BILATERAL TRADE FLOWS AND COMPARATIVE ADVANTAGE: 
DOES THE SIZE MATTER?1

MARTIN GRANČAY1 – TOMÁŠ DUDÁŠ2

1Faculty of Chemical and Food Technology, Slovak University of Technology, Bratislava, Slovakia 
E-mail: martin.grancay@stuba.sk

2Faculty of Economics and Entrepreneurship, Pan-European University, Bratislava, Slovakia

The paper applies a variant of the gravity model to test whether there is a positive link between the size of trade flows and the extent to which they follow the pattern of comparative advantage. Using UNCTAD’s 2016 trade data for every country in the world, and 255 merchandise items, we show that countries trading more with each other tend to follow the patterns of comparative advantages more than countries with smaller mutual trade flows. While smaller trade flows can be easily influenced by business decisions of individual companies or one-off trade contracts going against trade pattern predictions, this is not the case with larger flows. We also find signs that holding trade volume constant, more distant countries trade less than geographically proximate countries, in line with predictions from comparative advantage. The results are valid for the whole database of all country pairs in world trade, but the goodness of fit increases with the number of items these country pairs trade in. The paper is the first insight into the topic and can be expanded to a higher level of disaggregation and more variables in future research.

Keywords: international trade, gravity model, comparative advantage, trade volume

JEL-code: F10

1 We would like to express our sincere gratitude to the editors and two anonymous referees for their comments and useful suggestions on the previous version of this paper. The help of Jessie Krakosky with English-language editing is also greatly appreciated. As always, any remaining mistakes can be blamed on us.
1. INTRODUCTION

Two centuries have passed since Ricardo and Torrens discovered the principle of comparative advantage, and more than fifty years since Balassa developed its “revealed” version. While the former has become a staple of international economics, the latter is now probably the most widely-known (if not theoretically consistent) simple mathematical expression of specialization patterns in international trade. Numerous authors have used Balassa’s Revealed Comparative Index (RCA) and its derivatives to study global or local trade patterns. Papers on different industries (Bojnec – Ferto 2009; Han et al. 2009), regional aspects (Hoekman – Djankov 1997; Bojnec 2001; Worz 2005), alternative measures (Vollrath 1991; Hoen – Oosterhaven 2006; Yu et al. 2009), or statistical properties of the index (Hinloopen – Van Marrewijk 2001) abound. However, quite surprisingly, with the sole partial exception of Peterson (1988), to our knowledge there are no studies which would explore possible links between RCA and the volume of bilateral trade flows.

This paper follows a simple logic. Given the method of calculation of RCA as an ex-post numerical representation of a country’s trade patterns, each country’s total trade is necessarily fully in line with its RCA. That certainly does not mean that all the country’s trade flows are in line with its RCA. Some bilateral exports are small by volume. Some are much larger and constitute an important part of the country’s exports, sometimes even an important share of global exports. We expect to find that the latter flows tend to be more in line with RCA than the former – given the very specific method of calculating RCA, this is not mathematically self-evident. Wider flows normally include more commodities and more participants, and it can be expected that they should be governed by a principle of mutual benefit (in trade very often represented by comparative advantage); more so than narrower flows. Should this not be the case, one could argue that comparative advantages are an irrelevant concept in actual trade. Hence, the main hypothesis of the present paper is that there is a positive link between the size of trade flows and the extent to which they follow the pattern of comparative advantage.

The rest of the paper is structured as follows: in the next section a brief literature review is offered focusing on RCA. Section three details data and the methodology of the research. Section four presents the results and offers discussion. Finally, the last section concludes.
2. LITERATURE REVIEW

The concept of Revealed Comparative Advantage surely does not need to be introduced at length. Originally presented in Balassa’s 1965 paper, it has opened a topic which has been prominent in the literature in the last five decades. The index can be calculated as

\[
RCA = \frac{\sum_i^c x_i}{\sum_c\sum_i^c x_i},
\]

where \(x\) is country \(c\)'s total exports of product \(i\). Simply said, it is ratio of a product’s share in a country’s exports relative to the product’s share in global exports, and it can take values from zero to infinity. Values larger than 1 mean the country has a relative comparative advantage in the product, and vice versa.

While \(RCA\) is not exactly a flawless indicator, it has numerous advantages: It is simple, relatively easy to calculate at any level of aggregation, and it generates possible cross-country and cross-sector rankings (Balance et al. 1987). As a result, dozens of economists have adopted it for multitudes of uses. At first, the index was mostly utilized to study countries’ specialization patterns. It has probably already been applied to every single country in the world – starting with major industrial countries (Balassa 1977) and economic giants like China (Yue – Hua 2002) or Korea (Lee 1995), and ending with smaller economies such as Lithuania (Saboniene 2009) or Pakistan (Abbas – Waheed 2017), as well as in larger global (OECD 2011), regional (Fojtikova 2016) and cross-regional studies (Obadi – Korcek 2016).

The index has also been enormously popular in analyzing industry-level trade (Hardwick – Dou 1998; Bojnec – Ferto 2009; Han et al. 2009; Balogh – Jambor 2017; Boschma et al. 2017) and technological specialization (Meliciani 2002; Malerba – Montobbio 2003). More unorthodox applications include researching the role of \(RCA\) in investment (Feliciano – Lipsey 2017), climate policy (Boehringer – Alexeeva-Talebi 2013; Boehringer et al. 2017), or in quantitative studies of scientific output (Wang 2016). Even “trade sophistication indicators”, like \(\text{EXPY}\) and \(\text{PRODY}\), were partially derived using the index (Hausmann et al. 2007). The level of popularity of \(RCA\) can be easily demonstrated by more than 600 papers touching on the topic in the Web of Science database alone (2018).

Peterson (1988) has offered the only study we are aware of which is partially related to the present paper. He “examined the factors which can cause a coun-
try’s share of group exports of a particular product to be a potentially unreliable \textit{ex post} indicator of revealed comparative advantage” (Peterson 1988: 364) focusing on international travel services. He searched for a correlation between export shares and \textit{RCA} using a similar approach to what will be introduced in the next section. However, his interest was limited to theoretical issues of measuring comparative advantage and did not study possible links between \textit{RCA} and export shares in general.

A wide branch of literature is devoted to the shortcomings of the index which are already well known. They have been summarized succinctly for example in De Benedictis and Tamberi (2001) or Hinloopen and Van Marrewijk (2001). They include asymmetry, upper bound variability, mean variability, and as a result inconsistency across industries and countries (Yeats 1985). Consequently, numerous authors have tried to come up with alternative indices to overcome these issues. They include simple derivatives such as Symmetric Revealed Comparative Advantage (Laursen 1998), Weighted Revealed Comparative Advantage (Proudman – Redding 2000), Additive Revealed Comparative Advantage (Hoen – Oosterhaven 2006) or Normalized Revealed Comparative Advantage (Yu et al. 2009), as well as more complicated theoretically consistent alternatives to the \textit{RCA} (Costinot et al. 2012), and gravity-based estimators (French 2017a; 2017b).

There is no general consensus as to which alternative indicator is preferable. None of the indices is perfect – they all have pros and cons depending on the use (Sanidas – Shin 2010). In the present paper we will use Symmetric \textit{RCA} as it has the advantage of symmetric distribution, normality of error terms, and is simple to calculate.

3. DATA AND METHODOLOGY

For the sake of simplicity, clarity and comparability, the paper is based on data compiled by the United Nations Conference on Trade and Development (2018). These include export and import matrices for 2016 by product based on the three-digit level of the SITC commodity classification, Revision 3 for each pair of trading partners in the world. Curacao, Holy See, San Marino and South Sudan are excluded due to unavailability of data. The final database comprises 222 countries and regions and 255 merchandise items. Non-zero data are available for more than 2.5 million cells in the exporter-importer-merchandise matrix. Export and import figures are used “as is”; net exports are computed the usual way.

The authors convert available data to percentages, and hence calculate the importance of each merchandise for each exporter, the share of each aggregate
country-to-country export flow on global exports ($SHARE_{\text{WORLD}}$) and its share in total exports of the exporter ($SHARE_{\text{EXPORTER}}$). The same is done for the share in total trade defined as exports plus imports ($SHARE^2_{\text{WORLD}}$ and $SHARE^2_{\text{EXPORTER}}$). Obviously, a few country pairs play an important role in world trade while the majority of others are negligible. The top 10 international export flows (out of a possible maximum of 49,062) constituted over 15% of global exports in 2016 (Table 1).

Table 1. Top 10 international export flows (2016)

| Export flow               | Share  | Export flow            | Share  |
|---------------------------|--------|------------------------|--------|
| 1. China → United States  | 2.508% | 6. United States → Canada | 1.625% |
| 2. Mexico → United States | 1.945% | 7. United States → Mexico | 1.442% |
| 3. China → Hong Kong SAR | 1.869% | 8. China → Japan        | 0.840% |
| 4. Hong Kong SAR → China  | 1.858% | 9. Netherlands → Germany | 0.821% |
| 5. Canada → United States | 1.809% | 10. South Korea → China | 0.810% |

Note: Share in total global exports.
Source: authors, based on UNCTAD (2018).

Next, Revealed Comparative Advantage ($RCA$) indices are computed using the standard method (Balassa 1965) as seen in the previous section. While there has been a lot of criticism of the measure and several new indices have been proposed in the last five decades, these tend to suffer from similar or even more serious issues than the original. We will therefore only perform Laursen’s (1998; 2015) adjustment to make the index symmetric, i.e. the final Symmetric Revealed Comparative Advantage ($SRCA$) index is calculated as $(RCA-1)/(RCA+1)$, with values ranging from –1 to 1.

In addition to $SRCA$, relative $SRCA$ is also calculated. This can be defined as the difference between exporter’s $SRCA$ and importer’s $SRCA$ for each exported item. Its values can range from 2 for the highest relative comparative advantage of the exporter to –2 for the highest disadvantage. Values are reported only if both countries export the item to at least one country in the world.

To see how well exports from country A to country B match their global comparative advantages, Pearson correlation coefficients are taken between commodity structure of exports from A to B and the array of A’s $RCA$s for each exported item. This is done for exports and net exports on the one side, and $SRCA$ and relative $SRCA$ on the other side. Hence four matrices (222*222 countries) of Pearson coefficients are obtained which we label $COR_{X\times SRCA}$, $COR_{X\times SRCA}^*$, $COR_{NX\times SRCA}$ and $COR_{NX\times SRCA}^*$. Quite obviously, the informative value of these correlation coefficients depends on the number of products exported. For countries trading only 2 or 3 products, the Pearson coefficients are necessarily more accidental.
than for countries trading a wide range of items. Therefore, separate matrices are derived based on the number of exported items (calculated from UNCTAD 2018 data). While the original \( \text{COR}_{X/SRCA} \) matrix has just over 26,000 non-zero cells, the number drops to 6,476 when only export flows with at least 100 items are considered.

These correlation coefficients show how each trade flow is in line with the countries’ comparative advantages. The higher the coefficient, the higher the fit. In a hypothetical situation with a Pearson coefficient of 1.00, trade follows the structure of comparative advantages perfectly, i.e. the array of exported items ordered by trade volume is the same as the array of exported items ordered by \( \text{RCA} \).

The closest to this have been Iraqi exports to the Netherlands with the \( \text{COR}_{X/SRCA} \) approximately equal to 0.93 (Table 2).

### Table 2. Top 10 export flows in line with the comparative advantages, minimum 100 items (2016)

| Export flow       | \( \text{COR}_{X/SRCA} \) | Share  | Export flow       | \( \text{COR}_{X/SRCA} \) | Share  |
|-------------------|--------------------------|--------|-------------------|--------------------------|--------|
| 1. Iraq → Netherlands | 0.9307                  | 3.92 % | 6. Qatar → Iran   | 0.7410                  | 0.01 % |
| 2. Brunei → China  | 0.8606                  | 4.56 % | 7. Kuwait → Morocco| 0.7113                  | 0.06 % |
| 3. Nigeria → Zimbabwe | 0.7730                  | 0.01 % | 8. Brunei → Singapore | 0.7016          | 2.52 % |
| 4. Vanuatu → B. Faso | 0.7641                  | 4.58 % | 9. Greenland → Denmark | 0.6979   | 85.96 % |
| 5. Qatar → France  | 0.7556                  | 0.49 % | 10. Kuwait → Belgium | 0.6903          | 0.10 % |

*Note:* Share represents the share in the exporter’s total exports.

*Source:* authors, based on UNCTAD (2018).

Note that while \( \text{RCA} \) indices are computed using actual trade data, this does not invalidate the approach taken here of drawing conclusions from comparing trade volumes with \( \text{RCA} \) on a bilateral level. Owing to very different methods of calculation, even if a country exported all of its products to one partner only, this would not mathematically lead to the \( \text{COR}_{X/SRCA} \) coefficient of this trade necessarily being equal to 1. The key word with this type of comparative advantage indices is “relative”.

Once all the necessary calculations have been completed we can proceed to verify the main hypothesis. Should there be a link between the size of trade flows and the extent to which they follow the pattern of comparative advantage, there must be a correlation between \( \text{SHARE}_{\text{WORLD}} \) or \( \text{SHARE}_{\text{EXPORTER}} \) on one side of the equation and the Pearson coefficients introduced in the previous paragraph. In addition, gravity-type OLS regression analyses with country dummies will be performed to assess the statistical significance of the relationship. This is based on the assumption that distance – measured in kilometers between economic centers (not necessarily capital cities) and taken from the CEPII’s well-known GeoDist database (2018) – between trading partners might influence the importance of
comparative advantage in bilateral trade. However, we will not develop a full-fledged gravity model comprising all the standard variables as it would be far beyond the scope of the paper.

There are two opposite expectations from including distance into the model. On the one hand, it could be argued that transportation costs (and hence distance) work as a factor discouraging trade, and as a result distant countries will mostly trade goods with high levels of advantage for both of them, i.e. conforming to $RCA$. Conversely, geographically proximate countries face low transportation costs and their trade will therefore include a wide range of products regardless of their comparative advantages. As a result, the distance variable should have a positive sign in our models.

In a different line of argumentation, it might be the case that, ceteris paribus, distant countries’ mutual trade patterns are mainly formed by intergovernmental treaties and business decisions of a small number of companies; more so than trade patterns of geographically proximate countries. These treaties and decisions do not necessarily follow the country-level $RCA$ logic. Proximate countries are also often members of regional integration blocs which eliminate barriers to mutual trade, theoretically leading to better compliance with comparative advantage. Consequently, the role of distance might be negative.

A word of caution needs to be uttered about including the distance variable into our models. Numerous economists have shown that gravity models are generally valid, i.e. trade volumes as well as trade shares are influenced (among other factors) by distance. It follows that in our models distance is already implicitly included in the $SHARE$ variables. However, adding a separate distance variable does not constitute a serious econometric problem and is vital to ascertain its role in influencing whether trade is $RCA$-compliant.

4. RESULTS AND DISCUSSION

The correlation matrix (Table 3) shows that in the majority of cases there is only a very weak relationship between the studied variables. In line with logical expectations, the strength of the relationship is growing with the growing number of products exported. The highest correlation coefficients have been recorded when analyzing the share of a product’s export flows in a country’s total exports (as opposed to share of flows in global exports) in trade lines with at least 100 or 150 products.

The low correlation coefficients appear to suggest that the link between the size of trade flows and the extent to which they follow the pattern of comparative advantage is weak or non-existent. Obviously, this will need to be thoroughly
In regression analyses. Before we proceed, a simple exercise can be performed to show how the correlation is distributed.

In Figure 1 the whole dataset consisting of export flows with at least 100 exported products is divided into deciles based on the global share of the flow (left chart) or by flow’s share in the exporter’s exports (right chart). On the vertical axis the average correlation coefficient \( \text{COR}_{X/SRCA} \) is provided for each decile, along with ±1 standard deviation intervals. It can be seen that the average correlation coefficient has a growing tendency, i.e. bigger export flows are more in line with the RCA patterns than smaller export flows. This relationship is clearly visible in global exports (left) and is almost perfect when studying country-level shares (right) – the value of \( \text{COR}_{X/SRCA} \) increases decile by decile from an average of 0.18 for the lowest 10 % to 0.35 for the highest 10 %. While the correlation is not strong, the increase is substantial.

Net export variables \( \text{COR}_{NX/SRCA} \) and \( \text{COR}_{XvSRCA} \) have not been reported so far, as the results of the correlation analyses are unsatisfactory and show no traces of a strong relationship. Regression models are no exception; therefore we will limit further discussion to the export variables \( \text{COR}_{X/SRCA} \) and \( \text{COR}_{XvSRCA} \).

### Table 3. Correlation matrix

| Number of products exported | All     | Min. 5 | Min. 20 | Min. 50 | Min. 100 | Min. 150 |
|----------------------------|---------|--------|---------|---------|----------|----------|
| \( \text{COR}_{X/SRCA} \)  |         |        |         |         |          |          |
| Export flow’s share in global exports | 0.0052  | 0.0073 | 0.0188  | 0.0392  | 0.0618   | 0.0749   |
| Export flow’s share in exporter’s exports | 0.0722  | 0.1025 | 0.1457  | 0.1868  | 0.2344   | 0.2474   |
| Ln Export flow’s share in global exports | 0.0891  | 0.0816 | 0.0690  | 0.1094  | 0.1640   | 0.1871   |
| Ln Export flow’s share in exporter’s exports | 0.1897  | 0.2395 | 0.3076  | 0.3797  | 0.4247   | 0.4167   |

|              | \( \text{COR}_{X/vSRCA} \) |         |        |         |          |          |
|----------------------------|-----------------------------|---------|--------|---------|----------|----------|
| Export flow’s share in global exports | -0.0046                   | -0.0070 | -0.0084 | -0.0025 | 0.0044   | 0.0071   |
| Export flow’s share in exporter’s exports | 0.0663                    | 0.0913  | 0.1250  | 0.1634  | 0.2008   | 0.2150   |
| Ln Export flow’s share in global exports | 0.0537                    | 0.0485  | 0.0187  | 0.0428  | 0.0685   | 0.0632   |
| Ln Export flow’s share in exporter’s exports | 0.1504                    | 0.1998  | 0.2520  | 0.3098  | 0.3382   | 0.3163   |

Note: Pearson correlation coefficients between (1) indicators in rows and (2) correlation coefficients between SRCA and commodity structure of trade for each international export flow. The columns indicate that input data include only export flows with a certain number of products based on SITC rev. 3 classification. See section 3 of the text for a thorough explanation.

Source: authors, based on UNCTAD (2018).
The main regression models are reported in Table 4. All the regression coefficients are statistically significant. The independent variables of interest (SHARE) have a positive sign, which means that, other things being equal, increasing the importance of an export flow leads to its higher compliance with the predictions of comparative advantages. While small export flows do not really have to be

Table 4. Regression models (OLS with country dummies), dependent variable COR$_{X/SRCA}$

|                | Exported products > 50 | Exported products > 100 |
|----------------|------------------------|-------------------------|
| (1)            | (2)                    | (3)                     |
| SHARE$_{WORLD}$| 16.302***              | 13.188***               |
|                | (1.894)                | (1.645)                 |
| SHARE$_{EXPORTER}$ | 0.301***             | 0.313***               |
|                | (0.026)                | (0.028)                 |
| SHARE$_{WORLD}$ | -                     | 19.349***              |
|                | -                     | (2.039)                |
| SHARE$_{EXPORTER}$ | -                  | 0.350***               |
|                | -                     | (0.028)                |
| Ln DISTANCE    | -0.008***              | -0.009***              |
|                | (0.001)                | (0.001)                |
| Constant       | 0.466***               | 0.418***               |
|                | (0.032)                | (0.037)                |
| R²             | 0.332                  | 0.327                  |
| n              | 10539                  | 6476                   |

Note: Standard errors in parentheses. Significance: *** p < 0.01, ** p < 0.05, * p < 0.1. SHARE variables based on exports, SHARE2 variables based on total trade (X+M).

Source: authors, based on UNCTAD (2018).
aligned with comparative advantages and can be easily influenced by factors such as business decisions of individual companies or one-off trade contracts, larger export flows seldom go against RCA. In words inspired by Adam Smith, it appears that the countries are led by an invisible hand and by trading the goods in accordance with their own comparative advantage they also promote the general good.

Very interesting results can be seen when looking at the distance variable. Without exception, it is negative and statistically significant. Previously, we have noted that the expectations from including distance into the model are uncertain – there are reasons to believe the sign might be positive, but there are also reasons to believe it might be negative. The regression models suggest that the latter reasons prevail and that, holding trade volume constant, more distant countries trade less, in line with predictions from comparative advantages. This is an important conclusion that might be worth pursuing in future research.

Low values of $R^2$ (reaching 0.337 as a maximum) indicate that the models are far from explaining all the variability in the data around its mean. However, they do not make the models irrelevant, as the statistical significance of independent variables shows that they indeed influence the dependent variable. Moreover, the signs of the coefficients are not unexpected, and they have an economic interpretation that can be readily explained.

Table 5. Regression models (OLS without dummies), dependent variable COR$_{X/SRCA}^{\text{Exported products > 50}}$

|                      | Exported products > 50 | Exported products > 100 |
|----------------------|-------------------------|-------------------------|
|                      | (9)                     | (10)                    | (11)                     | (12)                     | (13)                     | (14)                     | (15)                     | (16)                     |
| SHARE$_{WORLD}$      | 7.019***                | –                        | –                        | 7.932***                | –                        | –                        | –                        | –                        |
|                      | (2.214)                 |                          |                          | (1.886)                 |                          |                          |                          |                          |
| SHARE$_{EXPORTER}$   | –                       | 0.503***                | –                        | –                        | 0.475***                | –                        | –                        | –                        |
|                      | (0.029)                 |                          |                          |                          | (0.028)                 |                          |                          |                          |
| SHARE$_{2WORLD}$     | –                       | –                       | 7.657***                | –                        | –                        | 8.708***                | –                        | –                        |
|                      |                          |                          | (2.380)                 |                          |                          | (2.028)                 |                          |                          |
| SHARE$_{2EXPORTER}$  | –                       | –                       | –                       | 0.596***                | –                        | –                        | –                        | 0.522***                |
|                      |                          |                          |                          | (0.030)                 |                          |                          |                          | (0.028)                 |
| Ln DISTANCE          | –0.016***               | –0.012***               | –0.016***               | –0.011***               | –0.018***               | –0.014***               | –0.018***               | –0.014***               |
|                      | (0.001)                 | (0.001)                 | (0.001)                 | (0.001)                 | (0.001)                 | (0.001)                 | (0.001)                 | (0.001)                 |
| Constant             | 0.405***                | 0.363***                | 0.405***                | 0.359***                | 0.417***                | 0.370***                | 0.417***                | 0.370***                |
|                      | (0.012)                 | (0.012)                 | (0.012)                 | (0.012)                 | (0.012)                 | (0.012)                 | (0.012)                 | (0.012)                 |
| $R^2$                | 0.014                   | 0.041                   | 0.014                   | 0.049                   | 0.029                   | 0.068                   | 0.029                   | 0.076                   |
| n                    | 10539                   | 6476                    |

Note: Standard errors in parentheses. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. SHARE variables based on exports, SHARE2 variables based on total trade (X+M).

Source: authors, based on UNCTAD (2018).

Society and Economy 41 (2019)
The results remain unchanged even when country dummies are dropped from the model (Table 5); yet the $R^2$ drops to almost zero, as could have been expected.

Apart from those displayed in Tables 4 and 5, dozens of other models have been estimated. Table 6 shows alternative regressions when relative SRCA is used to construct the dependent variable (hence the dependent variable is $COR_{X/RCA}$). The independent variables of interest have the same sign and similar value of parameters, but an important difference can be seen in all the models compared to the previous ones – distance becomes positive, yet is still statistically significant. The same is true when country dummies are dropped, when other variables are log-transformed or when we use dependent variables with a different limit on the number of exported items (not reported here). This equals a logical contradiction: on the one hand, we have models where distance positively influences $RCA$-compliance; on the other hand, we have models where distance affects $RCA$-compliance negatively. Obviously, both conclusions cannot hold simultaneously. Dropping the distance variable does not solve the issue either as the results remain virtually unchanged.

The solution to this “paradox” might rest in the fact that relative $SRCA$, which is used as the dependent variable in Table 6, is an indicator computed by subtracting the importer’s $SRCA$ from the exporter’s $SRCA$. For each country, $SRCA$ is calculated from a different base and using them in one equation might lead to

### Table 6. Regression models (OLS with country dummies), dependent variable $COR_{X/RCA}$

|                  | Exported products > 50 | Exported products > 100 |
|------------------|------------------------|-------------------------|
|                  | (17)                   | (18)                    | (19) | (20) | (21) | (22) | (23) | (24) |
| SHARE$_{WORLD}$  | 14.304***              | –                       | –    | –    | 11.696*** | –    | –    | –    |
|                  | (2.303)                |                         |      |      | (2.053) |      |      |      |
| SHARE$_{EXPORTER}$ | –                      | 0.376***               | –    | –    | –    | 0.391*** | –    | –    |
|                  |                        | (0.032)                |      |      |      | (0.035) |      |      |
| SHARE$_{WORLD}$  | –                      | –                      | 17.839*** | –    | –    | –    | 14.877*** | –    |
|                  |                        |                        | (2.480) |      |      |      | (2.213) |      |
| SHARE$_{EXPORTER}$ | –                      | –                      | –    | –    | 0.451*** | –    | –    | –    |
|                  |                        |                        |      |      | (0.033) |      |      |      |
| Ln DISTANCE      | 0.015***               | 0.017***               | 0.016*** | 0.017*** | 0.019*** | 0.020*** | 0.019*** | 0.020*** |
|                  | (0.002)                | (0.002)                | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Constant         | 0.128***               | 0.100***               | 0.126*** | 0.095** | 0.095** | 0.058 | 0.093** | 0.051 |
|                  | (0.039)                | (0.038)                | (0.039) | (0.038) | (0.046) | (0.046) | (0.046) | (0.046) |
| $R^2$            | 0.248                  | 0.256                  | 0.249 | 0.259 | 0.254 | 0.265 | 0.256 | 0.268 |
| n                | 10506                  |                         | 6465 |      |      |      |      |      |

Note: Standard errors in parentheses. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

SHARE variables based on exports, SHARE2 variables based on total trade (X+M).

Source: authors, based on UNCTAD (2018).
problems with comparability. This is a well-known criticism of revealed comparative advantage indices (Sanidas – Shin 2010) and hence relative SRCA indicators should not be used. (It is interesting to emphasize again that even in case of these regressions, the variables of interest have expected parameter signs and values.) The comparability problem is minimized in case of absolute SRCA, by calculating each country’s SRCA using only its own indicators, and by applying country dummies.

To further subject our results to robustness checks, we performed separate regression analyses by continent. Models 25 to 34 (Table 7) only include the exports of the respective continent’s to the whole world. All the parameters have the expected signs and values, even though in a few cases they are not statistically significant. The models for Asia, Africa and the Americas are fully in line with our main models; Europe and the Pacific region, however, show issues with statistical significance. In case of the Pacific this is not entirely surprising as it is by far the smallest region included in the analysis and is dominated by a single

Table 7. Regression models (OLS with country dummies) by exporter’s continent, dependent variable COR_{X/SRCA}

| Exported products > 50 | (25) Europe  | (26) Asia  | (27) Africa  | (28) Americas  | (29) Pacific  |
|-----------------------|--------------|------------|--------------|----------------|--------------|
| SHARE_{EXPORTER}      | 0.337***     | 0.102*     | 0.099        | 0.137*         | 0.024        |
|                       | (0.069)      | (0.058)    | (0.075)      | (0.074)        | (0.144)      |
| Ln DISTANCE           | -0.002       | -0.026***  | -0.033***    | -0.034***      | -0.027       |
|                       | (0.005)      | (0.004)    | (0.006)      | (0.006)        | (0.028)      |
| Constant              | 0.193***     | 0.391***   | 0.606***     | 0.622***       | 0.207        |
|                       | (0.053)      | (0.052)    | (0.058)      | (0.092)        | (0.300)      |
| R²                    | 0.269        | 0.244      | 0.236        | 0.241          | 0.504        |
| n                     | 4110         | 3279       | 1362         | 1500           | 288          |

| Exported products > 50 | (30) Europe  | (31) Asia  | (32) Africa  | (33) Americas  | (34) Pacific  |
|-----------------------|--------------|------------|--------------|----------------|--------------|
| SHARE2_{EXPORTER}     | 0.374***     | 0.122*     | 0.198**      | 0.051          | 0.654***     |
|                       | (0.073)      | (0.069)    | (0.088)      | (0.099)        | (0.200)      |
| Ln DISTANCE           | -0.002       | -0.025***  | -0.031***    | -0.036***      | -0.016       |
|                       | (0.005)      | (0.004)    | (0.006)      | (0.006)        | (0.027)      |
| Constant              | 0.189***     | 0.390***   | 0.592***     | 0.637***       | 0.105        |
|                       | (0.053)      | (0.052)    | (0.059)      | (0.093)        | (0.285)      |
| R²                    | 0.269        | 0.244      | 0.239        | 0.239          | 0.533        |
| n                     | 4110         | 3279       | 1362         | 1500           | 288          |

Note: Standard errors in parentheses. Significance: *** p < 0.01, ** p < 0.05, * p < 0.1.
SHARE variables based on exports, SHARE2 variables based on total trade (X+M).

Source: authors, based on UNCTAD (2018).
country – Australia. Indeed, the literature does not offer too many examples of papers successfully applying the gravity model of trade to the Pacific (e.g. Maiti – Kumar 2016). The European case has likely been caused by the single European market, which outweighs the role of distance in RCA compliance.

In what could be seen as a setback, models with \textit{SHARE\_WORLD} variables did not lead to any statistically significant results in robustness checks (not reported here). This equals a logical contradiction: on the one hand, export flows which have higher shares in a country’s total exports are more in line with \textit{RCA}; on the other hand, export flows which have higher shares in global exports appear not to be related with \textit{RCA}. Obviously, both shares should move in the same direction. However, the important difference consists in the fact that while \textit{SHARE\_EXPORTER} variables are calculated from a different total for each country’s exports, \textit{SHARE\_WORLD} variables are calculated from the same base, i.e. their sum gives total global trade, or less (if the number of flows in the model is limited by number of exported products). This means that country dummies have a vital role in models featuring the country-share variable (translating into a fixed-effect model) but are not necessary in the latter case. When robustness checks with \textit{SHARE\_WORLD} variables are re-run without country dummies, the results are similar in terms of signs and statistical significance to Table 7.

As is the case with every single research using statistical methods, the analysis is not straightforward and there are various issues that need to be considered. The most prominent issue in gravity models is the problem of zeroes, normally solved by Heckman correction or other alternative estimation methods. In our case, approximately 42\% of the data lines included a zero, i.e. there was no mutual trade between the countries. We chose to drop these data. However, this does not constitute a problem; “zeroes” bear no value for the present research as it is only interested in existing trade and does not aim to explain why countries do not trade with each other. It only tries to show how trading countries trade with each other. Henceforth we were able to drop the zeroes without the usual negative statistical consequences.

In addition to dropping the zeroes, we have deliberately put a lower limit on the number of products the countries trade with (50 or 100) to avoid very narrow trade flows where the values of correlation variables can be rather accidental. This means that even more lines are excluded from the database. Obviously, the excluded data are not randomly distributed and hence their exclusion influences the results of our analysis. Therefore, the results should not be interpreted generally, but have to be understood with care, applying only to bilateral trade flows which are “sufficiently diversified”. However, the results are valid even when all trading country pairs are considered; it is just the R-squared which becomes lower (see Table 8).
A possible variant of the present research would be to use the Spearman correlation coefficient instead of Pearson. The models (not reported here) are a bit more computation intensive, but the results remain unchanged in sign and statistical significance, and the coefficient of determination generally rises.

### Table 8. Regression models (OLS with country dummies) for different number of exported products, dependent variable COR$_{X/RCA}$

| SHARE$_{EXPORTER}$ | (35) > 20 | (36) > 10 | (37) > 5 | (38) > 3 | (39) All flows |
|-------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| (0.030) | 0.363*** | 0.405*** | 0.489*** | 0.538*** | 0.574*** |
| Ln DISTANCE | (0.001) | -0.005*** | -0.004** | -0.008*** | -0.011*** | -0.013*** |
| Constant | (0.027) | 0.466*** | 0.462*** | 0.452*** | 0.503*** | 0.475*** |
| | (0.027) | (0.032) | (0.032) | (0.035) | (0.042) | (0.042) |
| R$^2$ | 0.259 | 0.204 | 0.146 | 0.109 | 0.074 |
| n | 15590 | 19045 | 22362 | 24498 | 26076 |

| SHARE$_2{EXPORTER}$ | (40) > 20 | (41) > 10 | (42) > 5 | (43) > 3 | (44) All flows |
|-------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| (0.032) | 0.389*** | 0.433*** | 0.513*** | 0.562*** | 0.604*** |
| Ln DISTANCE | (0.001) | -0.005*** | -0.004** | -0.008*** | -0.011*** | -0.013*** |
| Constant | (0.027) | 0.466*** | 0.462*** | 0.452*** | 0.503*** | 0.474*** |
| | (0.027) | (0.032) | (0.032) | (0.035) | (0.042) | (0.042) |
| R$^2$ | 0.259 | 0.204 | 0.146 | 0.108 | 0.074 |
| n | 15590 | 19045 | 22362 | 24498 | 26076 |

Note: Standard errors in parentheses. Significance: *** p < 0.01, ** p < 0.05, * p < 0.1.
SHARE variables based on exports, SHARE2 variables based on total trade (X+M).

Source: authors, based on UNCTAD (2018).

5. CONCLUSIONS

The study of bilateral trade and its compliance with the predictions of comparative advantages based on the trade flow size has generated some interesting results:

- There is a positive link between the size of trade flows and the extent to which they follow the pattern of comparative advantages. Larger country-to-country export flows tend to be more in line with $RCA$ than smaller export flows.
Holding trade volume constant, more distant countries trade, less in line with comparative advantage predictions, than geographically proximate countries.

The results are valid for the whole database of all country pairs in world trade, but the goodness of fit increases with the number of items these country pairs trade with.

The first result might appear self-evident (due to the fact that $RCA$ indices are computed ex post from actual trade data, and the present paper uses $RCA$ indices in correlations with actual trade data), but mathematically this is not the case due to substantial differences in calculation. The finding means that wider trade flows can be considered more “predictable” in their composition than narrower ones, in the sense that they are more likely to conform to the expectations of comparative advantages. This fully complies with economic logic.

Another interesting finding regards distance. In usual gravity models, its negative role is interpreted as a proof that countries lying closer to each other trade more with each other. In our case the negative role means that countries lying closer to each other follow comparative advantage patterns in their mutual trade more than distant countries. This is opposite to what might have been expected. However, a combination of factors, such as regional integration, intergovernmental treaties or business decisions of a small number of companies can help explain the result.

The research can be further expanded and enhanced in multiple ways. A higher level of disaggregation can be used to calculate $RCA$, possibly using World Bank’s HS 6-digit classification. Alternative measures of comparative advantage can be explored, such as French’s gravity-based estimators (2017a; 2017b). The model can perhaps also be enlarged with other independent variables to achieve a higher goodness of fit. Any future research will necessarily be significantly more computation intensive.

REFERENCES

Abbas, S. – Waheed, A. (2017): Trade Competitiveness of Pakistan: Evidence from the Revealed Comparative Advantage Approach. Competitiveness Review 27(5): 462–475.

Balance, R. H. – Forstner, H. – Murray, T. (1987): Consistency Tests of Alternative Measures of Comparative Advantage. Review of Economics and Statistics 69(1): 157–161.

Balassa, B. (1965): Trade Liberalisation and “Revealed” Comparative Advantage. The Manchester School 33(2): 99–123.

Balassa, B. (1977): ‘Revealed’ Comparative Advantage Revisited: An Analysis of Relative Export Shares of the Industrial Countries, 1953–1971. The Manchester School 45(4): 327–344.

Society and Economy 41 (2019)
Balogh, J. M. – Jambor, A. (2017): The Global Competitiveness of European Wine Producers. *British Food Journal* 119(9): 2076–2088.

Boehringer, C. – Alexeeva-Talebi, V. (2013): Unilateral Climate Policy and Competitiveness: Economic Implications of Differential Emission Pricing. *World Economy* 36(2): 121–154.

Boehringer, C. – Garcia-Muros, X. – Cazcarro, I. – Arto, I. (2017): The Efficiency Cost of Protective Measures in Climate Policy. *Energy Policy* 104(1): 446–454.

Bojnec, S. (2001): Trade and Revealed Comparative Advantage Measures – Regional and Central and East European Agricultural Trade. *Eastern European Economics* 39(1): 72–98.

Bojnec, S. – Ferto, I. (2009): Agro-Food Trade Competitiveness of Central European and Balkan Countries. *Food Policy* 34(5): 417–425.

Boschma, R. – Martin, V. – Minondo, A. (2017): Neighbour Regions as the Source of New Industries. *Papers in Regional Science* 96(2): 227–246.

CEPII (2018): GeoDist. http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=6, accessed 10/06/2018.

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.

De Benedictis, L. – Tamberi, M. (2001): A Note on the Balassa Index of Revealed Comparative Advantage. *UNIVPM Working Paper* No. 158.

Feliciano, Z. M. – Lipsey, R.E. (2017): Foreign Entry into US Manufacturing by Takeovers and the Creation of New Firms. *Eastern Economic Journal* 43(1): 1–16.

Fojtikova, L. (2016): Trends in the Revealed Comparative Advantages of the EU Member States. *Economic Annals-XXI* 161(9-10): 7–11.

French, S. (2017a): A Gravity-Based Revealed Comparative Advantage Estimator. *UNSW Business School Research Paper* No. 2017 ECON 05.

French, S. (2017b): Revealed Comparative Advantage: What Is It Good for? *Journal of International Economics* 106(1): 83–103.

Han, X. – Wen, Y. – Kant, S. (2009): The Global Competitiveness of the Chinese Wooden Furniture Industry. *Forest Policy and Economics* 11(8): 561–569.

Hardwick, P. – Dou, W. (1998): The Competitiveness of EU Insurance Industries. *Service Industries Journal* 18(1): 39–53.

Hausmann, R. – Hwang, J. – Rodrik, D. (2007): What You Export Matters. *Journal of Economic Growth* 12(1): 1–25.

Hinloopen, J. – Van Marrewijk, C. (2001): On the Empirical Distribution of the Balassa Index. *Weltwirtschaftliches Archiv* 137(1): 1–35.

Hoekman, B. – Djankov, S. (1997): Determinants of the Export Structure of Countries in Central and Eastern Europe. *World Bank Economic Review* 11(3): 471–487.

Hoen, A. R. – Oosterhaven, J. (2006): On the Measurement of Comparative Advantage. *Annals of Regional Science* 40(3): 677–691.

Laursen, K. (1998): Revealed Comparative Advantage and the Alternatives as Measures of International Specialisation. *DRUID Working Paper* No. 98–30.

Laursen, K. (2015): Revealed Comparative Advantage and the Alternatives as Measures of International Specialization. *Eurasian Business Review* 5(1): 99–115.

Lee, J. (1995): Comparative Advantage in Manufacturing as a Determinant of Industrialization – The Korean Case. *World Development* 23(7): 1195–1214.

Maiti, D. – Kumar, S. (2016): Regional Agreements, Trade Cost and Flows in the Pacific. *Economia politica* 33(2): 181–191.
Malerba, F. – Montobbio, F. (2003): Exploring Factors Affecting International Technological Specialization: The Role of Knowledge Flows and the Structure of Innovative Activity. *Journal of Evolutionary Economics* 13(4): 411–434.

Meliciani, V. (2002): The Impact of Technological Specialisation on National Performance in a Balance-of-Payments-Constrained Growth Model. *Structural Change and Economic Dynamics* 13(1): 101–118.

Obadi, S. M. – Korcek, M. (2016): The “Revealed” Comparative Advantage and Competitiveness of the EU’s International Trade vis à vis the USA. *Ekonomicky casopis* 64(5): 397–422.

OECD (2011): *Globalisation, Comparative Advantage and the Changing Dynamics of Trade*. http://www.oecd.org/publications/globalisation-comparative-advantage-and-the-changing-dynamics-of-trade-9789264113084-en.htm, accessed 16/10/2018.

Peterson, J. (1988): Export Shares and Revealed Comparative Advantage. A Study of International Travel. *Applied Economics* 20(3): 351–365.

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

Saboniene, A. (2009): Lithuanian Export Competitiveness: Comparison with Other Baltic States. *Inzinerine ekonominia* 62(2): 49–57.

Sanidas, E. – Shin, Y. (2010): *Comparison of Revealed Comparative Advantage Indices with Application to Trade Tendencies of East Asian Countries*. Presented at the 9th Korea and the World Economy Conference. http://www.akes.or.kr/eng/papers%282010%29/24.full.pdf, accessed 10/10/2018.

UNCTAD (2018): *International Trade in Goods and Services*. UNCTADstat. http://unctadstat.unctad.org, accessed 8/10/2018.

Vollrath, T. L. (1991): A Theoretical Evaluation of Alternative Trade Intensity Measures of Revealed Comparative Advantage. *Weltwirtschaftliches Archiv* 127(2): 265–280.

Wang, L. (2016): The Structure and Comparative Advantages of China’s Scientific Research: Quantitative and Qualitative Perspectives. *Scientometrics* 106(1): 435–452.

Worz, J. (2005): Dynamics of Trade Specialization in Developed and Less Developed Countries. *Emerging Markets Finance and Trade* 41(3): 92–111.

Yeats, A. J. (1985): On the Appropriate Interpretation of the Revealed Comparative Advantage Index: Implications of a Methodology Based on Industry Sector Analysis. *Weltwirtschaftliches Archiv* 121(1): 61–73.

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

Yue, C. J. – Hua, P. (2002): Does Comparative Advantage Explains Export Patterns in China? *China Economic Review* 13(2-3): 276–296.

**Open Access.** This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (https://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted use, distribution, and reproduction in any medium for non-commercial purposes, provided the original author and source are credited, a link to the CC License is provided, and changes – if any – are indicated.