NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

Andrey A. Gnidchenko, Vladimir A. Salnikov

NET COMPARATIVE ADVANTAGE INDEX: OVERCOMING THE DRAWBACKS OF THE EXISTING INDICES

BASIC RESEARCH PROGRAM
WORKING PAPERS

SERIES: ECONOMICS
WP BRP 119/EC/2015

Andrey A. Gnidchenko[1], *Vladimir A. Salnikov*[2]

NET COMPARATIVE ADVANTAGE INDEX: OVERCOMING THE DRAWBACKS OF THE EXISTING INDICES[3]

In this paper, we examine a number of indices used for measuring comparative advantages of a country in international trade for a good and propose the new net comparative advantage index that has several strong features. First, it reflects net trade, and that's why it is more theoretically grounded than indices calculated only from export data. Second, it is consistent with *Kunimoto* (1977) theoretical framework that is highly appreciated among trade economists. Third, it is not totally focused on a single commodity (that is, it takes world trade structure into consideration), unlike net export index. Fourth, it accounts for economic openness, using GDP as a scale variable. Fifth, it is hardly exposed to structural distortions. Finally, its sign is consistent with the sign of the net trade. We compare the new index with CEPII theoretically grounded econometric indicator and show that the proposed index has better empirical characteristics and is much easier to calculate and interpret.

JEL Classification: F14.

Keywords: international trade, comparative advantages, goods, indices, net trade.

---

[1] Center for Macroeconomic Analysis and Short-Term Forecasting; Institute of Economic Forecasting, Russian Academy of Sciences; National Research University Higher School of Economics. Junior research fellow. E-mail: agnidchenko@forecast.ru

[2] Center for Macroeconomic Analysis and Short-Term Forecasting; Institute of Economic Forecasting, Russian Academy of Sciences; National Research University Higher School of Economics. Leading research fellow.

[3] This Working Paper is an output of a research project implemented as part of the Basic Research Program at the National Research University Higher School of Economics (HSE) in 2015. We are grateful to the anonymous referee for the comments that have helped to clarify the rationale underlying our position.

1. Introduction

During the latest decade, the Ricardian comparative advantage concept is getting much attention in economic research. The revival of interest in comparative advantages has started since *Eaton and Kortum* (2002) published a theoretical paper that successfully combined gravity variables[4] and technological factors.[5] This research was extended by a number of other authors.[6]

Two papers should be mentioned specially. *Costinot, Donaldson and Komunjer* (2012) proposed a new method to calculate comparative advantages that is based on the econometric equations consistent with their new version of Ricardian model in the spirit of *Eaton and Kortum* (2002). *Leromain and Orefice* (2013) developed this method into the database on the new RCA measure. The key characteristic of this measure is its theoretical consistency. We suppose that these two papers are going to discover a new direction of empirical research in comparative advantages, since the way of thinking that gives priority to theoretically consistent empirical measures[7] is a distinctive feature of modern economic studies.

However, we consider that such way of thinking isn't absolutely preferable[8]. In our view, the best comparative advantage index should meet several theoretical and empirical criteria developed in the literature (and not just be consistent with a single theoretical model). First, it should satisfy the *Kunimoto* (1977) theoretical framework – that is, one should be able to express it as the ratio of the expected-to-actual trade, as *Vollrath* (1991) shows. Second, it should have a stable distribution, so that one should be able to compare its values over time, industries and countries (*Hinloopen and Van Marrewijk, 2001*; *Hoen and Oosterhaven*, 2006). Third, it should reflect net trade rather than exports only (*Leamer*, 1984; *Balance, Forstner and Murray*, 1987). Fourth, it shouldn't be focused on a single commodity (*Vollrath*, 1991).

---

[4] Key gravity variables are distance between trade partners and their GDP. For details, see *Anderson* (2011).

[5] Classical definition of comparative advantages is connected with differences in production technologies between countries (these differences are reflected in labor costs necessary to produce one unit of a good). Later, E. Heckscher and B. Olin developed the definition that is connected with countries' factor abundance.

[6] *Chor* (2010) has shown econometrically that distance, factor abundance and productivity impact trade in a similar extent. He has also built a theoretical model that simultaneously accounts for factor abundance and productivity. *Fadinger and Fleiss* (2011), *Levchenko and Zhang* (2011) developed similar models. *Shikher* (2012) has extended the Eaton-Kortum model on the case of multiple industries producing final and intermediate goods. *Finicelli, Pagano and Sbracia* (2013) developed the concept of Ricardian selection: in contrast to self-selection of exporters (only low-productivity firms exit a market, as in *Melitz* (2003) model), this type of selection can force even a high-productivity firm to exit a market (if productivity in other country is higher).

[7] By theoretically consistent measure, we mean an empirical indicator that has a counterpart in a theoretical model (that is, it should reflect one of the elements of a theoretical equation).

[8] For example, in a recent study, *Hanson, Lind and Muendler* (2014) apply the data on the new RCA measure from CEPII and show that comparative advantages for most countries are very unstable: during 1997-2007, about 33% of goods with high comparative advantages have lost this status. However, the authors mention that this result does not hold when one calculates a standard Balassa index instead of the new RCA measure (the share of goods that have lost comparative advantages in this case is much lower). But what is striking is how they treat this fact! They simply state that Balassa index is influenced by various factors (they mention geographical factors as an example) that help a country to retain specialization even if productivity falls (while the new RCA index measures pure productivity). That could be true, of course, but also could be not. In fact, the authors adopt a latent assumption that a sketchy theoretical reasoning itself can be a decisive argument in favor of one or another indicator.

The motivation for our paper is to develop an index that would be consistent with these criteria (up to date, there is no a single index that satisfies these conditions). We seek an answer for the following question: what empirical form of an index is best suited for the comparative advantage concept? This question is important, since alternative comparative advantage measures are far from being consistent with each other (*Balance, Forstner and Murray*, 1987; *Bebek*, 2011). So, the choice of a particular measure affects the results severely, so that any empirical finding about comparative advantages is doubtful.

In the paper, we demonstrate that traditional way of calculating revealed comparative advantages is not inferior to new methods. Simultaneously applying several indices directly calculated from foreign trade data, one can provide a comprehensive description of a country's comparative advantages. However, it is hard to do the same with a single index, since each of the traditional indices has its own advantages and disadvantages. The index proposed in the paper integrates positive features of different indices and lacks most of their shortcomings; moreover, it is also more informative and easier to interpret than *Leromain and Orefice* (2013) measure.

The structure of the paper is the following. In Section 2, we overview different types of comparative advantage indices calculated from foreign trade data and detect their strong and weak features. In Section 3, we propose the net comparative advantage index to cope with most of the weak features. In Section 4, we describe a measure developed by *Leromain and Orefice* (2013) and show that this measure does not possess significant advantages over the indicator proposed in our paper. In Section 5, we do an empirical exercise to demonstrate some properties of the net comparative advantage index. Section 6 concludes.

2. Revealed comparative advantage indices

Ballance, Forstner and Murray (1987) state there are at least three interpretations of a revealed comparative advantage index – as a cardinal measure ("commodity-specific degree of comparative advantage enjoyed by one country vis a vis any other country")[9], as an ordinal measure ("commodity-specific ranking of countries by degree of comparative advantage") and as a dichotomous measure ("demarcation between countries that enjoy a comparative advantage in a particular commodity and those countries that do not").

A good index should provide information about the extent of comparative advantage (that is, to be a cardinal measure). *Yeats* (1985, p. 62) explains: "A cardinal index would be far more useful for most commercial and public policy applications since it would provide a measure of

[9] This is the traditional interpretation based on *Leamer* (1974).

the magnitude of the differences in a country's comparative advantage among industries. In contrast, an ordinal index would merely rank industries in terms of comparative advantage, but would not indicate whether the differences were large or small”.

2.1. Balassa index

A country’s comparative advantage in foreign trade for a certain good is traditionally estimated with Balassa index (*Balassa*, 1965) that reflects the level of a country’s export specialization in this good compared to the world average:

$$BI_{i,c,t} = \left(X_{i,c,t} \Big/ \sum_i X_{i,c,t} \right) \Bigg/ \left(\sum_c X_{i,c,t} \Big/ \sum_i \sum_c X_{i,c,t} \right), \qquad (1)$$

where $X_{i,c,t}$ is export value for good *i*, country *c* and year *t*.[10]

Balassa index is consistent with *Kunimoto* (1977) theoretical framework in *Vollrath’s* (1991) broad interpretation – that is, it is equal to the ratio of actual-to-expected export:

$$BI_{i,c,t} = X_{i,c,t} \big/ X^{E}_{i,c,t} \, . \qquad (2)$$

Expected export of a commodity by country is proportional to this country’s share in world export (it reflects some sort of neutral comparative advantage state of international trade without geographical specialization):

$$X_{i,c,t}{}^{E} = \sum_c X_{i,c,t} \frac{\sum_i X_{i,c,t}}{\sum_i \sum_c X_{i,c,t}} \, . \qquad (3)$$

A country has a comparative advantage in foreign trade for a good if Balassa index exceeds unity. Otherwise, a country does not specialize in this good.

The index compares the share of a good in a country’s total exports with the same share in total world exports (note that Balassa index is interpreted as the extent of comparative advantage because an ability to export is found to be an important feature of high-productivity firms in the literature[11]). The key advantages of Balassa index are simplicity of construction and interpretation and low requirements to data (only export data is necessary). It is also possible to calculate a similar index for imports (it would measure import dependence).

Nevertheless, the index has various disadvantages. First, it is sensitive to the number of exported goods. Specifically, for a country with few exported products, the share of each product in total exports would be higher than for a country with a diversified set of products in its export

---

[10] In fact, the numerator in (1) is an indicator proposed even earlier by *Liesner* (1958).
[11] *Wagner* (2007) summarizes 54 econometric papers on this issue.

basket.[12] Consequently, it is incorrect to compare heterogeneous countries via Balassa index (in fact, it is appropriate only for comparison of countries with similar level of development and participation in international trade). *Yeats* (1985) shows that the highest Balassa index for a particular industry across countries is not necessary the highest for the certain country across its industries. In other words, a country may be the world leader in industry A (that is, it may have the highest share in world exports), have higher Balassa index for industry B but simultaneously not being the world leader in industry B. It follows from the fact that Balassa index distribution is largely different for various countries and industries.

Second, Balassa index values for one good depend on values for other goods, since a high share of one good in total exports at the same time means a low share of other goods in total exports. This can be regarded as exposure to structural distortions. For example, high share of oil and gas (HS codes 2709, 2710, 2711) in Russia's exports results in understated values of Balassa index for other goods: if this share equals the world average, the index for each of the other goods would be higher by 2.8 times[13]. Of course, this result is generally in line with the principle of comparative advantages. However, note that comparative advantage index is intended to mark industries with higher productivity, not with higher resource abundance.

Third, the index is asymmetric: its values range from zero to infinity (in our view, not so important issue, but rather wide-spread in the literature on Balassa index). Goods with strong comparative advantages correspond to the interval [1; ∞), while goods without comparative advantages correspond to the interval [0; 1]. As a result, Balassa index in its original form should be used in econometric studies carefully (since it can take extreme values[14], and, more important, its distribution is highly unstable). However, this problem can be solved by taking the logarithm of Balassa index[15] or by normalizing it: standard deviation for the logarithm of Balassa index is much lower than one for the original index, and standard deviation for the normalized Balassa index is equal to 1 (*Appendix 1*, *Tab. 5*).

Finally, and most important: Balassa index does not account for import trade flows. In our view, a proper comparative advantage index should reflect net trade (see Section 2.4).

---

[12] Small island countries are natural examples of economies that export few goods. For example, in 2010 the Maldive Islands had non-zero exports only for 20 goods (from 1242, according to the 4-digit *1996 Harmonized System* classification). Moreover, only three goods (frozen fish, fish fillets and smoked fish) accounted for 70% of exports of the country. The direct consequence is high Balassa index values for these goods (greater than 100, while unity is enough to detect a comparative advantage). To compare, Norway's Balassa index for these goods ranges from 10 to 20. This leads us to the conclusion that comparing Balassa indices of large and small country causes invalid results.

[13] This means that a good with Balassa index exceeding 0.35 would become a good with a comparative advantage for Russia if we eliminate the factor of mineral resources.

[14] Extremely high values are most typical for small countries (high values of Balassa index for some goods), but not necessarily. Very small values of the index should be also treated as extreme values that can distort econometric estimations.

[15] *Vollrath* (1991) proposes exactly the same.

Balassa index is one of the most common comparative advantage indices based on trade flows data, but not the only one. There are a number of different indicators that also exploit the idea of comparative advantages.

2.2. Proudman-Redding index

Proudman and Redding (2000) noticed that original Balassa index "suffers from the disadvantage that its arithmetic mean is not necessary equal to one, and may vary both across economies and over time" (p. 395, note 2). They have modified Balassa index by dividing it on the index average across goods for a certain country[16]:

$$BI_{i,c,t}^{PR} = BI_{i,c,t} \Big/ \left(\frac{1}{N} \sum_i BI_{i,c,t} \right), \text{ or} \tag{4}$$

$$BI_{i,c,t}^{PR} = \left(X_{i,c,t} \Big/ \sum_c X_{i,c,t} \right) \Big/ \left[\frac{1}{N} \sum_i \left(X_{i,c,t} \Big/ \sum_c X_{i,c,t} \right) \right], \tag{5}$$

where N is number of exported goods.

This method, as authors state, corrects distortions arising from joint analysis of large and small countries. The distortions are caused by the fact that the average Balassa index value for small countries exceeds unity more often than its value for large countries. In support, the authors provide an example in the spirit of David Ricardo. In the example, England and France sell beer and wine. We modify it (by renaming goods and restating trading quantities) to be more close to the context of Russian economy (good illustration for an economy with structural distortions). We also provide a counter-example that demonstrates uselessness of Proudman-Redding approach for such cases (*Tab. 1*).

An example with two goods shows that Proudman-Redding index sometimes narrows Balassa index range of values (for instance, Balassa index for oil and gas for Russia is 5.24, while Proudman-Redding index is 1.87). But after generalizing this example on the case of multiple goods ("other goods" aggregate disintegrates into ten equal abstract items), one can see that the range of Balassa index values expands if the average Balassa index value doesn't exceed unity (for Russia, the average Balassa index value is 0.37[17]). So, this indicator doesn't possess an advantage over Balassa index. Besides, Proudman-Redding index is highly sensitive to the number of exported goods, especially in case of low export diversification.

---

[16] Their index measures "the extent of an economy's specialization in an individual sector" (p. 377).
[17] To calculate this, we averaged the index for all goods from the 4-digit *1996 Harmonized System* classification.

Tab. 1. Proudman-Redding index: an example and a counter-example

Good	*Export volumes (USD bln.)*			*Balassa index*		*Proudman-Redding index*	
	Russia	Other countries	World	Russia	Other countries	Russia	Other countries
			Modified example with two goods				
TOTAL	*355*	*13 508*	*13 863*				
oil and gas	246	1 588	1 834	**5,24**	0,89	**1,87**	0,93
other goods	109	11 920	12 029	0,35	1,02	0,13	1,07
indices' average value:				***2,80***	*0,95*	*1,00*	*1,00*
			Counter-example with multiple goods				
TOTAL	*355*	*13 508*	*13 863*				
oil and gas	246	1 588	1 834	**5,24**	0,89	**6,57**	0,88
good 1	11	1 192	1 203	0,35	1,02	0,44	1,01
…	…	…	…	…	…	…	…
good 10	11	1 192	1 203	0,35	1,02	0,44	1,01
indices' average value:				***0,80***	*1,01*	*1,00*	*1,00*

Source: Authors' calculations based on 2010 UN COMTRADE data.

2.3. Symmetric Balassa index modifications

Some modifications of Balassa index aim to correct its asymmetry (*Tab. 2*). But why one should warn about it? *Laursen* (1998) stated that "the use of the non-adjusted RCA in regression analysis gives much more weight to values above one, when compared to observations below one," since Balassa index ranges from zero to infinity. So, they consider it necessary to calculate symmetric version of the index (see *Tab. 2* for different symmetric modifications).

Hinloopen and Van Marrewijk (2001) demonstrated that "the distribution of the Balassa index differs considerably across countries, making comparisons of the index between countries problematic" (p. 3). *Hoen and Oosterhaven* (2006) showed that the distribution of Balassa index is highly unstable over time, industries and countries. They provided some specific comments as well. First, the mean of the sectoral Balassa index is well above one (in their interpretation, this means that countries tend to have a comparative advantage in their "average sector", though an "average sector" should be neutral). Second, the median of Balassa index over countries is well below one (that is, given that the mean exceeds one, the distribution has a long right tail). Moreover, this problem is getting more important with the increase in the disaggregation level (the mean depends on the number of industries positively, and the median – negatively).

Yu et al. (2009) argued that the sum of the ideal index by countries and by goods should equal to zero, that will reflect its comparability over countries and goods, and propose such an index. However, we are skeptical about this index, since they use differences in absolute export volumes, not shares (that's why it is not surprising that the sum of deviations in export volumes equals zero, but this is not a consequence of perfect comparability). As a result, values of this index simply reflect export volumes[18].

Tab. 2. Balassa index symmetric modifications

Papers	Balassa index modification formula	Range of values	Demarcation point[19]
Laursen (1998), *Dalum et al.* (1998)	$BI_{i,c}^{L}=\frac{BI_{i,c}-1}{BI_{i,c}+1}$	[-1; 1]	0
Hoen and Oosterhaven (2006)	$BI_{i,c}^{HO}=\frac{X_{i,c}}{\sum_i X_{i,c}}-\frac{X_{i,w}}{\sum_i X_{i,w}}$	[-1; 1]	0
Yu et al. (2009)	$\exists\hat{X}_{i,c}:\left(\hat{X}_{i,c}\Big/\sum_i X_{i,c}\right)\Big/\left(X_{i,w}\Big/\sum_i X_{i,w}\right)=1\Rightarrow$ $\Rightarrow BI_{i,c}^{Y}=\frac{X_{i,c}-\hat{X}_{i,c}}{\sum_i X_{i,w}}=\left(\frac{X_{i,c}}{\sum_i X_{i,w}}-\frac{X_{i,w}\sum_i X_{i,c}}{\sum_i X_{i,w}\sum_i X_{i,w}}\right)$	[-0,25; 0,25]	0
Notes: i – good; c – country; w – world; X – export.			

Source: Authors' calculations based on 2010 UN COMTRADE data.

We consider that the only advantage of symmetric Balassa index modifications is the presence of precise upper and lower limits (that prevents extremely high values of the index). But taking logarithms or normalizing Balassa index is a powerful alternative to get exactly the same effect. So, in our view, asymmetry of the Balassa index is not the most important issue.

2.4. Indices accounting for import data

All the indices discussed above use export data to reveal comparative advantage. But a country's strong positions in exporting a certain good may reflect also its deep involvement in international value added chains (for example, it may re-export a good with minor changes), that is especially important in the modern economy.

We consider that net export reflects comparative advantages. Our argument is confirmed in *de Ferranti et al.* (2002, p. 24), who stressed (relying on *Leamer's* (1984) framework) that net export is the right indicator of comparative advantages: first, it equals the difference between domestic production and consumption; second, "an exclusive focus on exports will ignore the possibility that countries import a substantial amount of goods that they also export" (p. 22). Moreover, *Balassa and Noland* (1989, p. 9) stated: "the use of the net export index is superior to the export index of RCA on trade-theoretical grounds", because "the former indicates the effects of comparative advantages on the relationship between exports and imports rather than on

[18] Correlation coefficient between absolute values of the index and export volumes equals 0.83. For the median country, this correlation is even higher (0.93). And for countries with high export concentration, the correlation equals 0.97-0.99!
[19] The threshold value of the index indicating the presence of comparative advantages.

exports alone". Finally, it has been shown theoretically by *Deardorff* (1980, p. 949) that net export results from relative autarky prices (high-autarky-priced items are generally imported, while low-autarky-priced items tend to be exported), and that's why *Balance, Forstner and Murray* (1987, p. 161) recommended to use RCA measures based upon net exports.

In addition to (1), *Balassa* (1965) also proposed the following index:

$$BI_{i,c,t}^{XM} = \left(X_{i,c}/M_{i,c}\right) \Big/ \left(\sum_i X_{i,c} \Big/ \sum_i M_{i,c}\right), \tag{6}$$

where $M_{i,c,t}$ is import value for good *i*, country *c* and year *t*.

Later (*Balassa*, 1977) he rejected this index due to import protection bias (since the degree of import protection differs from country to country). However, in our view, in a modern world of high intermediate imports, ignoring import flows may cause an even higher error than accepting equal degree of import protection over countries.[20]

A common index for the purpose of accounting for import trade flows is the relative net export index (*UNIDO*, 1982):

$$RNX_{i,c,t} = \frac{X_{i,c,t} - M_{i,c,t}}{X_{i,c,t} + M_{i,c,t}}. \tag{7}$$

The index takes values from -1 (non-zero imports, zero exports) to +1 (non-zero exports, zero imports). The advantage of this index is the absence of structural distortions: the calculation is done separately for every good. The number of exported and imported goods does not affect the results. Also, it's important that the index can be easily calculated not only for certain goods, but also for the whole economy.

The disadvantage of the relative net export index is its inability to identify importance of export and import flows for the economy (in terms of volumes). For example, the index may take the value of 1 (very high specialization) even if export is tiny but import is absent. Balassa index calculated from export data would correctly show very low specialization in this case.

Another index that accounts for import trade flows was proposed by *Donges and Riedel* (1977):[21]

---

[20] Moreover, export trade flows are also distorted by resource abundance problem, and this distortion is much higher than import-protection bias (this is reflected in the fact that export structure is much more concentrated than import structure).

[21] They used it as an instrument for observing changes in sectoral export structure.

$$DRI_{i,c,t} = \left(\frac{X_{i,c,t} - M_{i,c,t}}{X_{i,c,t} + M_{i,c,t}} / \frac{\sum_i X_{i,c,t} - \sum_i M_{i,c,t}}{\sum_i X_{i,c,t} + \sum_i M_{i,c,t}} - 1 \right) \cdot sign\left(\sum_i X_{i,c,t} - \sum_i M_{i,c,t} \right), \tag{8}$$

where $sign\left(\sum_i X_{i,c,t} - \sum_i M_{i,c,t} \right)$ equals 1 if $\sum_i X_{i,c,t} > \sum_i M_{i,c,t}$, -1 if $\sum_i X_{i,c,t} < \sum_i M_{i,c,t}$.

The values of this index range from -∞ to +∞ (depending on the net exports index for the country's total trade). If a country is a strong net exporter or importer (say, its net exports index is above 0,5 or below -0,5), than its extreme values are strongly limited (in this example, from -2 to 2). But if a country's exports and imports are close, the index can potentially take very large values, and this is a drawback of Donges-Riedel index.

Lafay (1992) developed the following index:

$$LFI_{i,c,t} = 100 \left(\frac{X_{i,c,t} - M_{i,c,t}}{X_{i,c,t} + M_{i,c,t}} - \frac{\sum_i (X_{i,c,t} - M_{i,c,t})}{\sum_i (X_{i,c,t} + M_{i,c,t})} \right) \frac{X_{i,c,t} + M_{i,c,t}}{\sum_i (X_{i,c,t} + M_{i,c,t})}. \tag{9}$$

Its values range from -50 to +50. Its positive feature is accounting for the importance of export and import trade flows for the country. However, the accounting method implies that index values depend mostly on commodities' weights in a country's trade turnover (that is, scale effect impacts its values dramatically[22]). This is a serious problem, because goods are not equal in trade volumes due to classification issues.[23]

Finally, in order to embody the demand dimension, *Vollrath* (1991) proposed two ways to account for export and import data jointly:

$$RCA^{V1}_{i,c,t} = RXA_{i,c,t} - RMA_{i,c,t}, \tag{10}$$

$$RCA^{V2}_{i,c,t} = \ln(RXA_{i,c,t}) - \ln(RMA_{i,c,t}), \tag{11}$$

where $RXA_{i,c,t}$ and $RMA_{i,c,t}$ are Balassa-type trade intensity measures transformed in order to avoid double counting:

$$RXA_{i,c,t} = \frac{X_{i,c,t}}{\sum_i X_{i,c,t} - X_{i,c,t}} \Bigg/ \frac{\sum_c X_{i,c,t} - X_{i,c,t}}{\sum_i \sum_c X_{i,c,t} - \sum_i X_{i,c,t} - \sum_c X_{i,c,t}}, \tag{12}$$

[22] Correlation coefficient between its values and the values of the net trade is above 0.9 for more than 100 countries and above 0.99 for more than 50 countries.
[23] For example, there are around 150 textile goods in the 4-digit *1996 Harmonized System* classification (all accounting for only 600 USD bln. of world exports), but barely 16 mineral products (accounting for 2000 USD bln).

$$RMA_{i,c,t} = \frac{M_{i,c,t}}{\sum_i M_{i,c,t} - M_{i,c,t}} \Bigg/ \frac{\sum_c M_{i,c,t} - M_{i,c,t}}{\sum_i \sum_c M_{i,c,t} - \sum_i M_{i,c,t} - \sum_c M_{i,c,t}}. \tag{13}$$

Vollrath believes that relative net export and Donges-Riedel indices do not measure comparative advantages, as “both focus on a single commodity and, therefore, do not fulfill the contrasting dimensions inherent in the principle of comparative advantage” (p. 272). Rather, he states, it measures intra-industry trade.[24] That’s why he proposes new export-to-import indices.

We also consider that relative net export index is per se a poor indicator of comparative advantages, since it does not account for world trade volumes and structure. However, in our view, the sign and the extent of the trade balance should be taken into account in comparative advantage estimations. In Section 3, we show how to do this.

Vollrath (1991) indicates that RCA^{V1} and RCA^{V2} have an important shortcoming: they are exposed to policy-induced distortions that arise from import protection (and, that is, in our view, even more important, to the difference between export and import structure[25]). Therefore, he admits that in some cases a simple modification of Balassa index would be preferable:

$$RCA^{V3}_{i,c,t} = \ln\left(RXA_{i,c,t}\right). \tag{14}$$

However, when abstracting from distortionary influences, the first two indices, Vollrath states, adhere more closely to actual comparative advantages.[26] Yet, we consider that they also reflect differences in export and import structure, so that their sign does not necessarily coincide with the sign of a trade balance. Moreover, RCA^{V2} can generate undefined results if a country does not export *or* import a commodity.[27]

So, there is no such a single indicator that is considered to be superior to other indices.

The common problem of indices that analyze exports and imports jointly is data quality. There are two important aspects. First, export and import flows data is presented in different prices (CIF for imports and FOB for exports[28]). On average, this results in the excess of imports over exports (i.e., the actual demarcation point is a bit lower than zero). Second, if one tries to

---

[24] We disagree with this opinion. To be a measure of intra-industry trade, UNIDO index should apply the absolute value of net exports in the numerator, as *Balassa* (1966) did.
[25] For instance, if oil accounts for 50% of a country’s exports, but is not imported by country at all, then export shares for all goods, in fact, would be halved. In this situation, a slight excess of exports over imports would result in negative values of these indices, while trade balance would be positive.
[26] Vollrath considers that the first index is preferable for highly disaggregated data, since it is more likely that a logarithm would be undefined at such level of disaggregation. The second index is preferable for aggregated data.
[27] Net comparative advantage index proposed in Section 3 of the paper is quite close to Vollrath’s RCA^{V1} and RCA^{V2}, but much more resistant to all these problems.
[28] CIF prices include transport, freight and insurance costs, FOB prices do not.

overcome this problem, another issue appears: mirror statistics that should be collected for this purpose differs from direct statistics in country coverage.[29] So, joint export-import estimations always contain a small error: the average CIF/FOB ratio for the world trade is around 1.06,[30] and the mean CIF/FOB ratio estimated econometrically is around 1.03 (*Gaulier and Zignano*, 2010, p. 15). However, we consider that the error would be much larger if one tries to explore comparative advantages without import data. Moreover, one could cope with these problems by making a deeper examination of the differences in the unit values and the quality of reporting quantities, like *Gaulier and Zignano* (2010), or by using data on a country's trade with the whole world instead of data on bilateral trade flows (that helps to diminish this error).

2.5. Indices accounting for openness of the economy

All indices discussed above have one common disadvantage – they do not account for openness of the economy. This is an important factor: for example, if a country's trade turnover is insignificant relative to its GDP (it is typical for the largest countries, and the countries that are not highly integrated in the world trade, such as Nepal, Gambia, Cyprus, Greece, etc.), then Balassa index would poorly reflect the importance of trade flows for the economy.

Bowen (1983) proposed the index that accounts for this factor, although he didn't address the problem of economic openness in his original paper. His concern was rather to find a proper scale variable for net trade, and he used consumption as a theoretically grounded variable (this logics is clearly stated in his further comments – see *Bowen* (1986)). *Bowen* (1983) started with the following claim: "theoretically, a fundamental difficulty with the preceding indices is that they treat exports and imports separately when comparative advantage is properly a net trade concept" (p. 468). Bowen index accounts simultaneously for net trade and economic openness, since it can be decomposed into two multipliers:

$$BowI = \frac{X_{i,c,t} - M_{i,c,t}}{C_{i,c,t}} = \frac{X_{i,c,t} - M_{i,c,t}}{X_{i,c,t} + M_{i,c,t}} \cdot \frac{X_{i,c,t} + M_{i,c,t}}{C_{i,c,t}}, \qquad (15)$$

where $C_{i,c,t}$ is the estimated level of consumption of good *i* by country *c*.

[29] Sometimes these differences boil down to the quality of custom service, sometimes – to methodological issues. Specifically, imports is usually recorded by the country of origin, while exports – by the country of last known destination (*UN*, 2013, Chapter XVI). For example, country A exports a good to country B and records it as the country of destination, but country C that imports the same good from country B records country A as the country of origin. In real world, such countries as Netherlands and Hong Kong (i.e., counties with large ports) usually play the role of country B. *Martin* (2013) provides an interesting exploration of the differences in trade data for the case of US-China bilateral trade.

[30] This figure is obtained simply by dividing the value of total imported goods on the value of total exported goods. Interestingly, *Carrere and Grigoriou* (2015) calculate a median CIF-FOB ratio for country pairs that report both export and import data and get a close figure of 1.057.

Bowen takes an assumption that countries have identical homothetic preferences. As a result, each country's consumption of a good should be proportional to the world's consumption (production) of this good:

$$C_{i,c,t} = \frac{GDP_{c,t}}{GDP_{W,t}} Q_{i,W,t}, \tag{16}$$

where $Q_{i,W,t}$ is the world production of good *i*, $GDP_{c,t}$ is GDP of country *c*, $GDP_{W,t}$ is the world GDP.

He also reformulates the net trade:

$$C_{i,c,t} = Q_{i,c,t} + M_{i,c,t} - X_{i,c,t} \Rightarrow X_{i,c,t} - M_{i,c,t} = Q_{i,c,t} - C_{i,c,t}, \tag{17}$$

where $Q_{i,c,t}$ is production of good *i* by country *c*.

Given (15), (16) and (17), he obtains:

$$BowI = \frac{Q_{i,c,t} - C_{i,c,t}}{C_{i,c,t}} = \frac{Q_{i,c,t} - \left(GDP_{c,t} / GDP_{W,t}\right) \cdot Q_{i,W,t}}{\left(GDP_{c,t} / GDP_{W,t}\right) \cdot Q_{i,W,t}}. \tag{18}$$

This index has been criticized by *Balance, Forstner and Murray* (1985, 1986) for the assumption that countries have identical homothetic preferences. *Bowen* (1986) responded that the only reason to make this assumption was simply minimizing data requirements (in order to be able to estimate consumption with production data).

In general, we appreciate Bowen's idea to use production and GDP as scale variables, because this is a simple way to account for economic openness. However, two problems appear. First, it is impossible to calculate the index for detailed commodity groups, since the data on production is rather aggregated. Second, Bowen's index, like relative net export index, does not take the world trade structure into consideration (in Vollrath's words, they both "focus on a single commodity"). Onwards, we present the index that does not totally focus on a single commodity, but uses GDP as a scale variable.

3. Net comparative advantage index

We propose an index that simultaneously accounts for export and import data and can be expressed within the *Kunimoto* (1977) theoretical framework. It is calculated as the net trade normalized by expected trade turnover, or relative net export index multiplied by the ratio of actual and expected trade turnover (that is, trade turnover expected in a hypothetical neutral comparative advantage world with no geographical specialization). We call this indicator *net comparative advantage index*:

$$NCA_{i,c,t} = \frac{X_{i,c,t} - M_{i,c,t}}{\left(X_{i,c,t} + M_{i,c,t}\right)^E} = RNX_{i,c,t} \frac{X_{i,c,t} + M_{i,c,t}}{\left(X_{i,c,t} + M_{i,c,t}\right)^E}, \tag{19}$$

where RNX is relative net export index, $(X_{i,c,t} + M_{i,c,t})^E$ is the expected trade turnover for country c and good i that is calculated as follows:

$$\left(X_{i,c,t} + M_{i,c,t}\right)^E = \left(\sum_c X_{i,c,t} + \sum_c M_{i,c,t}\right) \frac{GDP_{c,t}}{\sum_c GDP_{c,t}}. \tag{20}$$

An important feature of this index is its consistency with *Kunimoto* (1977) theoretical framework, where actual trade is compared to expected trade, with minor changes. In the traditional *Kunimoto* (1977) framework, world export of a commodity is distributed among countries in proportion of their share of world exports. In our version, world trade turnover of a commodity is distributed among countries in proportion of their share of world GDP.

Combining (19) and (20), one can obtain:

$$NCA_{i,c,t} = \underbrace{\frac{X_{i,c,t} - M_{i,c,t}}{X_{i,c,t} + M_{i,c,t}}}_{RNX_{i,c,t}} \times \underbrace{\frac{X_{i,c,t} + M_{i,c,t}}{GDP_{c,t}} \Bigg/ \frac{\sum_c X_{i,c,t} + \sum_c M_{i,c,t}}{\sum_c GDP_{c,t}}}_{RTO_{i,c,t}}, \tag{21}$$

where $RTO_{i,c,t}$ is relative trade openness of country c by good i in year t. This indicator allows us to simultaneously account for economic openness and importance of a trade flow of a certain good for the economy, since it can be rewritten as:

$$RTO_{i,c,t} = \underbrace{\frac{\left(\dfrac{X_{i,c,t} + M_{i,c,t}}{\sum_i X_{i,c,t} + \sum_i M_{i,c,t}}\right)}{\left(\dfrac{\sum_c X_{i,c,t} + \sum_c M_{i,c,t}}{\sum_i \sum_c X_{i,c,t} + \sum_i \sum_c M_{i,c,t}}\right)}}_{RT_{i,c,t}} \times \underbrace{\frac{\left(\dfrac{\sum_i X_{i,c,t} + \sum_i M_{i,c,t}}{GDP_{c,t}}\right)}{\left(\dfrac{\sum_i \sum_c X_{i,c,t} + \sum_i \sum_c M_{i,c,t}}{\sum_c GDP_{c,t}}\right)}}_{RO_{c,t}}, \tag{22}$$

where $RT_{i,c,t}$ is relative trade intensity of good i, and $RO_{c,t}$ is relative openness of the economy of country c.

Thus, we use net trade data, but at the same time overcome the critique of *Vollrath* (1991) that net trade indices focus on a single commodity.

Relative trade intensity of a good is the most volatile element of this index: its average value and the whole distribution are not stable due to the presence of extreme values (however, it is more stable than Balassa index – see footnote 32). Therefore, we also propose a symmetric version of the index based on normalized trade intensity (normalized trade intensity ranges from zero to unity):

$$SNCA_{i,c,t} = RNX_{i,c,t} \times RO_{i,c,t} \times SRT_{i,c,t}, \quad (23)$$

where $SRT_{i,c,t} = \frac{RT_{i,c,t}}{RT_{i,c,t} + 1}$. (24)

Note that the proposed procedure impacts extreme values primarily: for moderate values of relative trade intensity, there is a near-linear relationship between relative trade intensity and its symmetric modification (*Fig. 1*).

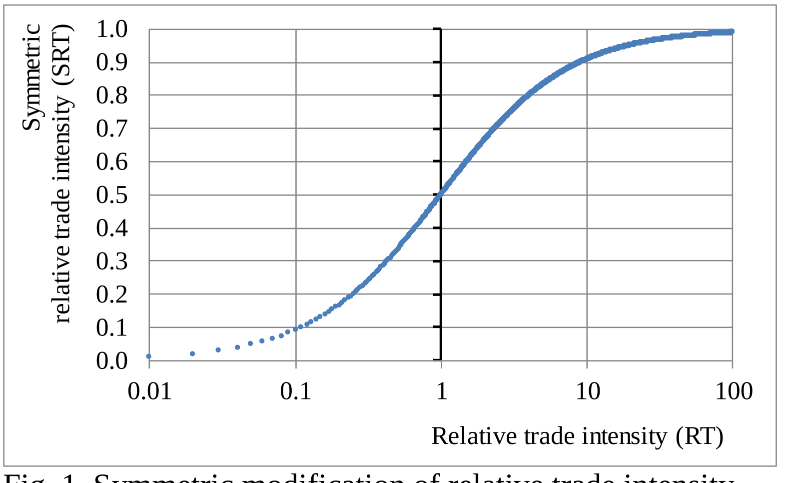


Fig. 1. Symmetric modification of relative trade intensity
Source: Authors' calculations.

Net comparative advantage index has a number of advantages:

- the sign of this index shows the state of the trade balance, exactly as the sign of original relative net export index; at the same time, the index takes the volumes of trade flows and GDP into account[31];
- the impact of structural distortions on the indicator's value is strongly smoothed (as both export and import data appear in calculations[32]);

[31] *RT* and *RO* components in (22).
[32] Recall the discussion of Balassa index disadvantages. In Section 2.1, we stated that Balassa index is exposed to structural distortions, in terms of distortionary influence of resource abundance on its values for *all industries* of a country. In the proposed net comparative advantage index, we use trade turnover instead of exports to calculate trade intensity. The implied economic

- the index values range from -∞ до +∞, with rare extreme values (in symmetric version, there are no extreme values);
- the index value for one commodity is almost independent from values for other commodities[33].

For every country, it is possible to approximate this index by two indicators – export- and import-driven Balassa indices. The former is the standard Balassa index calculated as in (1), and the latter is Balassa index calculated with import data. A simple difference between the two (Vollrath index RCA^{V1}) is enough for rough approximation, while the two-factor regression with the same indicators as variables provides a very good approximation (*Fig. 2*).

Note that regression results also contain information about the contribution of the two indicators. This contribution is different from country to country, showing the extent of structural differences between exports and imports. Let's consider the case of Russian economy. If import-driven Balassa index equals unity, then export-driven Balassa index value of 0.61 (not unity!) would reflect zero trade balance.[34] That's why *Vollrath* (1991) index is a biased measure, while the proposed index is fully consistent with the sign of the trade balance.

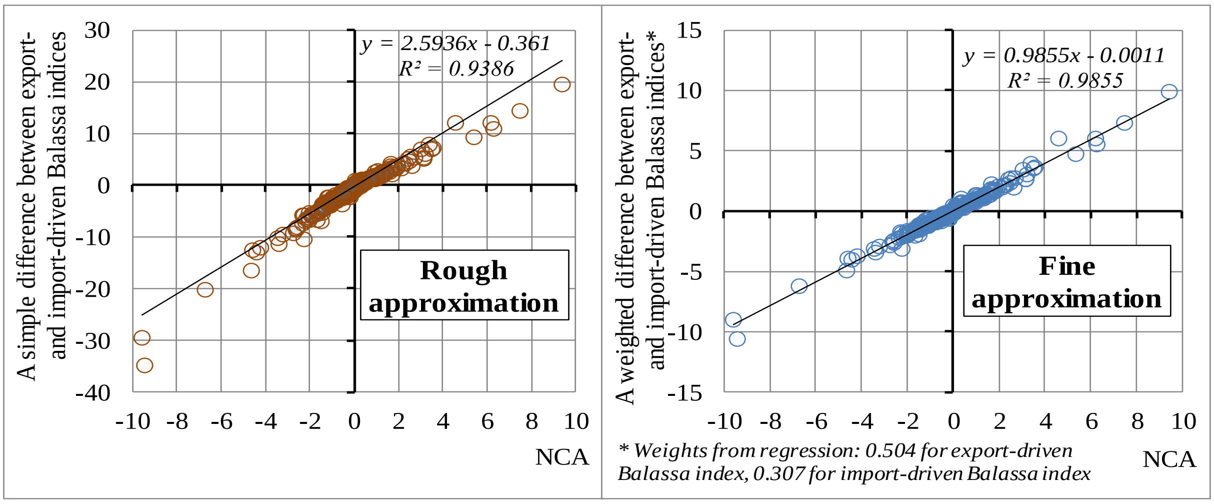


Fig. 2. Net comparative advantage index: approximations for Russia
Source: Authors' calculations based on 2010 UN COMTRADE data.

4. Theoretically grounded index

Economists from CEPII have recently proposed a new approach to detecting comparative advantages, based on econometric methods (*Leromain and Orefice*, 2013). Relying on *Costinot, Donaldson and Komunjer* (2012) model, they develop a method to separate the contribution of

---

interpretation is the following. Even if export structure is heavily distorted by resource abundance problem, import structure is finely diversified (moreover, in countries with high resource abundance import is more concentrated in processed commodities, and that weakens the distortions significantly).

[33] There is only a slight indirect influence through GDP – see (21).

[34] The deviation of this figure from unity reflects the power of structural distortions.

three groups of factors: characteristics of importers' domestic markets ("good-importer" fixed effects), patterns of bilateral trade ("exporter-importer" effects) and productivity of exporters ("good-exporter" effects). In fact, they treat the link between bilateral trade flows variation and exporters' characteristics on every market as the country's comparative advantage level.

They start from the theoretical model of *Costinot, Donaldson and Komunjer* (2012):

$$\ln(x_{i,c1,c2}) = \delta_{c1,c2}\delta_{i,c2}\theta\ln(z_{i,c1}) + \varepsilon_{i,c1,c2}, \quad (25)$$

where $x_{i,c1,c2}$ are bilateral trade flows (c_1 stands for exporter, and c_2 stands for importer), $\delta_{c1,c2}$ and $\delta_{i,c2}$ are country-pair fixed effects and importer-industry fixed effects, respectively, $z_{i,c1}$ is an exporter's fundamental productivity level in industry *i* (i.e., technological coefficient), and θ is a measure of productivity dispersion which is country invariant (θ = 6.53, according to the Costinot-Donaldson-Komunjer estimation).

Then they estimate $z_{i,c1}$ as a proxy for Ricardian comparative advantage (that is exporter-industry specific). Their first step is estimating $\delta_{i,c1}$ from the following empirical equation:

$$\ln(x_{i,c1,c2}) = \delta_{c1,c2}\delta_{i,c2}\delta_{i,c1} + \varepsilon_{i,c1,c2}, \quad (26)$$

where $\delta_{i,c1}$ are exporter-industry fixed effects.

Combining (25) and (26), they obtain:

$$z_{i,c1} = e^{\delta_{i,c1}/\theta}. \quad (27)$$

Their final comparative advantage index (we further call it theoretically grounded index, or econometric index) is the following:

$$RCA^{LO}_{i,c1} = \frac{z_{i,c1}z_{..}}{z_{i.}z_{.c1}}, \quad (28)$$

where $z_{..}$ is the average of $z_{i,c1}$ across industries and countries, $z_{i.}$ is the average of $z_{i,c1}$ for the country *i* across sectors, and $z_{.k}$ is the average of $z_{i,c1}$ for the sector *k* across exporters.

Leromain and Orefice (2013) have created a database on this indicator[35] and compared it with Balassa index. They have concluded that econometric indicator outperforms Balassa index: it is more stable over time, has a lower variation and is more close to normal distribution.

However, there are also counterarguments showing that an indicator proposed in our paper is preferable to econometric index due to the following facts:

[35] Their database is available at: http://www.cepii.fr/cepii/en/bdd_modele/presentation.asp?id=26.

- calculation of the econometric index is rather complicated (it requires introducing additional models and using data on bilateral trade flows that contain much more improper observations than unilateral trade flows);
- due to calculation difficulties, database country coverage is small (20 countries);
- comparison of econometric and Balassa indices done by *Leromaine and Orefice* (2013) is hardly correct[36];
- econometric index does not account for imports, in contrast to the proposed index;
- though econometric index is close to normal distribution graphically (*Appendix 3, Fig. 8*), statistical tests reject the normal distribution hypothesis[37].

The latter conclusion is important, since Leromaine and Orefice unreasonably emphasize normality of their theoretically grounded index while comparing it with other indices. It is true that their index has a better bell-shaped form than other indices (see *Appendix 3, Fig. 8*), but it doesn't mean that distribution of this index is normal.On the empirical side, the indicator proposed in our paper is better suited for estimating comparative advantages than Leromaine and Orefice index (*Tab. 3*). For example, econometric index indicates that Indonesia does not possess a comparative advantage in trade for nickel, while Balassa index and the proposed indicator show the opposite. A natural experiment conducted by the Indonesian government in early 2014 (it has banned export of unprocessed nickel ores) has demonstrated that this country impacts world nickel prices heavily (see, for instance, *Hume and Rice*, 2014). That looks like the evidence of high comparative advantage of Indonesian economy in trade for nickel.

The bulk of the difference between the two indices results from the fact that the proposed index accounts for import flows, while econometric index doesn't (pay attention to the values of import-driven Balassa index in *Tab. 3*). However, this is not a common rule: for some goods, econometric and Balassa indices (that are both calculated from export data) differ in their

[36] Three considerations should be mentioned specially. First, in econometric equations used to construct the index developed by *Leromaine and Orefice* (2013), trade flows are taken *in logarithms*. But then Leromaine and Orefice compare empirical characteristics of this index with those of Balassa index, not its logarithm! That's why their conlusion that the econometric index has lower variation and is more close to normal distribution is doubtful: characteristics of Balassa index logarithm, as shown in *Tab. 5* of the *Appendix 1*, are much better than characteristics of Balassa index. Second, it is necessary to understand that variation is not a good criterion for an indicator's relevance, since distribution of every index can be transformed to a standard normal distribution by subtracting its mean value and dividing by the standard deviation (see *Tab. 5*, right side). Finally, *Leromain and Orefice* (2013) calculate the percent change of mean values for Balassa index and the econometric index and show that it is much lower for the econometric index. Nevertheless, note that the percent change of the mean value severely depends on the mean value itself: if it is close to zero, the percent change would be much larger. Moreover, one can easily show that instability of Balassa index mean value is basically generated by the extreme values: excluding them makes the mean values for every index fairly stable. Balassa index instability over time is also a consequence of structural distortions – so, it can be made more stable by calculating it excluding mineral resources (one should also use the NCA index instead of Balassa index to cope with this problem).

[37] We performed a Skewness/Kurtosis test for normality in STATA. The normality hypothesis was rejected at 1% significance level for all indices examined in the paper, including *Leromaine and Orefice* (2013) index.

conclusion about the presence of comparative advantages (Canadian ferrous metals, Russian fur, Mexican plastics and some others). Interestingly, these goods are typically characterized by low share in world exports, that is consistent with negative correlation of a country's Balassa index rank and its share in world exports of the good (*Fig. 3*).

Tab. 3. The proposed and the econometric index values by selected goods

Country	Good (2-digit HS codes)	Leromaine and Orefice econometric index (RCA)		Net comparative advantage index (NCA)		Module of rank diffe-rences	Symmetric net comparative advantage index (SNCA)		Export-driven Balassa index		Import-driven Balassa index	Country's share in a good's export
		Value	Rank	Value	Rank		Value	Rank	Value	Rank		
Excessively optimistic values of the econometric indicator												
Turkey	manmade fibres	1.37	2	-1.74	20	*18*	-0.31	18	3.31	3	5.46	3%
India	precious stones	1.12	2	-1.68	20	*18*	-0.23	20	5.02	2	7.55	8%
Turkey	cotton	1.50	2	-1.62	20	*18*	-0.29	17	3.25	2	5.95	3%
Russia	railway vehicles	1.31	1	-1.16	19	*18*	-0.36	18	0.61	10	5.77	2%
Canada	articles of iron/steel	1.15	2	-0.52	20	*18*	-0.26	18	0.55	15	1.50	1%
India	silk	1.40	1	-0.25	19	*18*	-0.03	7	6.61	1	6.50	10%
Canada	printed books	1.13	3	-1.03	20	*17*	-0.36	20	0.88	8	2.98	2%
Mexico	plastics	1.10	3	-0.75	20	*17*	-0.35	20	0.59	14	1.76	1%
Canada	toys, games	1.20	3	-0.69	20	*17*	-0.32	20	0.55	9	1.55	1%
Canada	headgear	1.16	3	-0.61	20	*17*	-0.31	20	0.44	10	1.49	1%
Turkey	manmade filaments	1.33	3	-0.52	20	*17*	-0.10	14	4.12	1	4.15	3%
France	cork	1.05	3	-1.80	19	*16*	-0.46	19	1.12	2	4.36	4%
Indonesia	cotton	1.46	3	-1.28	19	*16*	-0.34	18	1.20	8	5.06	1%
Netherlands	tin	1.17	3	-0.93	19	*16*	-0.40	19	0.90	2	1.73	3%
Russia	furskins and fur	1.30	3	-0.71	19	*16*	-0.32	19	0.26	11	3.27	1%
Canada	musical instruments	1.26	3	-0.50	19	*16*	-0.25	17	0.55	10	1.38	1%
Australia	pharmaceuticals	1.22	1	-0.31	17	*16*	-0.15	14	0.59	9	1.50	1%
Excessively pessimistic values of the econometric indicator												
Indonesia	copper	0.69	20	0.63	1	*19*	0.26	2	1.93	2	0.80	2%
Argentina	precious stones	0.81	19	0.46	3	*16*	0.27	3	1.13	7	0.07	1%
Indonesia	nickel	0.58	19	2.21	3	*16*	0.63	3	4.71	3	0.11	5%
Brazil	clocks and watches	0.68	20	-0.08	4	*16*	-0.07	4	0.01	19	0.44	0%

Source: Authors' calculations based on 2010 UN COMTRADE and RCA CEPII data.

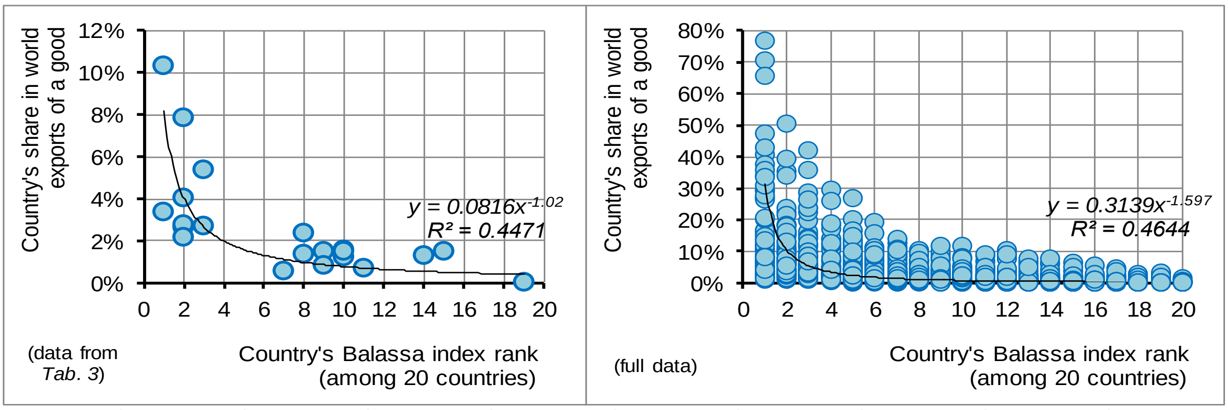


Fig. 3. Countries' Balassa index rank and their shares in world exports of a good
Source: Authors' calculations based on 2010 UN COMTRADE data.

So, the econometric index may signal about a country's strong comparative advantages even if the country has low volumes of trade, while the proposed indicator takes trade intensity of a good into account (*RT* component in (22)). This means that the proposed indicator not only eliminates the shortcomings of comparative advantage indices but also retains such an important feature of Balassa index (and similar indices) as accounting for trade flow volumes.

5. Empirical demonstration

NCA index allows one to reduce the number of dimensions, accounting simultaneously for several factors. Implying a set of indices is sometimes also a competitive strategy. However, it is much less convenient than using a single NCA index.

Consider an example of Russian frozen herrings (*Tab. 4*). This commodity is intensively exported and imported by Russia. Vollrath's index shows that there is a disadvantage in trade for frozen herrings. However, net export is positive: relative net export index provides one with the information that exports slightly exceeds imports. Nevertheless, the values of this index are far from top 10% threshold – this indicates that there is a weak comparative advantage in trade for frozen herrings. NCA index, on the contrary, shows that Russia has a strong comparative advantage in trade for frozen herrings, since low values of relative net export index are offset by really high values of trade intensity. None of the other indices can account for such different factors simultaneously. Another example with maize (see also *Tab. 4*) demonstrates the opposite situation: though relative net export index is extremely high, NCA index shows that the proper degree of comparative advantage is substantially lower for maize than for frozen herrings, since trade intensity for the former is several times lower than for the latter.

Tab. 4. Index values for selected commodities (Russia, 2012)

Commodity name	HS 2007 code	Balassa index (export)	Balassa index (import)	Vollrath's RCA^{V1}	Relative net export index	NCA index
Herrings, frozen	030351	6.0	8.0	-2.0	0.1	0.8
Maize (corn), other than seed	100590	0.6	0.0	0.5	1.0	0.3
Top 10% threshold		*0.6*	*0.0*	*0.1*	*0.7*	*0.1*
Lower 10% threshold		*0.0*	*3.1*	*-2.9*	*-1.0*	*-1.0*

Source: Authors' calculations based on 2012 UN COMTRADE data.

Empirically, one can easily see that both export-driven and import-driven Balassa indices influence the net comparative advantage index. In a general case, the net comparative advantage index is very close to Vollrath indices (that is, difference between export-driven and import-driven Balassa indices or their logarithms). This also holds true for OECD countries (see *Fig. 4*).

However, even for OECD countries, there are several cases in which these two indices disagree about the strength of comparative advantage. For example, while the NCA index shows that OECD countries are most competitive in "Ships, boats and floating structures" (green point at *Fig. 4*) and are not competitive in "Organic chemicals" (orange point at *Fig. 4*), the Vollrath index considers these two commodity groups as being nearly equally competitive. The reason for this difference is straightforward: the Vollrath index ignores the fact that export of ships for

OECD countries exceeds their import substantially (relative net export index equals 0.45), while trade balance for organic chemicals is negative (relative net export index equals -0.02).

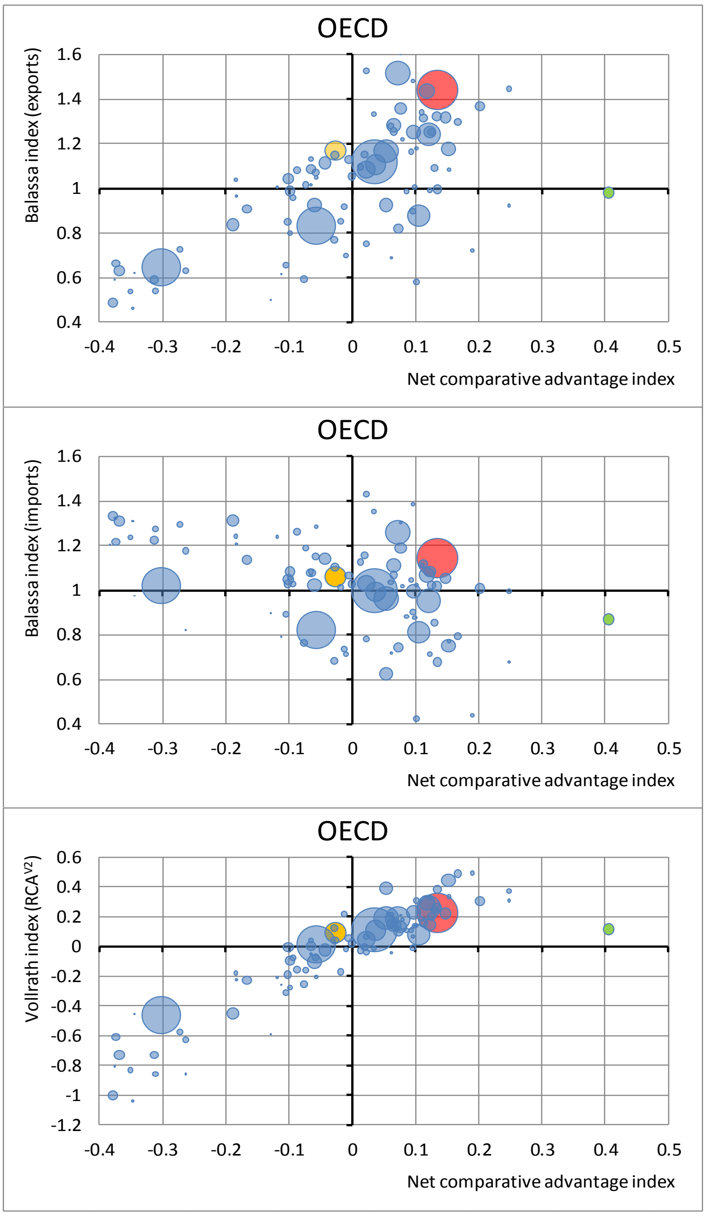


Fig. 4. NCA vs. Balassa and Vollrath indices for OECD countries
Source: Authors' calculations based on 2012 UN COMTRADE data.

For BRICS countries, there is a great variation in the lower tail (see *Fig. 5*): for instance, "Pharmaceutical products" (yellow point at *Fig. 5*) and "Ores, slag and ash" (purple point at *Fig. 5*) are close in the values of Vollrath index but differ drastically in the values of the NCA index. In this case, relative net export indices are relatively close (-0.35 for pharmaceutical products and -0.46 for ores), but the RTO component differs much (0.30 for pharmaceutical products and 2.01 for ores), reflecting substantial differences in trade intensity for these products.

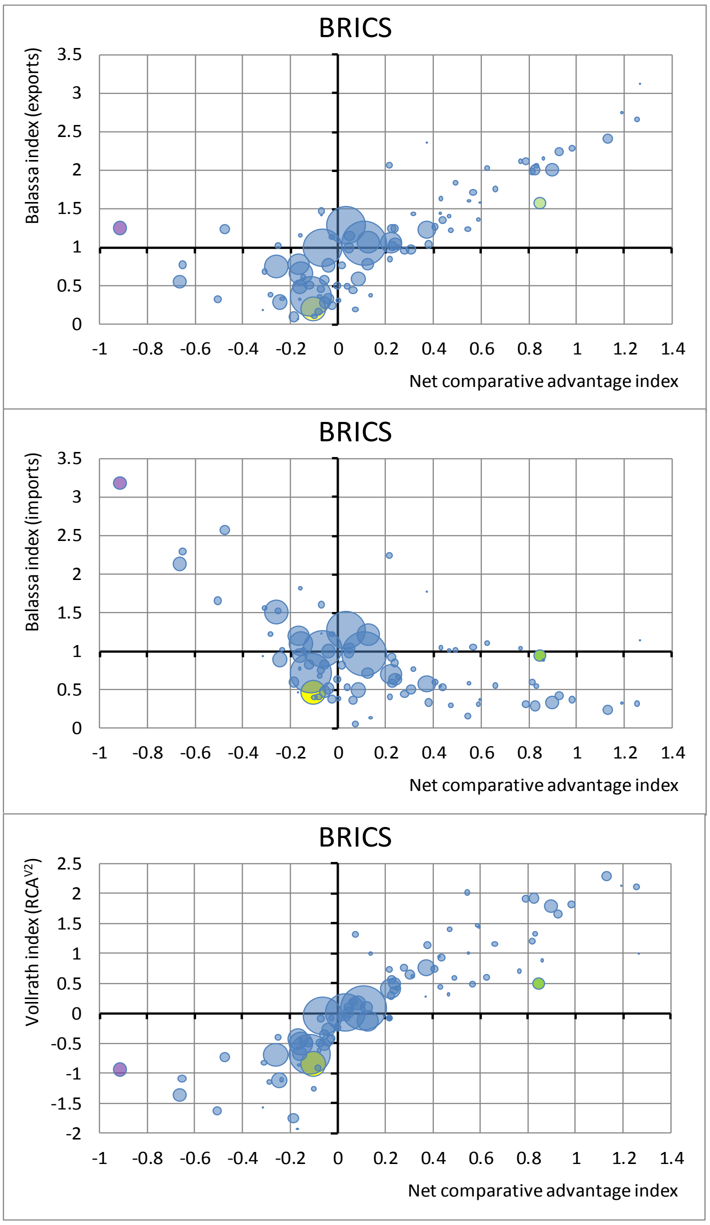


Fig. 5. NCA vs. Balassa and Vollrath indices for BRICS countries
Source: Authors' calculations based on 2012 UN COMTRADE data.

Looking at the first two graphs at *Fig. 4* and *Fig. 5* is also a clear empirical strategy to show that indices constructed only from export or import data are not proper instruments to speculate about comparative advantages.

Consider "Ships, boats and floating structures" (green point at *Fig. 4*) as the first example. Export-oriented Balassa index value (0.98) states that OECD countries do not have a comparative advantage in this commodity group. However, import-oriented Balassa index is even smaller (0.87). The NCA index captures this nicely, showing that OECD countries have a net comparative advantage in "Ships, boats and floating structures". The origin of distortions is high export concentration of this commodity group. In 2012, 67% of world exports (98.8 USD bln) were shipped by 3 countries: China, South Korea and Germany, with nearly equal export values (38.8, 37.8 and 22.2 USD bln, respectively). Import of this commodity, on the contrary, has a quite smoothed distribution across countries.

Now consider another example. OECD countries import “Vehicles other than railway or tramway rolling-stock, and parts and accessories thereof” (red point at *Fig. 4*) 1.15 times more actively than the world average, but this doesn’t imply that OECD countries are not competitive at the domestic market, since they export goods of this commodity group 1.44 times more actively than the world average. So, by looking only at the export- or import-driven Balassa index, one overlooks both the role of import inputs in exporting (that is extremely important for such countries as Hungary[38]) and the existence of intra-industry trade (or trade in differentiated products, that is important for automobile industry in Sweden, UK, France and some other countries[39]). The NCA index accounts for these aspects. A country that is able to produce automobile parts and a wide set of brands with high quality and unique consumer characteristics (for example, Germany that has only 34% import content in automobile exports and a Grubel-Lloyd index value of 0.43) is regarded to be more competitive than countries that are not able to do it (UK, Hungary, and so on).

Note that the NCA index is only a rough analogue of estimations based on value added: it complements exports in an industry with imports in this very industry, while it is possible to use goods from every other industry as intermediates (and some of them may embody a really high share of foreign value added). However, a good piece of news for the index is the fact that the greatest element in input-output tables is usually an intermediate consumption in the same industry that produces a good. So, the NCA index does not account for all intermediate inputs, but accounts for some part of it, in contrast to Balassa index. Moreover, it would demonstrate a better performance in the situation of potential switch to value added statistics.

To make this speculation empirically grounded, we provide calculations of Balassa index for trade flows in value added, relying on OECD-WTO *Trade in Value Added Database*[40] (the level of aggregation, however, is rather poor). We show that Balassa index for export value added does not deviate strongly from Balassa index for export volumes (*Appendix 2*, *Tab. 7*). The main reason is the fact that the share of value added in gross exports for certain industries

---

[38] According to *OECD* (2011), import contents of Hungarian exports equaled 56% in 2005. OECD STAN database records a figure of 63% for manufacturing industries (74% for “Motor vehicles, trailers and semi-trailers”, 79% for “Office, accounting and computing machinery”, and 82% for “Radio, television and communication equipment”).

[39] OECD STAN database reports that import contents of automobile exports in these countries is less than 45%, while Grubel-Lloyd index (the fraction of intra-industry trade in trade turnover by commodity) is over 0.80. This implies that trade in motor vehicles in these countries is driven by factors different from input-output relations (most likely, commodity differentiation in quality, brands, characteristics).

[40] The database is available here: http://stats.oecd.org/Index.aspx?DataSetCode=TIVA_OECD_WTO.

usually does not differ radically from the same share for the whole economy (*Appendix 2*, *Tab. 8*)[41]. Second, we calculate the NCA index for different countries from gross export (import) data and from value added data. Third, we complement the analysis with comparative advantage index of *Timmer et al.* (2013) that is based on global value chain (GVC) income (value added in the production of final manufacturing goods[42]) data[43]:

$$RCA_{i,c,t}^{GVC} = \left(GVC_{i,c,t} \Big/ \sum_i GVC_{i,c,t} \right) \Big/ \left(\sum_c GVC_{i,c,t} \Big/ \sum_i \sum_c GVC_{i,c,t} \right). \quad (29)$$

Consider the example of Hungary (*Fig. 6*). The potential change in the NCA index after switching to value added data for machinery industries is larger than changes in export- and import-driven Balassa indices, reflecting the fact that a high foreign value added is embodied in intermediate inputs from other industries that are used to produce machinery goods. So, in the context of potential switch to value added statistics, the NCA index outperforms Balassa index. Note that the difference between export- and import-driven Balassa indices would also provide a biased result. This is clearly showed at *Fig. 6* for electrical machinery (DL).[44]

Now let's examine Italian comparative advantages. From *Fig. 6*, one can easily see that Italy is highly competitive in textile products, leather and footwear (DB, DC, DN). The GVC-based RCA index shows that this industry is even more important in terms of factor incomes. However, for wood and printing (DD, DE), high values of the GVC-based index are combined with low values of the NCA index. This means that factor incomes in the industry are higher than Italian average, but import exceeds export, even in value added terms. The same situation is observed for Russian food industry. However, for Russian wood and printing industry, situation is quite the opposite: factor incomes are relatively low, while net trade calculated from value added data is positive.

In the ideal world with full information, one should use GVC-based index to estimate a country's comparative advantages. In the real world with poor data quality[45], the NCA index that accounts for net trade is a reasonable alternative.

---

[41] For details, see *Appendix 2* ("How to adjust comparative advantage indices for value added data").
[42] They define it as "the income of all production factors that have been directly and indirectly used in the production of final manufacturing goods."
[43] We use the data from: http://www.wiod.org/new_site/gvc.htm.
[44] Though export-driven Balassa index values calculated from value added data for this industry strongly exceed import-driven Balassa index values, the NCA index calculated from value added data is negative.
[45] The first problem is the absence of GVC data for detailed commodity groups. The second problem is the proportionality assumption: "the share of imports in any product consumed directly as intermediate consumption or final demand (except exports) is the same for all users" (*Ahmad*, 2013, p. 100). *Timmer et al.* (2013) relaxed this problem a bit by dividing all products into three end-use categories (intermediate use, final consumption use and investment use). Within each end-use category, however, they again imply the proportionality assumption.

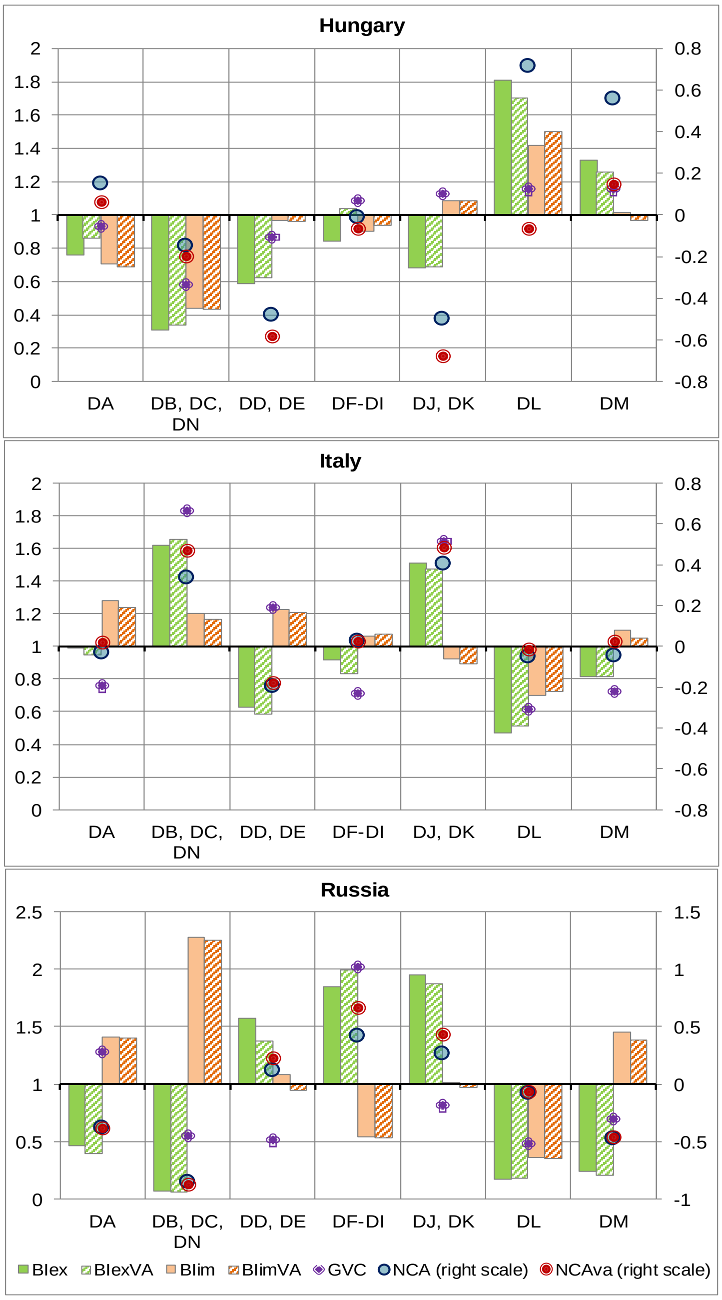


Fig. 6. Revealed comparative advantage indices calculated from gross and value added data
Source: Authors' calculations based on 2009 Trade in Value Added data.

6. Conclusion

The paper considers wide range of indicators used for detecting comparative advantages. Each of them is shown to have some of the following shortcomings: exposure to structural distortions, the focus on a single commodity, ignoring trade openness or import trade flows, instability of the distribution, presence of the extreme values, inconsistency with the sign of the net trade, etc. Special attention is given to the econometric index developed by economists from CEPII (*Leromain and Orefice*, 2013).

We propose the NCA index that is adjusted for the major shortcomings of other indices. Two important advantages of the index are simplicity of its decomposition to two or three meaningful parts (relative net exports, a good's trade intensity relative to the GDP which can be decomposed to good's trade intensity relative to the trade and openness of an economy) and possibility to calculate it for every level of disaggregation. Our index is the only index that simultaneously meets the following criteria: it accounts for trade flow volumes; it is neutral to the number of traded goods; it is hardly exposed to structural distortions; it is based on net trade data; and its sign is fully consistent with the sign of the net trade.

We show that CEPII econometric index doesn't have obvious advantages over the NCA index, since it doesn't match several criteria for comparative advantage indices (such as accounting for trade volumes, net trade and economic openness).

On the whole, we demonstrate that the NCA index has good empirical characteristics and at the same time does not require doing additional calculations and building extra models, unlike econometric indices. So, NCA index is easily calculated from foreign trade data and can be of a high-quality analytical tool if to use them properly.

References

Ahmad N. (2013). "Estimating Trade in Value-Added: Why and How?" in D.K. Elms, P. Low, eds., *Global Value Chains in a Changing World* (Geneva: WTO).

Anderson J. (2011). The Gravity Model // *Annual Review of Economics*, 3, 133-160.

Balassa B. (1965). Trade Liberalization and Revealed Comparative Advantage // *Manchester School of Economic and Social Studies*, 33(2), 99-123.

Balassa B. (1966). Tariff Reductions and Trade in Manufacturers among the Industrial Countries // *American Economic Review*, 56(3), 466-473.

Balassa B. (1977). "Revealed" Comparative Advantage Revisited: An Analysis of Relative Export Shares of the Industrial Countries, 1953-1971 // *Manchester School of Economic and Social Studies*, 45(4), 327-344.

Balassa B., Noland M. (1989). "Revealed" Comparative Advantage in Japan and the United States // *Journal of International Economic Integration*, 4(2), 8-22.

Ballance R., Forstner H., Murray T. (1985). On Measuring Comparative Advantage: A Note on Bowen's Indices // *Review of World Economics*, 121(2), 346-350.

Ballance R., Forstner H., Murray T. (1986). More on Measuring Comparative Advantage: A Reply // *Review of World Economics*, 122(2), 375-378.

Ballance R., Forstner H., Murray T. (1987). Consistency Tests of Alternative Measures of Comparative Advantage // *Review of Economics and Statistics*, 69(1), 157-161.

Bebek U. (2011). Consistency of the Proposed Additive Measures of Revealed Comparative Advantage // *Economics Bulletin*, 31(3), 2491-2499.

Bowen H. (1983). On the Theoretical Interpretation of Indices of Trade Intensity and Revealed Comparative Advantage // *Review of World Economics*, 119(3), 464-472.

Bowen H. (1986). On Measuring Comparative Advantage: Further Comments // *Review of World Economics*, 122(2), 379-381.

Carrere C., Grigoriou C. (2015). Can Mirror Data Help to Capture Informal International Trade? FERDI Working Paper 123.

Chor D. (2010). Unpacking Sources of Comparative Advantage: A Quantitative Approach // *Journal of International Economics*, 82(2), 152-167.

Costinot A., Donaldson D., Komunjer I. (2012). What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas // *Review of Economic Studies*, 79(2), 581-608.

Dalum B., Laursen K., Villumsen G. (1998). Structural Change in OECD Export Specialisation Patterns: de-specialisation and 'stickiness' // *International Review of Applied Economics*, 12(3), 423-443.

De Ferranti D., Perry G., Lederman D., Maloney W. (2002). *From Natural Resources to Knowledge Economy: Trade and Job Quality* (Washington, DC: World Bank).

Deardorff A. (1980). The General Validity of the Law of Comparative Advantage // *Journal of Political Economy*, 88(5), 941-957.

Donges J., Riedel J. (1977). The Expansion of Manufactured Exports in Developing Countries: An Empirical Assessment of Supply and Demand Issues // *Review of World Economics*, 113(1), 58-87.

Eaton J., Kortum S. (2002). Technology, Geography, and Trade // *Econometrica*, 70(5), 1741-1779.

Fadinger H., Fleiss P. (2011). Trade and Sectoral Productivity // *The Economic Journal*, 121(555), 958-989.

Finicelli A., Pagano P., Sbracia M. (2013). Ricardian Selection // *Journal of International Economics*, 89(1), 96-109.

Gaulier G., Zignano S. (2010). BACI: International Trade Database at the Product-level. The 1994-2007 Version. CEPII Working Paper 2010-23.

Hanson G., Lind N., Muendler M.-A. (2014). The Dynamics of Comparative Advantage. Unpublished Working Paper.

Hinloopen J., Van Marrewijk C. (2001). On the Empirical Distribution of the Balassa Index // *Review of World Economics*, 137(1), 1-35.

Hoen A., Oosterhaven J. (2006). On the Measurement of Comparative Advantage // *The Annals of Regional Science*, 40(3), 677-691.

Hume N., Rice X. (2014). Indonesia Export Ban Turns Nickel into a Star. The Financial Times. 20 March 2014.

Kunimoto K. (1977). Typology of Trade Intensity Indices // Hitotsubashi Journal of Economics, 17, 15-32.

Lafay G. (1992). The Measurement of Revealed Comparative Advantages, in M.G. Dagenais and P.A. Muet, eds., *International Trade Modeling*. London: Chapman & Hill.

Laursen K. (1998). Revealed Comparative Advantage and the Alternatives as Measures of International Specialisation. DRUID Working Paper 98-30.

Leamer E. (1974). The Commodity Composition of International Trade in Manufactures: An Empirical Analysis // *Oxford Economic Papers*, 26(3), 350-374.

Leamer E. (1984). *Sources of International Comparative Advantage: Theory and Evidence* (Cambridge, Mass.: MIT Press).

Leromain E., Orefice G. (2013). New Revealed Comparative Advantage Index: Dataset and Empirical Distribution. CEPII Working Paper No 2013-20.

Levchenko A., Zhang J. (2011). The Evolution of Comparative Advantage: Measurement and Welfare Implications. NBER Working Paper 16806.

Liesner H. (1958). The European Common Market and British Industry // *Economic Journal*, 68(270), 302-316.

Martin M. (2013). What's the Difference? – Comparing U.S. and Chinese Trade Data. Congressional Research Service Report RS22640.

Melitz M. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity // *Econometrica*, 71(6), 1695-1725.

Michaely M. (1962). *Concentration in International Trade*. Amsterdam: North-Holland Publishing Company.

OECD (2011). Import content of exports, in *OECD Science,Technology and Industry Scoreboard 2011* (OECD Publishing).

Proudman J., Redding S. (2000). Evolving Patterns of International Trade // *Review of International Economics*, 8(3), 373-396.

Shikher S. (2012). Putting Industries into the Eaton-Kortum Model // *The Journal of International Trade and Economic Development*, 21(6), 807-837.

Timmer M., Los B., Stehrer R., de Vries G. (2013). Fragmentation, Incomes and Jobs: An Analysis of European Competitiveness // *Economic Policy*, 28(76), 613-661.

UN (2013). *International Merchandise Trade Statistics: Compilers Manual, Revision 1 (IMTS 2010-CM)*. New York: United Nations.

UNIDO (1982). *Changing Patterns of Trade in World Industry: An Empirical Study on Revealed Comparative Advantage*. New York: United Nations.

Vollrath T. (1991). A Theoretical Evaluation of Alternative Trade Intensity Measures of Revealed Comparative Advantage // *Review of World Economics*, 127(2), 265-280.

Wagner J. (2007). Exports and Productivity: A Survey of the Evidence from Firm-level Data // *The World Economy*, 30(1), 60-82.

Yeats A. (1985). On the Appropriate Interpretation of the Revealed Comparative Advantage Index: Implications of a Methodology Based on Industry Sector Analysis // *Review of World Economics*, 121(1), 61-73.

Yu R., Cai J., Leung P. (2009). The Normalized Revealed Comparative Advantage Index // *The Annals of Regional Science*, 43(1), 267-282.

Appendix 1

Tab. 5. Comparative advantage indices: summary statistics

Original variables					*Normalized variables*				
		RCA					**RCA_n**		
	Percentiles	Smallest				Percentiles	Smallest		
1%	.6319315	.115832			1%	-1.875149	-4.436053		
5%	.7344125	.3467774			5%	-1.366634	-3.290093		
10%	.7947356	.3823719	Obs	22069	10%	-1.067309	-3.113472	Obs	22069
25%	.8880679	.4019697	Sum of Wgt.	22069	25%	-.6041902	-3.016227	Sum of Wgt.	22069
50%	.9855606		Mean	1.00983	50%	-.1204278		Mean	-1.90e-11
		Largest	Std. Dev.	.2015301			Largest	Std. Dev.	1
75%	1.095399	2.632444			75%	.4245945	8.051467		
90%	1.2463	2.707547	Variance	.0406144	90%	1.173372	8.424133	Variance	1
95%	1.379534	2.755848	Skewness	1.363121	95%	1.834482	8.663806	Skewness	1.363121
99%	1.68779	2.854494	Kurtosis	7.814495	99%	3.364061	9.153288	Kurtosis	7.814495

		lnRCA					**lnRCA_n**		
	Percentiles	Smallest				Percentiles	Smallest		
1%	-.4589743	-2.155614			1%	-2.374253	-11.31815		
5%	-.3086844	-1.059072			5%	-1.581995	-5.537691		
10%	-.2297459	-.9613616	Obs	22069	10%	-1.165867	-5.022606	Obs	22069
25%	-.1187071	-.9113786	Sum of Wgt.	22069	25%	-.5805219	-4.759119	Sum of Wgt.	22069
50%	-.0145447		Mean	-.0085832	50%	-.0314259		Mean	1.34e-10
		Largest	Std. Dev.	.189698			Largest	Std. Dev.	1
75%	.0911187	.9679126			75%	.5255824	5.147634		
90%	.2201794	.9960431	Variance	.0359853	90%	1.205931	5.295925	Variance	1
95%	.3217456	1.013725	Skewness	.2302551	95%	1.741341	5.389138	Skewness	.2302551
99%	.52342	1.048895	Kurtosis	5.255208	99%	2.804475	5.574533	Kurtosis	5.255209

		lnBIex					**lnBIex_n**		
	Percentiles	Smallest				Percentiles	Smallest		
1%	-10.41117	-19.33843			1%	-2.769649	-5.853205		
5%	-7.665322	-18.50678			5%	-1.821211	-5.565948		
10%	-6.275634	-18.38515	Obs	119686	10%	-1.3412	-5.523935	Obs	119686
25%	-4.125155	-17.51556	Sum of Wgt.	119686	25%	-.5984057	-5.223571	Sum of Wgt.	119686
50%	-2.031522		Mean	-2.392699	50%	.1247538		Mean	5.61e-11
		Largest	Std. Dev.	2.89512			Largest	Std. Dev.	1
75%	-.3361135	9.87243			75%	.7103628	4.236484		
90%	.8799586	10.18298	Variance	8.381718	90%	1.130405	4.343752	Variance	1
95%	1.627446	10.53391	Skewness	-.6042962	95%	1.388594	4.464965	Skewness	-.6042962
99%	3.361621	12.64275	Kurtosis	3.690736	99%	1.987593	5.193376	Kurtosis	3.690736

		BIex					**BIex_n**		
	Percentiles	Smallest				Percentiles	Smallest		
1%	.00003	0			1%	-.0068718	-.0068718		
5%	.0004684	3.99e-09			5%	-.0068713	-.0068718		
10%	.0018814	9.17e-09	Obs	119687	10%	-.0068698	-.0068718	Obs	119687
25%	.0161591	1.04e-08	Sum of Wgt.	119687	25%	-.0068541	-.0068718	Sum of Wgt.	119687
50%	.1311356		Mean	6.250526	50%	-.0067277		Mean	-5.35e-11
		Largest	Std. Dev.	909.5859			Largest	Std. Dev.	1
75%	.714542	19388.4			75%	-.0060863	21.30876		
90%	2.4108	26449.25	Variance	827346.6	90%	-.0042214	29.07147	Variance	1
95%	5.090855	37568.05	Skewness	330.1216	95%	-.0012749	41.2955	Skewness	330.1216
99%	28.83589	309509.9	Kurtosis	112040.8	99%	.0248304	340.2688	Kurtosis	112040.8

		PRex					**PRex_n**		
	Percentiles	Smallest				Percentiles	Smallest		
1%	.0000202	0			1%	-.1036865	-.1036879		
5%	.0003324	3.88e-09			5%	-.1036656	-.1036879		
10%	.0013679	6.88e-09	Obs	119687	10%	-.1035962	-.1036879	Obs	119687
25%	.0124877	1.76e-08	Sum of Wgt.	119687	25%	-.1028512	-.1036879	Sum of Wgt.	119687
50%	.1097564		Mean	1.547618	50%	-.0963344		Mean	-1.06e-10
		Largest	Std. Dev.	14.92574			Largest	Std. Dev.	1
75%	.6275369	1104.554			75%	-.0616439	73.89958		
90%	2.08747	1179.875	Variance	222.7777	90%	.0361691	78.946	Variance	1
95%	4.234059	1208.118	Skewness	42.57636	95%	.1799871	80.83824	Skewness	42.57636
99%	20.40912	1227.424	Kurtosis	2551.352	99%	1.26369	82.1317	Kurtosis	2551.352

Tab. 5. Comparative advantage indices: summary statistics (continued)

Original variables					*Normalized variables*				
	LAURex					**LAURex_n**			
	Percentiles	Smallest				Percentiles	Smallest		
1%	-.9999398	-1			1%	-.8760217	-.8761299		
5%	-.9990628	-1			5%	-.8744456	-.8761299		
10%	-.9962439	-1	Obs	119686	10%	-.8693792	-.8761299	Obs	119686
25%	-.968192	-.9999999	Sum of Wgt.	119686	25%	-.8189627	-.8761298	Sum of Wgt.	119686
50%	-.7681342		Mean	-.5125193	50%	-.4594067		Mean	6.49e-11
		Largest	Std. Dev.	.5564023			Largest	Std. Dev.	1
75%	-.1664923	.9998968			75%	.6219007	2.718206		
90%	.4136273	.9999244	Variance	.3095835	90%	1.664527	2.718256	Variance	1
95%	.6716388	.9999468	Skewness	1.06914	95%	2.128241	2.718296	Skewness	1.06914
99%	.9329666	.9999936	Kurtosis	2.944488	99%	2.597915	2.71838	Kurtosis	2.944488

	HOex					**HOex_n**			
	Percentiles	Smallest				Percentiles	Smallest		
1%	-.0085016	-.0668426			1%	-.654065	-5.006554		
5%	-.0024006	-.0668426			5%	-.198905	-5.006553		
10%	-.0013349	-.0668426	Obs	119689	10%	-.1193954	-5.006552	Obs	119689
25%	-.0004383	-.0668426	Sum of Wgt.	119689	25%	-.0525045	-5.006552	Sum of Wgt.	119689
50%	-.0001095		Mean	.0002655	50%	-.0279775		Mean	-8.01e-11
		Largest	Std. Dev.	.0134041			Largest	Std. Dev.	1
75%	-9.94e-06	.7502708			75%	-.0205504	55.95358		
90%	.0003389	.782474	Variance	.0001797	90%	.005478	58.35607	Variance	1
95%	.0014811	.802609	Skewness	32.88029	95%	.0906864	59.85823	Skewness	32.88029
99%	.0135925	.845933	Kurtosis	1427.731	99%	.9942487	63.09039	Kurtosis	1427.731

	DRI					**DRI_n**			
	Percentiles	Smallest				Percentiles	Smallest		
1%	-408.6794	-643.0815			1%	-5.472766	-8.717187		
5%	-56.46507	-643.0409			5%	-.597677	-8.716624		
10%	-21.89257	-643.0314	Obs	115158	10%	-.1191501	-8.716492	Obs	115158
25%	-7.627281	-642.9655	Sum of Wgt.	115158	25%	.0782994	-8.715581	Sum of Wgt.	115158
50%	-1.991098		Mean	-13.28424	50%	.1563113		Mean	-4.57e-11
		Largest	Std. Dev.	72.24777			Largest	Std. Dev.	1
75%	.2161794	636.6459			75%	.1868628	8.99585		
90%	5.414406	636.7332	Variance	5219.74	90%	.2588128	8.997058	Variance	1
95%	12.24178	640.2721	Skewness	-4.297053	95%	.3533122	9.046041	Skewness	-4.297053
99%	69.82108	641.0815	Kurtosis	39.46425	99%	1.150282	9.057245	Kurtosis	39.46425

	VolI					**VolI_n**			
	Percentiles	Smallest				Percentiles	Smallest		
1%	-9.579689	-19.13131			1%	-2.614146	-5.783738		
5%	-6.799311	-17.994			5%	-1.691511	-5.406335		
10%	-5.471604	-17.58671	Obs	114772	10%	-1.250927	-5.271179	Obs	114772
25%	-3.45605	-17.54972	Sum of Wgt.	114772	25%	-.58209	-5.258906	Sum of Wgt.	114772
50%	-1.496563		Mean	-1.701911	50%	.0681422		Mean	-5.41e-11
		Largest	Std. Dev.	3.013519			Largest	Std. Dev.	1
75%	.0585096	16.84259			75%	.5841745	6.153772		
90%	1.559869	17.33555	Variance	9.081295	90%	1.082383	6.317352	Variance	1
95%	2.898597	18.64771	Skewness	-.0229362	95%	1.526623	6.752779	Skewness	-.0229362
99%	6.309129	23.13908	Kurtosis	4.699561	99%	2.658367	8.243186	Kurtosis	4.699561

	YUex					**YUex_n**			
	Percentiles	Smallest				Percentiles	Smallest		
1%	-.0000561	-.0074815			1%	-.6250129	-83.11657		
5%	-.0000124	-.0057136			5%	-.1400895	-63.47615		
10%	-5.31e-06	-.0054742	Obs	119689	10%	-.0611916	-60.81701	Obs	119689
25%	-1.07e-06	-.0044045	Sum of Wgt.	119689	25%	-.0141442	-48.93309	Sum of Wgt.	119689
50%	-1.30e-07		Mean	2.02e-07	50%	-.0036912		Mean	-1.58e-10
		Largest	Std. Dev.	.00009			Largest	Std. Dev.	1
75%	-4.54e-09	.0058468			75%	-.0022974	64.95185		
90%	1.01e-06	.0076033	Variance	8.10e-09	90%	.0089232	84.46533	Variance	1
95%	6.30e-06	.0081028	Skewness	33.09813	95%	.0677856	90.0135	Skewness	33.09813
99%	.0000635	.0124544	Kurtosis	5341.467	99%	.703257	138.3567	Kurtosis	5341.467

Tab. 5. Comparative advantage indices: summary statistics (continued)

	Original variables **NCA**					*Normalized variables* **NCA_n**			
	Percentiles	Smallest				Percentiles	Smallest		
1%	-12.76676	-151844.9			1%	-.0283238	-394.9009		
5%	-3.474227	-13556.83			5%	-.0041566	-35.25263		
10%	-1.885214	-11455.13	Obs	160845	10%	-.000024	-29.7867	Obs	160845
25%	-.7113369	-8756.695	Sum of Wgt.	160845	25%	.0030289	-22.76882	Sum of Wgt.	160845
50%	-.2060176		Mean	-1.875993	50%	.0043431		Mean	-9.21e-11
		Largest	Std. Dev.	384.5091			Largest	Std. Dev.	1
75%	-.0119355	1568.628			75%	.0048479	4.08444		
90%	.2931182	2155.875	Variance	147847.3	90%	.0056412	5.611702	Variance	1
95%	1.258688	5424.455	Skewness	-383.5042	95%	.0081524	14.11236	Skewness	-383.5042
99%	8.553424	5848.897	Kurtosis	151219	99%	.027124	15.21621	Kurtosis	151219

	SNCA					**SNCA_n**			
	Percentiles	Smallest				Percentiles	Smallest		
1%	-1.601831	-7.070049			1%	-2.845744	-14.06256		
5%	-.9803729	-6.622752			5%	-1.570963	-13.14503		
10%	-.7412691	-6.059665	Obs	160844	10%	-1.080496	-11.98999	Obs	160844
25%	-.4144756	-5.968553	Sum of Wgt.	160844	25%	-.4101526	-11.80309	Sum of Wgt.	160844
50%	-.1558704		Mean	-.2145255	50%	.1203177		Mean	-3.20e-11
		Largest	Std. Dev.	.4875018			Largest	Std. Dev.	1
75%	-.0113062	5.044369			75%	.4168584	10.78744		
90%	.1645549	5.06219	Variance	.237658	90%	.777598	10.82399	Variance	1
95%	.4796612	5.211906	Skewness	-.2001549	95%	1.423967	11.1311	Skewness	-.2001549
99%	1.223833	5.452081	Kurtosis	14.64177	99%	2.950469	11.62377	Kurtosis	14.64177

Source: Authors' calculations in STATA (for abbreviations of variables, see *Tab. 6*).

Tab. 6. Abbreviations for comparative advantage indices

Abbreviation	Indicator
RCA	econometric indicator (*Leromaine and Orefice* , 2013)
lnRCA	econometric indicator, log
Blex	Balassa index
lnBlex	Balassa index, log
PRex	Proudman-Redding index
LAURex	Laursen index
HOex	Hoen-Oosterhaven index
DRI	Donges-Riedel index
Voll	Vollrath index (RCA^{V2})
YUex	Yu index
NCA	net comparative advantage index
SNCA	symmetric net comparative advantage index
…_n	normalized verion of an indicator under consideration*

** subtract the mean value and divide by the standard deviation*

Source: Compiled by the authors.

Appendix 2

Tab. 7. The ratio of Balassa indices calculated from gross and value added export data

	Me-dian	Food products	Textiles, leather, footwear	Wood, pulp, paper	Chemicals, construc-tion mater.	Metals	Machinery, equipment	Electrical equipment	Transport equipment	Manuf., n.e.s.
share in exports, %:		*7*	*7*	*5*	*22*	*10*	*10*	*21*	*14*	*4*
NACE codes		*DA*	*DB, DC*	*DD, DE*	*DF-DI*	*DJ*	*DK*	*DL*	*DM*	*DN*
Median		**1.02**	**1.00**	**0.99**	**1.03**	**0.99**	**0.99**	**1.05**	**0.97**	**1.03**
Australia	**0.96**	0.96	0.99	0.96	1.05	0.94	0.89	1.10	0.99	0.94
Austria	**1.01**	1.05	0.96	1.01	1.01	1.00	1.01	1.02	0.81	0.89
Argentina	**0.96**	0.94	0.96	0.89	1.01	0.96	0.94	0.98	0.82	0.98
Belgium	**1.01**	1.01	1.01	1.01	1.00	0.95	1.05	1.26	0.91	0.95
Bulgaria	**1.04**	1.04	1.16	0.94	0.81	1.00	1.02	1.14	1.09	1.07
Brazil	**0.97**	0.91	0.96	0.91	1.01	0.99	0.95	1.04	0.97	0.98
Brunei	**0.97**	0.75	0.93	0.91	1.20	0.70		1.01	1.03	1.04
United Kingdom	**0.99**	0.99	1.00	0.99	1.03	0.98	0.96	1.09	0.93	1.00
Hungary	**1.05**	1.13	1.05	1.06	1.23	1.02	1.00	0.94	0.95	1.13
Viet Nam	**0.89**	1.18	0.73	1.06	1.07	0.76	0.89	0.84	0.91	0.83
Germany	**0.97**	0.97	0.96	1.00	0.96	0.91	1.03	1.18	0.95	1.06
Hong Kong	**1.25**	1.25	1.28	1.27	1.04	1.14	1.05	1.02	1.36	1.25
Greece	**1.08**	1.04	1.12	1.07	0.96	0.93	1.09	1.08	1.18	1.11
Denmark	**0.94**	0.90	0.92	0.91	1.07	0.96	0.94	1.06	0.88	0.95
Israel	**1.12**	1.12	1.05	1.10	1.12	1.13	1.06	1.30	1.22	0.35
India	**1.08**	1.07	1.08	1.06	1.09	1.09	1.04	1.21	1.09	0.68
Indonesia	**0.97**	0.97	0.86	0.90	1.09	1.01	0.71	0.97	1.02	0.99
Ireland	**0.89**	0.81	0.97	0.76	1.19	1.25	0.98	0.72	0.61	
Iceland	**1.19**	1.39	1.19	0.77	0.94	1.18	1.22	1.03	1.20	1.19
Spain	**1.03**	0.99	1.03	0.99	0.96	1.04	1.03	1.09	0.95	1.03
Italy	**0.99**	0.96	1.03	0.93	0.91	0.94	0.99	1.10	1.00	1.02
Cambodia	**0.85**	1.22	0.73	1.14	1.07	0.71	0.85	0.82	0.94	0.80
Canada	**1.02**	1.02	0.99	1.04	1.05	0.92	0.97	1.06	0.90	1.05
China	**0.99**	1.01	1.14	0.88	0.96	0.99	0.92	0.97	1.03	1.11
Chinese Taipei	**1.11**	1.24	1.14	1.09	0.89	1.01	1.06	1.11	1.20	1.16
Latvia	**0.92**	0.92	0.95	0.95	1.00	0.83	0.87	0.79	0.84	0.94
Lithuania	**1.20**	1.16	1.12	1.20	0.69	1.12	1.21	1.35	1.31	1.25
Luxembourg	**1.04**	0.97	0.86	0.82	1.04	0.95	1.06	1.34	1.26	1.19
Malaysia	**0.97**	1.19	1.06	1.25	1.28	0.92	0.77	0.88	0.97	0.96
Malta	**1.12**	1.12	1.28	0.97	1.13	0.95	1.10	0.94	1.26	1.19
Mexico	**1.12**	1.16	1.13	1.12	1.36	1.18	1.03	0.75	1.06	1.03
Netherlands	**1.08**	1.02	1.10	1.18	0.85	1.08	1.12	1.02	1.05	1.39
New Zealand	**0.91**	0.91	0.89	0.91	0.91	0.97	0.94	1.01	0.83	0.95
Norway	**1.04**	0.97	1.04	1.04	1.19	0.78	0.95	1.07	0.90	1.05
Poland	**0.98**	1.03	0.95	0.98	1.03	0.97	0.97	1.05	0.90	1.02
Portugal	**1.06**	1.06	1.09	1.10	1.03	1.02	1.11	0.87	0.73	1.16
Russia	**0.92**	0.87	0.86	0.88	1.08	0.99	0.92	1.06	0.88	0.93
Romania	**1.00**	1.04	0.92	1.00	1.00	0.93	0.95	1.08	1.03	1.04
Saudi Arabia	**0.72**	0.75	0.74	0.69	1.08	0.55	0.72	0.85	0.63	0.69
Singapore	**1.07**	0.98	0.98	1.20	1.08	1.12	0.95	0.98	1.29	1.07
Slovak Republic	**1.03**	1.24	1.01	1.20	1.03	1.17	0.99	0.92	0.82	1.18
Slovenia	**0.96**	1.06	0.88	0.94	1.12	0.96	0.94	1.07	0.86	1.04
United States	**0.99**	0.91	0.90	0.95	0.99	1.02	0.96	1.13	0.99	1.00
Thailand	**1.01**	1.21	1.13	0.98	1.16	1.02	0.87	0.81	0.91	1.01
Turkey	**0.97**	1.02	1.04	0.93	0.97	0.92	0.97	1.06	0.97	0.94
Philippines	**1.26**	1.52	1.25	1.35	1.29	1.16	1.17	1.04	1.26	1.41
Finland	**1.08**	1.08	1.13	1.15	1.01	0.89	1.11	0.87	0.99	1.18
France	**1.01**	1.01	0.92	1.02	1.01	1.01	1.02	1.11	0.87	1.10
Czech Republic	**1.11**	1.16	0.96	1.11	1.18	1.09	1.12	0.73	0.95	1.11
Chile	**0.96**	0.93	0.91	0.96	0.77	1.12	0.92	1.50	1.03	1.12
Switzerland	**0.98**	1.02	0.82	1.03	0.94	1.04	0.98	1.14	0.93	0.98
Sweden	**1.06**	1.08	1.09	1.08	0.88	1.01	1.03	1.06	0.89	1.09
Estonia	**0.96**	0.96	0.97	0.93	1.05	0.85	0.94	0.99	1.08	0.95
South Africa	**1.01**	0.99	1.02	0.96	1.10	0.97		1.07	0.83	1.03
Korea	**1.14**	1.12	1.14	1.21	0.74	1.01	1.16	1.05	1.15	1.27
Japan	**0.97**	0.94	0.96	0.96	1.00	0.96	1.01	1.08	1.04	0.97

Source: Authors' calculations based on 2009 Trade in Value Added data.

Tab. 8. Domestic value added embodied in gross exports, %

	TOTAL	Food products	Textiles, leather, footwear	Wood, pulp, paper	Chemicals, construc-tion mater.	Metals	Machinery, equipment	Electrical equipment	Transport equipment	Manuf., n.e.s.
share in exports, %:		*7*	*7*	*5*	*22*	*10*	*10*	*21*	*14*	*4*
NACE codes		*DA*	*DB, DC*	*DD, DE*	*DF-DI*	*DJ*	*DK*	*DL*	*DM*	*DN*
TOTAL	**69**	**80**	**75**	**79**	**66**	**70**	**73**	**63**	**69**	**73**
Australia	**81**	89	86	89	81	77	76	81	80	81
Austria	**60**	72	62	70	58	61	64	56	48	56
Argentina	**85**	92	87	86	81	82	84	76	69	88
Belgium	**58**	67	63	67	55	56	64	67	52	59
Bulgaria	**60**	72	75	65	46	60	65	62	65	68
Brazil	**89**	93	92	92	85	89	89	85	86	92
Brunei	**63**	55	63	67	72	46		58	64	69
United Kingdom	**75**	85	81	85	73	74	76	75	69	79
Hungary	**54**	70	61	65	63	55	57	46	51	64
Viet Nam	**48**	64	37	58	48	36	44	36	43	42
Germany	**69**	77	72	79	64	63	75	75	66	78
Hong Kong	**51**	72	70	73	50	58	56	47	68	66
Greece	**69**	83	83	85	63	65	80	68	81	81
Denmark	**74**	77	74	77	76	72	74	71	65	75
Israel	**61**	79	69	77	65	70	68	73	74	22
India	**70**	87	82	85	73	78	77	78	76	51
Indonesia	**82**	91	76	85	85	84	61	72	83	86
Ireland	**58**	54	61	51	66	73	60	38	35	
Iceland	**52**	83	67	46	46	62	67	49	62	66
Spain	**73**	83	80	82	66	76	79	72	69	79
Italy	**77**	85	85	82	67	73	81	77	77	83
Cambodia	**48**	67	38	63	49	35	43	36	45	41
Canada	**71**	83	76	85	71	67	73	69	64	79
China	**65**	75	79	65	59	65	63	57	67	76
Chinese Taipei	**54**	77	67	68	46	56	60	55	65	66
Latvia	**68**	72	70	74	65	57	63	49	57	68
Lithuania	**56**	75	68	77	36	63	71	69	72	74
Luxembourg	**33**	37	31	31	33	32	37	41	42	42
Malaysia	**55**	75	62	78	66	51	44	44	53	55
Malta	**54**	69	74	59	58	52	62	46	67	68
Mexico	**63**	84	77	81	82	75	69	43	67	69
Netherlands	**53**	62	63	72	43	58	63	49	55	78
New Zealand	**80**	83	76	83	69	78	79	73	65	79
Norway	**70**	78	78	83	78	54	70	68	62	77
Poland	**69**	81	70	77	67	67	71	66	61	74
Portugal	**58**	70	68	73	56	59	68	46	41	70
Russia	**91**	91	84	91	93	91	88	88	79	89
Romania	**72**	86	72	83	69	68	73	71	74	79
Saudi Arabia	**90**	77	71	71	92	50	68	70	56	66
Singapore	**43**	49	46	59	44	49	44	39	56	49
Slovak Republic	**51**	73	56	71	50	61	54	43	42	64
Slovenia	**61**	74	57	65	65	59	60	59	52	67
United States	**85**	88	82	92	79	87	86	87	83	89
Thailand	**61**	84	74	68	67	63	56	45	55	65
Turkey	**73**	85	81	78	67	68	75	70	70	72
Philippines	**52**	91	71	80	64	61	65	50	66	78
Finland	**58**	72	70	76	56	52	68	46	57	72
France	**69**	81	69	81	67	71	75	70	60	81
Czech Republic	**55**	73	56	69	61	60	65	37	51	64
Chile	**70**	74	68	76	51	79	67	95	71	82
Switzerland	**65**	76	57	76	58	69	67	68	60	67
Sweden	**59**	73	69	73	49	61	64	57	52	68
Estonia	**62**	69	65	66	62	54	62	56	67	62
South Africa	**75**	85	82	82	78	73		73	62	81
Korea	**56**	71	68	77	39	56	68	53	64	74
Japan	**84**	90	86	91	79	81	89	82	86	86
Rest of the World	**73**	87	68	77	71	82	63	73	83	78

Source: Authors' calculations based on 2009 Trade in Value Added data.

How to adjust comparative advantage indices for value added data:

First, one should calculate domestic value added embodied in gross exports and foreign value added embodied in gross imports (see *Tab. 8* for exports data). That is vital for comparing the reaction of Balassa index and relative net exports index on switching to value added data.

The reaction of Balassa index in a certain industry depends on the extent of matching between the shares of domestic value added in gross exports for this industry and for the whole economy. If the mismatch is small (as for Singapore's exports of machinery and equipment, for example), then Balassa index would not change much after one switches to value added data. The reaction of Balassa index analogue for imports depends on the extent of matching between the shares of foreign value added in gross imports for the industry and for the whole economy.

Denote the share of value added in gross exports of country c in industry i as $x_{i,c,t}$, the share of value added in total gross exports of country c as $x_{c,t}$, the share of value added in world exports in industry i as $x_{i,W,t}$, and the share of value added in total world exports as $x_{W,t}$. Then Balassa index adjusted for value added should be calculated as follows:

$$BI_{i,c,t}^{VA} = \frac{X_{i,c,t}x_{i,c,t}}{\left(\sum_i X_{i,c,t}\right)x_{c,t}} \Bigg/ \frac{\left(\sum_c X_{i,c,t}\right)x_{i,W,t}}{\left(\sum_i \sum_c X_{i,c,t}\right)x_{W,t}} . \qquad (30)$$

Given (1), it is straightforward to show that the growth of Balassa index after switching to value added data is described by the following equation:

$$\frac{BI_{i,c,t}^{VA}}{BI_{i,c,t}} = \frac{x_{i,c,t}}{x_{c,t}} \Bigg/ \frac{x_{i,W,t}}{x_{W,t}} . \qquad (31)$$

Since the denominator of (31) does not depend on the country of interest, the growth of Balassa index for a certain industry across countries (columns in *Tab. 7*) is inversely proportional to the ratio $x_{i,W,t} / x_{W,t}$ (the ratio of columns for industries and the "Total" column in *Tab. 8*). That is, it is simple to represent this growth as the linear equation with the variable $x_{i,c,t}$ / $x_{c,t}$ and the constant $x_{i,W,t} / x_{W,t}$ (*Fig. 7*).

The reaction of relative net export index does not depend on the changes in world trade patterns. Denote the share of value added in gross imports of country c in industry i as $m_{i,c,t}$. Then relative net export adjusted for value added data should be calculated as follows:

$$RNX_{i,c,t}^{VA} = \frac{X_{i,c,t}x_{i,c,t} - M_{i,c,t}m_{i,c,t}}{X_{i,c,t}x_{i,c,t} + M_{i,c,t}m_{i,c,t}} . \qquad (32)$$

Thus, the reaction of the indicator would solely depend on the values of $x_{i,c,t}$ and $m_{i,c,t}$.

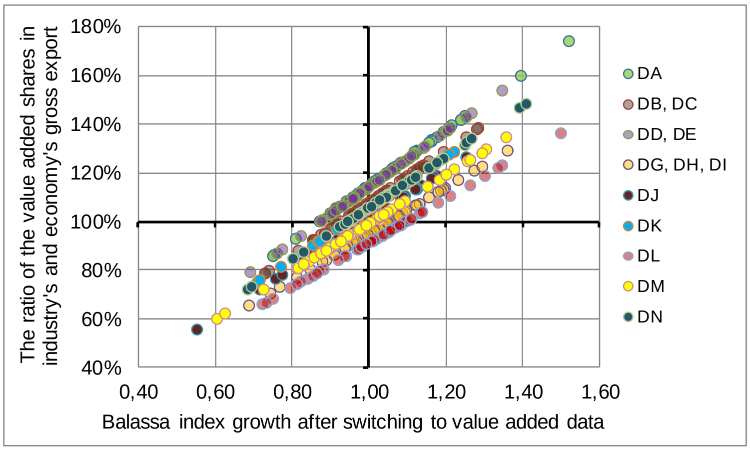
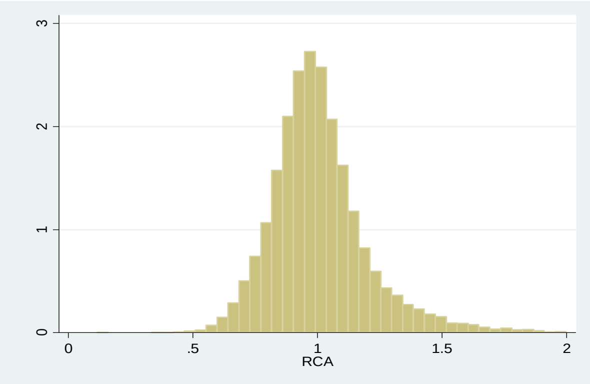


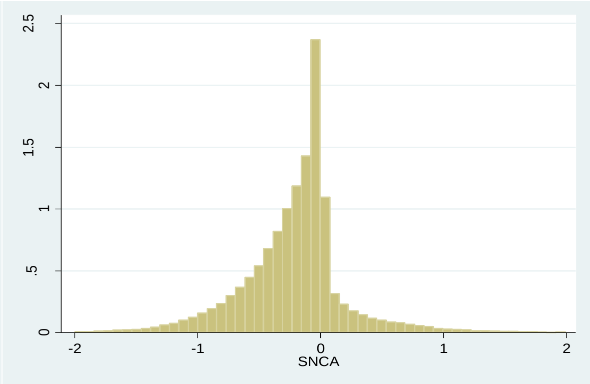
Fig. 7. Dependence of Balassa index growth after switching to value added data on the ratio of the value added shares in industry's and economy's gross export for commodity aggregates

Source: Authors' calculations based on 2009 Trade in Value Added data.

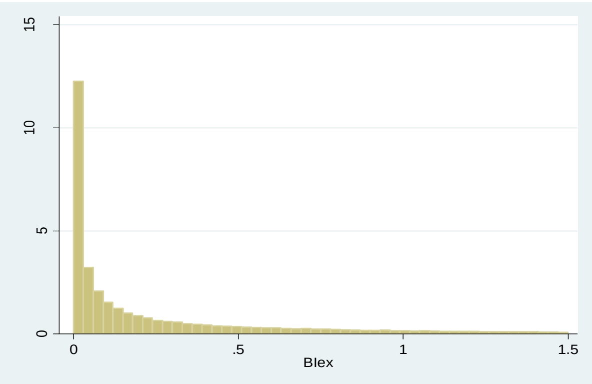
Appendix 3



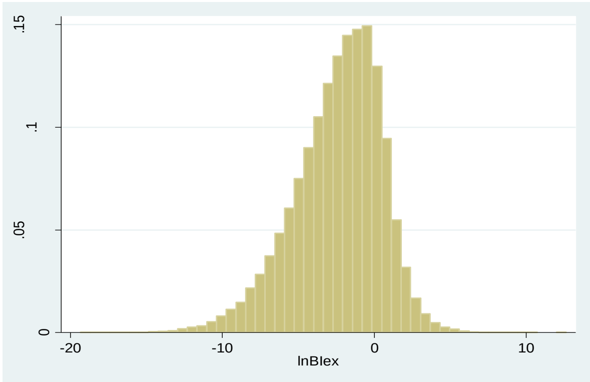
Leromain and Orefice (2013) index



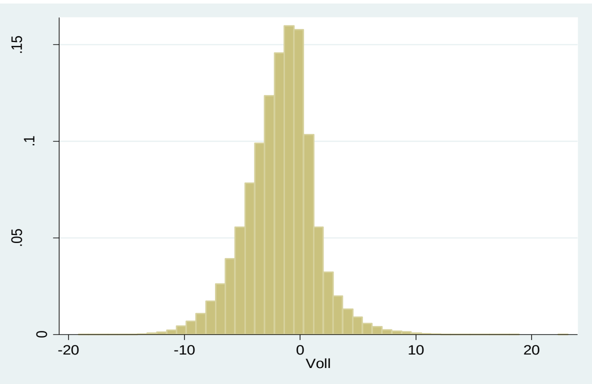
Symmetric NCA index



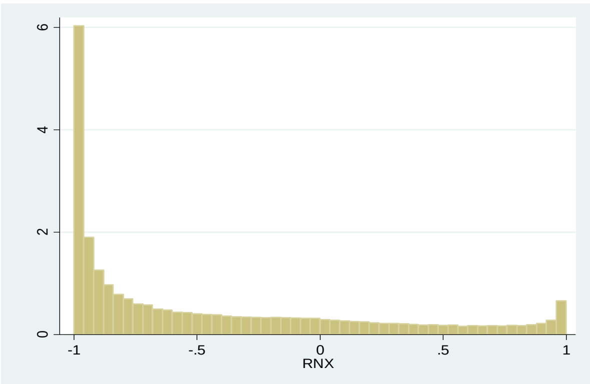
Balassa (1965) index



Logarithm of Balassa index



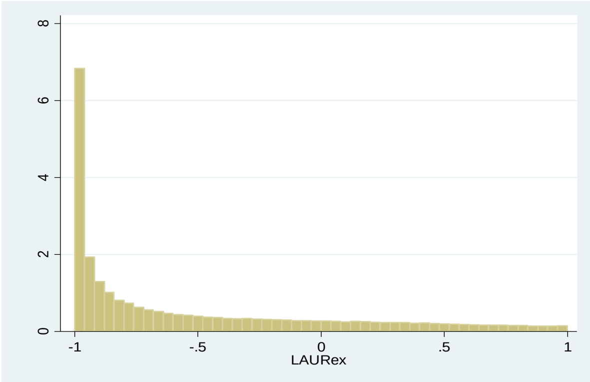
Vollrath (1991) RCA^{V2} index

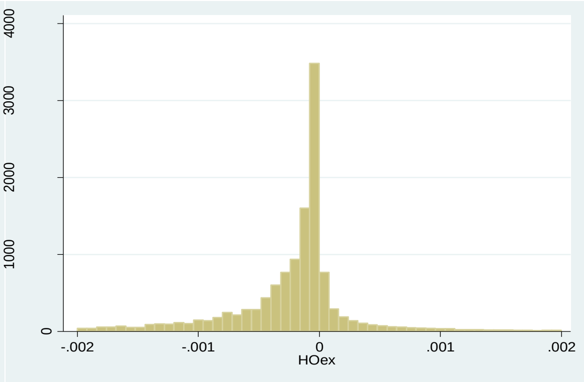


Relative net export index

Fig. 8. Probability distribution for revealed comparative advantage indices

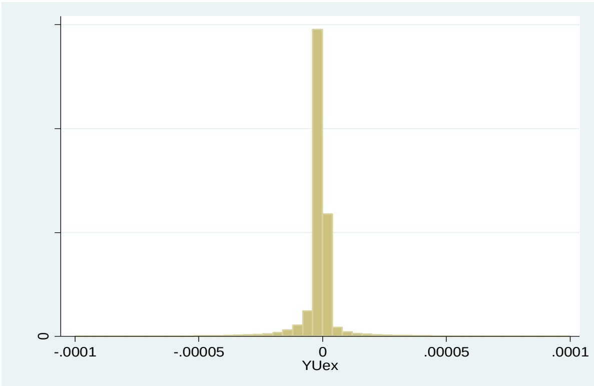
Source: Authors' calculations in STATA based on UN COMTRADE and RCA CEPII data.

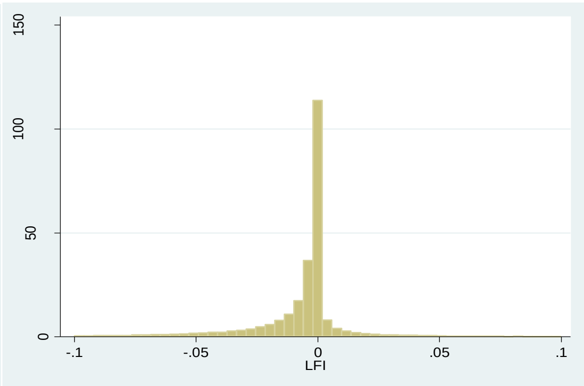




Laursen (1998) index

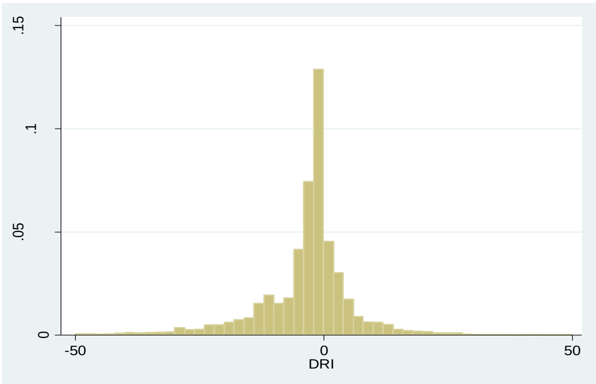
Hoen and Oosterhaven (2006) index

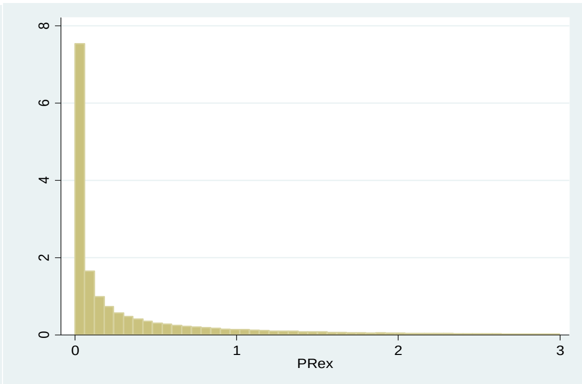




Yu et al. (2009) index

Lafay (1992) index





Donges and Riedel (1977) index

Proudman and Redding (2000) index

Fig. 8 (continued). Probability distribution for revealed comparative advantage indices

Source: Authors' calculations in STATA based on 2010 UN COMTRADE and RCA CEPII data.

Corresponding author:

Andrey Gnidchenko
Center for Macroeconomic Analysis and Short-Term Forecasting; Institute of Economic Forecasting, Russian Academy of Sciences; National Research University Higher School of Economics. Moscow, Russia
E-mail: agnidchenko@forecast.ru