# Which revealed comparative advantage index to choose? Theoretical and empirical considerations

Rémi Stellian and Jenny Paola Danna-Buitrago

**Abstract**

This article proposes guidelines for choosing an index of revealed comparative advantages (RCA) to analyse a given configuration of countries, products and periods. Following a systematic review of the main theoretical strengths and weaknesses inherent in RCA indices, a standardized method is designed to gauge the quality of their empirical measurement. This method is illustrated for Colombia and the Northern Triangle. The guidelines formalize a number of theoretical and empirical trade-offs that need to be weighed to provide a better basis for choosing an RCA index.

**Keywords**

Economic integration, intraregional trade, comparative advantage, measurement, economic indicators, evaluation, trade policy, Colombia

**JEL classification**

F13, F14, F15

**Authors**

Rémi Stellian is Associate Professor in the Department of Business Administration, Faculty of Economics and Administrative Sciences, Pontificia Universidad Javeriana, Bogotá, Colombia. Email: rstellian@javeriana.edu.co.

Jenny Paola Danna-Buitrago is Research Director at the Research Department of Fundación Universitaria Los Libertadores, Bogotá, Colombia. Email: jenny.danna@libertadores.edu.co.

# I. Introduction

The concept of comparative advantage refers to a country's capacity to produce a good or service with higher productivity and greater differentiation in its characteristics —quality, brand and after-sales service— than its trading partners (Jaimovich and Merella, 2015). These differentials in productivity and characteristics are key to explaining the potential of international trade and specialization to improve resource use and enhance welfare (Lassudrie-Duchêne and Ünal-Kesenci, 2001). Accordingly, public entities need to understand the national economy's comparative (dis)advantages to adjust the country's specialization pattern and thus gain the benefits of international economic integration (Chanteau, 2007). This is particularly relevant in contemporary times, which are characterized by continuous trade liberalization endeavours, both regional and multilateral (Menon, 2014).

As comparative advantages originate in a purely theoretical situation with no international trade, they cannot be observed directly (Lafay, 1987). Accordingly, since the pioneering work of Balassa (1965), the conventional approach has been to infer them indirectly through trade patterns. Since trade reflects such advantages, it is then argued that trade can be used to calculate an index of revealed comparative advantages (RCA). This index is a number that summarizes the comparative advantage of a given country for a given product in a given period of time. If the number is higher (lower) than some neutral value, comparative advantage (disadvantage) exists (Danna-Buitrago, 2017).[1]

The literature discusses various RCA indices without specifying which should be used to study a given configuration of countries, products and periods. Several RCA indices have emerged in response to criticisms of the index initially proposed by Balassa (1965) that seek to overcome its various weaknesses. Nonetheless, published studies do not agree on why one index is preferable to another, and the Balassa measure remains a reference in the literature.

The only certainty is that any empirical case of comparative advantage should be studied through an RCA index chosen to strike a balance between the inherent strengths and weaknesses of the different indices available in the literature, without neglecting the quality of the measurements. In this context, "inherent" refers to the theoretical strengths and weaknesses of the index formula itself, which must be analysed before applying it to a given configuration of countries, products and periods. Thus, the strengths and weaknesses in question are the result of the variables used to calculate the RCA index and the way in which these variables are combined in the formula. However, analysing theoretical strengths and weaknesses is not sufficient because it is possible for an RCA index to measure comparative advantages in a contradictory way when applied to specific countries, products and periods, even if, theoretically, it offers all possible strengths and has no weaknesses.

This article proposes guidelines to systematize the theoretical considerations that need to be taken into account and the empirical assessment that should be made when selecting an RCA index for use in an empirical study of comparative advantages. First, a systematic review is performed of the main strengths and weaknesses inherent in the Balassa (1965) RCA index. An assessment is also made of the extent to which 12 other RCA indices that are among the most frequently mentioned in the literature share the same strengths, make it possible to overcome the same weaknesses, and offer other strengths. Second, a standardized method is constructed to evaluate RCA indices based on the quality of their empirical measurements. This method systematizes several paths already examined in the literature, and the fact that it is standardized makes it applicable to any combination of countries, products and periods. The method in question is illustrated by applying the 13 RCA indices to the trade zone comprising Colombia and the Northern Triangle (El Salvador, Guatemala and Honduras). The purpose of these guidelines is to provide a more solid foundation for the balance to be struck, taking various theoretical and empirical considerations into account.

[1] Another methodology for measuring comparative advantage involves calculating the domestic resource cost (Cai, Leung and Hishamunda, 2009), which can provide a complementary point of view to that of RCA indices.

The article is organized in four sections including this introduction. Section II presents the theoretical considerations outlined in the previous paragraphs, and section III constructs and illustrates the standardized empirical evaluation method. Section IV summarizes the results obtained and suggests a future line of research.

# II. Inherent strengths and weaknesses of the RCA indices

In the rest of this article, the following notations will be used:

- $J$ denotes a group of countries that make up a trade zone. For example, $J$={COL, SLV, GTM, HND}; $i$ denotes a country in $J$.
- $K$ represents a set of products or product categories, for example, the third revision of the three-digit Standard International Trade Classification (SITC Rev.3), where $K=\{001, 011, 012, 016, \cdots, 898, 899, 971\}$ ($\#K=255$); $k$ denotes a product category in $K$.
- $T$ is a set of time periods. For example, $T=\{1995, 1996, 1997, \cdots, 2017\}$; $t$ denotes a period in $T$.
- $X_{ikt} \in \mathbb{R}_+$ represents country $i$'s exports to trade zone $J$ of product category $k$ in period $t$. Similarly, $M_{ikt} \in \mathbb{R}_+$ represents country $i$'s imports of product $k$ from trade zone $J$ in time period $t$.
- $x_{ijkt} \in \mathbb{R}_+$ represents the flow of trade in $k$ from origin country $i$ to destination country $j$ in time period $t$. $x_{iikt} = 0$ (a country cannot be simultaneously the origin and destination of a given trade flow), $\sum_{j\in J} x_{ijkt} = X_{ikt}$ and $\sum_{j\in J} x_{jikt} = M_{ikt}$. Some RCA indices are based on $x_{ijkt}$ instead of on $X_{ikt}$ and $M_{ikt}$.
- $Y_{it} \in \mathbb{R}_+$ is the gross domestic product (GDP) of country $i$ in period $t$. Some RCA indices incorporate GDP in their calculations.

## 1. The starting point: Balassa (1965)

The Balassa (1965) RCA index is the benchmark measure of comparative advantages (Konstantakopoulou and Tsionas, 2019). The recent literature continues to use this index, for example, Brakman and Van-Marrewijk (2017), Abbas and Waheed (2017), Esquivias (2017), Halilbašić and Brkić (2017), Hoang and others (2017), and Shaul Hamid and Aslam (2017). Balassa (1965) is based on the following idea: if in some period $t$ the share of product $k$ in the total exports of country $i$—namely, $X_{ikt} / \sum_{p\in K} X_{ipt}$—is larger than the equivalent share in all other countries in the trade zone ($\sum_{j\in J} X_{jkt} / \sum_{j\in J}\sum_{p\in K\, jpt}$), then $i$ has a better capacity to export $k$ than the other countries in the zone in period $t$. This situation would reveal comparative advantages of country $i$ with respect to product group $k$ at time $t$. Balassa (1965) constructs his RCA index by dividing the first ratio by the second. Assuming that $B_{ikt}$ is the Balassa (1965) RCA index calculated for $\langle i, k, t\rangle$, then:

$$B_{ikt} = \left(\frac{X_{ikt}}{\sum_{p\in K} X_{ipt}}\right) / \left(\frac{\sum_{j\in J} X_{jkt}}{\sum_{j\in J}\sum_{p\in K} X_{jpt}}\right) \quad (1)$$

Values of $B_{ikt} \in ]1,+\infty[$ reveal the existence of comparative advantages, whereas values of $B_{ikt} \in [0,1[$ reveal comparative disadvantages; $B_{ikt}=1$ is the neutral value (absence of advantages/disadvantages). If $\sum_{j\in J} X_{jkt}=0$ then the denominator of $B_{ikt}$ is equal to zero, and it is impossible to perform the corresponding division. In this case, the RCA index must be equal to its neutral value. If no country exports $k$ (that is $\sum_{i\in J} X_{kit}=0$), then no country has either advantages or disadvantages in $k$ in period $t$.

The Balassa $B$ has the merit of being compatible with the Kunimoto (1977) principle as extended by Vollrath (1991), hereinafter referred to as the "Kunimoto-Vollrath principle". Kunimoto (1977) states that the specialization of a country $i$ relative to another country $j$ is measured by comparing the value of exports from $i$ to $j$ with a theoretical value. If the observed value deviates from the theoretical value, then the specialization in question exists. The theoretical value is equal to country $i$'s global exports weighted by country $j$'s share in world trade. Vollrath (1991) builds on this to suggest another principle: there is a theoretical value of exports that reveals the absence of comparative advantages and disadvantages for country $i$ in terms of product $k$ (in $t$). If country $i$'s exports exceed (are below) the theoretical value, it has comparative advantages (disadvantages) in $k$. In this case, the theoretical value is equal to country $i$'s total exports weighted by the share of $k$ in the total exports of trade zone $J$; in other words, the product of $\sum_{p\in K} X_{ipt}$ and $\sum_{j\in J} X_{jkt} / \sum_{j\in J} \sum_{p\in K} X_{jpt}$. Consequently, $B$ can be rewritten as the observed value of exports divided by the theoretical value such that, under the Kunimoto-Vollrath principle, $B < 1$ reveals comparative advantages, while $0 \leq B < 1$ reveals comparative disadvantages.

Another strength of $B$ is that it considers the full structure of trade flows and not just flows corresponding to the product and country under consideration, in keeping with the relative nature of comparative advantages (Stellian and Danna-Buitrago, 2019). However, there are five weaknesses in the way $B$ measures comparative advantages. The first is its asymmetry: $B$ reveals comparative advantages by a number in the interval $]1,+\infty[$ and comparative disadvantages by a number in the interval $[0,1[$. The interval of comparative disadvantages has an upper bound that does not exist in the case of comparative advantages. Thus, comparative advantages and disadvantages are measured differently (Yu, Cai and Leung, 2009).

The second weakness is known as the "small-country bias". Paradoxically, a country that exports little—$\sum_{p\in K} X_{ipt} \rightarrow 0$— tends to have high values of $B$. This can be seen by rewriting $B_{ikt}$ as follows:

$$B_{ikt} = \frac{X_{ikt} \times \sum_{j\in J} \sum_{p\in K} X_{jpt}}{\sum_{p\in K} X_{ipt} \times \sum_{j\in J} X_{jkt}} \quad (2)$$

Thus, a low value of $\sum_{p\in K} \mathrm{X}_{ipt}$ means that the denominator tends to zero, which generates high values of $B$, even if this value should not reveal major comparative advantages (Yeats, 1985).

As a third weakness, $B$ does not make it possible to apply a flexible concept of comparative advantages that corresponds not only to productivity differentials (the traditional concept of comparative advantages) but also to a country's capacity to differentiate a product qualitatively with respect to its foreign counterparts. If $i$ has the capacity to manufacture $k$ with higher productivity than other countries, it will be able to sell it at a lower price, which will have a positive impact on its exports. However, this does not mean that country $i$ does not import $k$ at all. The other countries could offer differentiated versions of $k$ that $i$ might demand, even at a higher price. Thus, when calculating an RCA index, it is necessary to take both exports and imports into account (Lafay, 1987; Lassudrie-Duchêne and Ünal-Kesenci, 2001). In other words, considering exports and imports simultaneously makes it possible to capture comparative advantages in relation to both supply and demand (Vollrath, 1991). Accordingly, as $B$ is based solely on exports, it is not compatible with this flexible conception of comparative advantages.

The fourth weakness of $B$ is that it ignores countries' GDP, even though GDP provides a major theoretical foundation for measuring comparative advantages. In particular, if a country has a higher GDP and thus higher income, its demand for higher quality products will increase. If its trading partners supply that demand, they could gain comparative advantages (Jaimovich and Merella, 2015). Similarly, two countries could share the same $B$ value without possessing the same level of comparative advantages. Exports of the product in question will represent a larger share of the national economy that is smaller in terms of GDP. Because of this higher degree of specialization, comparative advantages are greater (Stellian and Danna-Buitrago, 2019).

The fifth weakness of $B$ is its lack of additivity. With respect to countries, additivity refers to the possibility of adding two or more RCA indices from different countries to ascertain the RCA index for a country grouping. For example, the RCA index of the Northern Triangle relative to Colombia for wood pulp and its chemical derivatives is equal to the sum of the RCA indices of Guatemala, Honduras and El Salvador relative to Colombia for that product. Similarly, additivity across products means that Colombia's RCA index relative to another country for wood pulp and its chemical derivatives will be equal to the sum of the RCA indices of Colombia for cellulose acetates, cellulose nitrates and cellulose ethers, respectively. This double additivity allows different country and product classifications with different levels of disaggregation to be used without affecting the measurement of comparative advantages (Yu, Cai and Leung, 2009).

The following subsections show the extent to which other RCA indices address these five weaknesses, without ignoring the Kunimoto-Vollrath principle and the relative nature of comparative advantages, in addition to other possible strengths.

## 2. Transformations of the Balassa (1965) RCA index

There are three RCA indices that transform $B$. First, Hoen and Oosterhaven (2006) calculate the difference between $X_{ikt}/\sum_{p\in K} X_{ipt}$ and $\sum_{j\in J} X_{jkt} / \sum_{j\in J} \sum_{p\in K} X_{jpt}$ instead of the ratio between them. Denoting the additive version of $B$ as $BA$, then:

$$BA_{ikt} = \frac{X_{ikt}}{\sum_{p\in K} X_{ipt}} - \frac{\sum_{j\in J} X_{jkt}}{\sum_{j\in J} \sum_{p\in K} X_{jpt}} \tag{3}$$

$BA$ can be rewritten as the difference between observed exports and their theoretical value normalized by country $i$'s total exports (in $t$):

$$BA_{ikt} = \frac{X_{ikt} - \sum_{p\in K} X_{ipt} \frac{\sum_{j\in J} X_{jkt}}{\sum_{j\in J} \sum_{p\in K} X_{jpt}}}{\sum_{p\in K} X_{ipt}} \tag{4}$$

Second, Laursen (2015) proposes a symmetric version of $B$, denoted $BS$:

$$BS_{ikt} = \frac{B_{ikt} - 1}{B_{ikt} + 1} \tag{5}$$

Third, Yu, Cai and Leung (2009) propose a normalized RCA index, denoted here as $N$. $N$ is almost identical to $BA$, since its starting point is the difference between observed exports and their theoretical value. The only change is that $N$ uses the total exports of trade zone $J$ to normalize the difference instead of country $i$'s total exports. Consequently:

$$N_{ikt} = \frac{X_{ikt} - \sum_{p\in K} X_{ipt} \frac{\sum_{j\in J} X_{jkt}}{\sum_{j\in J} \sum_{p\in K} X_{jpt}}}{\sum_{j\in J} \sum_{p\in K} X_{jpt}} \tag{6}$$

The indices $BA$, $BS$ and $N$ are all compatible with the Kunimoto-Vollrath principle. The compatibility of $BA$ and $N$ results from the calculation of the difference between the observed value of exports and their theoretical value. $BS$ is compatible with the Kunimoto-Vollrath principle because it is a logarithmic approximation of $B$. Similarly, by using the same variables as $B$, the transformations of $B$ are compatible with the relative nature of comparative advantages.

However, the transformations of $B$ do not fully resolve its five major weaknesses. The problem of asymmetry is solved since $BA$, $BS$ and $N$ are symmetric around zero. These three indices also avoid the small-country bias since they have upper bounds ($1$ for $BA$ and $BS$ and ¼ for $N$). However, none considers either imports or GDP. Moreover, Yu, Cai and Leung (2009) show that only $N$ is additive with respect to both countries and products.

## 3. The Balassa (1986) RCA index

Balassa (1986) proposes another RCA index, referred to here as $B2$:

$$B2_{ikt} = \frac{X_{ikt} - M_{ikt}}{X_{ikt} + M_{ikt}} \tag{7}$$

If $i$ records a positive trade balance (numerator) in its trade in $k$ in period $t$ it is assumed to have comparative advantages, represented by the index $B2_{ikt}>0$. Conversely, country $i$ will have comparative disadvantages if it registers a trade deficit in k, i.e. $B2_{ikt}<0$. Similarly, $B2_{ikt}=0$ is the neutral value of the index. The trade balance is normalized by country $i$'s total trade in $k$ in period $t$ (denominator). Zero is assigned as the value of $B2_{ikt}$ if $X_{ikt}+M_{ikt}=0$ (denominator equal to zero). This means that if $i$ does not trade in $k$, it has neither advantages nor disadvantages.

$B2$ is distinguished from the four RCA indices discussed above by its formula, which is based on imports associated with $\langle i, k, t\rangle$. The rationale for $B2$ is as follows: if country $i$ exports more of $k$ than it imports in period $t$, its combination of productivity and differentiation with respect to $k$ can be deemed superior to that of the other countries in the trade zone in question. Therefore, $B2_{ikt}>0$ ($B2_{ikt}<0$) reveals comparative advantages (disadvantages), which makes $B2$ compatible with a flexible concept of comparative advantages. $B2$ also solves the symmetry problem: its lower bound is -1, its neutral value is 0 and its upper bound is 1. avoids the small-country bias.

However, unlike $B$ and its transformations, $B2$ is not compatible with the Kunimoto-Vollrath principle, nor does it uphold the relative nature of comparative advantages because it is based solely on the trade flows associated with the country and product in question. Moreover, like $B$ and its transformations, $B2$ does not take GDP into account. Lastly, $B2$ is not additive.

## 4. Transformations of the Balassa (1986) RCA index

Just as some RCA indices transform $B$, two RCA indices transform $B2$. The first is the RCA index of Donges and Riedel (1977), denoted as $B2D$:

$$B2D_{ikt} = \left( \frac{B2_{ikt}}{\frac{\sum_{p\in K}(X_{ipt} - M_{ipt})}{\sum_{p\in K}(X_{ipt} + M_{ipt})}} - 1 \right) \cdot \operatorname{sign}\left( \sum_{p\in K} (X_{ipt} - M_{ipt}) \right) \tag{8}$$

As noted above, $B2_{ikt}$ is the balance of trade in $k$ normalized by total trade in $k$ for $\langle i, t\rangle$. $B2D$ divides $B2$ by the same type of variable for all commodities before subtracting 1 and weighting by sign ($\sum_{p\in K}(X_{ipt}-M_{ipt})$). The latter expression is equal to 1 if $\sum_{p\in K} (X_{ipt}-M_{ipt})\geq 0$ which means that country total $i$'s trade balance is positive, or -1 if it is strictly negative. $B2D_{ikt}\in ]0,+\infty[$ reveals comparative advantages. Thus:

- If $\sum_{p\in K}(X_{ipt}-M_{ipt})\geq 0$, then $B2D_{ikt}>0$ results from $B2_{ikt}>\sum_{p\in K}(X_{ipt}-M_{ipt})/\sum_{p\in K}(X_{ipt}+M_{ipt})$. In period $t$, if country $i$'s total trade balance is zero or in surplus, then it has comparative advantages for $k$ when its trade balance in $k$, normalized by its total trade in $k$ is greater than the equivalent magnitude for all products.

- Reciprocally, if $\sum_{p\in K}(X_{ipt}-M_{ipt})<0$, then $B2D_{ikt}>0$ results from $B2_{ikt}<\sum_{p\in K}(X_{ipt}-M_{ipt})/\sum_{p\in K}(X_{ipt}+M_{ipt})$. In period $t$, if $i$ has an overall trade deficit, then it has comparative advantages in $k$ when its balance of trade in $k$, normalized by its total trade in $k$, is less than the equivalent magnitude for all products.

By the same token, $B2D_{ikt}<0$ reveals comparative disadvantages for $\langle i,k,t\rangle$, which derives from:

- $B2_{ikt}<\sum_{p\in K}(X_{ipt}-M_{ipt})/\sum_{p\in K}(X_{ipt}+M_{ipt})$ if $\sum_{p\in K}(X_{ipt}-M_{ipt})\geq 0$.
- $B2_{ikt}>\sum_{p\in K}(X_{ipt}-M_{ipt})/\sum_{p\in K}(X_{ipt}+M_{ipt})$ if $\sum_{p\in K}(X_{ipt}-M_{ipt})<0$.

Measuring revealed comparative advantages by $B2D$ is less intuitive than using $B2$. However, $B2D$ takes into account trade flows involving country $i$ for all products in $K$. This is possible through country $i$'s trade balance and its total trade. As a result, $B2D$ is closer to the relative nature of comparative advantages than $B2$ is. $B2D$ also maintains the symmetry of $B2$ (symmetry around zero). Lastly, when considering exports and imports, $B2D$ is compatible with a flexible concept of comparative advantages.

However, $B2D$ is not compatible with the Kunimoto-Vollrath principle because it does not include a theoretical value for $B2_{ikt}$ (or at least a notional value of $X_{ikt}-M_{ikt}$), which must be calculated from trade flows associated with all countries and products. Moreover, $B2D$ is affected by the small-country bias. On the one hand, it is possible to show that if $\sum_{p\in K}X_{ipt}\rightarrow\sum_{p\in K}M_{ipt}$, then $B2D_{ikt}\rightarrow\pm\infty$. On the other hand, $\sum_{p\in K}X_{ipt}\rightarrow\sum_{p\in K}M_{ipt}$ is compatible with small values of $\sum_{p\in K}X_{ipt}$ and $\sum_{p\in K}M_{ipt}$. Lastly, $B2D$ ignores GDP and is not additive.

Gnidchenko and Salnikov (2015) propose another transformation of $B2$, denoted $B2G$:

$$B2G_{ikt} = B2_{ikt} \times \frac{\frac{X_{ikt}+M_{ikt}}{Y_{it}}}{\frac{\sum_{j\in J}(X_{jkt}+M_{jkt})}{\sum_{j\in J}Y_{jt}}} \qquad (9)$$

$B2G$ is calculated by weighting $B2$ by the degree of openness of $i$ with respect to $k$ in time $t$, normalized by the degree of openness of countries belonging to $J$. If $k$ has a greater weight in the economy of $i$ (in period $t$) than in the economy of $J$, the comparative advantages of $i$ in $k$ at time $t$, previously represented by $B2_{ikt}\in]0;1]$, will be greater. Similarly, the comparative disadvantages of $i$ in $k$ at time $t$, previously represented by $B2_{ikt}\in[-1;0[$, will also be greater. A merit of $B2G$ is that it takes into account the size of the different economies $J$ in through each country's GDP. As noted above, a country with a higher GDP and thus higher income will demand both more and higher-quality products; if its trading partners supply this demand, this can generate comparative advantages for them. $B2G$ captures this mechanism, as an increase in $\sum_{j\in J}Y_{jt}$ raises $B2G$.

Accordingly, $B2G$ overcomes three of the five weaknesses of $B$. It is symmetric around zero, it avoids the small-country bias, and its measurement of comparative advantages does not depend solely on trade flows for the product and country in question. However, the flow structure used to measure comparative advantages is not the same as in $B2D$, which includes the flows to and from $i$ for all products in $k$. By contrast, $B2G$ captures the flows of $k$ for all countries belonging to trade zone $J$. Thus, $B2D$ reflects the relative nature of comparative advantages across products, whereas $B2G$ does so across countries (Yu, Cai and Leung, 2009). Ideally, both types of flows should be considered for all countries and all products.

Lastly, although Gnidchenko and Salnikov (2015) suggest that $B2G$ is compatible with the Kunimoto-Vollrath principle, $B2G$ does not include a theoretical value of $X_{ikt}-M_{ikt}$ or $B2_{ikt}$ calculated from the full set of trade flows, nor is it additive.

## 5. An RCA index *à la* Vollrath (1991)

Vollrath (1991) proposes the following index, denoted as $V_{ikt}$:

$$V_{ikt} = \left(\frac{X_{ikt}}{\sum_{p\in K\setminus\{k\}} X_{ipt}}\right) / \left(\frac{\sum_{j\in J\setminus\{i\}} X_{jkt}}{\sum_{j\in J\setminus\{i\}} \sum_{p\in K\setminus\{k\}} X_{jpt}}\right) - \left(\frac{M_{ikt}}{\sum_{p\in K\setminus\{k\}} M_{ipt}}\right) / \left(\frac{\sum_{j\in J\setminus\{i\}} M_{jkt}}{\sum_{j\in J\setminus\{i\}} \sum_{p\in K\setminus\{k\}} M_{jpt}}\right) \quad (10)$$

The first term in $V$ is similar to $B$ except that in calculating the corresponding ratios, no account is taken of exports associated with $i$ or $k$. The second term is constructed like the first but from imports instead of exports.

A problem arises with $V$ if $i$ is the only exporter of $k$ (in period $t$), so $\sum_{j\in J}\setminus_{\{i\}} X_{jkt} = 0$. In this case, the denominator of the first term is equal to 0, and $V_{ikt}$ cannot be calculated. The same is true for the second term if $i$ is the only importer of $k$.[2] Because of this, the literature (for example, Hadzhiev, 2014) recommends against ignoring the trade flows associated with $i$ or $k$. This modification of $V$ is denoted by $V'$:

$$V'_{ikt} = \left(\frac{X_{ikt}}{\sum_{p\in K} X_{ipt}}\right) / \left(\frac{\sum_{j\in J} X_{jkt}}{\sum_{j\in J} \sum_{p\in K} X_{jpt}}\right) - \left(\frac{M_{ikt}}{\sum_{p\in K} M_{ipt}}\right) / \left(\frac{\sum_{j\in J} M_{jkt}}{\sum_{j\in J} \sum_{p\in K} M_{jpt}}\right) \quad (11)$$

The first term is then $B_{ikt}$. It is possible that $\sum_{j\in J} X_{jkt} = 0$. In this case, $\sum_{j\in J} M_{jkt} = 0$ (if no country exports $k$, then no country imports $k$), and none of the terms can be calculated because both have a zero denominator. However, this case means attributing the neutral value to $V'_{ikt}$, namely zero. In the absence of trade flows associated with $k$ in $J$, theoretically no country has either advantages or disadvantages.

$V'$ extends the Kunimoto-Vollrath principle because it takes into account the theoretical value of exports and is also based on the theoretical value of imports, calculated analogously. Moreover, since it is based on $M_{ikt}$ and $\sum_{p\in K} M_{ipt}$ simultaneously in $\sum_{j\in J} M_{jkt}$ and in $\sum_{j\in J}\sum_{p\in K} M_{jpt}$, the index reflects the doubly relative nature of comparative advantages and is compatible with a flexible concept of them. However, $V'$ does not avoid the small-country bias because it contains $B$; in addition, it does not take GDP into account, and it is not additive. In other words, like the RCA indices discussed above, $V'$ does not address the five major weaknesses of $B$.

## 6. RCA indices based on hypothetical trade balances

There is a class of RCA indices similar to that are constructed from the contribution to the trade balance (CTB). These are based on a modification of the Kunimoto-Vollrath principle in which exports are replaced by the trade balance and the share of each product in the zone's trade (Lafay, 1992 and 1987). If the share of $k$ in the area's trade in period $t$ is denoted by $w_{kt}$:

$$w_{kt} = \frac{\sum_{i\in J}(X_{kit} + M_{kit})}{\sum_{i\in J} \sum_{p\in K}(X_{pit} + M_{pit})} \quad (12)$$

[2] Vollrath (1991) also suggests calculating the logarithm of the first term in $V_{ikt}$ or the difference between the logarithms of each term. These indices do not solve the problem noted for $V_{ikt}$. Moreover, if $X_{ikt}=0$ or $M_{ikt}=0$ then one term or even both are equal to zero, which makes it impossible to calculate the logarithm.

The standard CTB index, denoted as $C$, is calculated as follows:

$$C_{ikt} = \frac{X_{ikt} - M_{ikt} - w_{kt} \sum_{l \in K}(X_{ilt} - M_{ilt})}{\sum_{i \in J} \sum_{p \in K}(X_{pit} + M_{pit})} \tag{13}$$

$X_{ikt}$–$M_{ikt}$ is the observed trade balance associated with $\langle i, k, t \rangle$, and $w_{kt} \sum_{l \in K}(X_{ilt} - M_{ilt})$ corresponds to the theoretical trade balance, calculated from the country's trade balance weighted by the share of the product in the zone's trade. To reveal comparative advantages (disadvantages), the observed balance must be greater (less) than the theoretical balance, which can be expressed as $C_{ikt}>0$ ($C_{ikt}<0$) by calculating the difference between the two. This index is normalized to total trade (not to total exports, as in $N$).

By nature $C$, is compatible with the Kunimoto-Vollrath principle. Like $V'$, $C$ takes into account the entire structure of exports and imports and is thus in keeping with the relative nature of comparative advantages. Of the five weaknesses identified, the only one $C$ does not solve is the absence of GDP in its calculation. In particular, $C$ is compatible with a flexible concept of comparative advantages. According to the logic of $C$, the trade balance must be large enough to reflect a better combination of productivity and differentiation. For this purpose, the theoretical trade balance determines whether the actual trade balance can be considered sufficiently high or not. Unlike $B2$, it will not always be a zero balance that determines whether there are advantages or disadvantages, but an individualized balance for each $\langle i, k, t \rangle$.

There are a number of variants of $C$. First, it is possible to normalize the index on the GDP of the country in question instead of on its total trade. This gives rise to the RCA index denoted $CY$:

$$CY_{ikt} = \frac{1}{Y_{it}} \left[ X_{ikt} - M_{ikt} - w_{kt} \sum_{l \in K} (X_{ilt} - M_{ilt}) \right] \tag{14}$$

$CY$ makes it possible to take the size of the economy into account (but not the size of each economy in $J$, unlike $B2G$). If the difference between the observed trade balance and the actual balance is the same in two countries, $CY$ will be greater for the country with the smaller GDP. In fact, when GDP is smaller, trade in $k$ is a larger share of a country's economy. Because of this higher degree of specialization, an identical difference between the two balances should reveal greater comparative advantages, and normalization based on GDP makes it possible to generate this effect. However, normalization means that additivity across countries is lost.

Second, in addition to normalization based on GDP, a procedure has been proposed for adjusting trade flows to reflect comparative advantages more accurately. This is because trade flows are subject to short-term fluctuations, which do not imply a change in the RCA index. One solution is to assume that in some benchmark period, denoted as $r$, the trade share of $k$, $\{w_{kr}; k \in K\}$ is associated with a minimization of the cyclical bias affecting trade flows. Hence, $X_{ikt}$ and $M_{ikt}$ must be multiplied by $w_{kr}/w_{kt}$ so that $w_{kt}=w_{kr}$ $\forall k \in K$ and the structure of trade flows in each period corresponds to the structure in period $r$ (Stellian and Danna-Buitrago, 2017). This results in the RCA index denoted as $CY^r$:

$$CY^r_{ikt} = \frac{1}{Y_{it}} \left[ \frac{w_{kr}}{w_{kt}} (X_{ikt} - M_{ikt}) - w_{kt} \sum_{l \in K} \frac{w_{lr}}{w_{lt}} (X_{ilt} - M_{ilt}) \right] \tag{15}$$

The trade flow adjustment procedure is a strength that no other RCA index has, although it entails the loss of product additivity. This is an additional weakness, since normalization based on GDP, which is also applied to $CY^r$, means the loss of additivity across countries.

## 7. The RCA index of Leromain and Orefice (2014)

This RCA index is constructed differently from the others and is based on the estimation of the following equation:

$$ln\ x_{ijkt} = \delta_{ijt} + \delta_{ikt} + \delta_{jkt} + \varepsilon_{ijkt} \tag{16}$$

This equation decomposes the flow of exports from country $i$ to country $j$ of product $k$ in time $t$ ($x_{ijkt}$) into four parts:

(i) $\delta_{ijt}$ is the share of $x_{ijkt}$ relative to $\langle i, j\rangle$ irrespective of the product under consideration.

(ii) $\delta_{ikt}$ is the share of $x_{ijkt}$ relative to $\langle i, k\rangle$ irrespective of the country of destination.

(iii) $\delta_{jkt}$ is the share of $x_{ijkt}$ relative to $\langle j, k\rangle$ irrespective of the country of origin.

(iv) $\varepsilon_{ijkt}$ is the error term.

**Observation 1**: Equation (16) is an additive decomposition of $x_{ijkt}$. It is also possible to decompose $x_{ijkt}$ multiplicatively as follows:

$$x_{ijkt} = \phi_{ijt}\ \phi_{ikt}\ \phi_{jkt} + \varepsilon_{ijkt}$$

In this case, $ln\ \phi_{ijt}$, $ln\ \phi_{ikt}$ and $ln\ \phi_{jkt}$ have the same meaning as $\delta_{ijt}$, $\delta_{ikt}$ and $\delta_{jkt}$. The difference lies in the estimation method. If the decomposition is additive, the estimation can be performed by ordinary least squares. If the decomposition is multiplicative, the estimation must be done differently, such as by using the method of moments (French, 2017).

Once equation (16) has been estimated, the Costinot, Donaldson and Komunjer (2012) model[3] can be used to write the following:

$$\delta_{ikt} = \theta \ln z_{ikt} \tag{17}$$

where $z_{ikt}$ is a proxy for the basic productivity of $i$ in terms of $k$ at time $t$. The coefficient $\theta>1$ regulates the influence of $z_{ikt}$ on $x_{ijkt}$. A higher value of $\theta$ implies a higher value of $x_{ijkt}$. In the Costinot, Donaldson and Komunjer (2012) model, $\theta$ captures potential deviations in basic productivity in the different varieties of $k$. A higher value of $\theta$ means that these deviations are smaller. Consequently, from one variety of $k$ to the next, productivity is less dispersed around the basic productivity.

Equation (17) gives the following expression for $z_{ikt}$:

$$z_{ikt} = e^{\frac{\delta_{ikt}}{\theta}} \tag{18}$$

Two variables can then be calculated:

(i) Productivity, $z_{ikt}$, normalized on the average productivity of $i$ in $t$, namely $(1/\#K)\ \sum_{l\in K} z_{ilt}$.

(ii) The countries' average productivity in terms of $k$ in period $t$, namely $(1/\#J)\ \sum_{j\in J} z_{jkt}$, before normalizing to the average productivity of the countries for all products in the same period, namely $(1/\#J\#K)\ \sum_{j\in J}\sum_{l\in K} z_{jlt}$.

3 This model conceptualizes a world economy with a single factor of production (labor) that is perfectly mobile within a country but immobile from one country to another. The other features of the model are constant returns, heterogeneous productivity among different varieties of a product, perfect markets and frictions in international trade (Costinot, Donaldson and Komunjer, 2012 and French, 2017).

These two variables reveal comparative advantages in the case of $\langle i, k, t\rangle$ if the first variable exceeds the second. This means that in period $t$, country $i$ has higher productivity, on average, than other countries belonging to trade zone $J$. Leromain and Orefice (2014) construct their RCA index by dividing the first variable by the second, similar to the calculation of $B$, so as to extend the Kunimoto-Vollrath principle by calculating a theoretical productivity from the average productivities (for countries, for products, and for countries and products simultaneously). This index is denoted as $Z$:

$$Z_{ikt} = \left( \frac{z_{ikt}}{\frac{1}{\#K} \sum_{l \in K} z_{ilt}} \right) / \left( \frac{\frac{1}{\#J} \sum_{j \in J} z_{jkt}}{\frac{1}{\#J\#K} \sum_{j \in J} \sum_{l \in K} z_{jlt}} \right) \tag{19}$$

Calculation of the ratio gives rise to the asymmetry problem. To avoid this, it is possible to calculate the difference according to the same rationale used for $BA$, which gives rise to the RCA index denoted as $ZA$:

$$ZA_{ikt} = \frac{z_{ikt}}{\frac{1}{\#K} \sum_{l \in K} z_{ilt}} - \frac{\frac{1}{\#J} \sum_{j \in J} z_{jkt}}{\frac{1}{\#J\#K} \sum_{j \in J} \sum_{l \in K} z_{jlt}} \tag{20}$$

The main strength of $Z$ and $ZA$ is that they are based on a theoretical model of the world economy. In addition, both consider the complete structure of flows at the disaggregated level through variables of the type $x_{ijkt}$ instead of $X_{ikt}$ and $M_{ikt}$. However, neither $Z$ nor $ZA$ possesses additivity, and they only capture comparative advantages in the form of productivity differentials. Moreover, neither is compatible with a flexible concept of comparative advantages.

# III. Standardized method for assessing the quality of empirical measurements of RCA indices

Table 1 summarizes the analysis of the RCA indices discussed in the previous section. This synthesis clearly shows that there is no RCA index that does not suffer from at least one of the five weaknesses of $B$ and is also compatible with the Kunimoto-Vollrath principle, reflects the relative nature of comparative advantages and has additional strengths, such as the possibility of adjusting trade flows to correct for cyclical bias, of using disaggregated trade flows, or of being underpinned by a theoretical model.

In this sense, the RCA-CTB indices are the most consistent, but they pose a number of dilemmas that should not be ignored. As explained above, additivity across countries is lost when taking the GDP of the country into account, and additivity across products is lost when adjusting for trade flows. Moreover, even if $Z$ and $ZA$ are based on a theoretical model and use disaggregated flows, they are not additive, are not linked to a flexible concept of comparative advantages and do not take GDP into account.

Consequently, analysing the formulae through which the RCA indices measure comparative advantages is not sufficient to find one that is superior to the others. For this reason, it is necessary to strike a balance between the strengths and weaknesses inherent in the main RCA indices available in the literature. However, as noted in the introduction, the quality of the empirical measurements of the various RCA indices must also be taken into account in this search. In the following subsection, three criteria are used to assess quality.

**Table 1**
Synthesis of the comparative analysis of the RCA indices

| | *B* | *BA* | *BS* | *N* | *B2* | *B2D* | *B2G* | *V'* | *C* | *CY* | $CY^r$ | *Z* | *ZA* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Compatibility with the Kunimoto-Vollrath principle | ✓ | ✓ | ✓ | ✓ | | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Compatibility with the relative nature of comparative advantages: | | | | | | | | | | | | | |
| • Across products | ✓ | ✓ | ✓ | ✓ | | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| • Across countries | ✓ | ✓ | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Symmetry | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ |
| Absence of small-country bias | | ✓ | ✓ | ✓ | ✓ | | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Flexible concept of comparative advantages | | | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | |
| RCA index based on GDP | | | | | | | ✓ | | | ✓ | | | |
| Additivity across products | | | | ✓ | | | | | ✓ | ✓ | | | |
| Additivity across countries | | | | ✓ | | | | | ✓ | | | | |
| Adjustment of trade flows to correct for cyclical bias | | | | | | | | | | | ✓ | | |
| RCA index based on disaggregated trade flows | | | | | | | | | | | | ✓ | ✓ |
| RCA index supported by a theoretical model | | | | | | | | | | | | ✓ | ✓ |

**Source:** Prepared by the authors.

## 1. Criteria for assessing the quality of empirical measurements of an RCA index

The first criterion is trend-stationarity through time. For a given $\langle i, k\rangle$ the value taken by $VCR_{ikt}$ should not tend to change significantly from one period to another, as (dis)advantages generally change only over long-term horizons (Lafay, 1987; Leromain and Orefice, 2014). If $\langle VCR_{ikt}: t \in T\rangle$ exhibits a certain degree of volatility over time, this is unlikely to reflect changes in comparative advantages alone. Accordingly, volatility prevents trade flows from revealing comparative advantages correctly. This does not mean that comparative advantages are immutably fixed, but when choosing between two RCA indices, preference should be given to the one that offers the highest trend-stationarity over time to avoid overestimating genuine changes in comparative advantages (Danna-Buitrago, 2017).

The second criterion is symmetry in the distribution of comparative advantages and disadvantages. For example, $N$ is symmetric around zero. This section considers another type of symmetry. By definition, a country will always have comparative advantages for some products and comparative disadvantages for others (Yu, Cai and Leung, 2009). From this perspective, symmetry can be conceptualized in two different ways:

(i) In terms of quantity: for a given $\langle i, t\rangle$ and denoting $K^+ \subseteq K$ as the set of product categories with advantages and $K^- \subseteq K$ as the set with disadvantages, symmetry results from $\#K^+ = \#K^-$ (or $\#K^+ = \#K^- \pm 1$ if one of the sets contains an even number of elements and the other has an odd number).

(ii) In terms of value: for a given $\langle i,t\rangle$, symmetry exists if the total of the distances between the values of an RCA index revealing comparative advantages and the neutral value is equal to the total of the distances between the values revealing comparative disadvantages and the neutral value.

In general, symmetry is unlikely to be observed. However, when choosing between two RCA indices, the index with greater symmetry should be preferred.

The third criterion is consistency in the rankings generated by an RCA index (Yeats, 1985; Leromain and Orefice, 2014). This takes two forms. The first is consistency between the intercountry and intracountry rankings: if the value of an RCA index ranks a country as the first among several for a certain product (intercountry rank), this same value should be among the highest among the values calculated for the country in question (intracountry rank). Conversely, if the value of an RCA index ranks a country last among several for a certain product, this same value should also be among the lowest of all values calculated for the country in question, and likewise for any intermediate intercountry rank.

The second form is consistency in the intercountry rankings. If, in period $t$, country $i$ is ranked first for $k$ among several countries because $VCR_{ikt} \geq VCR_{jkt}\ \forall j \neq i$, then $i$ should also be ranked as the first country for any other product $l$ for which $VCR_{ilt}$ is considered sufficiently close to $VCR_{ikt}$. Similarly, if $i$ is classified in $t$ as second among several countries for $k$ because $VCR_{ikt} < VCR_{jkt}\ \exists j \neq i$, then $i$ should also be classified as the second country for any $l \neq k$ for which $VCR_{ilt}$ is considered sufficiently close to $VCR_{ikt}$; and so on for all possible positions in the ranking. Between two RCA indices, preference should be given to the one that classifies the countries most consistently.

Next, variables are calculated that measure each of the three criteria presented above. This process is illustrated using the following universe:

- $J$: Northern Triangle and Colombia.
- $K$: The 255 product categories in SITC Revision 3 (United Nations, 1986).
- $T$: Every year from 1995 to 2017.

For this purpose, data on trade flows are taken from UNCTADstat (UNCTAD, n/d), and GDP data are taken from *World Economic Outlook* (IMF, 2022). The reference year used to adjust the trade flows in the case of $CY^r$ is 2015, the year in which the free trade agreement between Colombia and the Northern Triangle came into force. The reference year may be altered in future research. To calculate $Z$ and $ZA$ the value of $\theta$ is set at $\theta = 6.534$, which is an estimate from Costinot, Donaldson and Komunjer (2012) (see Leromain and Orefice (2014) for a discussion of this). Descriptive statistics for each RCA index are given in table 2.[4]

**Table 2**
RCA indices for the Northern Triangle and Colombia, 1995–2017, SITC nomenclature

| | Mean | Mode | Standard deviation | Median | Minimum | Maximum |
|---|---|---|---|---|---|---|
| $B$ | 1.071238 | 0.123252 | 1.453552 | 0.626614 | 0 | 12.44475 |
| $BA$ | 1.15e-19 | -0.000350 | 0.010443 | -2.7e-05 | -0.13392 | 0.485808 |
| $BS$ | -0.27102 | 0.086298 | 0.543078 | -0.22955 | -1 | 0.851243 |
| $B2$ | 0.00887 | -0.003420 | 0.694176 | 0 | -1 | 1 |
| $B2D$ | -0.28968 | 0.516978 | 21.36610 | -0.38261 | -245.074 | 247.0740 |
| $B2G$ | -0.40082 | 0.024104 | 2.876211 | 0 | -11.5670 | 11.56699 |
| $V'$ | -0.163642 | 0.109415 | 4.930400 | 0 | -139.200 | 12.44475 |
| $N$ | 5.38e-20 | -0.000540 | 0.002057 | -5.4e-06 | -0.03967 | 0.092898 |
| $C$ | -3e-20 | 0.001191 | 0.001496 | -5.8e-07 | -0.04187 | 0.047364 |
| $CY$ | -1e-20 | 0.000799 | 0.000215 | -2.3e-08 | -0.00643 | 0.005074 |
| $CY^r$ | 1.09e-20 | -0.000880 | 0.000203 | -1.2e-08 | -0.00226 | 0.003442 |
| $Z$ | 1.751084 | 0.975745 | 1.108850 | 1.566902 | 0.07682 | 9.811509 |
| $ZA$ | 0.750231 | 0.097259 | 1.307500 | 0.383862 | -2.52129 | 13.54330 |

**Source:** Prepared by the authors.

[4] The calculations can be requested from author Rémi Stellian.

## 2. Trend-stationarity through time

The first way to measure trend-stationarity through time is to use the standard deviation (Leromain and Orefice, 2014). It is possible to calculate the standard deviation for the $\#T$ values of an RCA index for a given country and a given product category. A lower standard deviation means that for the $\langle i, k\rangle$ pair considered, the RCA index is less dispersed around its mean through time, so the stationarity is greater. Denoting $\sigma_{ik}$ as the standard deviation of $\langle VCR_{ikt}$:$t \in \mathrm{T}\rangle$, the first variable used to evaluate the trend-stationarity over time of an RCA index is the mean of $\sigma_{ik}$:

$$\bar{\sigma} = \frac{1}{\#J \times \#K}\sum_{i\in J}\sum_{k\in K}\sigma_{ik} \tag{21}$$

Table 3 summarizes the value of $\bar{\sigma}$ for each RCA index. The RCA-CTB indices, as well as $BA$ and $N$, each have a minimum standard deviation.

**Table 3**
Average standard deviation of each RCA index

| *B* | *BA* | *BS* | *B2* | *B2D* | *B2G* | *V'* |
|---|---|---|---|---|---|---|
| 0.73 | 2.67e-03 | 0.32 | 0.40 | 12.25 | 1.56 | 1.73 |
| *N* | *C* | *CY* | *CY$^r$* | *Z* | *ZA* | |
| 5.74e-04 | 4.28e-04 | 5.16e-05 | 4.33e-05 | 0.77 | 0.82 | |

**Source:** Prepared by the authors.

We suggest extending the measurement of trend-stationarity through time using the variable $\beta$ as estimated in the following equation:

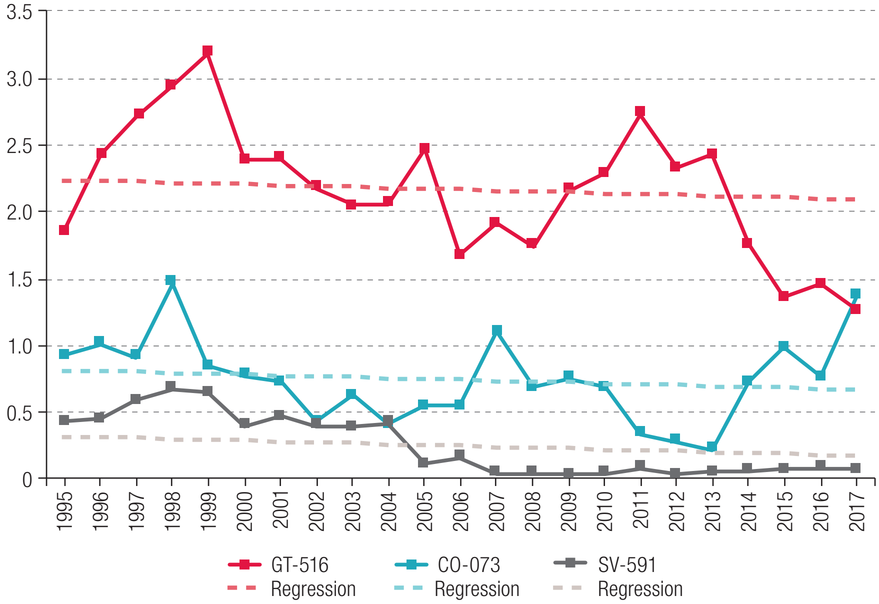
$$VCR_{ikt} = \alpha + \beta t + \gamma_{ik} + \varepsilon_{ikt} \tag{22}$$

An RCA index calculated for a given $\langle i, k, t\rangle$ is the dependent variable explained by $t$ where $t = 0$ for the first available year (1995 in this case), $t = 1$ for the second available year (1996), and so on up to $t = \#T$ (23 years in this study); $\alpha$ is a constant, $\gamma_{ik}$ is a fixed effect for each country-product combination, and $\varepsilon_{ikt}$ is the error term. Trend-stationarity over time is at a maximum when $\beta = 0$. In this case, the equation is rewritten as $VCR_{ikt} = \alpha + \gamma_{ik} + \varepsilon_{ikt}$, so that $\langle VCR_{ikt} : t \in T\rangle$ tends to stay around a constant long-run value given by $\alpha+\gamma_{ik}$. However, it is possible that the estimation results in $\beta \neq 0$. In this case, if $\beta$ is closer to zero, the RCA index will change less over time and will ultimately be more compatible with the criterion of trend-stationarity through time.

Equation (22) is a modified version of the one suggested by Yu and others (2010): $VCR_{ikt}=\alpha_{ik}+\beta_{ik}\, t+\varepsilon_{ikt,}$ where $\alpha_{ik}$ and $\beta_{ik}$ perform an individualized estimation for each $\langle i, k\rangle$. However, it is advisable to calculate a coefficient $\beta$ that encompasses all countries and all product categories. This allows for a more synthetic reading of trend-stationarity through time without losing the specificity of each country-product combination $\langle i, k\rangle$ through the fixed effect $\gamma_{ik}$. Similarly, Laursen (2015) proposes to estimate $VCR_{ikt1} = \alpha_i + \beta_i VCR_{ikt0} + \varepsilon_{ik}$. According to this equation, there is greater stationarity between the initial period ($t_0 = 1995$) and the final period ($t_1$=2017) for country $i$, with $\beta_i$-$1|\rightarrow 0$ and $|\alpha|\rightarrow 0,$ because in this case $VCR_{ikt1} \rightarrow VCR_{ikt0} + \varepsilon_{ik}$. By contrast, equation (22) has the merit of considering all periods and not just the initial and final ones.

**Example 1**: Figure 1 plots the RCA index calculated by $B$ for the following country-product combinations: Colombia – 073 (chocolate), El Salvador – 591 (insecticides) and Guatemala – 516 (other organic chemicals). The estimation of equation (22) results in $\alpha = 0.139$ and $\beta = -0.00657$. The values of the fixed effects are 0.679, 0.185 and 2.097, respectively.

**Figure 1**
Estimation of equation (22): Balassa (1965) RCA index




**Source:** Prepared by the authors.

Table 4 reports the estimation of equation (22). Only $BA$, $N$ and the RCA-CTB indices result in $\beta = 0$, with the null hypothesis accepted with a p-value equal to 1.

**Table 4**
Estimation of equation (22)

| | *B* | *BA* | *BS* | *B2* | *B2D* | *B2G* | *V'* |
|---|---|---|---|---|---|---|---|
| $\beta$ | -0.00657*** | 1.40e-21 | -0.000335 | -0.00225*** | 0.00648 | -0.0153*** | -0.00656 |
| | (-6.69) | (0.00) | (-0.95) | (-4.98) | (0.32) | (-7.24) | (-1.76) |
| $\alpha$ | 0.139 | -0.00257 | -0.877*** | 0.758*** | -0.109 | 0.173 | 0.111 |
| | -0.67 | (-1.62) | (-11.69) | -7.9 | (-0.03) | -0.38 | (0.14) |
| | *N* | *C* | *CY* | *CY^r* | *Z* | *ZA* | |
| $\beta$ | 4.97e-22 | -1.45e-21 | 8.61e-23 | 6.47e-23 | 0.0502*** | 0.0503*** | |
| | (0.00) | (-0.00) | (0.00) | (0.00) | (-64.56) | (-54.47) | |
| $\alpha$ | -0.00027 | -0.00012 | -1.5e-06 | -1.2e-06 | -0.117 | -1.185*** | |
| | (-0.85) | (-0.54) | (-0.05) | (-0.05) | (-0.71) | (-6.04) | |

**Source:** Prepared by the authors.
**Note:** *p<0.05, **p<0.01 and *** p<0.001; t-statistics in parentheses.

## 3. Symmetry in the distribution of advantages and disadvantages

In the case of symmetry with respect to quantity,[5] the absolute difference between the number of product categories with comparative advantages and the number with comparative disadvantages, $D_{it}$, is first calculated for $\langle i, t \rangle$:

$$D_{it} = |\#\{k: VCR_{ikt} > v\} - \#\{k: VCR_{ikt} < v\}| \tag{23}$$

where $v \in \{0,1\}$ gives the neutral value.

[5] Leromain and Orefice (2014) measure symmetry using the coefficient of skewness and the meanmedian difference. These statistics are useful for measuring symmetry around the mean of $\{VCR_{ikt}; k \in K\}$. However, the mean in question does not always correspond to the neutral value, although the neutral value cannot be ignored when measuring symmetry.

If $D_{it} = 0$, each category of products with comparative advantages corresponds to a category of products with comparative disadvantages for $\langle i, t\rangle$. In this case, symmetry with respect to quantity would be complete. If $D_{it} \neq 0$ there is asymmetry in quantity, and $D_{it}$ represents the number of product categories that have no counterpart in terms of advantages or disadvantages.

The average value of $D_{it}$ is then calculated as:

$$\bar{D} = \frac{1}{\#J \times \#T} \sum_{i \in J} \sum_{t \in T} D_{it} \tag{24}$$

A $\bar{D}$ value closer to zero means that the index has a greater capacity to be compatible with quantity symmetry.

In the case of symmetry in terms of value, it is necessary to calculate the variable $E_{it}$ for each $\langle i, t\rangle$:

$$E_{it} = \left| \sum_{k \in K} (VCR_{ikt} - v) \right| \tag{25}$$

The difference between $VCR_{ikt}$ and the neutral value for each product category is calculated, and $E_{it}$ is the sum of these differences in absolute value terms. If $E_{it} = 0$, then the negative differences (disadvantages) are balanced by the positive differences (advantages), and value symmetry would be complete. If $E_{it} > 0$, there is asymmetry with respect to value.

The average value of $E_{it}$ is then calculated as:

$$\bar{E} = \frac{1}{\#J \times \#T} \sum_{i \in J} \sum_{t \in T} E_{it} \tag{26}$$

Table 5 presents the results for the universe studied. $B2D$ generates the greatest symmetry with respect to quantity, followed by the RCA-CTB indices. By contrast, the latter offer the greatest symmetry with respect to value. This is also true for $N$.

**Table 5**
Symmetry with respect to quantity and value for each RCA index

| | *B* | *BA* | *BS* | *B2* | *B2D* | *B2G* | *V'* |
|---|---|---|---|---|---|---|---|
| $\bar{D}$ | 70.02 | 70.02 | 70.02 | 83.07 | 36.83 | 83.07 | 54.20 |
| $\bar{E}$ | 60.02 | 3.03e-17 | 69.30 | 69.92 | 349.23 | 160.86 | 84.28 |
| | *N* | *C* | *CY* | *CY^r* | *Z* | *ZA* | |
| $\bar{D}$ | 70.02 | 44.91 | 44.91 | 41.22 | 203.83 | 203.83 | |
| $\bar{E}$ | 5.84e-18 | 3.77e-18 | 4.80e-19 | 6.53e-19 | 235.58 | 235.07 | |

**Source:** Prepared by the authors.

**Observation 2**: $B$ and its transformations ($BA$ and $BS$), along with $N$ share the same value of $\bar{D}$. These indices are all based on exports alone, which suggests that the asymmetry is the same regardless of how an RCA index is calculated from exports. The same is true for $C$ and $CY$ but not $CY^r$. This suggests that the normalization variable does not influence asymmetry with respect to quantity if the trade flows are not adjusted. Lastly, the additive version of $Z$ does not influence the quantity asymmetry.

## 4. Consistency in the country rankings

Based on the guidelines proposed by Yeats (1985) and Leromain and Orefice (2014), consistency between the intercountry and intracountry ranking positions is measured by the correlation coefficient between two variables:

(i) The average of the values taken by an RCA index that ranks country $i$ in position $x \in \{1,2,\cdots,\#J\}$.

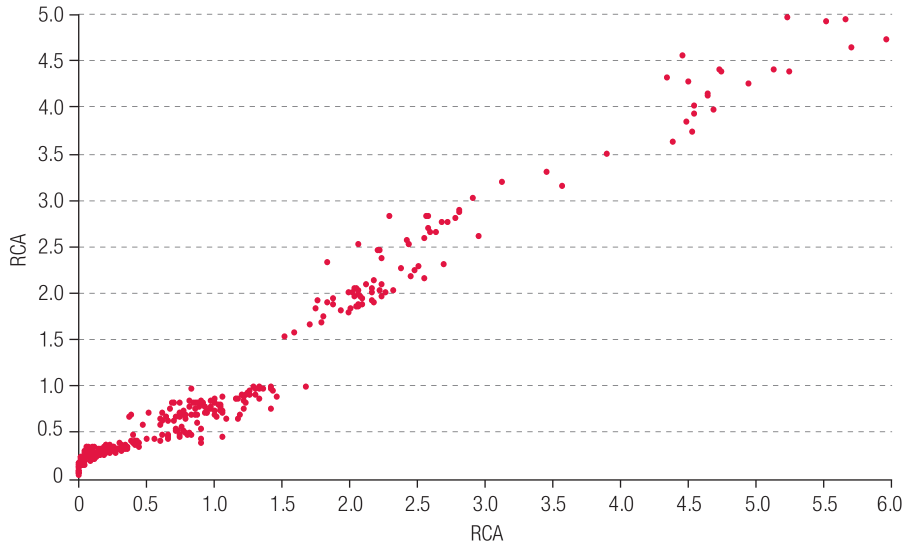
$$VCR_{it}^{x} = \frac{1}{\#K_{it}^{x}} \sum_{k\in K_{it}^{x}} VCR_{ikt} \quad \text{with:}$$
$$K_{it}^{x} = \{k: \#\{j: VCR_{jkt} \le VCR_{ikt}\} = \#J - x, \#\{j: VCR_{jkt} \ge VCR_{ikt}\} = x - 1\}\} \qquad (27)$$

(ii) The average of the values taken by an RCA index between percentile $(100/\#J) \cdot (\#J{-}x)$ —not included— and percentile $(100/\#J)\cdot(\#J{+}1{-}x)$ (or values of $\langle VCR_{ikt}{:}k{\in}K\rangle$ less than or equal to percentile $100/\#J$ if $x = \#J$), where $p_{it}\,(Y)$ is the $Y$th percentile of $\{VCR_{ikt}{:}k \in K\}$:

$$vcr_{it}^{x} = \frac{1}{\#\kappa_{it}^{x}} \sum_{k\in \kappa_{it}^{x}} VCR_{ikt}, \quad \text{with:}$$
$$\kappa_{it}^{x} = \begin{cases} \left\{k: p_{it}\left(\frac{100}{\#J}(\#J - x)\right) < VCR_{ikt} \le p_{it}\left(\frac{100}{\#J}(\#J + 1 - x)\right)\right\} & \text{if } x < \#J \\ \{k: VCR_{ikt} \le p_{it}(100/\#J)\} & \text{if } x = \#J \end{cases} \qquad (28)$$

Thus, the correlation coefficient of $\{\langle VCR_{it}^{x}; vcr_{it}^{x}\rangle: i \in J, x \in \{1,2,\cdots,\#J\}, t \in T\}$ is calculated. A coefficient closer to 1 indicates greater consistency between the intercountry and intracountry rankings. Figure 2 illustrates the dispersion of the points, and table 6 presents the correlation coefficient for each RCA index. *B2* has the lowest coefficient, while *B* has the highest.

**Figure 2**
Consistency between intercountry and intracountry rankings,
Balassa (1965) RCA index




**Source:** Prepared by the authors.

**Table 6**
Correlation coefficient between the intercountry and intracountry rankings for each RCA index

| *B* | *BA* | *BS* | *B2* | *B2D* | *B2G* | *V* |
|---|---|---|---|---|---|---|
| 0.9849 | 0.9517 | 0.9116 | 0.8307 | 0.9666 | 0.9745 | 0.7775 |
| *N* | *C* | *CY* | *CY*ʳ | *Z* | *ZA* | |
| 0.9012 | 0.9356 | 0.9526 | 0.9576 | 0.8523 | 0.9109 | |

**Source:** Prepared by the authors.

To achieve consistency in the intercountry rankings, a measurement is suggested based on the following accounting for bias:

- For each $k \in K_{it}^{2} \cup K_{it}^{3} \cup \cdots \cup K_{it}^{\#J}$, a bias exists if the corresponding RCA index is above the third quartile of $K_{it}^{1}$.
- For each $k \in K_{it}^{3} \cup K_{it}^{4} \cup \cdots \cup K_{it}^{\#J}$, a bias exists if the corresponding RCA index is above the third quartile of $K_{it}^{2}$, and for each $k \in K_{it}^{1}$, there is a bias if the corresponding RCA index is below the first quartile of $K_{it}^{2}$.
- For each $k \in K_{it}^{4} \cup K_{it}^{5} \cup \cdots \cup K_{it}^{\#J}$, a bias exists if the corresponding RCA index is above the third quartile of $K_{it}^{3}$, and for each $k \in K_{it}^{1} \cup K_{it}^{2}$, there is a bias if the corresponding RCA index is below the first quartile of $K_{it}^{3}$.
- ... and so on successively, to count how many elements $k$ in $K_{it}^{1} \cup K_{it}^{2} \cup K_{it}^{3} \cup \cdots \cup K_{it}^{\#J-1}$ correspond to an RCA index below the first quartile of $K_{it}^{\#J}$.

The rationale for this accounting is as follows. If, in any period, an RCA index ranks a country $i$ in $x$th place for certain products, then:

- For all ranking positions below $x$, $i$ should not have an RCA index above three quarters of the lowest RCA indices (quartile 3) that classify it as country $x$.
- For all positions above $x$, $i$ should not have an RCA index higher than three quarters of the highest RCA indices (quartile 1) that rank it as country $x$.

In the present analysis, quartiles 1 and 3 are used as a starting point. Future work could analyse the extent to which the accounting for bias changes if quartiles 1 and 3 are replaced by other magnitudes, such as the fifth and ninety-fifth percentiles.

If $Q_1(K_{it}^{x})$ and $Q_3(K_{it}^{x})$ are the first and third quartiles of the values associated with $K_{it}^{x}$, the number of biases in the classification of $i$ as country number $x$ in time period $t$ denoted as $s_{it}^{x}$, is calculated as follows:

$$s_{it}^{x} = \#\left\{k \in \bigcup_{y=1}^{x-1} K_{it}^{y} : VCR_{ikt} < Q_1(K_{it}^{x})\right\} \cup \left\{k \in \bigcup_{y=x+1}^{n} K_{it}^{y} : VCR_{ikt} > Q_3(K_{it}^{x})\right\} \tag{29}$$

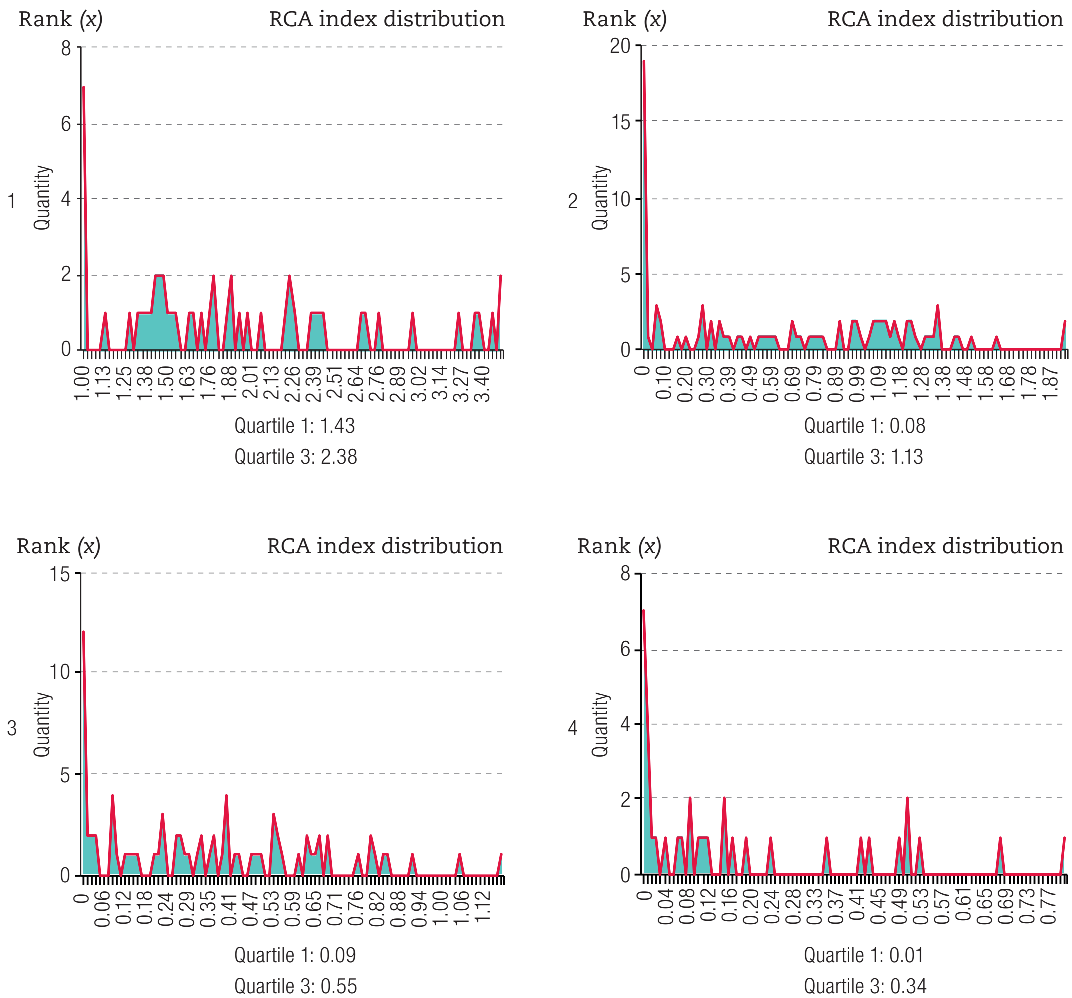
**Observation 3**: If $VCR_{ikt} = VCR_{ilt}$, then $k$ and $l$ share the same intracountry rank. If $VCR_{ikt} = VCR_{jkt}$, then $i$ and $j$ share the same intercountry rank.

Lastly, the total biases are calculated for each $\langle i, t\rangle$ before inferring the average of these totals:

$$\bar{S} = \frac{1}{\#J \times \#T} \sum_{i \in J} \sum_{t \in T} \left( \sum_{x=1}^{\#J} s_{it}^{x} \right) \tag{30}$$

**Example 2**: Figure 3 shows the $B$ values that led to El Salvador in 2017 being ranked first with respect to a number of products and in the second, third and fourth places with respect to others. Data on quartiles 1 and 3 are also included. Thus, for example, quartile 1 of rank 3 is 0.09, but El Salvador is ranked as the second country (in other words, one place higher), with an RCA index below 0.09, in 25 product categories. In addition, quartile 3 of rank 3 is 0.55, but El Salvador ranks fourth (one rank lower), with an RCA index greater than 0.55, in two product categories. Consequently, $B$ generates 27 biases by ranking El Salvador third. Figure 4 shows the total cumulative biases for each country in each year.

**Figure 3**
Ranking of El Salvador relative to Colombia, Guatemala and Honduras according to the Balassa (1965) RCA index in 2017




**Source:** Prepared by the authors.

Table 7 presents the results for the universe studied. *V′* generates the least bias (6.04), followed by *B2G* (7.66) and *B2* (13.60).

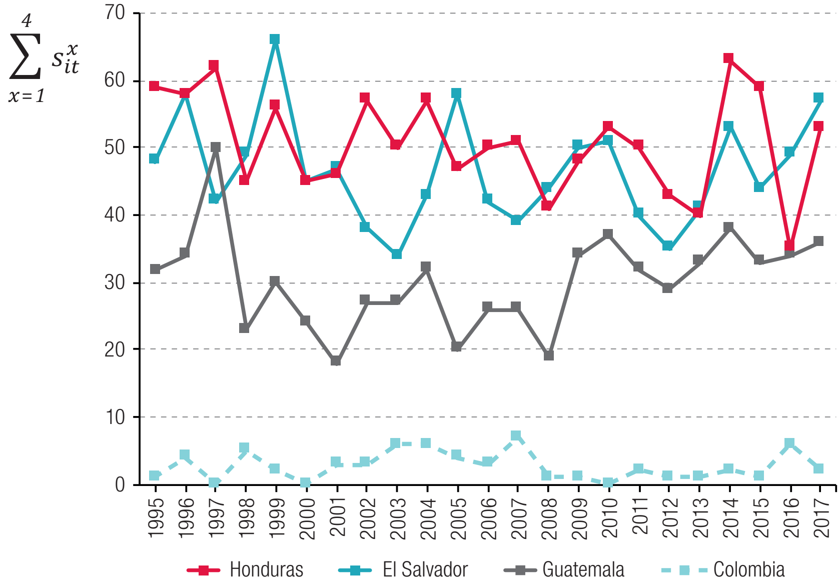
**Table 7**
Average biases in the intercountry rankings according to each RCA index

| *B* | *BA* | *BS* | *B2* | *B2D* | *B2G* | *V′* |
|---|---|---|---|---|---|---|
| 32.57 | 32.57 | 30.60 | 13.60 | 20.96 | 7.66 | 6.04 |
| *N* | *C* | *CY* | *CY*$^r$ | *Z* | *ZA* | |
| 32.11 | 28.26 | 41.53 | 36.73 | 23.61 | 34.45 | |

**Source:** Prepared by the authors.

**Figure 4**

Biases in the classification of countries in the Northern Triangle and Colombia zone according to the Balassa (1965) RCA index




**Source:** Prepared by the authors.

## 5. Summary

Table 8 presents the two measurements obtained for each of the RCA indices studied and the average of these six variables. This average is a synthetic variable that indicates the extent to which an RCA index measures comparative advantages adequately in a given context. By calculating this average with the absolute value of $\beta$ (distance between $\beta$ and 0) and the distance between the correlation coefficient and 1 ($|cc-1|$ in the case of the third criterion), the comparative advantage measurement is of higher quality when the final average is closer to 0.

**Table 8**

Summary of empirical measurements

| | *B* | *BA* | *BS* | *B2* | *B2D* | *B2G* | *V'* |
|---|---|---|---|---|---|---|---|
| $\bar{\sigma}$ | 0.73 | 2.67e-03 | 0.32 | 0.40 | 12.25 | 1.56 | 1.73 |
| $\lvert\beta\rvert$ | 6.57e-03 | 1.40e-21 | 3.35e-04 | 2.25e-03 | 6.48e-03 | 1.53e-02 | 6.56e-03 |
| $\bar{D}$ | 70.02 | 70.02 | 70.02 | 83.07 | 36.83 | 83.07 | 54.20 |
| $\bar{E}$ | 60.02 | 3.03e-17 | 69.30 | 69.92 | 349.23 | 160.86 | 84.28 |
| $\lvert cc-1\rvert$ | 0.02 | 0.05 | 0.09 | 0.17 | 0.03 | 0.03 | 0.22 |
| $\bar{S}$ | 32.57 | 32.57 | 30.60 | 13.60 | 20.96 | 7.66 | 6.04 |
| Average | 27.23 | 17.11 | 28.39 | 27.86 | 69.88 | 42.19 | 24.41 |
| | *N* | *C* | *CY* | *CY^r^* | *Z* | *ZA* | |
| $\bar{\sigma}$ | 5.74e-04 | 4.28e-04 | 5.16e-05 | 4.33e-05 | 0.77 | 0.82 | |
| $\lvert\beta\rvert$ | 4.97e-22 | 1.45e-21 | 8.61e-23 | 6.47e-23 | 5.02e-02 | 5.03e-02 | |
| $\bar{D}$ | 70.02 | 44.91 | 44.91 | 41.22 | 203.83 | 203.83 | |
| $\bar{E}$ | 5.84e-18 | 3.77e-18 | 4.80e-19 | 6.53e-19 | 235.58 | 235.07 | |
| $\lvert cc-1\rvert$ | 0.0988 | 0.0644 | 0.0474 | 0.0424 | 0.1477 | 0.0891 | |
| $\bar{S}$ | 32.11 | 28.26 | 41.53 | 36.73 | 23.61 | 34.45 | |
| Average | 17.04 | 12.21 | 14.42 | 13.00 | 77.33 | 79.05 | |

**Source:** Prepared by the authors.

Table 8 illustrates this average in the case studied. The best measurements result from the RCA indices that relate to contribution to the trade balance. The measurement is best if $C$ is preferred; however, $C$ does not take into account GDP, and no adjustments are made to correct for the cyclical bias in trade flows. Therefore, one could select $CY$, for which the final average, although not as good as that of $C$, is still better than the averages achieved by the other indices. However, as noted above, when going from $C$ to $CY$, additivity across countries is lost. Following the same logic, it would be possible to select $CY^r$, but then additivity across products would be lost. It is worth noting that other RCA indices might provide better measurements for different trade zones. In this regard, section II showed that it is impossible to give a definitive answer as to which RCA index is theoretically most appropriate (inherent strengths or weaknesses of RCA index formulae), and there is no *a priori* reason why the same RCA index should provide the best empirical measures for a large sample of universes $J \times K \times T$.

# IV. Conclusion

Which index of revealed comparative advantage (RCA) should be applied to certain countries, products and time periods? To help answer this fundamental question, this article systematically reviewed the inherent strengths and weaknesses of RCA index formulae and then devised a standardized method for assessing the quality of an RCA index's empirical measurements. By combining these two contributions, a series of theoretical and empirical considerations were formalized to weigh the pros and cons of different RCA indices more effectively and ultimately help choose one. The example of the trade zone formed by the Northern Triangle and Colombia suggests that RCA indices related to the contribution to the trade balance should be preferred, although other indices cannot be ruled out *a priori* for other trade zones.

Once an RCA index has been chosen, it can be used in different ways. For example, Stellian and Danna-Buitrago (2017) investigate the capacity of an RCA index to stay above a critical value over time and thus reveal significant comparative advantages. Another option is to study whether the international specialization pattern actually corresponds to that revealed by the structure of comparative advantages (Konstantakopoulou and Tsionas, 2019). How to apply an RCA index could also be the subject of a standardized method and would constitute a future line of research. Thus, it would be possible to encourage the generalization of protocols for the selection and use of RCA indices, in order to have tools available that are part of the common language of researchers working on issues in international economics. This could facilitate debate and discussion on these issues with a view to providing a more solid basis for international integration policies.

# Bibliography


Abbas, S. and A. Waheed (2017), "Trade competitiveness of Pakistan: evidence from the revealed comparative advantage approach", *Competitiveness Review*, vol. 27, No. 5.

Balassa, B. (1986), "Comparative advantage in manufactured goods: a reappraisal", *The Review of Economics and Statistics*, vol. 68, No. 2.

___(1965), "Trade liberalization and 'revealed' comparative advantage", *The Manchester School*, vol. 33, No. 2.

Brakman, S. and C. Van Marrewijk (2017), "A closer look at revealed comparative advantage: gross-versus value-added trade flows", *Papers in Regional Science*, vol. 96, No. 1.

Cai, J., P. Leung and N. Hishamunda (2009), "Assessment of comparative advantage in aquaculture: framework and application on selected species in developing countries", *FAO Fisheries and Aquaculture Technical Paper*, No. 528, Rome, Food and Agriculture Organization of the United Nations (FAO).

Chanteau, J.-P. (2007), "La notion d'avantage comparatif peut-elle expliquer les choix de spécialisation industrielle?", *Économies et Sociétés: Relations Économiques Internationales*, P, No. 38.

Costinot, A., D. Donaldson and I. Komunjer (2012), "What goods do countries trade? A quantitative exploration of Ricardo's ideas", *The Review of Economic Studies*, vol. 79, No. 2.

Danna-Buitrago, J. P. (2017), "La Alianza del Pacífico+4 y la especialización regional de Colombia: una aproximación desde las ventajas comparativas", *Cuadernos de Administración*, vol. 30, No. 55.

Donges, J. B. and J. Riedel (1977), "The expansion of manufactured exports in developing countries: an empirical assessment of supply and demand issues", *Weltwirtschaftliches Archiv*, vol. 113, No. 1.

Esquivias, M. A. (2017), "The change of comparative advantage of agricultural activities in East Java within the context of ASEAN economic integration", *Agris on-line Papers in Economics and Informatics*, vol. 9, No. 1.

French, S. (2017), "Revealed comparative advantage: what is it good for?", *Journal of International Economics*, vol. 106.

Gnidchenko, A. A. and V. A. Salnikov (2015), "Net comparative advantage index: overcoming the drawbacks of the existing indices", *Higher School of Economics Research Paper*, No. WP BRP 119/EC/2015.

Hadzhiev, V. (2014), "Overall revealed comparative advantages", *Eurasian Journal of Economics and Finance*, vol. 2, No. 1.

Halilbašić, M. and S. Brkić (2017), "Export specialization of South East European countries in their trade with the European Union", *Economic Review: Journal of Economics and Business*, vol. 15, No. 1.

Hoang, V. and others (2017), "Agricultural competitiveness of Vietnam by the RCA and the NRCA indices, and consistency of competitiveness indices", *Agris on-line Papers in Economics and Informatics*, vol. 9, No. 4.

Hoen, A. and J. Oosterhaven (2006), "On the measurement of comparative advantage", *The Annals of Regional Science,* vol. 40, No. 3.

IMF (International Monetary Fund) (2022), "GDP, current prices" [online] imf.org/external/datamapper/PPPGDP@WEO/OEMDC/ADVEC/WEOWORLD.

Jaimovich, E. and V. Merella (2015), "Love for quality, comparative advantage, and trade", *Journal of International Economics*, vol. 97, No. 2.

Konstantakopoulou, I. and M. G. Tsionas (2019), "Measuring comparative advantages in the Euro Area", *Economic Modelling*, vol. 76.

Kunimoto, K. (1977), "Typology of trade intensity indices", *Hitotsubashi Journal of Economics*, vol. 17, No. 2.

Lafay, G. (1992), "The measurement of revealed comparative advantages", *International Trade Modelling*, M. G. Dagenais and P.-A. Muet (eds.), London, Chapman & Hall.

___(1987), "Avantage comparatif et compétitivité", *Economie Prospective Internationale*, No. 29.

Lassudrie-Duchêne, B. and D. Ünal-Kesenci (2001), "L'avantage comparatif, notion fondamentale et controversée", *L'économie mondiale 2002*, Paris, La Découverte.

Laursen, K. (2015), "Revealed comparative advantage and the alternatives as measures of international specialization", *Eurasian Business Review*, vol. 5, No. 1.

Leromain, E. and G. Orefice (2014), "New revealed comparative advantage index: dataset and empirical distribution", *International Economics*, vol. 139.

Menon, J. (2014), "From spaghetti bowl to jigsaw puzzle? Fixing the mess in regional and global trade", *Asia and the Pacific Policy Studies*, vol. 1, No. 3.

Shaul Hamid, M. F. and M. Aslam (2017), "Intra-regional trade effects of ASEAN free trade area in the textile and clothing industry", *Journal of Economic Integration*, vol. 32, No. 3.

Stellian, R. and J. P. Danna-Buitrago (2019), "Revealed comparative advantages and regional specialization: Evidence from Colombia in the Pacific Alliance", *Journal of Applied Economics,* vol. 22, No. 1.

___(2017), "Colombian agricultural product competitiveness under the free trade agreement with the United States: analysis of the comparative advantages", *CEPAL Review*, No. 122 (LC/PUB.2017/10-P), Santiago, Economic Commission for Latin America and the Caribbean (ECLAC).

UNCTAD (United Nations Conference on Trade and Development) (n/d), UNCTADstat [online database] https://unctadstat.unctad.org/wds/.

United Nations (1986), "Standard international trade classification, revision 3", *Statistical Papers*, series M, No. 34/Rev. 3 (ST/ESA/STAT/SER.M/34/Rev.3), New York.

Vollrath, T. L. (1991), "A theoretical evaluation of alternative trade intensity measures of revealed comparative advantage", *Review of World Economics*, vol. 127, No. 2.

Yeats, A. J. (1985), "On the appropriate interpretation of the revealed comparative advantage index: implications of a methodology based on industry sector analysis", *Review of World Economics*, vol. 121, No. 1.

Yu, R. and others (2010), "Assessing the comparative advantage of Hawaii's agricultural exports to the US mainland market", *The Annals of Regional Science*, vol. 45, No. 2.

Yu, R., J. Cai and P. Leung (2009), "The normalized revealed comparative advantage index", *The Annals of Regional Science*, vol. 43, No. 1.