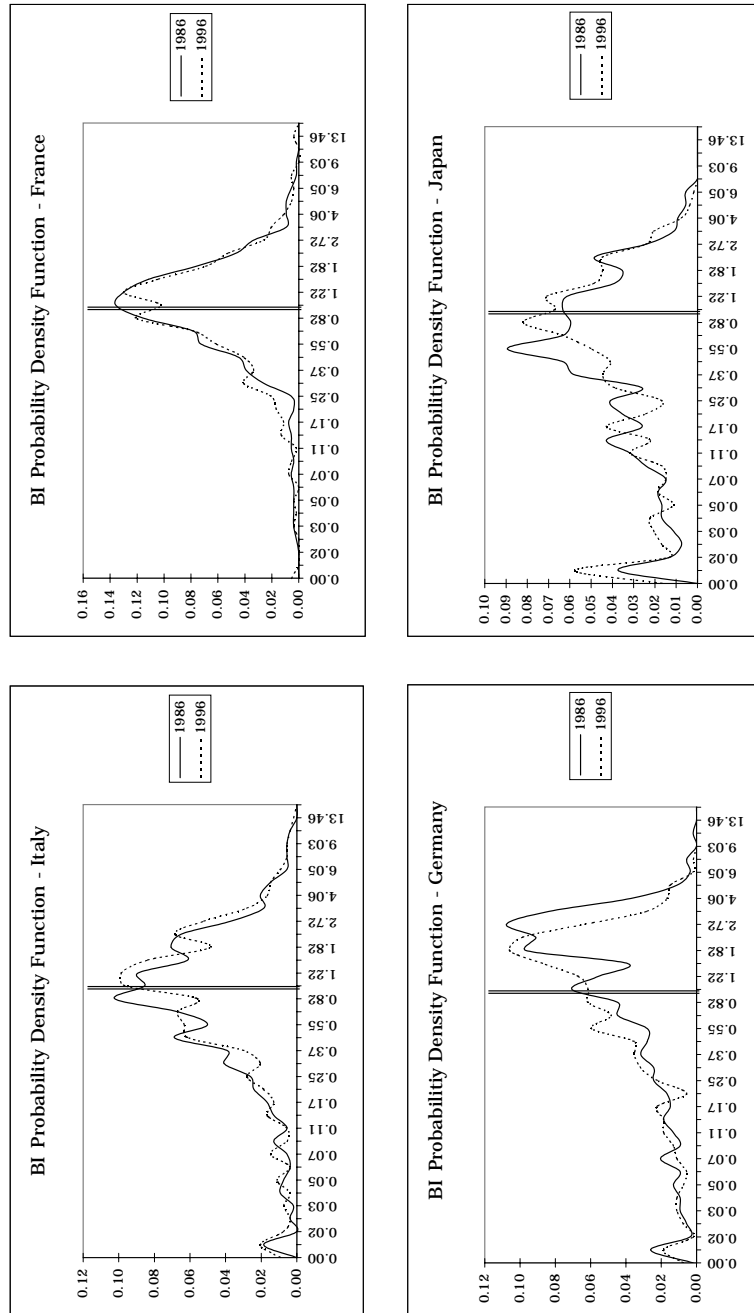
<sup>13</sup>We would emphasize that the lack of clarity is associated *not* with the cross-country relation but with the informational content of the  $BI$  in different countries. As far as the former is concerned, as stated in footnote 8 and fully developed in Appendix 2, the cross-country relation associated with the  $BI$  is very precise:  $BI_{2s} > BI_{1s} \iff \frac{X_{2s}}{X_2} > \frac{X_{1s}}{X_1}$ .

<sup>14</sup>Formally, if we write  $BI = \frac{X_{cs}}{X_c} = \frac{a}{b(a)}$ , the total derivative of  $BI$  is  $dBI = \frac{1}{b^2} \cdot (b \cdot da - a \cdot \frac{db}{da})$ .

<sup>15</sup>The most significant cases are the UK, Japan and the US, although also Korea, Mexico, Thailand, and Taiwan all have a mean value below 1.

<sup>16</sup>The shape of the Probability Density Function crucially depends on: (1) the weight of

Figure 1: PDF - Italy (a), France (b), Germany (c) and Japan (d)



density is higher in the central classes (from  $0.3 < BI < 1$  and  $1 < BI < 4$ ).

Since *the* characteristic of the  $BI$  distribution is that  $\mu(BI) > m(BI)$ , changes in sectors characterized by high values of  $BI$  will be over-weighted in terms of the effect on the overall sign of the comparative advantage dynamic. Since the  $BI$  provides a demarcation between *below 1 sectors* and *above 1 sectors*, and since the demarcation is not symmetric, the relative weight attached to specialized sectors is excessive compared to the one attached to non specialized sectors. If one infers the change in the economy's average extent of specialization through the dynamics of the sectoral arithmetic mean,  $\mu(BI)$  (Proudman and Redding, 1998, p. 84), one numerically overestimates changes in sectors characterized by  $BI > 1$  with respect to changes occurring in sectors characterized by  $BI < 1$ . The evolution of a country will tend to depend mostly<sup>17</sup> on the evolution of the *above 1 sectors*.

A third issue related to the statistical properties of the distribution of the residuals when the dynamic stability of a national specialization pattern is tested by means of a *Galtonian regression*. When the final year  $BI$  is regressed on the initial year  $BI$ , as Laursen (1998) points out, the estimates tend to be biased given the lack of normality of the error terms due to the skewed distribution of the  $BI$ . We now discuss this issue in the light of the appropriateness of the alternative normalization proposed for the  $BI$  (in section 2.2.2).

---

the zero class, which itself depends on the database, the countries considered and the level of disaggregation used. With respect to the 27.5% of zeros in the Hinloopen and Marrewijk (2001) database, the average percentage of zeros is 1.2% for industrialized countries and 3.2% for LDCs; (2) the size of the classes of the PDF. This latter point requires further explanations. In figure 1 we have used 35 equiproportional classes - using a logarithmic transformation - for positive values of the  $BI$  plus a standing alone zero class. This procedure has the great advantage of generating a classification invariant across countries and across time. When the Scott algorithm was used the result was 15 classes, but the first class was between 0 and 0.7, the second one was between 0.7 and 1.4, and (for Italy) 90% of the observations were in the first four classes. When the Friedman and Diaconis algorithm was used the result was 40 classes, but only the first three classes were below 1. The cumulative density function reached 90% in the ninth class, and the first four classes contain 60% of the observation (in the case of Italy). The Doane algorithm, which is used for asymmetric distributions, gave results similar to ours.

<sup>17</sup>This tendency does not imply absolute dependence, since, as already discussed, the  $BI$  - and, therefore  $\mu(BI)$  - does not depend exclusively from the sectoral shares, but rather on the combination of sectoral effect (the change in  $\frac{X_{cs}}{X_{ws}}$ ) and overall effect (the change in  $\frac{X_c}{X_w}$ ), and the two may work in opposite directions.

### 2.1.3 Point 2: Compositional effect and Mean variability.

Let us further examine the issue of  $BI$  asymmetry in connection with the problem of across-time ranking, doing so with the help of Table 2.

Sitc (rev.2)	Italy 1986			Italy 1996		
	$\frac{X_w}{X_c}$	$\frac{X_{cs}}{X_{ws}}$	$BI$	$\frac{X_w}{X_c}$	$\frac{X_{cs}}{X_{ws}}$	$BI$
<b>6543</b>	17.92	0.6	10.8	21.17	0.53	11.4
<b>5530</b>	17.92	0.039	0.71	21.17	0.048	1.01
<b>6997</b>	17.92	0.068	1.21	21.17	0.060	1.27
$\mu(BI)$	17.92	0.064	1.15	21.17	0.058	1.23
$\sigma(BI)$	0	0.07	1.28	0	0.066	1.40

Table 2: Balassa Index Decomposition

Table 2 comprises (columns 1 and 4) the values of  $\frac{X_w}{X_c}$  and (columns 3 and 6) the values of the  $BI$  in Table 1 (row 5 and columns 1 and 2, respectively). Columns 2 and 5 now show the values of  $\frac{X_{cs}}{X_{ws}}$ . We concentrate on one single country, Italy, in 1986 and 1996.

Following our previous discussion of how sectoral and overall effects code-terminate the value of  $BI$ , we have decomposed the original index into two different elements. The first,  $\frac{X_{cs}}{X_{ws}}$ , is the national share of sectoral world trade, which varies across sectors. The second element,  $\frac{X_w}{X_c}$ , is the reciprocal of the national share of total world trade, or the upper bound of  $BI$ , which is fixed across sectors for each country. In the case of Italy it was 17.92 in 1986 (the Italian share of world exports was 5.6%) and 21.17 in 1996 (4.7% of world export). In terms of across-sector comparison, the value of  $BI$  only depends on the value of the first element (the higher the national share in sectoral world trade, the higher the value of  $BI$ ); while in terms of across-country or across-time comparison, the value of  $BI$  depends on the compositional effect of both elements (see Appendix 2 for a formal proof). The fact that between 1986 and 1996 Italy, France and Germany decreased their respective shares of world exports in different proportions influences on the values of the  $BI$ . This may eventually give rise to apparently paradoxical results. The Italian Textiles sector and Iron and Steel sector cases (Table 2, first and third row) are illuminating. Since the two are similar, we shall discuss only the latter.

As shown in Table 1, between 1986 and 1996 Italy reduced its share of world exports from 5.6% to 4.7% while its share of Iron and Steel exports fell from 6.8% to 6.0%. These events induced opposite effects on the  $BI$  value (the former increased the  $BI$  value while the latter one reduced it) that, in

this case, can be considered as paradoxical. In spite of the reduction in  $\frac{X_{cs}}{X_{ws}}$ , the *BI* value increased, and this was due precisely to the increase in  $\frac{X_w}{X_c}$ . The former effect prevailed over the latter, so that the increase in specialization in the Iron and Steel sector revealed by the Italian *BI* had little to do with the international competitiveness of the Italian Steel sector but instead reflected the pattern of overall Italian exports : competitiveness and specialization are two different and possibly contrasting phenomena.

These apparently paradoxical cases - the explanation of which lies in the definition of (R)CA as a *relative* concept - are not infrequent. In our sample, we found that 14.5%, 12.83%, and 5.2% of the sectors - in Italy, France and Germany, respectively - in the 1986-1996 interval were characterized by a relatively smaller sectoral effect (the variation of the numerator of *BI* was relatively smaller with respect to the overall effect at the denominator).

The relevance of compositional considerations also has implications for the representation of a complex phenomenon by means of synthetic indicators like the mean value.

If the arithmetic mean can represent the average extent of specialization, bearing equation 1 in mind, the fact that the country's share of sectoral exports,  $\frac{X_{cs}}{X_{ws}}$ , is normalized in the *BI* by a weighted sum of export shares in all sectors,  $\sum_{s=1}^S \frac{X_{cs}}{X_{ws}} \cdot \frac{X_{ws}}{X_w} \equiv \frac{X_c}{X_w}$ , poses - as pointed out by Proudman and Redding (1998) - a specific problem in terms of across-time comparison.

This specific normalization implies that the mean of the *BI* is time variant, as can be noted in Table 1. In the Italian case, it was 1.15 in 1986 and increased to 1.23 in 1996; in the case of France, it was 1.09 in 1986 and it increased to 1.10 in 1996; while, in the German case, it was 1.42 in 1986 and decreased to 1.17 in 1996. The mean of the *BI* - as regards the upper bound - varies across countries, and is also variant across time for the same country.

This characteristic is strictly dependent on the structure of national exports. If the economic size of *c* is relatively small ( $\frac{X_w}{X_c}$  is relatively high) and the sectoral size of *c* (in some *s*) is relatively large ( $\frac{X_{cs}}{X_{ws}}$  is relatively high), then the value of the sectoral mean of the *BIs* will tend to be above the demarcation value. The higher the upper bound and the values of the *above 1 sectors*, the higher the mean value of the *BI*.

The variability of the arithmetic mean is again the result - when analysing specialization using the *BI* - of the coexistence of the overall country effect and the sectoral effects codetermining the dynamics of RCA. But specialization *is* the result of these two combined forces.

## 2.2 Two alternative normalizations.

The difficulties of interpreting the cardinal  $BI$  across countries and across time, the specific normalization of the  $BI$ , and the fact that the index is not symmetric and its mean value is not fixed, have induced some researchers to propose various ways to "re-normalize" the original index. We shall discuss just two of them here. Neither of them is satisfactory and we will show that the attempt to remedy the shortcomings of the original  $BI$  have given rise to further ones.

### 2.2.1 Fixing the mean: Proudman and Redding (1998).

In their analysis of specialization patterns in Europe, Proudman and Redding (1998) disregard the overall country effect and focus exclusively in the changes in RCA across sectors (*intra distribution dynamics*). They do this by filtering the effect of changes in the overall (or average) level of specialization of country  $c$  and by stabilizing the mean value of the  $BI$  across time.

Proudman and Redding (1998) propose a normalization around the sectoral mean according to the following formula<sup>18</sup>:

$$BI_{PR} = \frac{\frac{X_{cs}}{X_{ws}} \cdot \frac{X_w}{X_c}}{\sum_{s=1}^S \frac{X_{cs}}{X_{ws}} \cdot \frac{X_w}{X_c} \cdot \frac{1}{S}} \equiv \frac{\frac{X_{cs}}{X_{ws}}}{\frac{1}{S} \cdot \sum_{s=1}^S \frac{X_{cs}}{X_{ws}}} \quad (2)$$

The  $BI_{PR}$  is obtained by normalizing the numerator of the  $BI$ ,  $\frac{X_{cs}}{X_{ws}}$ , not by a *weighted sum* of the same export shares, as for the  $BI$ , but instead by the *arithmetic mean* of export shares, or, equivalently, by normalizing the original  $BI$  by its cross-sectoral mean (if we divide each sectoral value contained in table 1 by the cross-sectoral mean, we obtain the values contained in Table 3).

The  $BI_{PR}$  is therefore an *alternative* measure of RCA which ranges from 0 to  $\frac{S}{\sum_{s=1}^S \frac{X_{cs}}{X_{ws}}}$ <sup>19</sup> and is equal to 1 (its demarcation value) when  $\frac{X_{cs}}{X_{ws}} = \frac{1}{S} \cdot \sum_{s=1}^S \frac{X_{cs}}{X_{ws}}$ . Country  $c$  has a comparative disadvantage in sector  $s$  if  $0 < BI_{PR} < 1$ , while it has a comparative advantage in sector  $s$  if  $BI_{PR} > 1$ ; and the same country is said to be specialized in sector  $s$  if it shows a sectoral share

<sup>18</sup>The same normalization is proposed by Proudman and Redding (1999), and, in the contest of international technological specialization, by Mancusi (2000).

<sup>19</sup>As for the  $BI$ , in theory the  $BI_{PR}$  ranges from 0 to  $+\infty$ , but the effective upper bound tends to  $+\infty$  when  $X_c \rightarrow 0$ . The upper bound of the  $BI_{PR}$  is directly linked with the  $BI$  upper bound, since it is equal to the  $BI$  upper bound divided by the mean value of the  $BI$ . Depending on the value, below 1 or above 1, of the  $BI$  mean, the  $BI_{PR}$  normalization increases or reduces the degree of asymmetry of the distribution, exacerbating or attenuating the problems discussed in section 2.1.2

	Italy		France		Germany	
Sitc (rev.2)	1986	1996	1986	1996	1986	1996
<b>6543</b>	9.39	9.27	0.81	0.97	0.03	0.05
<b>5530</b>	0.62	0.81	5.38	5.12	0.01	0.02
<b>6997</b>	1.05	1.03	0.95	0.76	4.89	3.92
$\mu(BI_{PR})$	1	1	1	1	1	1
$\sigma(BI_{PR})$	1.11	1.4	0.79	0.76	0.86	0.87

Table 3: Balassa Index (Proudman and Redding, 1998)

greater than its average sectoral share<sup>20</sup>. Given the statistical properties of the arithmetic mean, the  $BI_{PR}$  has a time invariant mean.

Even if the  $BI$  and the  $BI_{PR}$  are absolutely equivalent in terms of statistical content (since only a change of scale is involved), we suggest this transformed index has three possible defects: the first two are related to the concept itself of average level of specialization, and the third to the consistency in terms of meaning between  $BI$  and  $BI_{PR}$ .

**Dismissing the country effect.** The first defect of the  $BI_{PR}$  concerns the identification of specialization with intra (sectoral) distribution dynamics, focusing on relative changes across sectors and dismissing the relevance of overall country effects. As already pointed out,<sup>21</sup> the normalization used in the  $BI_{PR}$  hides - by construction - uniform effects in sectoral exports. It only reveals changes in the sectoral composition of exports, not the change in national exports with respect to world exports. Hence, all the issues related to the *convergence debate*, such as the leadership, catching up and falling behind of countries in terms of per capita income and productivity, are uncoupled from analysis of trade specialization: a methodological choice which seems neither historically sound nor empirically justified.

**Giving too much emphasis to the first moment of the distribution.** The second defect of the  $BI_{PR}$  concerns the implicit emphasis given to the arithmetic mean of the  $BI$  (or alternatively, to the arithmetic mean of  $\frac{X_{cs}}{X_c}$ )

<sup>20</sup>When  $\frac{X_{cs}}{X_c} > \frac{1}{S} \cdot \sum_{s=1}^S \frac{X_{cs}}{X_{ws}}$ , then  $BI_{PR} > 1$ ; and when  $\frac{X_{cs}}{X_c} < \frac{1}{S} \cdot \sum_{s=1}^S \frac{X_{cs}}{X_{ws}}$ , then  $BI_{PR} < 1$ . Since the total number of sectors  $S$  explicitly enters in equation 2, the index is sensitive to data aggregation. The choice of  $S$  influences the sectoral  $BI_{PR}$  value, the upper bound, and the mean value of the  $BI_{PR}$ .

<sup>21</sup>See the discussion by Christopher Bliss and Peter Neary in Proudman and Redding (1998, pp.97-103).

used as the normalization term. Firstly, the arithmetic mean is a very poor synthetic indicator in cases when the underlying distribution is characterized by a pronounced skewness (Chew, 1990). Secondly, the economic meaning of  $\mu(BI)$  - as well as the concept of average level of specialization - is unclear.

Recalling the definition of the  $BI$  by equation 1 we may write the mean of  $BI$  is

$$\mu(BI) = \frac{\mu\left(\frac{X_{cs}}{X_{ws}}\right)}{\mu\left(\frac{X_c}{X_w}\right)} \equiv \frac{X_w}{X_c} \cdot \mu\left(\frac{X_{cs}}{X_{ws}}\right) \quad (3)$$

and then, that

$$\mu(BI) \leq 1 \iff \mu\left(\frac{X_{cs}}{X_{ws}}\right) \leq \frac{X_c}{X_w}. \quad (4)$$

But since  $\mu\left(\frac{X_{cs}}{X_{ws}}\right)$  has no clear meaning, the same applies to  $\mu(BI)$ <sup>22</sup> and to the concept itself of the average extent of specialization (Proudman and Redding, 1998, p.84). Saying that a country is specialized *on average* because  $\mu(BI) > 1$  is meaningless, or put otherwise - referring back to table 1 - it is meaningless to say that in 1996 Italy was more specialized than Germany *on average* because its  $\mu(BI)$  value was greater.

**Meaning consistency with respect to  $BI$ .** The third defect is the loss of consistency with respect to the  $BI$ . As said, from a statistical viewpoint the  $BI$  and  $BI_{PR}$  are equivalent but relevant problems arise because of the existence of a demarcation value in 1. The information derived from the  $BI_{PR}$  may have a completely different economic content with respect to the original Index, and, in particular, the  $BI_{PR}$  may express opposite status in terms of sectoral RCA. We can describe this eventuality as follows.

Figure 2 graphically represents the relation between  $BI$  and  $BI_{PR}$  in the two cases where  $\mu(BI) \geq 1$  and  $\mu(BI) \leq 1$ . Given that in general  $\mu(BI) \neq 1$ , all sectors between the mean itself and the demarcation value, in the original index, will change their status when transformed into  $BI_{PR}$ : if  $\mu(BI) \leq 1$  all the non-specialized sectors with an index  $1 > BI > \mu(BI)$ , will be classified as specialized using  $BI_{PR}$ , given that they show a  $BI_{PR} > 1$ . The opposite is true if  $\mu(BI) \geq 1$ . One can verify this confronting Table 1 and Table 3.

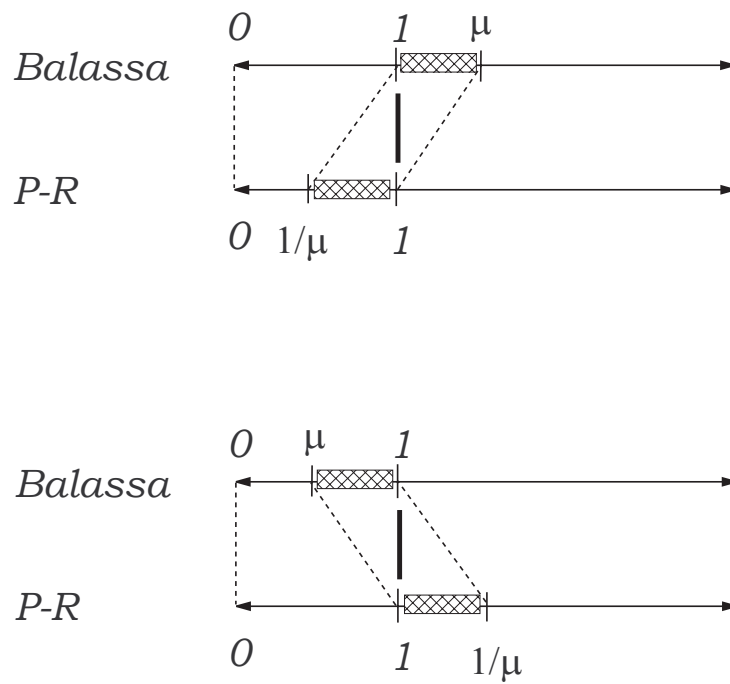
By way of example, in 1996 the mean for Italy was 1.23; 55 out of 538 sectors had a  $BI$  value lying in the range between 1 and the mean itself;

---

<sup>22</sup>Moreover, consider that, by construction, if some values of the BI of one country are above one, other values must be below one; as a consequence, it makes no sense to say that if all sectors have the same BI value (the meaning of  $\mu(BI)$ ) this value is greater or less than one.



Figure 2: Balassa normalizations -  $\mu(BI) \geq 1$  (on top) -  $\mu(BI) \leq 1$  (lower figure)



Germany had a mean value of 1.17, and 26 sectors between this value and 1; France had a mean value of 1.10 and 38 sectors between this value and 1. All these sectors would be classified as specialized using the standard  $BI$ , and as not specialized using the  $BI_{PR}$ . In short, if the  $BI_{PR}$  is used serious errors of interpretation may be committed, especially when the  $BI$  distribution is concentrated around the mean.

### 2.2.2 Symmetry: Laursen (2000).

In order to deal with the skewness of  $BI$ , a different normalization has been proposed - following Cantwell (1989) - by Laursen (1998, 2000) and Dalum and Villumsen (1996); Laursen and Drejer (1997); Dalum et al. (1998) and recently by Brasili et al. (2000):

$$BI_L = \frac{BI - 1}{BI + 1} \equiv \frac{X_{cs} \cdot X_w - X_{ws} \cdot X_c}{X_{cs} \cdot X_w + X_{ws} \cdot X_c} \quad (5)$$

The  $BI_L$  is an approximation of a log transformation of the  $BI$ <sup>23</sup>, which ranges from  $-1$  to  $+1$  and is equal to 0 when  $\frac{X_{cs}}{X_{ws}} = \frac{X_c}{X_w}$  is equal to 1 (when the  $BI$  is equal to 1). The index is therefore symmetric; its average value approximates the log of the geometric mean of  $BI$  (exactly because, as said, the  $BI_L$  is an approximation of a log transformation of the  $BI$ ); and  $BI_L$  has its demarcation value in 0. Country  $c$  has a comparative disadvantage in sector  $s$  if  $-1 < BI_L < 0$ , while it has a comparative advantage in sector  $s$  if  $0 < BI_L < +1$ ; and the same country is said to be specialized in sector  $s$  if it dispaly a sectoral share greater than the same sectoral share for the totality of the set of countries considered<sup>24</sup>.

With respect to  $BI$ , the  $BI_L$  has a (more) symmetric distribution and preserves the same original economic content. But apart from aesthetic considerations, Laursen (1998, 2000) shows, when dealing with the issue of changes in industrial countries specialization, that the residuals of a *Galtonian regression* on the transformed index are often normally distributed, a condition rarely met when the  $BI$  is used instead of  $BI_L$ .

Laursen (1998, 2000) tests the normality of the residuals - for 9 countries and 22 sectors from the OECD STAN database (1995 Edition) - by means

---

<sup>23</sup>As suggested by (Vollrath, 1991), one can (partly) deal with the problems associated with the asymmetry of the  $BI$  by taking a logarithmic transformation of the original  $BI$ . However, as pointed out by (Laursen, 1998), when the sectoral exports of a country are zero, the transformation is not defined. And this is often the case when the set of countries considered is very dishomogeneous, or when the level of sectoral disaggregation is high.

<sup>24</sup>When  $\frac{X_{cs}}{X_c} > \frac{X_{ws}}{X_w}$ , then  $BI_L > 0$ ; and when  $\frac{X_{cs}}{X_c} < \frac{X_{ws}}{X_w}$ , then  $BI_L < 0$ . If  $BI_L = +1$  this means that  $X_{cs} > 0$  and  $X_{ws} \rightarrow 0$ . Similarly, if  $BI_L = -1$  this means that  $X_{ws} > 0$  and  $X_{cs} \rightarrow 0$ .

	Italy		France		Germany	
<b>Sitc (rev.2)</b>	1986	1996	1986	1996	1986	1996
<b>6543</b>	0.83	0.84	-0.06	0.04	-0.91	-0.89
<b>5530</b>	-0.17	0.01	0.71	0.70	-0.98	-0.95
<b>6997</b>	0.09	0.12	0.02	-0.09	0.75	0.64
$\mu(BI_L)$	-0.13	-0.12	-0.07	-0.09	-0.05	-0.11
$\sigma(BI_L)$	0.42	0.45	0.33	0.35	0.49	0.45

Table 4: Balassa Index (Laursen, 1998)

of a Jarque-Bera test; the hypothesis of normality is rejected in 50 per cent of cases if a pure  $BI$  is used for the linear regressions.

We replicated the exercise performed by Laursen using both the new edition of the OECD STAN database (1999) and the ECLAC(UN)-World Bank *TradeCan* Database (1999). However, since the Jarque-Bera test is an asymptotic test, we give our results only for the latter, more disaggregated database<sup>25</sup>.

To avoid a selection bias both industrialized countries and LDCs were included. Our results are particularly clear: even when the  $BI_L$  is used, only with a low disaggregation level is it possible to have normally distributed residuals. This depends: a) on the greater dispersion of data at higher disaggregation levels, and b) on the asymptotic characteristics of the test used. Using an increasingly detailed disaggregation, the normality of residuals is progressively rejected. Table 5 gives the results.

At 2 digits, only 4 out of 13 cases are rejected, a result similar to that obtained by Laursen (1998), but when we move to a 3 digit level only 3 cases pass the normality test. At 4 digits level none pass.

Moreover, the cardinal interpretation of the  $BI_L$  is not easy. Being a quasi-logarithmic transformation of the  $BI$ , the  $BI_L$  has an intuitive interpretation: it shows a *proportional difference* in country  $c$ 's sectoral quota with respect to  $w$ . The problem is that the correspondence between  $BI_L$  and the  $\ln(BI)$  is perfect when  $BI = 1$ , but the two transformations tend progressively to diverge for extreme values of the distribution of  $BI$ . The divergence is even more marked when the disaggregation in the data is high, as shown in Figure 3 for Italy in 1996, or when the countries examined are

<sup>25</sup>The period covered by TradeCan (1999) is much shorter than that considered by STAN (ten years rather than a quarter of century), but trade data are presented at different levels of aggregation, from no more than a few sectors to many hundreds.

Table 5: Jarque-Bera test - Residuals of a "Galtonian" regression on the  $BI_L$ , 1986-96

Digits	2	3	4
# of sectors	38	157	538
Germany	70.3	135.3	718.5
Italy	21	50.6	764.5
France	1.2	2625.1	875.9
USA	0.9	2.2	728.7
UK	13.1	73.9	560.7
Japan	0.2	22.3	208.1
Argentina	3.6	4.3	53.7
China	0.2	6.4	59.4
Spain	3.6	18.7	132.7
Korea	3.8	18.2	40.2
Thailand	22.1	46.4	141
Taiwan	1.6	0.7	33.6
Brazil	0.6	52.3	78.6

*Galtonian regression:*  $BI_L^{96} = \alpha + \beta \cdot BI_L^{86} + \epsilon$ .  $\alpha$  and  $\beta$  are linear parameters and  $\epsilon$  is an error term.  $\chi^2$  critical values: 5.99 (2 d.f. 95% significance).

*Data source:* ECLAC- World Bank (1999).

LDCs<sup>26</sup>.

Figure 3:  $\ln(BI) - BI_L$  (2 and 4 digits)

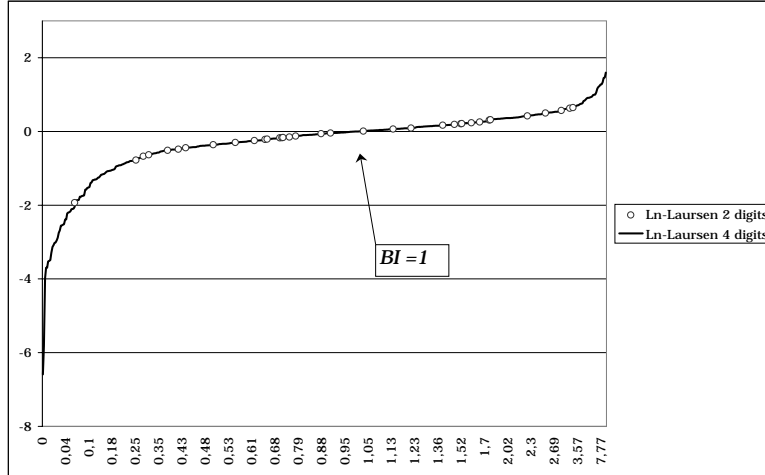


Figure 3 plots the difference between  $\ln(BI)$  and  $BI_L$  on the vertical axis and the  $BI$  on the horizontal axis; the continuous line is associated with 4-digit data, while the bullets are associated with 2-digit data. It is evident that the distortion induced by  $BI_L$  is higher for extreme values of the distribution, which are more relevant to finer levels of disaggregation.

We may therefore conclude that the alleged advantage of the transformed index is still to be demonstrated.

We believe that there are good reasons for using the original  $BI$ . The  $BI_L$  measure preserves most of the distribution characteristics of the original  $BI$ , and it can be considered a more elegant substitute for the original  $BI$ . Nevertheless, the "forced" symmetry may obscure some of the  $BI$  dynamics, especially when these are expressed by a change of kurtosis or of the symmetry itself. Its reduced asymmetry does not imply normality; and its use may induce a bias associated with extreme values of the distribution, when these are relevant, as they usually are for LDCs.

<sup>26</sup>While as regards industrial countries the value for  $\frac{\ln(BI)}{BI_L}$  on the extreme of the right tail of the distribution was between 2 and 3, in cases such as Thailand and Argentina the value was 3.3 and 4.3, respectively.

### 3 Back to the origin ... with some caution.

In the previous section we discussed the main shortcomings of the original *BI* and the defects of the alternative normalization of the Index by Proudman and Redding (1998) and Laursen (2000). A possible strategy would be to define a *new* index, but constructing a mean invariant symmetric RCA index consistent with the original *BI* is impracticable, and it would be hampered by the impossibility of bringing the mean value and the demarcation value of the *BI* to the same identical threshold without losing consistency with the definition of comparative advantage implicit in the *BI*. We shall therefore adopt a less original but, in our opinion, more useful strategy based on the information yielded by analysis of the characteristics of the *BI* distribution. We shall develop our analysis by focusing on the dynamics of RCA (in Italy, France, Germany, and Japan), taking account of the characteristics of the *BI* and its shortcomings in terms of across-country and across-time comparison highlighted in the previous sections and formally described in Appendix 2.

The analysis is carried out using a highly disaggregated set of data, since - as we have already stressed - operating at a more aggregate level may hide important features of a country's specialization pattern. For example, in a previous work (?) we found that the Italian *BI* structure is highly and positively correlated with that of some LDCs (as has been traditionally emphasized) only if the comparison is carried out at a high level of aggregation, but it completely disappears at a higher level of disaggregation. As a consequence, we prefer to work with the largest number of sectors possible. Moreover, we have found that many of the distribution parameters are systematically influenced by the level of aggregation.

#### 3.1 The dynamics of RCA

One of the more intuitive ways to analyse the dynamics of RCA is to use a synthetic indicator of dissimilarity like the Michaely Index (Krugman, 1991; Kim, 1995). The index is

$$MI_{t \in [1,2]} = \frac{1}{2} \cdot \sum_{s=1}^S \left| \frac{X_{cs2}}{X_{c2}} - \frac{X_{cs1}}{X_{c1}} \right|. \quad (6)$$

where  $t$  indicates the years of comparison, so that  $X_{cs2}$  are the sectoral  $s$  exports of country  $c$  at time  $t=2$ ; and  $X_{c2}$  are the total exports of country  $c$  at time  $t=2$ . The *MI* ranges from 0 (perfect stability of the specialization pattern) to 1 (absolute mobility of the specialization pattern) and indicates the intensity of change, but not the direction of change.

When deriving the index in equation 6, we used only the numerator of the  $BI$ , having verified that the effect of the change in the denominator of the  $BI$  was not relevant *at an aggregated* level (the correlation between  $BI$  and  $\frac{X_{cs}}{X_c}$  was positive and significant for the countries analyzed).

Surprisingly, the information obtained from the  $MI$  proves not to be very interesting. The result is an  $MI$  value of 0.08 for France, Germany and Italy, and 0.14 for Japan. What the  $MI$  shows is that Japan had a more mobile pattern between 1986 and 1996, while the three European countries displayed a fairly stable one. The striking feature is that France, Germany and Italy display exactly the same rate of mobility: an finding that should be checked using a more informative instrument.

We took up suggestions from non-parametric statistics (Lehmann, 1975) and estimated a probability density function for the  $BI$  of each of the three European countries plus Japan. The results are set out in Figure 1, and they are quite different from what one might expect from the results obtained with the  $MI$ . While the data for France reveal an *overall* stable density, in the case of Italy and Germany the estimated PDFs show a certain degree of mobility. Even more strikingly, between 1986 and 1996 the two countries move in opposite directions: the PDF for Germany shifts to the left, while the PDF for Italy shifts to the right. These changes seem to indicate a stable pattern of RCA in the case of France characterized by some mobility around the demarcation value; an increase in the number of *below 1 sectors* in the case of Germany; and an increase in the number of *above 1 sectors* in the case of Italy. As far as Japan is concerned, on the one hand, the kurtosis highlighted the larger proportion of low  $BI$  sectors; while on the other hand, the initial and final distributions are sufficiently different to confirm the higher  $MI$  value. Nevertheless, it is very difficult to gain a clear idea of the overall direction of change, as in the case of Italy and Germany.

In order to evaluate the statistical significance of these changes a two tailed Wilcoxon's signed rank test was performed. This test was chosen instead of a more traditional  $t$ -test because it did not require the assumption of normally distributed data. The null hypothesis was the absence of difference between the 1986 and the 1996  $BI$  series for each of the three countries. At a level of confidence of  $P = 0.05$  the test did not reject the null hypothesis in the case of France, while the  $H_0$  was rejected in the case of Japan, and strongly rejected in the case of Italy and Germany.

Some more information can be extrapolated from analysis of the parameters of the distributions. Table 6 shows the moments for the French, Italian, German and Japanese estimated PDF are shown. Remember that the latter three countries are usually clustered as the clearest examples of successful followers in the second half of the last century (with the exception of its last

decade). Our analysis shows that they have very different features in terms of  $BI$  distribution, not all of which are particularly obvious.

The following table sets out the typical parameters of a statistical distribution, except for a couple which are not relevant. It does not give average values (for the reasons already given) or minimum values (because they are always 0 or very close to it). The Herfindahl's concentration index has also been included.

	France		Italy		Germany		Japan	
	1986	1996	1986	1996	1986	1996	1986	1996
<b>Maximum</b>	7,48	11,47	10.76	11.42	9.14	7.15	5.7	5.08
$m(BI)$	1,09	1,10	0.78	0.88	1.14	0.95	0.46	0.49
$\sigma(BI)$	0,87	1,08	1.28	1.4	1.23	1.02	0.84	0.81
<b>Skewness</b>	2,89	4,60	3.25	3.12	1.35	1.51	2.38	1.78
<b>Kurtosis</b>	13,11	33,91	15.25	13.83	3.75	3.83	7.71	3.95
<i>Herfindahl</i> × 100	1.24	1.60	1.15	0.90	0.62	0.76	4.35	2.28

*Data source:* ECLAC- World Bank (1999).

Table 6: Distribution parameters of  $BI$

A glance at Table 6 suffices to reveal interesting differences among the specialization structures of the four countries. Italy maintains very strong specialization (strong competitiveness) in its "max" sectors, as demonstrated by a very high and growing maximum, while it increases the richness of its specialization structure, given the high and growing median,  $m(BI)$ . Almost the opposite emerges in the Japanese case: a permanently low and diminishing level of the maximum, and above all, a specialization structure where most sectors are characterized by a  $BI < 1$  (as indirectly indicated by the very low  $m(BI)$ ). Germany lies somewhere in between: its median is close to one, indicating that roughly half of its sectors are characterized by a  $BI > 1$ ; some of its specialized sectors are highly competitive (the high level of the maximum), although this characteristic is declining in importance. Italy resembles to Germany in terms of  $m(BI)$ , but the presence of highly specialized sectors appears to be more persistent. The Italian structure seems more dispersed, but, in particular, more asymmetric and with more extreme values (which depends mainly on the high level of the maximum). If Germany and Japan are compared, the two distributions appear to have very similar levels of (high) symmetry and of (few) extreme values, but the different median values show that the German structure is somewhat more solid than the



Japanese one, given that in the former case only half of the sectors are *below 1 sectors*. Accordingly, the Japanese structure seems more prone to asymmetric shocks (on its specialized sectors). We can also adopt a more "ad hoc measure" than the median, to extract more information from the structure of the *BI*: namely, the share of sectors with *BI* under unity. In this case one observes that 73% per cent of Japanese sectors are *below 1 sectors*, while the proportions in Italy, France and Germany are 56%, 57% and 51%. These results are confirmed by the Herfindahl index: Japan's export structure is much more concentrated than are Italy's, France's and Germany's, which means that Japan's competitive strength is based on a relatively few sectors. The polarization of the Japanese pattern of specialization may be the result of the sector-polarized economic policy of past years (Ryutaro et al., 1988).

### 3.2 Stationarity

Since the case of France warrants further investigation, more information can be obtained by analyzing the stationarity of Markovian transition matrices.

In previous section we examined the dynamics of the *BI* distribution on the basis of the *MI*; the shortcoming of this approach is that it yields only quantitative indications about the intensity of changes. No indication of the *direction* of these changes can be extrapolated.

In this section, further and more evident information on the dynamics of the *BI* distribution over time is yielded by the analysis of the stationarity of Markovian chains. Our transition matrices are built on a ten-years base; probabilities - that is relative frequencies of passing from one state to another - directly compare the initial with the final year<sup>27</sup>. In order to provide readable tables, we decided to have four classes built around 1, the demarcation value of the *BI*. Because of the limited number of classes, only the two central ones are proportional (from 0.5 to 1 and from 1 to 2), while the two extreme ones are residuals (from 0 to 0.5 and from 2 to the maximum). Nevertheless, the results of the analysis hold even for more detailed disaggregations. Since our matrices are all regular (all elements are positive), from the first stage onwards, the limit distribution will be characterized by equal lines. This limit distribution should be interpreted not as a "forecast" for the future, but merely as indicative of the drift of the time series.

The results differ greatly among the four countries in the example. France displays values for the limit distribution which are quite similar to the real values for both of the initial and the final year. In this case we find a high

---

<sup>27</sup>The alternative strategy of calculating all the passages from one year to the next over the entire period (1986-1987, 1987-1988, etc.) was excluded because it would have mixed long and short term elements.

ITALY	1996			
1986	0 – 0.5	0.5 – 1	1 – 2	> 2
0 – 0.5	0.76	0.18	0.06	0.00
0.5 – 1	0.13	0.55	0.26	0.06
1 – 2	0.05	0.13	0.57	0.24
> 2	0.08	0.05	0.14	0.73
<b>Initial distr.</b>	0.33	0.28	0.26	0.14
<b>Final distr.</b>	0.31	0.25	0.26	0.18
<b>Limit distr.</b>	0.26	0.22	0.25	0.27

Table 7: Markov Transition Matrix - Italy 1986-96.

FRANCE	1996			
1986	0 – 0.5	0.5 – 1	1 – 2	> 2
0 – 0.5	0.70	0.24	0.05	0.02
0.5 – 1	0.19	0.61	0.20	0.01
1 – 2	0.03	0.20	0.68	0.08
> 2	0.04	0.08	0.27	0.61
<b>Initial distr.</b>	0.20	0.36	0.34	0.09
<b>Final distr.</b>	0.22	0.34	0.34	0.09
<b>Limit distr.</b>	0.25	0.34	0.32	0.09

Table 8: Markov Transition Matrix - France 1986-96.

GERMANY	1996			
1986	0 – 0.5	0.5 – 1	1 – 2	> 2
0 – 0.5	0.87	0.12	0.01	0.01
0.5 – 1	0.33	0.45	0.21	0.01
1 – 2	0.03	0.31	0.59	0.06
> 2	0.02	0.02	0.45	0.50
<b>Initial distr.</b>	0.29	0.18	0.22	0.31
<b>Final distr.</b>	0.32	0.19	0.31	0.17
<b>Limit distr.</b>	0.59	0.22	0.16	0.03

Table 9: Markov Transition Matrix - Germany 1986-96.

JAPAN	1996			
1986	0 – 0.5	0.5 – 1	1 – 2	> 2
0 – 0.5	0.78	0.17	0.05	0.00
0.5 – 1	0.28	0.46	0.27	0.00
1 – 2	0.13	0.16	0.49	0.23
> 2	0.07	0.13	0.36	0.44
<b>Initial distr.</b>	0.53	0.23	0.10	0.14
<b>Final distr.</b>	0.50	0.23	0.12	0.14
<b>Limit distr.</b>	0.45	0.23	0.23	0.10

Table 10: Markov Transition Matrix - Japan 1986-96.

share in the central classes, summing to 66% (70% in 1986). Japan again has a large proportion of *below 1 sectors*, even if it has decreased (from 76% to 68%). Italy again has a large and indeed increasing share of *above 1 sectors* (from 39% to 52%). by contrast, Germany's position deteriorates markedly. Whereas in the case of France, Italy and Japan the limit distribution seems to be a continuation of the previous trend between the initial and final year, in that of Germany there is very strong amplification of the already evident worsening trend. The results in terms of stationarity therefore confirm the quantification of the dynamics described in figure 1.

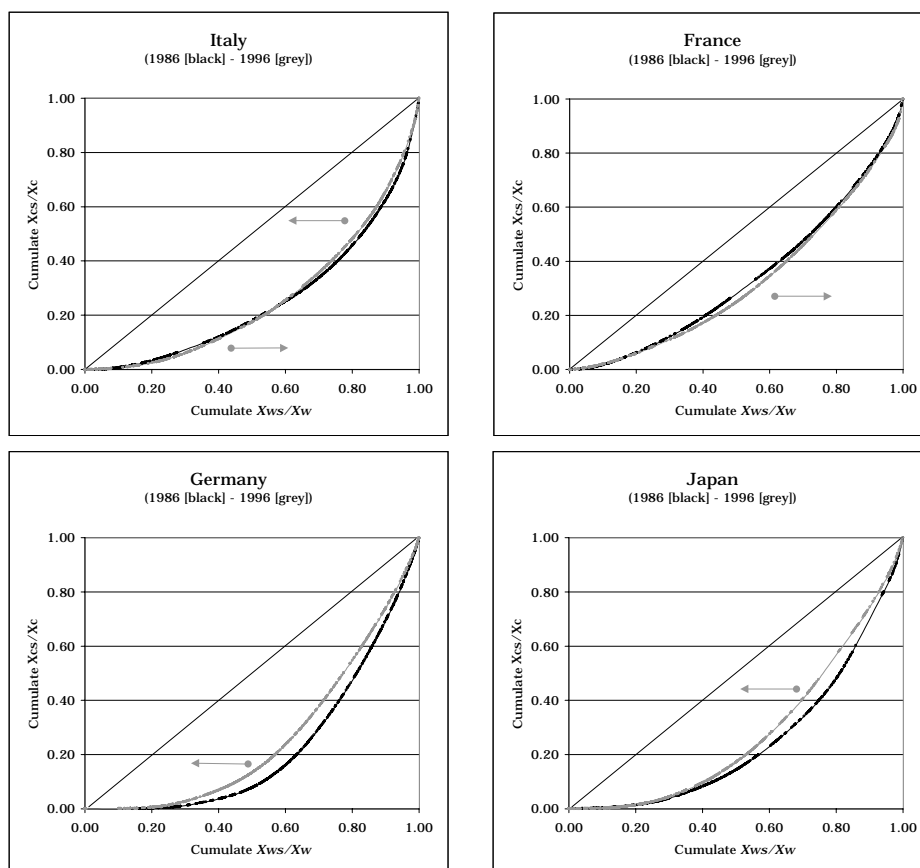
Caution is required in the case of Germany. On the one hand, the sharp deterioration is not very probable, on the other hand, the limit distribution appears not to be "unrealistic" in the sense that it is possible to find similar (real) situations in other countries (all of them LDCs). As said, results should be interpreted as merely indicative of the drift.

### 3.3 Concentration

If the analysis of the dynamics of RCA is conducted with the benchmark  $w$  as a term of reference, the methodology that can be applied derives from concentration analysis. Instead of calculating a Gini (concentration ratio) index, as is usually done (Amiti, 1998, 1999; Brulhart and Torstensson, 1996; Brulhart, 2001) we shall derive Lorenz curves from the decomposition of the  $BI$ . We prefer this graphical instrument because the Gini index can yield similar values for very different distributions and does not discriminate in cases like that of Italy (figure 4 - quadrant (a)), which we shall now discuss, together with those of Germany, Japan and France.

Plotted in the above figures are the cumulative value of the numerator of

Figure 4: Concentration - Italy (a), France (b), Germany (c) and Japan (d)



the  $BI, \frac{X_{cs}}{X_c}$ , on the vertical axis, and the cumulative value of the denominator of the  $BI, \frac{X_{ws}}{X_w}$ , on the horizontal axis. The same procedure has been followed - after ranking the series according to the  $BI$  values - for 1986 and 1996 data. The  $45^\circ$  line represents the *equidistributional locus* associated with a case in which country  $c$  has the same pattern of RCA as the benchmark  $w$ , in every year; a condition equivalent to a  $BI = 1$  in every sector  $s$ . Since we ranked the series according to the  $BI$  values, the specialization Lorenz curve will always be below the  $45^\circ$  line, and country  $c$ 's specialization pattern will be as similar to the benchmark as its specialization curve is close to the *equidistributional locus*.

In the case of France (figure 4 - quadrant (b)), between 1986 and 1996 the Lorenz curve moved to the right, in the direction of the arrow, indicating that the French pattern of RCA is moving away from the benchmark  $w$ . Or better - given the previously emphasized stability of its trade pattern - it indicates that  $w$  is moving away from France.

Germany (figure 4 - quadrant (c)) is the opposite case. Between 1986 and 1996, the Lorenz curve moved to the left, showing that the German pattern of RCA is approaching to the *equidistribution locus*.

Japan (figure 4 - quadrant (d)) is similar to Germany: between 1986 and 1996, the Lorenz curve moved to the left. But, whereas in the case of Germany the low  $BI$  sectors are mainly responsible for the change, in that of Japan the move to the left depends more on the high  $BI$  sectors.

Italy (figure 4 - quadrant (a)) is a different and more complex case. Between 1986 and 1996, the Lorenz curve twisted in the directions indicated by the arrows, which depicts an articulated dynamic. In its *below 1 sectors*, Italy is growing increasingly different from the rest of the world, but it is more similar to  $w$  in its *above 1 sectors*. It is as if the Italian pattern of specialization has been driven by two different forces pushing in opposite directions, towards or away from the benchmark pattern of RCA.

## 4 Summary and conclusion.

Structural transformation associated with economic growth and the process of international integration through trade both emphasize the capacity of regions to compete dynamically. Sources of advantages in specific industries differ in time and space, so that analysis of their pattern seems unavoidable.

In this paper we have discussed the main characteristics and shortcomings of the most widely used index of RCA, the Balassa Index. Some recent studies in applied international trade have directed attention to these shortcomings and tried to circumvent by undertaking some sort of transformation of the

original *BI*. We have shown that in doing so they have introduced further and important limitations. We have also shown that analysis of the statistical characteristics of the *BI* distribution can yield very interesting information about the state and dynamics of country advantages in international trade. In our case, France, Germany, Italy, and Japan have been used to exemplify the information on specialization dynamics that can be extracted from the distribution analysis of trade data.

## Appendix 1 : The *TradeCan* Database.

In the paper we use data from the *TradeCan* database, 1999 edition, by ECLAC and WB (ECLAC- World Bank, 1999). This database contains world imports of commodities and manufactured goods from 1985 to 1996 (for some countries the starting year was 1986). Exporter countries can be selected individually; the sector classification adopted is the SITC rev. 2, available at different aggregation levels up to 4 digits; *TradeCan* accounts for "well over 90% of world trade" (from *TradeCan* User Guide).

Our definition of "manufacture" differs somewhat from the conventional one, given that it is not available in *TradeCan* as a separate macro-sector: "manufacture" is the sum of sectors with codes from 5 to 9. With respect to the standard definition of manufactures, food sectors (codes 0, 1, 2, and 4) are not included because of the difficulties in separating raw material from manufactures. Our manufacture comprises 538 sectors.

More information about the sector classification and country set used can be found in UN (1976) .

## Appendix 2: The three dimensions of $BI$ : sectors, countries and time.

### Cross-sector analysis

Given the  $BI$  equation in 1 (now replicated in 7), at time  $t$  and given the chosen benchmark  $w$ ,

$$BI_{[cs]} = \frac{\frac{X_{cs}}{X_{ws}}}{\frac{X_c}{X_w}}. \quad (7)$$

If we consider only two hypothetical sectors - sector 1 and sector 2 of country  $c$  - the  $BI$  for each sector will be

$$BI_{[c1]} = \frac{\frac{X_{c1}}{X_{w1}}}{\frac{X_c}{X_w}} \quad (8)$$

and

$$BI_{[c2]} = \frac{\frac{X_{c2}}{X_{w2}}}{\frac{X_c}{X_w}}. \quad (9)$$

Then, given that the two equations have the term  $\frac{X_c}{X_w}$  in common, it is always true that  $BI_{c2} \equiv BI_{c1} \cdot \frac{X_{c2}}{X_{w2}} \cdot \frac{X_{w1}}{X_{c1}}$ . Therefore,

$$BI_{c2} > BI_{c1} \iff \frac{X_{c2}}{X_{w2}} > \frac{X_{c1}}{X_{w1}}. \quad (10)$$

Can we say that  $BI_{c2} > BI_{c1} \iff \frac{X_{c2}}{X_c} > \frac{X_{c1}}{X_c}$ , i.e. the export share of sector 2 is larger in comparison to sector's 1 share? The answer is *no*, in general, because this proposition would be true only if  $X_{w1} = X_{w2}$ , which would be fortuitous.

What we can say is that

$$BI_{c2} > BI_{c1} \iff \frac{X_{c2}}{X_c} > \frac{X_{c1}}{X_c} \cdot \frac{X_{w2}}{X_{w1}} \quad (11)$$

which implies that

$$BI_{c2} > BI_{c1} \iff \frac{X_{c2}}{X_c} > \frac{X_{c1}}{X_c} \implies X_{w2} < X_{w1}. \quad (12)$$



## Cross-country analysis

On the other hand, if only two hypothetical countries - country 1 and country 2 - and one sector  $s$  are considered, we have

$$BI_{[1s]} = \frac{\frac{X_{1s}}{X_{ws}}}{\frac{X_1}{X_w}} \quad (13)$$

and

$$BI_{[2s]} = \frac{\frac{X_{2s}}{X_{ws}}}{\frac{X_2}{X_w}}. \quad (14)$$

Since the two equations have  $\frac{X_{ws}}{X_w}$  in common, it is always true that  $BI_{2s} \equiv BI_{1s} \cdot \frac{X_{2s}}{X_2} \cdot \frac{X_1}{X_{1s}}$ . Therefore,

$$BI_{2s} > BI_{1s} \iff \frac{X_{2s}}{X_2} > \frac{X_{1s}}{X_1}. \quad (15)$$

## Dynamics

If we now allow time  $t$  to change and we consider the same hypothetical country  $c$  and one of the two sectors  $s \in (1, 2)$  (i.e. sector 1), and if we conduct an intertemporal comparison for the same sector (i.e. between 1986 and 1996), we have that

$$BI_{c1}^{96} > BI_{c1}^{86} \iff \frac{X_{c1}^{96}}{X_{w1}^{96}} > \frac{X_{c1}^{86}}{X_{w1}^{86}} \implies \frac{X_c^{96}}{X_w^{96}} < \frac{X_c^{86}}{X_w^{86}} \quad (16)$$

where, the last condition depends on  $X_c^{96} < X_c^{86} \implies X_w^{96} < X_w^{86}$ .

The violation of the condition expressed in (16) yields the apparently paradoxical results highlighted in section 2.1.3.

Hence when analysing the dynamics of RCA one should bear in mind that if the  $BI$  has increased this may be the result of several causes, of which only some concern the increase in the export sectoral share.

## References

- Mary Amiti. New Trade Theories and Industrial Location in the EU: a Survey of Evidence. *Oxford Review of Economic Policy*, 14(2):45–53, 1998.
- Mary Amiti. Specialization Patterns in Europe. *Weltwirtschaftliches Archiv*, 135(4):573–593, 1999.
- Bela Balassa. Trade Liberalisation and Revealed Comparative Advantage. *Manchester School of Economics and Social Studies*, 33:99–123, 1965.
- Robert H. Ballance, Helmut Forstner, and Tracy Murray. Consistency Tests of Alternative Measures of Comparative Advantage. *The Review of Economics and Statistics*, V(2):157–161, 1987.
- Harry P. Bowen. On the Theoretical Interpretation of Indices of Trade Intensity and Revealed Comparative Advantage. *Weltwirtschaftliches Archiv*, (119):464–472, 1983.
- Andrea Brasili, Paolo Epifani, and Rodolfo Helg. On the Dynamics of Trade Paterns. *De Economist*, 148(2):233–258, June 2000.
- Marius Brulhart. Growing Alike or Growing Apart ? Industrial Specialisation of the EU Countries. In Charles Wyplosz, editor, *EMU and Its Impact on Europe and Developing Countries*. Oxford University Press, Oxford, 2001. Forthcoming.
- Marius Brulhart and Jan Torstensson. Regional integration, scale economies and industry location in the european union. Cepr wp. 1435, 1996.
- John Cantwell. *Technological Innovation and Multinational Corporations*. Blackwell, Oxford, 1989.
- Soo Hong Chew. Mean value. In John Eatwell, Murray Milgate, and Peter Newman, editors, *The New Palgrave - Time Series and Statistics*, pages 149–158. MacMillan, 1990.
- Tito Cordella. Patterns of Trade and Oligopoly Equilibria: An Example. *Review of International Economics*, 6(4):554–53, 1998.
- Bent Dalum, Keld Laursen, and Gert Villumsen. Sturctural Change in OECD Export Specialisation Pattern: De-specialisation and 'Stickiness'. *International Review of Applied Economics*, 12:447–467, 1998.

- Bent Dalum and Gert Villumsen. Are oecd export specialisation patterns 'sticky' ? relations to the convergence - divergence debate. *Druid wp 96 - 3*, 1996.
- Donald R. Davies. Critical Evidence on Comparative Advantage? North-North Trade in a Multilateral World. *Journal of Political Economy*, 105 (5):1051–1060, 1997.
- Alan V. Deardorff. The General Validity of the Law of Comparative Advantage. *Journal of Political Economy*, 88:464–957, 1980.
- ECLAC- World Bank. *TradeCan database*. World Bank, 1999.
- Masahisa Fujita, Paul Krugman, and Anthony Venables. *The Spatial Economy: Cities, Regions and International Trade*. MIT Press, 1999.
- Oded Galor. A Two Sector Overlapping Generation Model: A Global Characterisation of the Dynamical System. *Econometrica*, 60:351–386, 1992.
- Arye L. Hillman. Observations on the Relation between 'Revealed Comparative Advantage' and Comparative Advantage as indicated by Pre-Trade Relative Prices. *Weltwirtschaftliches Archiv*, 116:315–321, 1980.
- Jeroen Hinloopen and Charles Van Marrewijk. On the Empirical Distribution of the Balassa Index. *Weltwirtschaftliches Archiv*, 137(1):1–35, 2001.
- Edgar M. Hoover. The measurement of Industrial Location. *Review of Economics and Statistics*, 18:162–171, 1936.
- Sukkoo Kim. Expansion of Markets and the Geographic Distribution of Economic Activity: The Trends in the U.S. Manufacture Structure - 1860-1987. *Quarterly Journal of Economics*, 110(4):881–908, 1995.
- Paul Krugman. *Geography and Trade*. MIT Press, Cambridge MA, 1991.
- Jean Lafay. The measurement of revealed comparative advantages. In M.G. Dagenais and P-A. Mouet, editors, *International Trade Modelling*. Chapman and Hall, London, UK, 1992.
- Keld Laursen. Revealed comparative advantage and the alternatives as measures of international specialisation. *Druid wp 98-30*, 1998.
- Keld Laursen. *Specialization, Trade and Growth*. Routledge, 2000.
- Keld Laursen and Ina Drejer. Do inter-sectoral linkages matter for international export specialisation. *Druid wp 97-15*, 1997.

- Erich L. Lehmann. *Nonparametrics: Statistical Methods based on Ranks*. Holden Day, S. Francisco CA, 1975.
- H. H. Liesner. The european common market and british industry. *Economic Journal*, 68:302–316, 1958.
- Maria Luisa Mancusi. International technological specialization in industrial countries: Patterns and dynamics. February 2000.
- Michael Michaely. *Concentration in International Trade*. North Holland, Amsterdam, NL.
- Andrew Mountford. Trade, Convergence and Overtaking. *Journal of International Economics*, 46, 1998.
- DSTI - STAN Industrial Database*. OECD, 1999.
- James Proudman and Stephen Redding. *Openness and Growth*. The Bank of England, 1998.
- James Proudman and Stephen Redding. Evolving patterns of international trade. mimeo, December 1999.
- David J. Richardson and Chi Zhang. Revealing Comparative Advantage: Chaotic or Coherent Patterns Across Time and Sector and U.S. Trading Partner ? In Magnus Blomstrom and Linda S. Goldberg, editors, *Topics in Empirical International Economics*, chapter 7. University of Chicago Press, July 2001.
- Komiya Ryutaro, Okuno Masahiro, and Suzumura Kotaro. *Industrial Policy of Japan*. Academic Press, 1988.
- André Sapir. The Effect of Europe's Internal Market Program on Production and Trade: A first assessment. *Weltwirtschaftliches Archiv*, 132(3):457–475, 1996.
- Henri Theil. *Economics and Information Theory*. North Holland, Amsterdam, NL, 1967.
- UN. *Classification by broad economic categories. Defined in terms of SITC, rev.2*. United Nations, statistical papers, series m, n.53, rev.1 edition, 1976.
- Thomas L. Vollrath. A Theoretical Evaluation of Alternative Trade Intensity Measures of Revealed Comparative Advantage. *Weltwirtschaftliches Archiv*, 127:265–280, 1991.

Adrian J. Yeats. On the Appropriate Interpretation of the Revealed Comparative Index. *Weltwirtschaftliches Archiv*, 121:61–73, 1985.