Scientometrics
https://doi.org/10.1007/s11192-021-04251-z

Revealed comparative advantages in academic publishing of "old" and "new" European Union Member States 1998–2018

Martin Grančay[1] · **Tomáš Dudáš**[2] · **Ladislav Mura**[2]

Received: 2 September 2019 / Accepted: 16 December 2021
© Akadémiai Kiadó, Budapest, Hungary 2021

Abstract
The paper uses a variant of the Revealed Comparative Advantage index to examine comparative advantages of the European Union (EU) countries in 254 research areas and their development from 1998 to 2018, focusing mostly on the split into the old and new EU members and their convergence. The results confirm that there has been convergence inside the groups as well as between the groups. EU researchers from different countries now publish papers in more similar research areas and the countries' research-area mix is more coherent than before. The EU has increased its comparative advantage in the majority of research areas in the field of Science and in almost all research areas in Social Sciences and Arts and Humanities in the last two decades. On average, the growth has been faster in new members than in old members. Science is still the dominant source of comparative advantage for the bloc, but Social Sciences and Arts and Humanities are set to challenge its lead.

Keywords Academic publishing · Web of science · Comparative advantage · RCA · European Union

JEL Classification I20 · I23

Mathematical Subject Classification 00–02

Introduction

For nearly two decades now the European Union has followed the goal of trying to "become the most competitive and dynamic knowledge-based economy in the world" (European Commission, 2000). It has set out to establish a European Area of Research and Innovation, harvesting the synergies of individual national research systems working

✉ Martin Grančay
martin.grancay@stuba.sk

[1] Faculty of Chemical and Food Technology, Slovak University of Technology in Bratislava, Radlinského 9, 81237 Bratislava, Slovak Republic

[2] Faculty of Economics and Entrepreneurship, Pan-European University, Tematínska 10, 85105 Bratislava, Slovak Republic

Published online: 10 February 2022

together. Numerous measures were taken to stimulate a catch-up process in countries which were soon to become new members of the bloc. Increased mobility of researchers via Erasmus and COST, or a separate "Widening and Strengthening" part of the framework programs are just the most well-known examples. Overall, the competitiveness strategy has had mixed success (Rogge, 2019), but studies suggest that an internal market for research has, at least partially, been achieved (Makkonen & Mitze, 2016).

While multiple papers have tried to assess the development of EU's research competitiveness (e.g. Leydesdorff, 2000; Pastor et al., 2015), only a handful have studied the intra-EU convergence of research between the "old"[1] and the "new"[2] member states. This is where our contribution finds its place in literature. The goal of the present paper is to quantify the scientific comparative advantages of the EU countries across various research areas and assess their development over time with a special focus on possible convergence between the old and the new member states. Such a research has a double importance. First, knowing advantages will allow policy-makers and research institutions to take measures to either further strengthen them or to support other desirable research areas with a current disadvantage. Second, measuring comparative advantage convergence will be an additional pebble in the mosaic of testing the "internal market for research" hypothesis. As will be seen later, theoretical arguments can be made for both intra-EU comparative advantage convergence and divergence.

The paper uses a variant of the Revealed Comparative Advantage index, applied originally in economics (Balassa, 1965) and introduced to scientometrics by Lattimore and Revesz (1996). The concept is mathematically similar to the more widely used activity index (Rousseau & Yang, 2012).

The paper is divided into seven sections adhering to the usual structure. Introduction is followed by two sections offering literature review of the revealed comparative advantages concept and research convergence in the European Union. Section ``Data and methodology'' focuses on data and methodology. The results are presented and discussed in sections ``The development of RCA in the European Union'' and ``Discussion: convergence between the old and new countries''. Finally, the last section concludes and provides ideas for further research.

Literature review: revealed comparative advantages

Revealed comparative advantages is a concept known mostly from international economics, but it has also found use in an array of other fields. Its original version was developed by Balassa (1965) trying to quantify the extent of comparative advantage the countries had in exporting individual products. The index can be calculated as

$$RCA_{ic} = \frac{\frac{x_i}{\sum_1^i x_i}}{\frac{\sum_1^c x_i^c}{\sum_1^c \sum_1^i x_i^c}}$$

[1] Countries which entered into the EU before 2004: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden and the UK. As the paper is based on data up to 2018, the UK is still considered a member of the EU.

[2] Countries which have entered into the EU since 2004: Bulgaria, Croatia, Cyprus, Czechia, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Romania, Slovakia and Slovenia.

where x_i^c is export of product i by country c. The numerator can be described as a product's share on country's total exports. The denominator is the product's share on global exports. The index takes values from zero to infinity, the crucial point being 1 – values larger than 1 mean the country has a revealed comparative advantage in the product, while values lower than 1 mean the country has a revealed comparative disadvantage in the product.

Despite its wide use, the acceptance of RCA is far from universal. Numerous shortcomings of the index have been identified, including but not limited to asymmetry, variability and inconsistency (Yeats, 1985). As a result, alternative indices have been proposed, such as Symmetric Revealed Comparative Advantage (Laursen, 2015), Weighted Revealed Comparative Advantage (Proudman & Redding, 2000), Additive Revealed Comparative Advantage (Hoen & Oosterhaven, 2006) or Normalized Revealed Comparative Advantage (Yu et al., 2009). These will be discussed in detail in ``Data and methodology'' section.

The RCA index has enjoyed considerable popularity in economics due to its simplicity, easy calculation and interpretation (Balance et al., 1987). Apart from its usual role in international trade (OECD, 2011; Yue & Hua, 2002, and hundreds of other studies) it has been used to analyze investment (Feliciano & Lipsey, 2017) or technological specialization (Malerba & Montobbio, 2003). Applications outside of its field of origin include climate policy (Boehringer et al., 2017), historical studies of technological innovations (Streb et al., 2006) as well as information science (Radosevic & Yoruk, 2014).

The use of RCA in scientometrics was introduced by Lattimore and Revesz (1996). They used numerous indicators to analyze the performance of Australian science by scientific field. They applied the usual RCA logics from economics implicating that "a country has a revealed comparative advantage in a field if the share of that country's papers or citations in that field is much higher than the world share of papers or citations in that field" (p. 9). They enumerated the fields where Australia had a comparative advantage, but also made an interesting conclusion that only a weak positive correlation existed between comparative advantage in a field and quality.

Since Lattimore and Revesz (1996), a few papers borrowed RCA for scientometric analyses. Chuang et al. (2010) conducted a study of 26 mostly developed nations, calculated their RCA in 24 scientific fields and applied clustering methods to observe and compare scientific performance of the nations. Harzing and Giroud (2014) published an analogous research based on a set of 34 countries and 21 scientific fields. Similar to the previous paper, the choice of the countries was somewhat random. Both studies came to some interesting conclusions regarding country clusters with different RCA patterns, and tried to explain the causes of these differences. Lee et al.'s (2011) database consisted of 25 countries. They used RCA for clustering purposes in their study of correlation between research and economic growth.

Arguably the most encompassing research effort applying RCA was undertaken by Radosevic & Yoruk (2014) who explored the changes in science-output patterns in several world regions for the period 1981–2011. Their RCA statistics was based on 21 "broad disciplines" and included publications as well as citations. On the whole, they concluded that the center of the scientific gravity has been shifting towards the Asia–Pacific, but also noticed some subtler changes in Europe, namely sings of convergence of the post-socialist countries of Central and Eastern Europe with the rest of the world and their divergence from declining science output of other former USSR countries.

Other papers employed RCA in a simple one-country setting. Wang (2016) used the index to observe the rising quantity and quality of China's scientific output between 2005 and 2013, thus confirming an earlier report by Huang and Tian (2014) which had applied a similar method. Wang and Wang (2017) were interested in scientific collaborations of

Chinese and European authors. They noticed that these collaborations lead to an improvement in the RCA scores of China, but not necessarily in the RCA scores of the European Union countries. Horta (2018) focused on Hong Kong and Singapore and found out their "scientific wealth" has been decreasing recently. Lee et al. (2012) studied Taiwan, but were interested in the bibliometric analysis of agriculture-related fields only. Most recently, Horta and Shen (2019) used RCA to assess the evolution and current situation of the Chinese research system. In general, a half of the papers applying RCA in scientometrics focus on research in Chinese-speaking regions.

Literature review: research convergence in the European Union

Dozens of research papers have been published on different aspects of research output of the European Union. Some of them offer a complex scientometric analysis (Pastor et al., 2015), others compare the bloc's scientific performance with USA (Albarran et al., 2010; Herranz & Ruiz-Castillo, 2013), China (Basu et al., 2018) or both (Leydesdorff et al., 2014), yet others focus on partial regional-level (Murashova & Loginova, 2017) or research-area-level issues (Micheli et al., 2011). Interestingly, despite the fact that in other fields of science, social sciences and humanities there have been hundreds of research attempts to study convergence between the "old" and the "new" EU members, i.e. to quantify whether there has been any catching-up process related to the new members' entry into the EU, this type of literature is scarce in the field of scientometrics.

Three years after the largest enlargement of the EU Glänzel and Schlemmer (2007) published the first study offering a glimpse into how the new EU members from Central and Eastern Europe fare on the EU scientific market. They analyzed changes in authorship structure, i.e. whether the old "socialist" pattern of high share of papers published by academies of sciences still prevails, or whether there has been a move closer to what they call the "Western mode" – a stronger collaboration with universities and industry. Even though all their data is pre-enlargement, they show that the latter is true and conclude that slow process of European homogenization and convergence has been taking place.

Teodorescu and Andrei (2011) focused on international co-authorship of papers and their impact on citations. They identified an increase in the share of intra-EU research collaborations and highlighted how joint publication with scholars from old EU members raise the citation impact of new countries. However, they did not calculate any specific convergence indicators.

Cecere and Corrocher (2013) studied international projects and have shown that R&D cooperation among old members is stronger and much more frequent than cooperation among new members or between the two groups. Contrary to the previously mentioned papers, they were not able to find any traces of convergence.

Pastor et al. (2015) proposed a new composite indicator of quantity and quality of research. They tested it on data for European higher education institutions for 1996–2010. They found that the situation among EU countries is "highly diverse", but they also indicated that to a certain extent convergence appears to be underway.

Makkonen and Mitze's (2016) paper is the most recent example. Their research concerns the level of co-publication intensity among authors from old and new EU members. They applied Difference-in-Difference approach using data for 1991–2012 and came to the conclusion that entry of the new countries into the EU increased the co-publication intensity between the old and new members (with the exception of Bulgaria and Romania) and

this increase was statistically significant. Moreover, the number of co-publications among new member countries has surged as well.

As can be seen, the papers usually focus on co-authorship of scientific publications or international projects. The majority of them compare results of the old and the new member states, the latter being either referred to as EU-12 (before Croatia's accession to the EU in 2013), EU-10 (before the 2007 enlargement) or CEE-8 (Central and Eastern Europe; before 2007 and excluding Malta and Cyprus). Only few papers study a single country, such as Hirv's (2018) research on Estonia or Szalavetz's (2014) analysis of innovation patterns in Hungary. Importantly, the studies use different methods and come to different conclusions. Predictions that "there is no doubt that entrance into [the] European Union will boost research cooperation of new member states with the core of EU" (Marshakova-Shaikevich, 2007, p. 14) have not been verifiably fulfilled.

We are not aware of any published paper which would investigate the development of old and new member states' comparative advantages in scientific publishing before and after the EU enlargement(s).

Data and methodology

The research presented in this paper is based on Web of Science (2019) Core Collection data for years 1998, 2008 and 2018, extracted from the database in June 2019. The years were chosen to be equidistant from each other, the first being well before the largest enlargement in the history of the European Union, the second right after Romania and Bulgaria joined the bloc to make it EU-27, and the third the last one available at the time of writing. We decided against using time series data for the whole period, as this would bring only marginal benefits not offsetting the disproportionately higher resource demands of the data analysis. We will mostly use data for 1998 and 2018. Document type was set to "article". We extracted papers published by 28 current EU members using the "countries/regions" menu. Papers with authors from multiple countries were assigned to each of the countries, hence appearing more than once in our database. As a result, the total number of articles in the database is higher than the real number of articles published.

We are aware that the Web of Science is not an ultimate database which would include all the research papers published anywhere in the world. Indeed, it is relatively highly selective and in the past used to focus mostly on "Western" journals. While it definitely cannot be considered "a seal of quality", it provides a useful filter of academic papers and in many EU countries academic career advancement is based on papers indexed in the WoS. Another alternative which could be used is Elsevier's Scopus – a database similar to the WoS with a significant overlap with it. We believe both of the alternatives would be suitable for the current research, but chose to use the WoS due to its (in our opinion) superior analytical capabilities and higher number of years in operation. One service which arguably includes a wider range of papers than both WoS and Scopus is Google Scholar, but it is not a database and hence does not offer the required analytical tools.

The dataset used is based on 254 research areas as defined by Thomson Reuters in the Web of Science (2019). Each one of these belongs to a citation index (Arts & Humanities – AH, Science – SC or Social Sciences – SS). In nine cases, one area is assigned to more than one index; these instances were treated separately and the authors of the paper

selected the principal index for each.[3] The final set consists of 28 AH, 176 SC and 50 SS research areas. Its basic descriptive statistics can be seen in Table 1.

Throughout the paper, we consistently consider all the EU members which entered the bloc before 2004 as "old members" (15, including the UK) and the ones which have entered since 2004 as "new members" (13). This is in line with the usual practice, and implicitly accepts that there are significant differences between the EU-15 and the countries admitted in the so-called Eastern enlargements. These differences encompass economic situation, political climate, cultural factors and – crucially for our paper and as also evidenced in the literature review section – research outputs and environment.

The calculation of RCA is relatively straightforward, using the approach known from international economics and presented in ``Literature review: revealed comparative advantages'' section of this paper. The share of a country's papers in a given research area is divided by the share of the world's papers in that research area, i.e.

$$RCA_{rc} = \frac{\frac{P_r}{\sum_1^r P_r}}{\frac{\sum_1^c P_r^c}{\sum_1^c \sum_1^r P_r^c}}$$

where P_r^c. is the number of articles in research area r published by country c in a given year. As can be seen from the equation, RCA is a relative measure which needs to be calculated separately for each country and each research area. It does not depend on country size, but rather on how each country divides its research activity compared to how the world does. For example, if 1% of a country's research output is in Astronomy, but the world average for Astronomy is just 0.2%, the country has a revealed comparative advantage in Astronomy with the respective RCA value being 5. Conversely, if just 0.1% of a country's research output is in Astronomy, then it has a revealed comparative disadvantage in the field with RCA of 0.5.

Even though the above form of the index has been used by multiple papers, it suffers from asymmetry, variability and inconsistency (Yeats, 1985). In economics there has been a lively debate on shortcomings of RCA and numerous new indices have been proposed. Interestingly, this has not been the case in scientometrics. None of the published papers, with the notable exception of Rousseau and Yang (2012), focuses on this issue. Rousseau and Yang (2012) subjected the so-called "activity index" (a concept very closely related to RCA) to a thorough mathematical scrutiny and point out some of its theoretical problems. However, they do not offer solutions. Therefore, we have to turn to economic literature again.

The RCA scrutiny follows two main directions: one branch of authors makes simple changes to the index to address selected statistical deficiencies; another group offers more robust alternatives. The first approach has led to several new indices. Weighted RCA was proposed by Proudman and Redding (2000) to cope with the fact that the original RCA's arithmetic mean across sectors (or research areas in our case) is not usually equal to one. Hoen and Oosterhaven (2006) achieved stability of the mean by a different approach called additive RCA, substituting the multiplicative character of the index for subtraction. Yu et al. (2009) tried to guarantee comparability of the index over time and across countries,

[3] For example, research area "History – Philosophy of Science" belongs to all three citation indices. We decided to place it in the AH category, taking into account that both "History" and "Philosophy" are in the same.

Table 1 Descriptive statistics (1998, 2008, 2018). *Source*: Own elaboration

		World	EU	EU-old	EU-new
Total publications		1,251,794	543,081	501,233	41,848
		2,034,831	852,651	762,767	89,884
		3,175,035	1,283,178	1,134,633	148,545
AH	No. of publications	32,740	7557	7194	363
		75,811	21,102	19,136	1966
		100,202	40,386	35,482	4904
	Average per research area (st. deviation)	1169 (1060)	270 (299)	257 (283)	13 (23)
		2708 (2689)	754 (874)	683 (776)	70 (105)
		3579 (3635)	1442 (1,513)	1,267 (1,306)	175 (227)
	Median	700	136	126	3
		1785	407	376	26
		2321	906	815	61
	Top research area by publications	History	History	History	Philosophy
		History	History	History	Philosophy
		History	History	History	Human. mult
SC	No. of publications	1,135,430	512,828	472,375	40,453
		1,739,530	750,539	667,957	82,582
		2,728,549	1,081,960	951,762	130,198
	Average per research area (st. deviation)	6451 (6609)	2914 (3167)	2684 (2875)	230 (350)
		9884 (9674)	4264 (4139)	3795 (3627)	469 (584)
		15,503 (16,324)	6148 (6120)	5408 (5371)	740 (830)
	Median	4406	1862	1725	105
		7436	2983	2757	304
		10,487	4508	3952	450
	Top research area by publications	Bioch. mol. bio	Bioch. mol. bio	Bioch. mol. bio	Chem. mult
		Mater. sc. mult	Mater. sc. mult	Mater. sc. mult	Mater. sc. mult
		Mater. sc. mult	Mater. sc. mult	Mater. sc. mult	Mater. sc. mult

Table 1 (continued)

		World	EU	EU-old	EU-new
SS	No. of publications	83,624	22,696	21,664	1032
		219,490	81,010	75,674	5336
		346,284	160,832	147,389	13,443
	Average per research area (st. deviation)	1672 (1339)	454 (484)	433 (459)	21 (33)
		4390 (3994)	1620 (1745)	1513 (1592)	107 (166)
		6,926 (6,218)	3,217 (2,991)	2,948 (2,660)	269 (343)
	Median	1177	338	319	9
		2835	1005	970	48
		4692	2151	2043	146
	Top research area by publications	Economics	Economics	Economics	Economics
		Economics	Economics	Economics	Economics
		Edu. educ. res	Economics	Economics	Economics
Share AH: SC: SS		3%: 91%: 6%	2%: 94%: 4%	2%: 94%: 4%	1%: 97%: 2%
		4%: 85%: 11%	2%: 88%: 10%	2%: 87%: 10%	2%: 92%: 6%
		3%: 86%: 11%	3%: 84%: 13%	3%: 84%: 13%	3%: 88%: 9%

AH Arts & Humanities, *SC* Science, *SS* Social Sciences, *Human. mult.* Humanities multidisciplinary, *Bioch. mol. bio.* Biochemistry and molecular biology, *Chem. mult.* Chemistry multidisciplinary, *Mater. sc. mult.* Materials science multidisciplinary, *Edu. educ. res.* Education and educational research. The rows in each cell represent data for 1998 (1st row), 2008 (2nd row) and 2018 (3rd row)

and developed normalized RCA. Finally, Laursen (2015) tackled the asymmetry problem by subtracting one from RCA and dividing it by RCA + 1. At the other end of the specter, Costinot et al. (2012) offered what they "humbly" referred to as a "theoretically consistent alternative" to Balassa's RCA. Their work diverged significantly from the original index and is based on OLS estimation of different types of bilateral-trade-flow fixed effects. Similarly, French (2017) ventured into the world of gravity modeling to deliver his version of the index.

Despite the relatively high number of alternatives to the original RCA index, none of them is perfect (Sanidas & Shin, 2010). Moreover, the more robust of them cannot be applied to scientometrics as they require country-level data on both sides (export and import); publishing, obviously, is not a bilateral flow. Consequently, in this paper we will use Laursen's (2015) symmetric RCA (SRCA). It is straightforward to compute, solves the asymmetry problem and allows easy comparison. It can be calculated as

$$SRCA_{rc} = \frac{RCA_{rc} - 1}{RCA_{rc} + 1}$$

Its values range from -1 to 1, where zero is the breaking point between comparative disadvantage and comparative advantage. SRCA is calculated for 15 old EU member states and 13 new EU member states for each of the above mentioned 254 research areas for 1998, 2008 and 2018. In further text the values of SRCA indices are analyzed and compared using standard statistical methods.

To study convergence, two approaches are taken. The first approach consists of calculating sigma-convergence of comparative advantages for each research area across all countries, and across old and new countries separately. The second approach is a bit more computation intensive. It requires us to construct arrays of SRCA indices by research area for each country. Spearman rank correlation is applied for each country pair and compared across time. Then sigma-convergence of the same arrays in different country blocs is calculated and compared across time. Unlike some other convergence indicators (such as beta-convergence) sigma-convergence is easy to calculate, standardized and does not lead to diametrically different results when different specifications are used (Quah, 1993). We follow the most frequent method of its calculation – the coefficient of variation defined as the standard deviation divided by the mean of the sample.

It is important to note that we do not predict in advance whether convergence or divergence will be found to be present due to the fact that two conflicting theories are possible. On one hand, the East–West integration of Europe can be expected to strengthen international scientific cooperation and hence should lead to convergence of research structures (and comparative advantages) among countries. On the other hand, however, integration can lead to a higher level of specialization, meaning countries will focus more on utilizing the comparative advantages they already possess, possibly creating regional research clusters; this would lead to divergence of comparative advantages on the EU-level. The present research will offer a hint indicating which one of the two conflicting theories prevails.

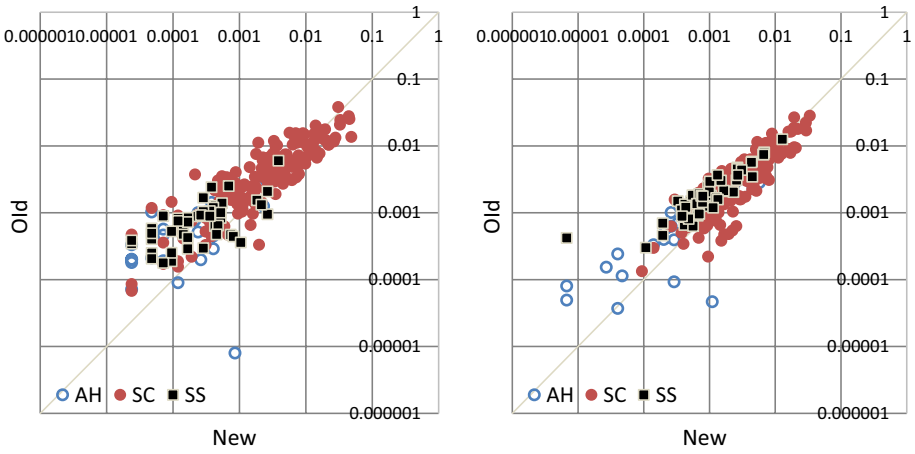

Fig. 1 Research areas by their share on old and new members' total publications in 1998 (left) and 2018 (right). *Note*: The diagonal line represents equal share on total publications in both new and old member states group. AH – Arts & Humanities, SC – Science, SS – Social Sciences. *Source*: Own elaboration

The development of RCA in the European Union

Between 1998 and 2018, the number of papers published by authors from the European Union has increased more than two-fold (Table 1). Similar to the situation in the USA, the growth has been slower than the world average, which can be explained by huge acceleration of publication efforts of developing countries.[4] Old and new EU members have followed different paths, the latter growing significantly faster (3.5 times compared to 2.3 times). Interestingly, on a country level, the best performer was an old member (Luxembourg, 17.7 times) and the worst performer was a new member (Bulgaria, 1.7 times), well in contrast with the general trend. Luxembourg's high growth rate was caused by its very low starting point and subsequent changes in research infrastructure in the last two decades, which can be witnessed by the establishment of the University of Luxembourg in 2003 and the Luxembourg Institute of Science and Technology in 2015, currently the largest research institutions in the country. Bulgaria's slow pace is due to its relatively high starting point linked to strong research tradition in physics and chemistry, several national journals indexed in Web of Science even in the 1990s, and meager changes to scientific career advancement requirements. In addition, Bulgaria has had problems with science funding and related corruption (Vesper, 2018) and it was found to lag behind even in other studies (Makkonen & Mitze, 2016).

The highest rate of growth in every single country studied was recorded in the fields of Arts & Humanities and Social Sciences. The reasons include a switch from traditionally published monographs (Hammarfelt, 2014) to publishing more scientific articles, and a higher number of newly founded journals in these two fields (Sánchez-Gil et al., 2018). This can be visually observed in Fig. 1, where boxes representing individual research areas

[4] In the same period, India's publication output has increased 5.6 times and China's a whopping 20.3 times. Their starting bases were much lower than the starting base of the EU or USA, hence this faster growth is not surprising.

Table 2 Top 10 research areas by SRCA in old and new EU members in 1998–2018 (overall). *Source*: Own elaboration

	Old members	New members
1998	1. Physics nuclear (SC)	1. Literature Slavic (AH)
	2. Physics particles fields (SC)	2. Physics nuclear (SC)
	3. Astronomy astrophysics (SC)	3. Chemistry multidisciplinary (SC)
	4. Allergy (SC)	4. Physics particles fields (SC)
	5. Rheumatology (SC)	5. Materials science textiles (SC)
	6. Hematology (SC)	6. Chemistry analytical (SC)
	7. Dermatology (SC)	7. Spectroscopy (SC)
	8. Anesthesiology (SC)	8. Physics multidisciplinary (SC)
	9. Physics atomic molecular chemical (SC)	9. Physics atomic molecular chemical (SC)
	10. Geography physical (SC)	10. Mathematics (SC)
2018	1. Astronomy astrophysics (SC)	1. Literature Slavic (AH)
	2. Physics particles fields (SC)	2. Physics particles fields (SC)
	3. Allergy (SC)	3. Physics nuclear (SC)
	4. Paleontology (SC)	4. Agricultural economics policy (SC)
	5. Evolutionary biology (SC)	5. Materials science paper wood (SC)
	6. Physics nuclear (SC)	6. Astronomy astrophysics (SC)
	7. Rheumatology (SC)	7. Folklore (AH)
	8. Archaeology (AH)	8. Forestry (SC)
	9. Hematology (SC)	9. Engineering marine (SC)
	10. Neuroimaging (SC)	10. Logic (SC)

AH Arts & Humanities, *SC* Science, *SS* Social Sciences

within SS and empty circles representing research areas within AH are gradually moving towards right and up, signifying higher share in total publications in 2018 (right chart) than in 1998 (left chart). In EU, the share of AH and SS in total publications has jumped from 5.6% to 15.7% in the respective period.

Another interesting thing to mention (Fig. 1) is that the vast majority of AH and SS markers lie above the diagonal line, i.e., Arts & Humanities and Social Sciences have a higher share on total publications in old EU members than in the new EU members. The finding is in line with literature which has noticed that the socialist publishing pattern of many of the CEE countries was skewed towards "hard science" (Glänzel & Schlemmer, 2007), or "fundamental science" as some have called it (Radosevic & Yoruk, 2014); much more so than in the West. Not only was this preferred by the governments, it also had an additional advantage for the authors of being less ideologically dependent and hence socially "safe". Consequently, 82% of AH research areas and 86% of SS areas had a higher importance in old EU members than in the new EU members in 1998. While the former has decreased to 68% in 2018, the latter has increased even further to 94%. It might therefore appear that the gap has not been closing and especially Social Sciences continue to play merely a minor role in the scientific landscape of the new members. However, this conclusion could prove to be shortsighted without further analysis. This is why we turn our attention to comparative advantages.

The list of research areas with the highest values of SRCA is dominated by Science (Table 2). Social Sciences and Arts & Humanities have only a very marginal presence both

Table 3 Top 10 research areas by SRCA in old and new EU members in 1998–2018 (country level). *Source*: Own elaboration

	Old members		New members	
1998	Limnology (SC)	0.92 LUX	Green sustainable sc. tech. (SC)	0.97 MLT
	Virology (SC)	0.92 LUX	Development studies (SS)	0.96 MLT
	Medical labor. technology (SC)	0.85 LUX	Anthropology (SS)	0.95 CRO
	Folklore (AH)	0.80 AUT	Materials sc. composites (SC)	0.94 LAT
	Mycology (SC)	0.75 LUX	Literature Slavic (AH)	0.92 CZE
	Medical ethics (SC)	0.75 BEL	Social issues (SS)	0.92 CRO
	Materials sc. paper wood (SC)	0.74 FIN	Materials science textiles (SC)	0.92 CRO
	Anesthesiology (SC)	0.74 AUT	Public administration (SS)	0.90 MLT
	Construction building tech. (SC)	0.71 LUX	Physics particles fields (SC)	0.90 CYP
	Forestry (SC)	0.69 FIN	Nursing (SC)	0.90 MLT
2018	Literature British Isles (AH)	0.80 IRE	Folklore (AH)	0.96 EST
	Poetry (AH)	0.78 IRE	Literature Afr. Austr. Can. (AH)	0.95 LAT
	Urban studies (SS)	0.74 LUX	Literature Slavic (AH)	0.93 CRO
	Logic (SC)	0.74 LUX	Literature Slavic (AH)	0.90 SLO
	Regional urban planning (SS)	0.72 LUX	Physics particles fields (SC)	0.89 LAT
	Literature romance (AH)	0.67 SPA	Literature Slavic (AH)	0.89 SVK
	Materials sc. paper wood (SC)	0.67 FIN	Agricultural econ. policy (SC)	0.88 ROM
	Physics particles fields (SC)	0.66 GRE	Obstetrics gynecology (SC)	0.87 MLT
	Social sc. math. methods (SS)	0.64 LUX	Materials sc. paper wood (SC)	0.86 SVK
	Folklore (AH)	0.64 FIN	Physics nuclear (SC)	0.85 LAT

Research area (citation index/field) RCA score (country code)

in the old and new EU members, occupying one (old) or two places (new) in top 10 in 2018, namely Slavic Literature, Folklore and Archaeology. This is not entirely unexpected considering the much higher number of SC research areas and their longer historical tradition, as shown previously. Several fields of physics stand out, supplemented by a few medical areas in the old and engineering in the new member states. Even if average values of SRCA are taken, SC scores far better in the European Union than SS and AH do.[5] This finding is similar to what Radosevic & Yoruk (2014) observed in their study of RCA in several global regions, where Science had higher comparative advantage than Social Sciences in every single region in the world (including EU-15 and CEE) with the sole exception of North America.

The scene is a bit different if SRCA values for all research areas and all countries are separated and ranked in a joint list (Table 3). The dominance of SC, which was still clearly visible in 1998, disappears in 2018 when AH comes to the forefront. It is perhaps not

[5] Average SRCA values in the European Union were the following: In 1998, AH -0.45; SC -0.10; SS -0.39. In 2018, AH -0.13; SC -0.02; SS -0.07.

Average SRCA values in the old members were the following: In 1998, AH -0.37; SC -0.02; SS -0.24. In 2018, AH -0.07; SC -0.01; SS 0.08.

Average SRCA values in the new members were the following: In 1998, AH -0.56; SC -0.18; SS -0.56. In 2018, AH -0.19; SC -0.03; SS -0.21.

Please note these are not weighted averages, but simple averages of all research areas within the field.

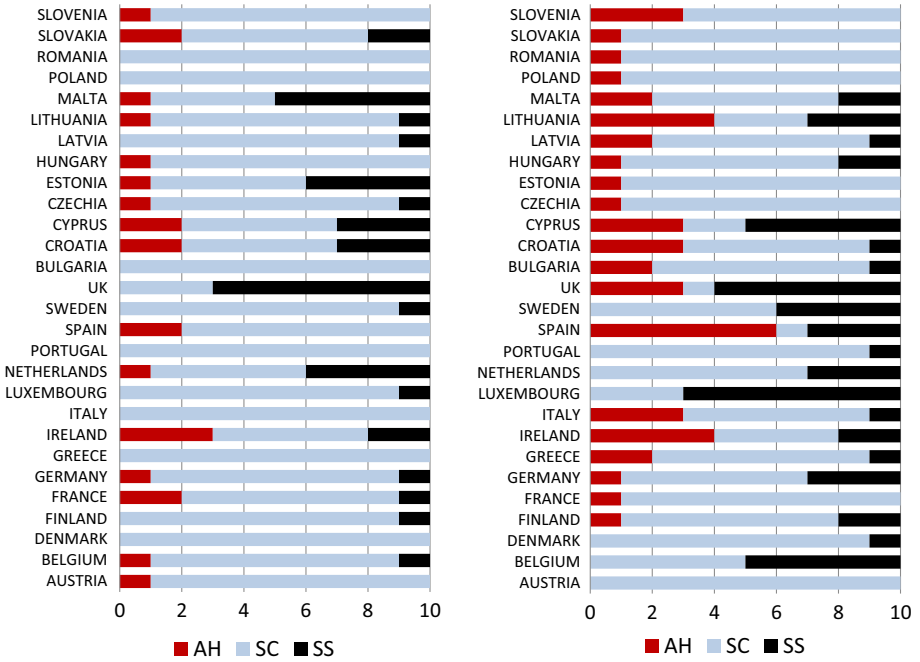

Fig. 2 Split of top 10 research areas by SRCA in 1998–2018. *Note*: AH – Arts & Humanities, SC – Science, SS – Social Sciences. *Source*: Own elaboration

surprising that different sub-areas of literature steal the show – who else should have a comparative advantage in Slavic Literature if not Slavic nations, who in Romance Literature if not Romance nations etc.? In fact, taking a summary table of top 10 research areas for each of the 28 countries in 2018, 16.4% of the items are AH (despite representing only 11.0% of the research areas). Twelve of them are literature, in 10 cases being a country's top research area by SRCA. Social Sciences have also increased their position, mostly in the old member states.

Country-level split of top 10 research areas by SRCA further shows that Austria is currently the only EU member country where all the positions are taken by SC (Fig. 2). SS and AH have the majority in six countries, compared to two countries in 1998. The United Kingdom – traditional mecca of modern-day Social Sciences, Arts & Humanities[6] – has been joined by Ireland, Luxembourg, Spain, Cyprus and Lithuania.

We have noticed earlier in the paper that the average non-weighted SRCA across all research areas remains negative in the EU. Despite this, the trend is positive (Fig. 3 left).It should be noted that negative average SRCA is not necessarily a problem and one cannot transform it to the conclusion that the EU's research is not competitive. It is logical that there are usually more areas with comparative disadvantage than with comparative advantage – a country simply cannot be good in too many things at once. Moreover, areas with disadvantage are often completely neglected, while in areas with advantage the competition

[6] Economics, Linguistics or Geography are just a few examples of research areas which are considered to have been born in the UK.

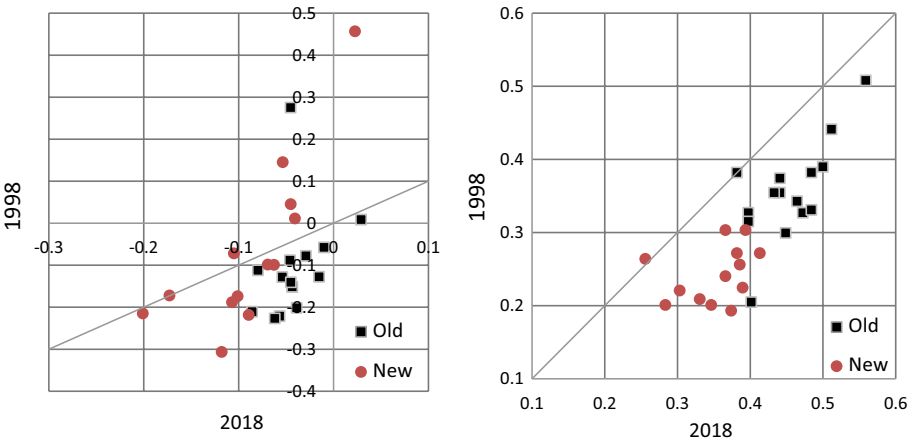

Fig. 3 Old and new EU members: average SRCA (left) and share of research areas with positive SRCA (right) 1998–2018. *Note*: The diagonal line represents reference line with equal values in 1998 and 2018. *Source*: Own elaboration

is usually fierce, hence negative SRCA scores tend to be more distant from zero than positive SRCA scores. This is a well-known fact from the trade literature, the home of the RCA. For example, in 2016, average SRCA of the world trade (222 countries/regions and 255 merchandise items) was -0.55. Even such a trade giant as Germany had a negative average SRCA of -0.15 (own calculations based on UNCTAD data).

Returning to application of SRCA in scientometrics, in 22 out of the 28 EU members there has been an increase in the average indicator value in the last two decades. The only two countries with a positive average SRCA are currently the UK and Malta. Interestingly, in 1998 there were five of these countries, which means that in several of them average SRCA must have turned negative since then. The explanation of this seemingly undesirable development lies in the fact that these are mostly small countries which were active only in a relatively low number of research areas in 1998, whereas now their publication patterns are much more diversified. As clarified before, higher number of research areas increases the number of areas with disadvantage (not their share), where SRCA values tend to be more distant from zero than in the ones with advantage, mathematically leading to lower overall average. It can also be argued that countries might be trying to establish themselves in new research areas, which is necessarily connected with low starting point and a hoped-for higher future SRCA values.[7]

The positive trend is even more obvious when looking at shares of research areas with positive SRCA (Fig. 3 right). This has increased in every single country except for Bulgaria where it is stagnant; the problems of Bulgarian science are well known (Vesper, 2018).

[7] The countries with positive SRCA values in 1998 are: Luxembourg (average SRCA 0.27 in 65 research areas, now -0.05 in 207), Cyprus (0.15 in 105, now -0.05 in 228), Estonia (0.01 in 154, now -0.04 in 226), Latvia (0.04 in 95, now -0.04 in 196) and Malta (0.46 in 57, now 0.02 in 181).

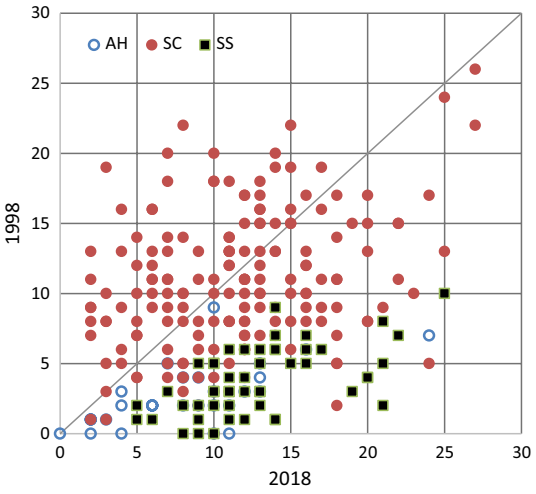

Fig. 4 Research areas by number of countries with positive SRCA in 1998–2018. *Note*: The diagonal line represents reference line with equal values in 1998 and 2018. *Source*: Own elaboration

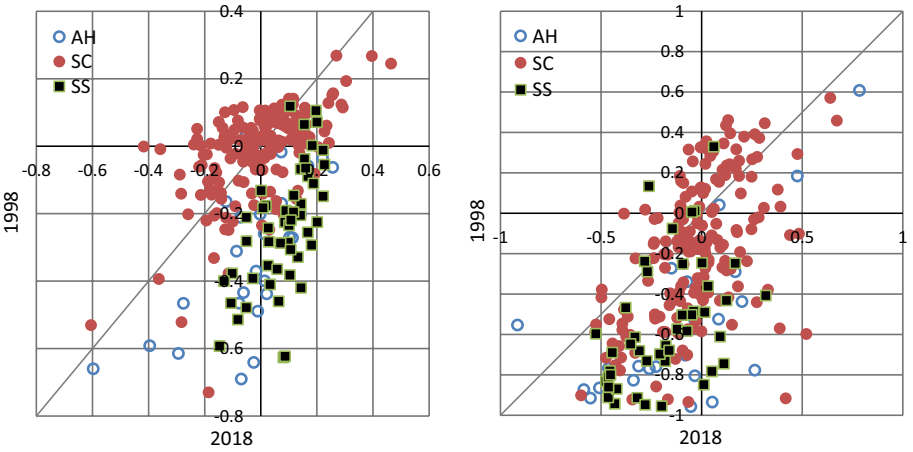

Fig. 5 SRCA by research area in old (left) and new (right) EU member states 1998–2018. *Note* The diagonal line represents reference line with equal values in 1998 and 2018. AH – Arts & Humanities, SC – Science, SS – Social Sciences. *Source*: Own elaboration

In 1998, there were nine research areas where no (current) EU country had a positive value of SRCA. In 2018, there was only one such research area left.[8] Also, in 1998, on average 8.5 countries had a revealed comparative advantage in a research area, and in 8 research areas the comparative advantage was shared by at least 20 countries (Fig. 4). In 2018, these numbers increased to 11.4 and 24. There are currently two research areas where only a single EU-28 country does not have a revealed comparative advantage:

[8] The research areas in 1998 were: Architecture, Art, Asian studies, Engineering marine, Family studies, Film radio television, Law, Literature American and Psychology clinical. In 2018 the sole one remaining was Literature American.

these are Astronomy astrophysics and Physics particles fields (in both cases the country is Luxembourg).

In line with some previous conclusions, especially AH and SS have gone a long way since 1998 with none of the research areas in these fields decreasing its number of EU members with positive SRCA in the last two decades. From among Social Sciences Economics and Business have reached the highest progress, jumping from 10 to 25 countries and from 2 to 21 countries. Presumably, in case of these two research areas the general positive trend has been further strengthened by the reforms of evaluation systems undertaken in numerous new member countries (Grančay et al., 2017). In Arts & Humanities, Archaeology has been the most successful area (from 7 to 24).

The rising comparative advantage of EU's research in SS and AH can be strongly evidenced by analyzing full SRCA data for each research area in 1998 and 2018 (Fig. 5). Among the old member states, there has been an increase in SRCA of every single AH research area except for Folklore and every single SS research area except for Ergonomics. Among the new member states there is only one AH area and five SS areas with negative development.[9] Comparative advantage of the majority of SC research areas has also grown, but the ratio has been much lower here: 51% SC research areas for old and 67% for new member states. Importantly, average SRCA growth was of a higher magnitude than average SRCA decrease, so the general trend is positive even in the field of Science.[10]

Several main patterns of scientific development in the EU could be clearly identified in the analyses presented so far:

- The number of papers published by EU authors is growing. The growth is faster in SS and AH than in SC, and faster in new members than in old members.
- The EU has enlarged its share in global publication activities in the majority of research areas in SC and in almost all research areas in SS and AH, meaning it has increased its revealed comparative advantage in these research areas. On average, the growth has been faster in new members than in old members.
- The EU has traditionally had the largest comparative advantage in research areas belonging to the field of SC. In the last two decades, SS and AH have challenged the SC's dominance.

None of these findings is unexpected. Global and regional increase in scientific publication activity is a well-researched fact, and even though literature does not offer a complex analysis of EU's scientific comparative advantages, increased dynamics of non-SC research areas has already been observed in the EU as well as in other countries (Radosevic & Yoruk, 2014; Wang & Wang, 2017).

It must be taken into account that our results are also a consequence of the WoS indexing policies. It is a well-known criticism that the database over-represents English-language journals (Mongeon & Paul-Hus, 2015), which it has reacted to by launching an "Eastern enlargement" (Grančay et al., 2017) in the recent years. To minimize the effect, we only included Core Collection articles in the analysis. However, during the 20 years studied the

[9] AH: Literary Reviews. SS: Anthropology, Ethics, Psychology Multidisciplinary, Sociology and Urban Studies.

[10] In old members, average growth of SRCA in SC research areas was 0.11 and average decrease -0.09. In new members, average growth of SRCA in SC research areas was 0.31 compared to average decrease of -0.16.

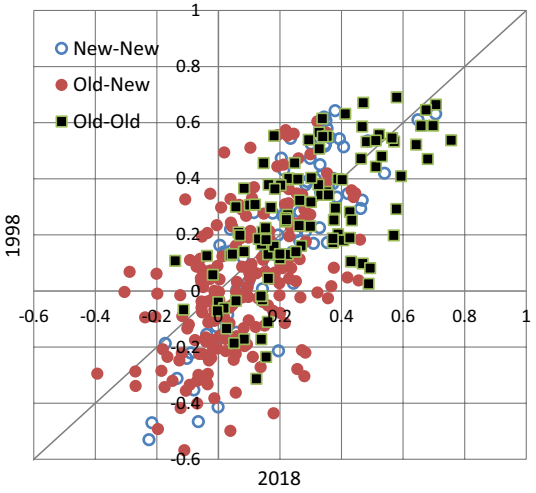

Fig. 6 Spearman correlations of SRCA arrays among all EU member states 1998–2018. *Note*: The diagonal line represents reference line with equal values in 1998 and 2018. *Source*: Own elaboration

number of journals covered by the Core Collection still increased by 278%. We argue this is not a concern for the current research, similarly as changes in global demand, technological progress or decisions of multi-national corporations are not a concern for the validity of the RCA concept in economics. Comparative advantage (or disadvantage) – irrespective of the field – is a consequence of numerous factors. In case of scientific publishing these include, among others, national research policies, university-level strategies and career advancement requirements as well as indexing decisions of database owners. Whether the academic community likes it or not, business strategy of the Web of Science significantly influences individual, institutional and national RCA and will continue to do so as long as the WoS is seen in many countries as the top scientific database which generates increased impact by means of visibility and citations.

The preliminary results presented above give a hint of whether there has been any sort of scientific convergence within the EU or the two groups of countries are moving in the opposite directions; however, a more robust statistical approach is needed to reach final conclusions. This will be provided in the next section.

Discussion: convergence between the old and new countries

Two of the most-widely used indicators of convergence are beta-convergence and sigma-convergence. While the former is directly related to the neo-classical growth theory and involves estimating a growth equation (Barro & Sala-i-Martin, 1992), hence being strongly dependent on specification adopted, data available and grouping of statistical units, the latter studies development of dispersion of an indicator over time, is easier to use and has lower dependence on specification adopted (Friedman, 1992; Quah, 1993). Moreover, beta-convergence is a necessary, but not a sufficient condition for sigma-convergence. As

Table 4 Spearman correlations of SRCA arrays among all EU member states 1998–2018. *Source*: Own elaboration

	Highest correlations		Lowest correlations	
1998	France-Italy	(OO) 0.6900	Germany-Malta	(ON) -0.5681
	Germany-Italy	(OO) 0.6709	Malta-Poland	(NN) -0.5306
	Austria-Germany	(OO) 0.6637	Austria-Malta	(ON) -0.4989
	Finland-Sweden	(OO) 0.6451	France-Malta	(ON) -0.4925
	Hungary-Poland	(NN) 0.6425	Czechia-Malta	(NN) -0.4702
	Italy-Spain	(OO) 0.6311	Hungary-Malta	(NN) -0.4660
	Czechia-Slovakia	(NN) 0.6311	Spain-Malta	(ON) -0.4363
	Bulgaria-Poland	(NN) 0.6197	Luxembourg-Poland	(ON) -0.4175
	France-Spain	(OO) 0.6142	Malta-Slovenia	(NN) -0.4141
	Czechia-Poland	(NN) 0.6098	Netherlands-Lithuania	(ON) -0.4016
2018	Belgium-Netherlands	(OO) 0.7566	UK-Poland	(ON) -0.3929
	Austria-Germany	(OO) 0.7069	Ireland-Poland	(ON) -0.3048
	Czechia-Slovakia	(NN) 0.7060	Ireland-Czechia	(ON) -0.2881
	Denmark-Sweden	(OO) 0.6978	UK-Czechia	(ON) -0.2693
	Netherlands-Sweden	(OO) 0.6804	Netherlands-Romania	(ON) -0.2686
	Finland-Sweden	(OO) 0.6750	Netherlands-Slovakia	(ON) -0.2379
	France-Germany	(OO) 0.6587	Malta-Poland	(NN) -0.2251
	Czechia-Poland	(NN) 0.6487	Czechia-Malta	(NN) -0.2157
	Denmark-Netherlands	(OO) 0.6429	Netherlands-Czechia	(ON) -0.2045
	Belgium-Sweden	(OO) 0.5936	Netherlands-Poland	(ON) -0.2004

OO Old and old member state. *NN* New and new member state. *ON* Old and new member state

a result, sigma-convergence calculated as the coefficient of variation will be used in this section.[11]

An important methodological decision is what sigma-convergence should be applied to. In the usual setting, researchers have a single value of an indicator for each country for each year and can perform simple calculations in a straightforward way. In our case, however, there is no single indicator, but an array of values of SRCA for 254 research areas for each country for each year.[12] We will tackle this situation using two approaches:

- First, similar to Grančay et al. (2015), for each EU member, we will rank all research areas by SRCA in the decreasing order. We will then compute Spearman correlation coefficients for all the country pairs and in the last step calculate sigma-convergence of the SRCA arrays.
- Second, we will compute sigma-convergence for each research area individually. Due to the nature of SRCA which consists of a mix of both positive and negative values, using it would not be statistically sound, hence we will need to go back to applying the original RCA index (where the values are positive by definition).

[11] Some other common methods are Theil index, Atkinson index, Gini index or standard deviation.

[12] A single indicator would be helpful for making the calculation easier, but is not really desired as it would necessarily not be able to hold all the information a full array of SRCA values can hold.

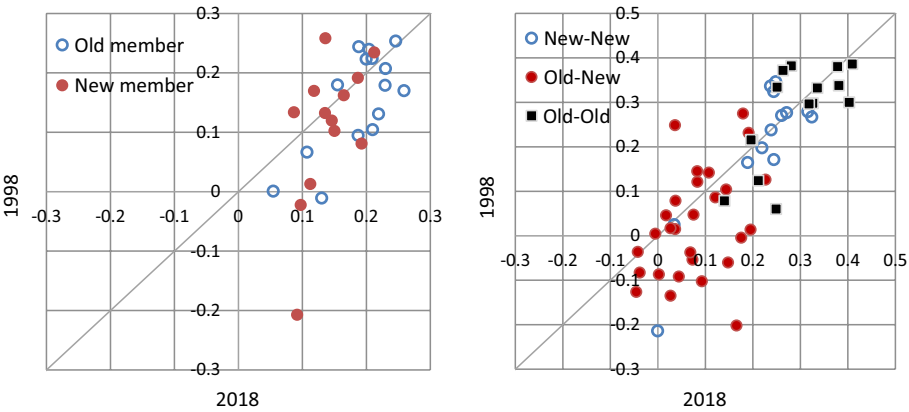

Fig. 7 Average Spearman correlation of SRCA by country (left) and by country and country-pair type (right) 1998–2018. *Note*: The diagonal lines represent reference lines with equal values in 1998 and 2018. *Source*: Own elaboration

Out of the total of 378 country pairs within the EU,[13] the Spearman correlation coefficient between the SRCA arrays increased in 218 cases (58%) in the studied period (Fig. 6). The growth was much more frequent in country pairs consisting of one new and one old member state (64%) than between old members (50%) or between new members (53%). At 0.168 average correlation growth was slightly higher than average correlation decrease (0.145).

Without exception, both in 1998 and 2018, the top 10 highest correlations were between country pairs from the same group, often reaching values higher than 0.60 (Table 4). Apart from the old couples, such as Austria-Germany, Belgium-Netherlands, France-Italy, or the Scandinavian countries, which have traditionally had similar comparative advantages and focus in science, the new member states with similar comparative advantages all belong to the Visegrad region. This shows that regional research clusters are present. On the other end of the specter, the majority of country pairs with the lowest values of correlation are between the groups, or involve Malta. As an island state which is in many ways – including science – significantly different from the rest of the continent, this is not surprising.

Average Spearman correlation of SRCA by country – calculated for each country as simple average of 27 correlation values with all other member states – has increased in the majority of cases in the last two decades (Fig. 7 left). This indicator can be used to show which EU country has comparative advantages in research which are the closest to (or which are the furthest away from) the other member states. Austria, Spain, Belgium and Denmark belong to the first group while Luxembourg, Romania, Malta and Cyprus belong to the second one.

As for the first group, Austria is a very interesting case. Being a neighboring country of two old and four new members, it has an ideal position to be the unifying element in many areas of life. Indeed, keeping its traditional scientific proximity to Germany intact, it has increased its SRCA correlation with new members in the last two decades. This is partially a result of the numerous grant schemes that support cross-border research between

[13] $27+26+25+\ldots+1$.

Table 5 Sigma-convergence of SRCA arrays among EU member states 1998–2018. *Source*: Own elaboration

	All countries	Old members	New members
1998	0.7901	0.5426	1.1190
2018	0.3171	0.2838	0.2726

Table 6 Sigma-convergence of RCA for each research area 1998–2018. *Source*: Own elaboration

	All countries	Old members	New members	Total areas
All areas	185 (73%)	183 (72%)	154 (67%)	254
AH	12 (43%)	13 (46%)	28 (50%)	28
SC	132 (75%)	122 (69%)	176 (69%)	176
SS	41 (82%)	48 (96%)	50 (64%)	50

Share on total in brackets. Some research areas had no publications in new countries in 1998, hence cannot show any convergence. These are excluded from the calculation of share. *AH* Arts & Humanities, *SC* Science, *SS* Social Sciences

Austrian scientists and their post-socialist foreign partners (such as Action Austria-Slovakia, Action Czech Republic-Austria etc.) and its advantageous geographical location. In a similar fashion, Spain has been increasing its dynamics of scientific networking with new EU members as well. This supports the results of Teodorescu and Andrei (2011) who found that Spain has come from zero to becoming one of the most important collaborators of East European nations between 1989 and 2009. On the contrary, Belgium and Denmark have been maintaining their higher level of alignment with research in other old member states without any significant overlap with the new members.

Countries with the lowest degree of SRCA similarity with other EU members are the two small island states, Luxembourg and Romania. The first three of them are the smallest EU members, and consequently have a tiny scientific output mostly concentrated in a few areas. In Malta and Cyprus, some of the highest ranked SRCA areas are related to marine sciences, where overlap with landlocked countries is hard to find. We have also discussed the specifics of Luxembourgian science earlier already. An interesting case is Romania whose position here does not have any ready explanation. Its comparative advantages in research have low levels of correlation with other new members, and almost no correlation exists with the old members. The situation has been worsening since 1998. Finding the exact answer to why this is the case is beyond the scope of the present paper, but previous research of co-authorship has shown similar pattern, namely gradual decrease of Romania's international collaboration (Makkonen & Mitze, 2016; Teodorescu & Andrei, 2011).

In general, if country averages are split into two – i.e. each country will have a separate average SRCA correlation with old members and a separate average SRCA correlation with new members – old members have a higher correlation of SRCA with other old members than new members with other new members, and old-new country pairs lag behind both of these groups; however, the general trend is positive (Fig. 7 right).

The anecdotal hints of convergence presented up to now can be put to test by applying sigma-convergence to RCA and SRCA arrays among the EU members. Using the first method and calculating sigma-convergence of average country-level correlation of SRCA arrays we find out that convergence exists (Table 5). The coefficient of variance has decreased considerably and the EU countries now have a much more coherent set of

comparative advantages in research than twenty years ago. The convergence has been faster within the new-member group than within the old-member group. This is not surprising given a much more varied composition of the former, consisting of several distinct regional groups such as the Baltic nations, the Visegrad countries or the small island states. The Visegrad group has the lowest level of variation at 0.16.

The second method – computing sigma-convergence for each research area across all countries – shows that 185 out of the 254 research areas have a lower value of sigma-convergence in 2018 than in 1998, i.e. there is lower variance now than before. The higher share of converging than diverging research areas applies to the EU as a whole, but also to the old member states (183 areas converging) and the new member states (154 areas converging) separately (Table 6). Strong convergence can be seen in the areas of Social Sciences and Science. On the other hand, Arts & Humanities seem to be diverging slightly. The reason for this can be found in the low number of publications in AH (as seen in Table 1). Some of the AH research areas are very narrow, i.e. only a handful of countries publish in them and even then they only publish a handful of papers. Average number of papers per research area per country in AH was 10 in 1998, increasing to 52 in 2018. This compares very unfavorably to SS (16; 115) or SC (104; 220). Consequently, SRCA values have a much higher cross-national volatility. Another reason of the lower AH convergence is – as we have already noticed – a higher share of region-specific research areas than in other fields, such as different types of Literature (Romance, Slavic, German etc.).

On the overall level, convergence was confirmed to exist using both approaches. This supports the conclusions of Makkonen and Mitze (2016) who found convergence of co-publication intensity in the EU, or Pastor et al. (2015) who observed traces of convergence in quantity and quality of research output of European higher education institutions.

Having established that there is convergence, a different question should be asked: To what extent is convergence necessary? Obviously, creation of a single European research area does not mean that all the countries will have comparative advantages in the same research areas. This would not be desirable. On one hand, similar comparative advantages strengthen the respective areas and bring synergies. On the other hand, however, as is the case in trade and numerous other fields, different regions have different history and underlying potential, hence comparative advantages which take into account regional specifics should be sought. It follows that divergence of SRCA is not necessarily a bad sign.

Caution has to be taken when trying to reach conclusions about the causality of the relationship. We have found evidence that convergence exists but we have not used any statistical methods which could confirm that it was induced by the countries' membership in the European Union. It appears likely, but is far from certain. Other factors have played an important role in the world of science, such as the general boom of Social Sciences and Arts & Humanities, "Eastern enlargement" of the citation databases (Grančay et al., 2017), strategic decisions by the governments or changes in economic structure which necessarily lead to changes in scientific focus. Exogenous factors, such as changes in research competitiveness of third countries have surely also had an impact on the EU as well. In the end, it is the nature of the comparative advantages (and other relative indicators) that they are determined both internally and externally and a single change in a small country anywhere in the world influences everyone else.

Co-authorship has also certainly affected our data. In our approach, a paper which is authored by three researchers from three different countries is counted three times, once for each country. Robust evidence exists that collaboration in the European Union and resulting number of co-authored papers have increased significantly in the last two decades (Makkonen & Mitze, 2016; Teodorescu & Andrei, 2011). Consequently, the higher number

of joint papers translates into a higher similarity of comparative advantages. However, this is a natural development connected to globalization of science which should not be of any concern for the soundness of our concept.

Conclusions

Seventeen years have passed since the biggest enlargement of the European Union. Between 2004 and today thirteen new members have joined the bloc, increasing its population by a quarter and its size by a third. The number of research outputs has not been boosted so much. Today the 13 new members participate on the EU's scientific publications with a 12%-share, which is just a small increase of 1 percentage point since 2008 (Web of Science, 2019; own calculations). While this might appear slow, a faster speed can't really reasonably be expected. Research systems of many of the new EU members were directed by socialist principles for decades, and well into the beginning of the twenty-first century tenure positions in several countries did not even require journal publications (Grančay et al., 2017). The inertia is strong and the catch-up process will take longer time.

The story of comparative advantages is slightly different. In the paper we have assessed the development of scientific comparative advantages of the old and new EU members using the SRCA indicator borrowed from international trade and adjusted for our purposes. The results confirm that there has been convergence inside the groups as well as between the groups regarding both (1) research areas with comparative advantage and (2) average country-level SRCA correlation. EU researchers from different countries now publish papers in more similar research areas and the countries' research-area mix is more coherent than before. The EU's comparative advantages have grown in the majority of research areas since 1998, faster in Social Sciences and Arts & Humanities than in Science, and faster in new members than in old members.

The results also confirm what Glänzel and Schlemmer (2007) stated a decade ago: Although convergence is taking place, individual countries have maintained their particularities. This is a positive sign. "United in diversity" is the EU's multidimensional motto, which in science can be translated to mean that one does not create competitiveness by suppressing regional differences and supporting uniformity. Quite the opposite, the richness and long-term success lie in respecting local comparative advantages. The EU has understood it and supports multinational education and research, and interdisciplinary approaches combining Science, Social Sciences and Arts & Humanities through programs such as Horizon Europe, Erasmus + etc.

The paper presented the first research of scientific comparative advantages of the EU and their development on national level. As such, it is of invaluable importance for policy-makers and research institutions who can verify whether the current status quo is in line with their projections, and based on a detailed analysis adjust their policies and adopt measures to either further strengthen the strong areas or support the weak ones. Governments can use the scientific RCA to assess whether there is a match between a country's economic and research structures and invest into research areas that could lead to fulfilment of long-term economic goals. While they can arguably do that even without calculating RCA, it provides a useful tool for measuring progress and can therefore be considered a vital key performance indicator for strategic plans. On the EU level, the paper's insights can be utilized in planning of future research framework

programs, for example to adjust the so-called widening and strengthening policies and to direct them towards specific research areas (as opposed to the current setting encompassing all fields of SC, SS an AH).

The research could be further expanded in numerous ways. First, it could use alternative indicators of comparative advantage such as Weighted RCA or Additive RCA. Second, a longer time period with annual data might be taken into account. Third, the focus could be put not only on quantity, but also on quality of research, defined either by number of citations, papers in top journals or other frequently used measure – as Lattimore and Revesz (1996) noticed, correlation between comparative advantage and quality tends to be weak. A different database than Web of Science might be used as the main data source to test whether the results are also valid in a set of papers with different research area structure. Finally, other, perhaps more complex indicators of convergence could be applied.

References

Albarran, P., Crespo, J. A., Ortuno, I., & Ruiz-Castillo, J. (2010). A comparison of the scientific performance of the US and the European union at the turn of the 21st century. *Scientometrics, 85*(1), 329–344. https://doi.org/10.1007/s11192-010-0223-7

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. https://doi.org/10.2307/1937915

Balassa, B. (1965). trade liberalisation and „Revealed" comparative advantage. *The Manchester School, 33*(2), 99–123. https://doi.org/10.1111/j.1467-9957.1965.tb00050.x

Barro, R. J., & Sala-i-Martin, X. (1992). Convergence. *Journal of Political Economy, 100*(2), 223–251. https://doi.org/10.1086/261816

Basu, A., Foland, P., Holdridge, R., & Shelton, R. D. (2018). China's rising leadership in science and technology: quantitative and qualitative indicators. *Scientometrics, 117*(1), 249–269. https://doi.org/10.1007/s11192-018-2877-5

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. https://doi.org/10.1016/j.enpol.2017.01.007

Cecere, G., & Corrocher, N. (2013). The intensity of interregional cooperation in information and communication technology projects: an empirical analysis of the framework programme. *Regional Studies, 49*(2), 204–218. https://doi.org/10.1080/00343404.2012.759651

Chuang, Y., Lee, L., Hung, W., & Lin, P. (2010). Forging into the innovation lead—a comparative analysis of scientific capacity. *International Journal of Innovation Management, 14*(3), 511–529. https://doi.org/10.1142/S1363919610002763

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. https://doi.org/10.1093/restud/rdr033

European Commission. (2000). *Presidency Conclusions of the Lisbon European Council 23 and 24 March 2000*. 1.8.2019. https://www.consilium.europa.eu/ueDocs/cms_Data/docs/pressData/en/ec/00100-r1.en0.htm.

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. https://doi.org/10.1057/eej.2015.4

French, S. (2017). Revealed comparative advantage: what is it good for? *Journal of International Economics, 106*(1), 83–103. https://doi.org/10.1016/j.jinteco.2017.02.002

Friedman, M. (1992). Do old fallacies ever die? *Journal of Economic Literature, 30*(4), 2129–2132.

Glänzel, W., & Schlemmer, B. (2007). National research profiles in a changing Europe (1983–2003) an exploratory study of sectoral characteristics in the Triple Helix. *Scientometrics, 70*(2), 267–275. https://doi.org/10.1007/s11192-007-0203-8

Grančay, M., Vveinhardt, J., & Šumilo, Ē. (2015). Trade in central and Eastern European countries ten years after their EU accession — Is there convergence? *Society and Economy, 37*(4), 443–460. https://doi.org/10.1556/204.2015.37.4.2

Grančay, M., Vveinhardt, J., & Šumilo, Ē. (2017). Publish or perish: how central and Eastern European economists have dealt with the ever-increasing academic publishing requirements 2000–2015. *Scientometrics, 111*(3), 1813–1837. https://doi.org/10.1007/s11192-017-2332-z

Harzing, A. W., & Giroud, A. (2014). The competitive advantage of nations: an application to academia. *Journal of Infometrics, 8*(1), 29–42. https://doi.org/10.1016/j.joi.2013.10.007

Hammarfelt, B. (2014). Using altmetrics for assessing research impact in the humanities. *Scientometrics, 101*(2), 1419–1430. https://doi.org/10.1007/s11192-014-1261-3

Herranz, N., & Ruiz-Castillo, J. (2013). The end of the "European Paradox." *Scientometrics, 95*(1), 453–464. https://doi.org/10.1007/s11192-012-0865-8

Hirv, T. (2018). Effects of European union funding and international collaboration on Estonian scientific impact. *Journal of Scientometric Research, 7*(3), 181–188. https://doi.org/10.5530/jscires.7.3.29

Hoen, A. R., & Oosterhaven, J. (2006). On the measurement of comparative advantage. *Annals of Regional Science, 40*(3), 677–691. https://doi.org/10.1007/s00168-006-0076-4

Horta, H. (2018). The declining scientific wealth of Hong Kong and Singapore. *Scientometrics, 117*(1), 427–447. https://doi.org/10.1007/s11192-018-2845-0

Horta, H., & Shen, W. Q. (2019). Current and future challenges of the Chinese research system. *Journal of Higher Education and Policy Management, Online First.* https://doi.org/10.1080/1360080X.2019.1632162

Huang, X., & Tian, C. (2014). International competitiveness of the Chinese publishing industry. *Publishing Research Quarterly, 30*(1), 104–114. https://doi.org/10.1007/s12109-014-9344-9

Lattimore, R., & Revesz, J. (1996). *Australian Science — Performance from Published Papers*. Bureau of Industry Economics, Australian Government Printing Office.

Laursen, K. (2015). Revealed comparative advantage and the alternatives as measures of international specialization. *Eurasian Business Review, 5*(1), 99–115. https://doi.org/10.1007/s40821-015-0017-1

Lee, L.-C., Lee, Y.-Y., & Liaw, Y.-C. (2012). Bibliometric analysis for development of research strategies in agricultural technology: the case of Taiwan. *Scientometrics, 93*(3), 813–830. https://doi.org/10.1007/s11192-012-0833-3

Lee, L.-C., Lin, P.-H., Chuang, Y.-W., & Lee, Y.-Y. (2011). Research output and economic productivity: a Granger causality test. *Scientometrics, 89*(2), 465–478. https://doi.org/10.1007/s11192-011-0476-9

Leydesdorff, L. (2000). *Scientometrics, 47*(2), 265–280. https://doi.org/10.1023/a:1005639026643

Leydesdorff, L., Wagner, C. S., & Bornmann, L. (2014). The European Union, China, and the United States in the top-1% and top-10% layers of most-frequently cited publications: competition and collaborations. *Journal of Informetrics, 8*(3), 606–617. https://doi.org/10.1016/j.joi.2014.05.002

Makkonen, T., & Mitze, T. (2016). Scientific collaboration between "old" and "new" member states: did joining the European Union make a difference? *Scientometrics, 106*(3), 1193–1215. https://doi.org/10.1007/s11192-015-1824-y

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. https://doi.org/10.1007/s00191-003-0162-z

Marshakova-Shaikevich, I. (2007). The visualization of scientific collaboration of 15 "old" and 10 "new" EU countries in the field of social sciences. *COLLNET Journal of Scientometrics and Information Management, 1*(1), 9–16. https://doi.org/10.1080/09720502.2007.10700948

Micheli, A., Di Salvo, F., Lombardo, C., Ugolini, D., Baili, P., & Pierotti, M. A. (2011). Cancer research performance in the European Union: a study of published output from 2000 to 2008. *Tumori, 97*(6), 683–689. https://doi.org/10.1177/030089161109700601

Mongeon, P., & Paul-Hus, A. (2015). The journal coverage of web of science and Scopus: a comparative analysis. *Scientometrics, 106*(1), 213–228. https://doi.org/10.1007/s11192-015-1765-

Murashova, E., & Loginova, V. (2017). University-industry interaction trends in the Baltic sea region: a bibliometric analysis. *Baltic Journal of European Studies, 7*(2), 28–58. https://doi.org/10.1515/bjes-2017-0009

OECD. (2011). *Globalisation, Comparative Advantage and the Changing Dynamics of Trade*. 16.7.2019. http://www.oecd.org/publications/globalisation-comparative-advantage-and-the-changing-dynamics-of-trade-9789264113084-en.htm.

Pastor, J. M., Serrano, L., & Zaera, I. (2015). The research output of European higher education institutions. *Scientometrics, 102*(3), 1867–1893. https://doi.org/10.1007/s11192-014-1509-y

Proudman, J., & Redding, S. (2000). Evolving patterns of international trade. *Review of International Economics, 8*(3), 373–396. https://doi.org/10.1111/1467-9396.00229

Quah, D. (1993). Galton's fallacy and tests of the convergence hypothesis. *The Scandinavian Journal of Economics, 95*(4), 427–443. https://doi.org/10.2307/3440905

Radosevic, S., & Yoruk, E. (2014). Are there global shifts in the world science base? Analysing the catching up and falling behind of world regions. *Scientometrics, 101*(3), 1897–1924. https://doi.org/10.1007/s11192-014-1344-1

Rogge, N. (2019). Regional productivity growth in the EU since 2000: something is better than nothing. *Empirical Economics, 56*(2), 423–444. https://doi.org/10.1007/s00181-017-1366-7

Rousseau, R., & Yang, L. (2012). Reflections on the activity index and related indicators. *Journal of Informetrics, 6*(3), 413–421. https://doi.org/10.1016/j.joi.2012.01.004

Sánchez-Gil, S., Gorraiz, J., & Melero-Fuentes, D. (2018). Reference density trends in the major disciplines. *Journal of Informetrics, 12*(1), 42–58. https://doi.org/10.1016/j.joi.2017.11.003

Sanidas, E., & Shin, Y. (2010). Comparison of revealed comparative advantage indices with application to trade tendencies of East Asian Countries. In *9th Korea and the World Economy Conference*. 15.7.2019. http://www.akes.or.kr/eng/papers%282010%29/24.full.pdf.

Streb, J., Baten, J., & Yin, S. X. (2006). Technological and geographical knowledge spillover in the German empire 1877–1918. *Economic History Review, 56*(2), 347–373. https://doi.org/10.1111/j.1468-0289.2006.00347.x

Szalavetz, A. (2014). Innovation in Hungary - the impact of EU accession and integration into global value chains. *International Journal of Management and Economics, 42*(1), 40–59. https://doi.org/10.2478/ijme-2014-0042

Teodorescu, D., & Andrei, T. (2011). The growth of international collaboration in East European scholarly communities: a bibliometric analysis of journal articles published between 1989 and 2009. *Scientometrics, 89*(2), 711–722. https://doi.org/10.1007/s11192-011-0466-y

Vesper, I. (2018). Bulgaria in the cold as European Union freezes its innovation funding. *Nature, 554*(7691), 156–157. https://doi.org/10.1038/d41586-018-01374-x

Wang, L. (2016). The structure and comparative advantages of China's scientific research: quantitative and qualitative perspectives. *Scientometrics, 106*, 435–452. https://doi.org/10.1007/s11192-015-1650-2

Wang, L., & Wang, X. (2017). Who sets up the bridge? Tracking scientific collaborations between China and the European Union. *Research Evaluation, 26*(2), 124–131. https://doi.org/10.1093/reseval/rvx009

Web of Science. (2019). www.webofknowledge.com.

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. https://doi.org/10.1007/BF02705840

Yu, R., Cai, J. C., & Leung, P. S. (2009). The normalized revealed comparative advantage index. *Annals of Regional Science, 43*(1), 267–282. https://doi.org/10.1007/s00168-008-0213-3

Yue, C. J., & Hua, P. (2002). Does comparative advantage explain export patterns in China? *China Economic Review, 13*(2–3), 276–296. https://doi.org/10.1016/S1043-951X(02)00073-1