Hungarian Educational Research Journal 13 (2023) 2, 276–297
DOI: 10.1556/063.2022.00192

AKADÉMIAI KIADÓ

International co-authored publications: The effect of joining the European Union or being part of the European Research Area

GIULIO MARINI*

Quantitative Social Science, Social Research Institute, UCL, London, United Kingdom

RESEARCH ARTICLE

Received: July 5, 2022 • Accepted: September 27, 2022
Published online: October 26, 2022

ABSTRACT

The paper investigates the increasing number of international co-authored publications, comparing countries that accessed the European Union (EU) in 2004 (EU04) against other Central-Eastern European Countries (othEast-ERA), adopting a scientometric approach. This comparison looks at whether to be part of the EU is different from being part of the European Research Area (ERA) – given that both entities aim at fostering more international collaborations. The hypothesis is that EU might convey more opportunities for the sake of international publications, although ERA assures access to European funding schemes anyway. Analysing the census of internationally co-authored publications from 1995 to 2015, difference-in-differences regressions show that Countries that joined EU in 2004 performed better than other Central-Eastern ones. Implications for the public policies in science are discussed.

KEYWORDS

scientometric, European Research Area, European Union, funding agency, Central Eastern Europe, international collaborations

* Corresponding author. 55–59 Gordon Square, WC1H 0NU London, United Kingdom; E-mail: g.marini@ucl.ac.uk

INTRODUCTION AND RESEARCH QUESTION

Although the construction of a European level in research is far from being simple and linear in its development (Lepori, Reale, & Larédo, 2014), with still strong national constraints (EC 2017; Nedeva, 2012), the European Research Area (ERA) has attempted some homogenisation and internationalisation. As known, ERA is a projection and an emanation of the EU (Ulnicane, 2015), but many more countries may benefit from its initiatives. From one side, this achievement is a commitment by European Union member states, whose road map is analysed periodically (EC 2016; EC 2017; EC 2020; Smith et al., 2015). On the other side, ERA has the ambition to facilitate not only specific freedoms such as internal market for researchers, but also a sort of boundaryless space. Prominently, ERA goes geographically beyond the membership to the European Union (EU) as such, allowing other European countries to benefit from collaborative schemes and opportunities. The openness guaranteed by ERA is in fact reflected also in the opportunity to join the main EU funding schemes by academics based in other countries. There is a cogent theoretical ground to compare EU-member countries and other European countries. The benefits of collaborations might possibly be achieved by being part of ERA without necessarily joining the EU. Accessing EU in fact would bring regulative and financial constraints, not only political representation. This means that comparing ERA and EU raises a research policy problem under a rational choice theory: is it worth joining EU if a country is already benefitting from ERA?

In Central-Eastern Europe 10 countries joined EU in 2004 (Poland, Slovak Republic, Latvia, Lithuania, Estonia, Czech Republic, Slovenia, Hungary, Cyprus and Malta – "EU04"). Romania and Bulgaria joined later in 2007, and Croatia in 2015. Many other non-Western countries are notwithstanding part of ERA – they are here labelled "Other Central-Eastern ERA countries" (othEast-ERA). Both of these two groups of countries ("EU04" and "othEast-ERA") share a similar historical legacy, although the latter group is not (yet) part of the EU. It makes sense hence to compare these two sets of countries checking whether having joined the EU has made any difference, or viceversa if being part of ERA is enough for securing more internationalisation.

The comparison this paper provides is pursued via a quasi-experimental research design (Gopalan, Rosinger, & Ahn, 2020). In particular, the paper provides a comparison between the two groups of countries (those that joined and those that did not join the EU in 2004), and by years before and after the 2004 EU enlargement. This procedure is particularly suitable for cases when a phenomenon under observation such as international co-authorships tends to grow anyway, and an understanding of possible marginal gains are useful to detect net advantages. This research design considers EU04 countries those that went under a "treatment"; other Central-Eastern European countries that continued to be non-members of EU, regardless of being in some cases in a process of considering to access EU. The latter group is commonly defined in these research designs as "control group". These countries are: Turkey, Republic of Serbia, Bosnia and Herzegovina, Montenegro, North Macedonia, Albania, Ukraine, Belarus, Russian Federation, Moldova, plus other former countries comprising the USSR – Georgia, Armenia, Azerbaijan, Kazakhstan, Kyrgyzstan, Turkmenistan, Tajikistan, and Uzbekistan, grouped together and labelled "othCIS". We use "othEast-ERA" label to identify these countries.

In more details, why the comparison of EU and non-EU Members is relevant? ERA status might be a sufficient condition to boost a country's research performance at international level.

ERA ambitions have been so long twofold. On one side ERA's rationale is to coordinate and to integrate research from national or sporadic regional collaborations into a durably regional one. On the other side, ERA provides conditions for "excellence" to be achieved "across the [Innovation] Union" (EC 2007). This detail identifies the absence of a clear border of ERA, which goes beyond the EU (Yegorov, 2009). To consider funding agencies is also essential, as European schemes have been the only worldwide agencies piloting the otherwise autopoietic development of international science (Marini, 2018; Wagner & Leydesdorff, 2005a). In other words, (international) collaborations happen by themselves, without necessarily any push factor. When talking of countries (instead of, say, single universities or single researchers), the issue of spontaneus web of collaborations mark the "strong" and the "weak" countries – both denoted by their *position* in the network, rather than by other attributes. The strong countries are not more *central*, or "core ones"; the weak are not more *peripheral* (Leydesdorff & Wagner, 2008). The dynamic between respective "core" and "peripheral" countries is in fact an expected output ERA should contribute to. As the European Commission stated: "with neighbouring countries, the objective should be to establish a borderless 'broader ERA', which would underpin and benefit from other elements of the European neighbourhood Policy" (EC 2007).

This paper aims at understanding which membership status (inside or outside the EU, though still inside ERA) has been a better condition when it comes to talk about the extent to which international co-authored publications are signed by a larger number of countries (hypothesis 1); and whether EU membership to facilitates to have higher influence (hypothesis 2) – measured by a normalised indicator of citations (CNCI).[1] In other words, this paper faces the question whether the ten EU04 countries have been better off than the latter group. For the sake of simplicity, the paper excludes any publication co-authored also by scholars based in Romania, Bulgaria and Croatia – leaving the quasi-experimental research design simplified to only one key moment: May 1st 2004.

The paper develops as follows: 'literature review' exposes the contributions in the topic of co-authored international publications, with emphasis on Central-Eastern Europe. 'Dataset and Variables' explains how data have been retrieved. It also describes the variables included in the dataset. 'Descriptive statistics by time and treatment, and difference-in-differences hypotheses' section fosters descriptive statistics. It also shows the basic difference-in-differences tests and the equations along with covariates for both hypotheses. 'Results' section exposes the main interpretation. 'Discussion and conclusion' section wraps up the finding with some policy implications. Limitations are also listed.

LITERATURE REVIEW

Science at the aftermath of the fall of Real Socialism

In terms of research policies, the preoccupation of the legacy of the Eastern European bloc arose already in a first stream of research developed much before the enlargement of EU in 2004, precisely at the wake of the collapse of Real Socialism. For instance, Braun and Glänzel (1996)

[1]CNCI figures are publicly delivered by Clarivate Analytic. This indicator permits to compare outputs published in different years and in different disciplines to have a fair understanding of relevance of each publication.

suggested that the steeped increase of co-authored publications at the wake of collapse of the Council for Mutual Economic Assistance (COMECON), partially anticipated by Hungary and Poland in the 1980s, was more a result of a *substitution process* for insufficient national of regional funding opportunities, rather than a systematic shift toward (Western) Europe. This trend was probably favoured by the high concentration of publications in top disciplines throughout these Central-Eastern countries (Kozlowski, Radosevic, & Ircha, 1999). When analysing only the publications funded by the European Union, number and co-authoring countries changed in a remarkable way (Braun & Glänzel, 1996) for the Central-Eastern countries.

In more recent years, Kozak, Bornmann, and Leydesdorff (2015) look at Eastern countries, but without disentangling by: *a*) status of the country in relation to the European Union, and *b*) funding agency. These authors conclude that Central-Eastern Europe is still not completely fulfilled in its potential.

International co-authorships and European funding

Some literature aims at demonstrating the importance of European schemes for research (Ovalle-Perandones, Gorraiz, Wieland, Gumpenberger, & Olmeda-Gómez, 2013). However, studies usually ignore the specific funding agencies. The issues of funding agency is either merely exploratory (Wang, Liu, Ding, & Wang, 2012), or is focusing in multi-funding schemes in a specific field, like nanotechnology (Wang & Shapira, 2011). In addition, the centrality in these networks can reveal how a position of a country changes across time. In particular, the fragmentation of the European schemes (Georghiou, 2001) – and the lack of reliable secondary data until recent times – makes assessment of transnational funding agencies even more compelling.

The Framework Programmes along their waves have contributed to expanding the ray of collaborations. The frequency of international collaborations has intensified as well (Scherngell & Lata, 2011). Another study in the field of nanotechnology publications provides insights about the rate of growth by each EU member country, the number of collaborations among countries, and the role of European schemes (Ovalle-Perandones et al., 2013).

Eligibility to funding schemes

On top of the above factor, the specific condition of eligibility for either EU members or non-member States encourages further reasons to understand if, and the extent to which, membership per se yields advantages. Yet, there is dearth of attention in literature about Countries of Central Eastern Europe that are not in the European Union, whether they are Associate Countries of the main EU schemes (i.e. Republic of Serbia), or not (i.e. Russian Federation), but still eligible in a "just-in-case" mode according to specific projects. A notable exception is a study about Israel and EU funding schemes: Israel's increased engagement in science production in collaboration with Europe at the expense of the US has started already in the 1990s (Zimmerman, Glänzel, & Bar-Ilan, 2009). An equally dated study about Switzerland, a country that as well participates in European schemes as an external country, showed that benefits existed, but they were inferior to those yielded by other Western EU member States (Reger, Bührer, Balthasar, & Bättig, 1998). A recent analysis comparing Switzerland and the UK at institutional level confirms some constraints for Switzerland, but a more prominent role played by uncertainty triggered by Brexit (Cavallaro & Lepori, 2021). Nevertheless, Israel and Switzerland

might be defined as Western Countries from a geopolitical and economic point of view. Therefore these findings are not necessarily similar for Central or Eastern European countries.

European Union membership and international collaborations

Many contributions analysed the role of networks (Hoekman, Frenken, & Tijssen, 2010; Kozak, et al., 2015; Luukkonen, Tijssen, Persson, & Sivertsen, 1993; Pajić, 2015; Wagner & Leydesdorff, 2005b), as it is implicit in the concept of (international) collaborations between authors affiliated with different European Countries. To this regard, the enlargement of the European Union towards Central and Eastern Europe has already been analysed from the co-authorships point of view (Makkonen & Mitze, 2016), including Social Network Analysis applications in the global web of outputs (Kozak et al., 2015). Other works focus on the global networks and their centrality indicators, highlighting a space for further competition (Wagner & Leydesdorff, 2005a; Wagner, Whetsell, & Leydesdorff, 2017). In the specific case of European Union partnerships, Frenken (2002), Glänzel, Schubert, and Czerwon (1999), Moed, De Bruin, Nederhof, and Tijssen (1991) and Tijssen (2008) explored some "Europeanisation" via co-authorships, but these studies need updates and preferably the use of less aggregated data.

The issue of the convenience in joining the EU has been underexplored in the topic. An exception is Mattsson (Mattsson, Laget, Nillson, & Sundberg, 2008), giving however no empirical definitive findings. Another limitation is that those figures are also dating back to no later than 2004.

Nevertheless, membership in the European Union is associated with an increased rate of publications, as already discussed for some Eastern European Countries (Teodorescu & Andrei, 2011). Policy makers have realised the necessity to overcome the national perspective, resulting in finding the regional component of Central Eastern EU still unfulfilled in its potential (Zgaga, 2014). Other studies pinpoint another side of the coin. Geodesic distance-related analyses at subnational collaborations also showed that new member States of the EU are catching-up with the "West", finding that the distance factor is being reduced over time (Hoekman et al., 2010). It is also reported that the first 15 EU States to form the EU (EU15) keep higher performances (Hoekman et al., 2010).

The condition of being an Associate Country is potentially interesting. Ülle Must (2006) studied the pace of increase of publications, concluding that Turkey, an Associate Country, was not particularly performing.

The issue of EU membership as possible "treatment" has also deep legacy patterns for countries not in (immediate) prospect of joining the EU such as Ukraine (Davydchyk, Mehlhausen, & Priesmeyer-Tkocz, 2017; Yegorov, 2009).

International co-authorships and influence of research

The influence of research (i.e. number of received citations) when science has an international dimension is manifest and longstanding (Narin, Stevens, & Whitlow, 1991). The issue of influence performed by publications authored in Central-Eastern European countries also resonates in previously discussed studies, highlighting in the first instance West-East collaborations (Braun & Glänzel 1996; Glänzel, Schuert, Czerwon, 1999). Another interesting term of comparison is that between international co-authorships and domestic peripheral circuits (Pajić, 2015). A third example is the analysis of different intramural R&D expenditures (Vinkler, 2008).

Similarly, empirical studies – as Allik (2013) does for the Baltic States or Inzelt et al. for Hungary (2009) – quantify the advantage in getting an international dimension. Mali, Pustovrh, Platinovšek, Kronegger, and Ferligoj (2017) demonstrate for Slovenian publications that productivity and excellence are more likely to happen in co-authored publications, whereas fragmented domestic funding schemes might not secure both. Artificially boosted impact factors of research is also a byproduct of circles of research within Eastern Europe (Teodorescu & Tudorel, 2014), which is detrimental to actual relevance and connectedness to other more influential parts of the world. All these studies are consistent in supporting the possible contribution of a design aimed at testing if non-membership to EU per se plays a positive role or not. This literature also supports the construction of the necessary confounding variables, usually called covariates in quasi-experimental research designs, to set up the best possible tests in answering the research question.

DATASET AND VARIABLES

The original dataset includes the entire set of univocal internationally co-authored publications from 1995 until 2015 (both years included) by any of the EU04 countries or the "control group" (othEast-ERA), unless articles have been published by any possible pair of countries belonging to both of these groups. The total number of these observations is 569,243. Publications that are not internationally co-authored are ignored.

Preparation of dataset

This dataset is an original extraction from Clarivate InCites Web of Science. International co-authored publications are identified by authors' affiliations. The number of publications are the observations, or units of analysis, and they are identified by pairs of countries. In order to answer the research questions, these publications are dived into two main groups. The first group is international co-authored publications by any EU04 Countries. The second group is made up of publications co-authored by "othEast-ERA" Countries as already defined.

For more details about the construction of the dataset, an annex is available.

Variables

Time. Binary time variable refers to publications from 1995 up to 2003 as "before treatment". From 2004 until 2015 publications refer to during "treatment" period. This variable in essential to run a difference-in-differences regression. We refer to 2004–2015 period as a "treatment" one also for othEastERA Some 71% of publications occurred in 2004 or later (see Table 1). Although EU04 Countries joined in the middle of 2004, we select this moment arbitrarily on the evidence that findings don't change if 2003 or 2005 were selected.

Treat. This binary variable comprises co-authored publications with at least one of EU04 countries when the value equals "1". Any co-authored publication with at least one of Eastern-ERA non-member State of EU equals "0". Publications with any country belonging to both sets are excluded as previously stated. As Table 1 shows, around 50% of considered publications are "treated" and other half belongs to "non-treated" publications.

Number of countries in each publication (*No_coll*). This variable takes into account how many Countries (or regions – see further) are present in each internationally co-authored

Table 1. Descriptive statistics

	Obs.	Mean	*S.D*	Min	Max	Kurtosis
Time	569,243	0.7130	*0.4524*	0	1	
Treat	569,243	0.5045	*0.5000*	0	1	
no_coll	569,243	2.5531	*1.279*	2	27	29.38189
CNCI	569,236	1.1459	*3.725*	0	429.5881	2163.117
RDGDP_pub	569,243	0.9619	*0.309*	0.0185	1.9532	
dom_avs	569,243	0.6229	*0.185*	0.2184	0.9870	
fEU	569,243	0.0208	*0.143*	0	1	
fESA	569,243	0.0004	*0.021*	0	1	
fCOST	569,243	0.0006	*0.025*	0	1	
fESF	569,243	0.0017	*0.041*	0	1	
fJRC	569,243	0.0010	*0.032*	0	1	
fERC	569,243	0.0026	*0.050*	0	1	
fEC	569,243	0.0021	*0.046*	0	1	
US	569,277	0.2708	*0.444*	0	1	
EU15	569,277	0.6228	*0.485*	0	1	
DE	569,243	0.2196	*0.414*	0	1	
UK	569,243	0.1242	*0.323*	0	1	
FR	569,243	0.1189	*0.324*	0	1	
IT	569,243	0.0866	*0.281*	0	1	
ES	569,243	0.0533	*0.225*	0	1	
NL	569,243	0.0490	*0.216*	0	1	
SE	569,243	0.0470	*0.212*	0	1	
AT	569,243	0.0370	*0.189*	0	1	
FI	569,243	0.0317	*0.175*	0	1	
BE	569,243	0.0362	*0.187*	0	1	
DK	569,243	0.0230	*0.150*	0	1	
GR	569,243	0.0197	*0.139*	0	1	
PT	569,243	0.0134	*0.115*	0	1	
IE	569,243	0.0081	*0.090*	0	1	
LU	569,243	0.0010	*0.031*	0	1	
ASIA(1)	569,243	0.1411	*0.348*	0	1	
West_AC(2)	569,243	0.0721	*0.259*	0	1	
AFR	569,243	0.0172	*0.130*	0	1	
LAM(3)	569,243	0.0322	*0.176*	0	1	
MID (4)	569,243	0.0163	*0.127*	0	1	
CA	569,243	0.0459	*0.209*	0	1	
EU04_within	569,243	0.0320	*0.176*	0	1	

Source: Own elaboration on InCites WoS.

(1) Countries of Oceania and Asia with exclusion of Middle East and all former Soviet Union Republics.

(2) Israel, Switzerland, Norway, and Iceland.

(3) Latin American Countries, Mexico, and Central America included.

(4) Middle East Countries minus Turkey and Israel.

publication (See Annex for more details). Minimum value is 2 because in order to be internationally co-authored at least two countries should have been found in a single article. This variable is relevant according to several studies. Following Lemarchand's (2012), number of co-authoring countries is determinant on self-organizing networks. More "peripheral" countries are more likely to be present in co-authored publications when networks become more dense (Breschi & Cusmano 2004) – this being in part confuted by Wagner et al. (2017) who discovered a hierarchical pattern within exponential growth and saturation of countries. Large projects covariate with higher probabilities to have more likely, among co-authoring countries, those countries that on average are less performing (Kahn, 2018). Less favoured countries or territories, when co-authoring under EU related schemes, might display a lower rate of growth for other reasons (Lewison, 1993).

CNCI. The Category Normalised Citation Index represents the main bibliometrical indicator that this dataset considers about influence of research. This indicator is released by Clarivate Analytics and normalises citations by time lag and discipline – making citations, from outputs published in different disciplines and different years, comparable. Range is 0 (no citations gained) to infinite. This variable is supposed to provide more fine-grained results when testing the second hypothesis, if compared to similar studies (Khor & Yu, 2016; Persson, Glänzel, & Danell, 2004). *CNCI* allows to disentangle good research (often European level one) from the utmost one (more likely to be led by US based scientists, but also from some of the Western European Countries), avoiding biases highlighted by Rodriguez-Navarro and Narin (2018). Although by definition CNCI equals 1 on average, Table 1 shows different values by set of publications. This figure can be explained by the nature of co-authored publications, which are notoriously on average more successful in gaining citations. In this dataset, the average is in fact above 1 (1.14). The distribution of this variable is exponential as the kurtosis value shows, which is expected considering the aforementioned literature.

RDGDP_pub (Research & Development expenditure as a percentage of GDP). This variable measures the average investments of countries and takes into account different degrees of investments which might be considerably different by country (See Annex for more). Since on average EU04 countries invest more resources for research in comparison to othEast-ERA, the results about whether joining the EU pays off are more consistent by checking for this variable. As shown in Table 1, on average these international publications have a R&D investments as a percentage of GDP that is equal to less than 1% (ERA's aim across these years was to reach 3%).

Dom_avs (Average of domestic, or non-internationally co-authored, publications by country – for more details about source and rationale for including this variable refer to Annex). The membership to the EU in these two hypotheses is expected to increase both values (no_coll and CNCI) after 2004 for EU04 countries, at parity of this variable. Also, this variable considers the averages by country and by the two spans of time under investigation to satisfy the principle of differing conditions not only by groups (treated vs. non-treated), but also by period (pre- vs. post-treatment). Observing Table 1, dom_avs has an average of 0.62 in terms of CNCI, considerably below the average value of the universe of publications, and also much lower than the average of international co-authored publications, which is almost double.

Funding Schemes. InCites Web of Science source can detect several funding agencies acknowledged for every single article. They are: European Research Council (fERC); European Community (fEC); European Commission Joint Research Centre (fJRC); European Social Fund (fESF); European Cooperation in Science and Technology (fCOST); European Space Agency

(fESA); European Union (fEU). By each of these schemes, binary variables have been computed in order to know if each single co-authored paper has been acknowledged as funded by any of these seven schemes. In terms of scientific outputs, fEU is the main contributor that the European Commission has put in place to propel the European Research Area (Commission of the European Communities, 2000; EC, 2000; EC 2012) and to favour the rise in the number of publications. As Table 1 also shows, fEU is present in some 2% of the co-authored publications in the dataset, followed by fERC with 0.5%. All these schemes are present in InCites WoS repositories from early 2000s years, especially from 2007 onwards. These variables are essential as confounding variables because schemes can be accessed by any of the ERA country, and any scientist in the world may be involved in co-authored publications funded by any of these schemes.

As Table 1 shows, in publications by EU04 countries and by other Eastern ERA countries (othEast-ERA), more than 60% of these publications are signed with at least one of EU15 countries. The therein list of the 15 countries is sorted in descending order, listing Germany, UK, France, and Italy as the first four ones – the total exceeds the figure for EU15 due to possible multiple co-authorships. In 27% of the cases, these countries co-authored with the US. In 14% of the cases they co-authored with one of the Australasian countries. Much less frequently East-ERA countries (othEast-ERA) co-authored with other countries. The case of co-authorships only within EU04 countries occurs only 4.5% of the times.

DESCRIPTIVE STATISTICS BY TIME AND TREATMENT, AND DIFFERENCE IN DIFFERENCES HYPOTHESES

This section provides some further descriptive statistics (see Fig. 1 and Table 2).

From Fig. 1 it is possible to note that for both dependant variables there is a notable rise for countries having joined the EU in 2004, and slightly less for East-ERA ones. Yet, at a glance it is hard to understand which of the two lines had a steeper growth, once selecting 2004 year as a diriment point in time. Table 2 shows figures by each of the countries composing both groups. For each of the 10 accessing countries in 2004, there has been a growth both between 1995–2003

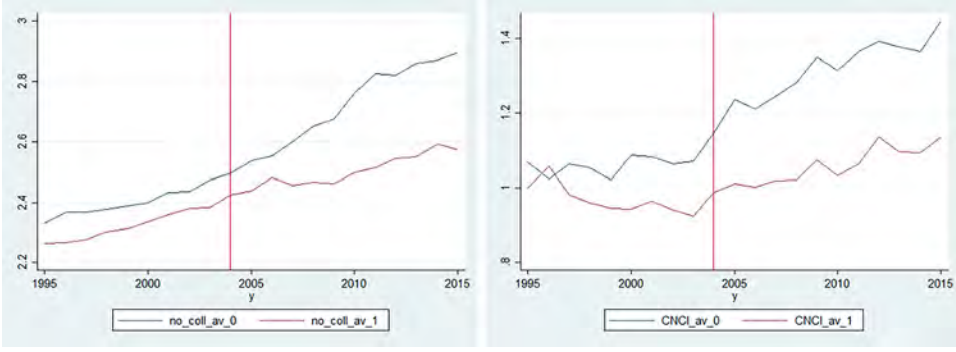

Fig. 1. Time series (year) of two dependant variables by each of the two groups of variable *treat*, 2004 year is the commencing of "treatment"
Source: own elaboration on InCites Web of Science.

Table 2. Totals by treatment variable ("treat") with absolute numbers and respective averages by the two periods. Details of single countries composing "treat" variable

	N		*no_coll*		*CNCI*	
	1995–2003	2004–2015	1995–2003	2004–2015	1995–2003	2004–2015
PL	30,373	79,117	2.4235	2.8321	1.0220	1.4063
LT	1,951	7,045	2.4464	2.7766	0.8803	1.1343
EE	2,336	8,423	2.5364	3.2557	1.1347	1.6686
CZ	16,126	53,927	2.4256	2.8139	1.0665	1.3929
SK	6,998	17,741	2.4021	2.6175	0.9095	0.9794
LV	1,354	3,077	2.4387	3.0208	0.9261	1.1713
SI	3,790	14,629	2.6003	2.8437	1.1610	1.2650
HU	17,500	38,983	2.5090	2.8860	1.1681	1.3322
MT	196	1,368	2.5969	3.1031	1.7894	1.3299
CY	884	5,987	2.4106	2.7697	1.0524	1.2879
Treat = 1	*77,147*	*210,029*	*2.4033*	*2.7414*	*1.0610*	*1.3280*
RU	64,640	109,630	2.3457	2.5466	0.9639	1.0585
BY	3,194	5,415	2.3087	2.4334	0.8902	0.6964
TR	10,322	50,512	2.2627	2.4743	1.0475	1.1717
UA	10,676	20,297	2.3300	2.5092	0.7454	0.8452
BA	129	1,648	2.3101	2.3259	1.6023	0.6251
ME	nil	939	nil	2.3163	nil	0.6435
MK	309	1,529	2.3689	2.6181	1.0422	0.9800
AL	216	891	2.2500	2.4265	0.6971	0.6968
MD	711	1,645	2.4219	2.5739	0.6579	1.1138
othCIS(1)	2,592	7,394	2.7998	2.9478	0.7597	0.8754
RS	660	12,489	1.4939	2.4623	0.8149	0.9802
Treat = 0	*86,237*	*195,823*	*2.3280*	*2.5091*	*0.9639*	*1.0641*

(1) For variable CNCI in post 2004 time series some 7 values are missing.
Source: own elaboration on InCites Web of Science.

and 2004–2015. Nevertheless, similar patterns are present in the other East-ERA countries, though some of them have improved much more than others – the case of Serbia for number of collaborating countries and Moldova for the influence of its research are points in case of higher increases if compared also with countries that now are part of the EU. Notably, Bosnia-Herzegovina decreased sensibly its influence in international co-authored publications between these two sets of years. These figures encourage the adoption of a difference-in-differences data analysis.

A basic model of difference-in-difference for both dependant variables confirms the hypothesis that having acquired the status of member of EU yielded advantages both in terms of average number of co-authoring countries, and also in terms of average influence of research in co-authored publications. Table 3 shows the values for these tests whose averages are the same as reported in Table 2.

Considering the exposed literature, further analyses are required, namely checking by the main covariates the current dataset of microdata offers. In particular, hypotheses are tested following these equations:

Table 3. Difference-in-Difference tests for both hypotheses without covariates

	Time	Treat 1–Treat 0	S.err.	|t|	P>|t|
No_coll	t_0	0.075	0.006	11.97	0.000
	t_1	0.232	0.004	58.26	0.000
	Diff-in-diff	0.157	0.007	21.08	0.000
CNCI	t_0	0.097	0.018	5.26	0.000
	t_1	0.264	0.012	22.57	0.000
	Diff-in-diff	0.167	0.022	7.64	0.000

Source: Own elaboration on InCites Web of Science.

$$\text{No\_coll} = \beta_0 + \beta_1 \text{RDGDP\_pub} + \beta_2 \text{dom\_avs} + \beta_3 \text{treat*time} + \beta_4 \text{time} + \beta_5 \text{treat} + \sum_i^{world} \beta \text{CU} + \sum_j^{n} \beta \text{Fund} + \varepsilon_i \quad (1)$$

$$\text{CNCI} = \beta_0 + \beta_1 \text{RDGDP\_pub} + \beta_2 \text{dom\_avs} + \beta_3 \text{no\_coll} + \beta_4 \text{treat*time} + \beta_5 \text{time} + \beta_6 \text{treat} + \sum_i^{world} \beta \text{CUi} + \sum_j^{n} \beta \text{Fund} + \varepsilon_i \quad (2)$$

*Treat*time* is the combination of the two binary variables *treat* and *time* – one of the usual ways to connotate the "interaction" between "treatment and time" in these tests. *CU* defines the Countries, or Regions, in analysis to predict the extent to which the identity of co-authoring countries is relevant to predict the dependant variables in question. *Fund* defines the series of seven binary variables regarding the EU funding schemes. RDGDP_pub and dom_avs are the variables already explained in the previous section. The second hypothesis includes also the dependant variable of the first hypothesis on the ground that influence of research is arguably predictable also by the increase of countries that are present in co-authored publications.

RESULTS

Table 4 provides the results for hypothesis (1). The main difference in differences interaction variable (EU04#time) displays a statistically significance interaction as also found in Table 3 that had no covariates. It is possible to observe that "domestic research at country level" (dom_avs) and investment in research (RDGDP_pub) have both significant and negative coefficients, meaning that the less domestic research in a country is influent, the more likely those countries will find themselves in publications co-authored by a larger number of co-authoring countries. Since marginal countries are more likely to be in co-authored publications whenever articles are published by many countries at the same time (Leydesdorff & Wagner, 2008), this result is to some extent expected. Similarly, this is true also for the other variable concerning expenditure in RD: the effect of joining the EU is stronger for those countries that are weaker in funding research at national level. Remarkably, these two variables do not interfere with the validity of the difference-in-differences regression, considering that EU04 countries have on average higher influence in domestic publications (dom_avs). At the same time, EU04 countries increased expenditure in RD, whereas other Central Eastern European countries composing the control group had their expenditure in RD as a percentage of GDP decreased.

Table 4. Difference in differences regression for hypothesis regarding number of countries in co-authored international papers with confounding variables. Margins of treat*time computed at the bottom of the table

				OBS.	569,243
Source	SS	df	Ms	F (21, 569,221)	34462.90
Model	521073.46	21	24813.023	Prob > F	0.0000
Residual	409834.74	569,22	71999230	R^2	0.5597
Total	930909.0	569,24	1.63535	Adj R^2	0.5597
				Root MSE	0.84852

no_coll	Coef.	Std.Err	t	$P>\|t\|$	[95% Conf. Int.]	no_coll
1.EU04	−0.024	0.006	−4.320	0.000	−0.035	−0.013
1.time	0.064	0.004	18.370	0.000	0.058	0.071
EU04#time	0.140	0.005	27.050	0.000	0.130	0.150
dom_avs	−0.027	0.011	−2.410	0.016	−0.049	−0.005
RDGDP_pub	−0.054	0.004	−13.910	0.000	−0.061	−0.046
fEU	0.031	0.008	3.930	0.000	0.016	0.047
fERC	0.071	0.022	3.160	0.002	0.027	0.114
fEC	−0.029	0.025	−1.170	0.243	−0.077	0.020
fJRC	0.036	0.035	1.040	0.299	−0.032	0.105
fCOST	0.003	0.045	0.060	0.955	−0.085	0.090
fESA	0.096	0.054	1.780	0.075	−0.010	0.201
fESF	−0.038	0.027	−1.390	0.163	−0.092	0.015
US	1.265	0.003	441.350	0.000	1.260	1.271
EU15	1.534	0.003	542.100	0.000	1.529	1.540
ASIA	1.295	0.003	381.210	0.000	1.288	1.302
West_AC	1.701	0.004	383.610	0.000	1.693	1.710
MID	1.181	0.009	130.770	0.000	1.163	1.198
LAM	1.360	0.006	212.430	0.000	1.348	1.373
AFR	1.240	0.009	142.300	0.000	1.223	1.257
CA	1.495	0.005	275.520	0.000	1.485	1.506
EU04_within	0.664	0.007	92.380	0.000	0.650	0.678
_cons	0.758	0.007	108.410	0.000	0.744	0.771

	Margin	S.err.	t	$P>\|t\|$	[95% Conf. Int.]	
0 0 (nonEU04; t_0)	2.468	0.003	706.830	0.000	2.461	2.475
0 1 (nonEU04; t_1)	2.532	0.003	995.420	0.000	2.527	2.537
1 0 (EU04; t_0)	2.443	0.004	682.370	0.000	2.436	2.450
1 1 (EU04; t_1)	2.648	0.003	1036.060	0.000	2.643	2.653

Source: own elaboration on InCited WoS data.

Regarding the funding schemes, only two schemes over seven are statistically significant: ERC and EU – being both positive. ESA would be statistically significant if threshold were considered at $P < 0.1$ level. However, keeping constant the publications stemmed from participation in EU funding schemes projects, it is visible that funding schemes as a whole do not predict the extent to which publications are co-authored by more countries at the same time.

On the other hand, according to the definition of co-authored publications and number of co-authoring countries in each publication, the set of variables regarding co-authoring countries

displays statistically valid slopes. They also contribute to a large extent in reaching more than 50% of variance explained in the model ($R^2 = 0.5597$). In particular, having joined the EU by the accessing ten countries in 2004 made them have more collaborations, especially with Western Associate Countries (1.70) and then with any of EU15 ones (1.53). With a lower coefficient, EU04 countries' increased likelihood to co-author with US-based scholars emerges as well (1.27), whereas also publications co-authored only within EU04 has a positive statistical coefficient, despite being much lower (0.66).

Table 5. Difference in differences regression for hypothesis regarding influence of research with confounding variables. Margins of "treat*time" computed at the bottom of the table

				OBS.	569,236
Source	SS	df	Ms	F (22, 569,213)	784.63
Model	232525.68	22	10569.35	Prob > F	0.0000
Residual	7667531.0	569,2	13.47040	R^2	0.0294
Total	7900056.7	569,2	13.87837	Adj R^2	0.0294
				Root MSE	3.6702

CNCI	Coef.	Std.Err.	t	$P>\|t\|$	[95% Conf.	Int.]
1.EU2004	−0.059	0.024	−2.440	0.015	−0.107	−0.012
1.time	0.006	0.015	0.390	0.696	−0.024	0.036
EU2004#time	0.099	0.022	4.450	0.000	0.056	0.143
no_coll	0.369	0.006	64.440	0.000	0.358	0.381
dom_avs	0.388	0.048	8.010	0.000	0.293	0.483
RDGDP_pub	0.044	0.017	2.620	0.009	0.011	0.077
fEU	0.083	0.034	2.420	0.016	0.016	0.151
fERC	−0.051	0.097	−0.520	0.600	−0.240	0.139
fEC	−0.039	0.107	−0.360	0.717	−0.248	0.171
fJRC	0.099	0.152	0.650	0.514	−0.198	0.397
fCOST	−0.240	0.194	−1.240	0.215	−0.619	0.139
fESA	−0.178	0.233	−0.760	0.445	−0.634	0.278
fESF	0.085	0.118	0.720	0.470	−0.146	0.316
US	0.472	0.014	32.860	0.000	0.444	0.500
EU15	0.252	0.015	16.710	0.000	0.222	0.281
ASIA	0.212	0.016	12.870	0.000	0.180	0.244
West_AC	0.154	0.022	7.180	0.000	0.112	0.197
MID	0.075	0.040	1.900	0.058	−0.002	0.153
LAM	0.174	0.029	6.030	0.000	0.117	0.230
AFR	−0.075	0.038	−1.960	0.050	−0.150	0.000
CA	0.376	0.025	15.050	0.000	0.327	0.425
EU04_within	0.126	0.031	4.030	0.000	0.065	0.188
_cons	−0.446	0.031	−14.610	0.000	−0.506	−0.386

	Margin	Std.Err.	t	$P>\|t\|$	[95% Conf.	Int.]
0 0 (notEU04; t_0)	1.103	0.015	72.800	0.000	1.074	1.133
0 1 (notEU04; t_1)	1.133	0.011	102.610	0.000	1.111	1.155
1 0 (EU04; t_0)	1.035	0.016	66.590	0.000	1.005	1.066
1 1 (EU04; t_1)	1.216	0.011	109.600	0.000	1.194	1.238

Source: own elaboration on InCite WoS data.

Overall, this first hypothesis shows that the number of co-authoring countries for EU04 countries increased more than those in other Eastern ERA countries since 2004 on, when compared between 1995 and 2003 years against 2004–2015.

Table 5 shows the result of the difference–in-differences for equation (2). In general, this model, though consistent (Prob > F = 0.0000), is poor in explaining the variance (R^2 is lower than 3%). Although the difference-in-differences is statistically significant, the extent to which EU04 countries have been able to increase their influence in co-authored publications in comparison to other Eastern ERA countries is small, and possibly dependant by unobserved variables. In this case the main covariates are: the presence of the US (coefficients equals 0.47 in Table 5) as a co-authoring country, the presence of Canada as a co-authoring country (0.38), and co-authoring with any EU15 country (0.25). The average of domestic influence (dom_avs: 0.39) is one of the best predictors, and number of co-authoring countries (no_coll: 0.37) as well – confirming literature already discussed. European funding schemes do not give particular indication in terms of influence, probably due to the presence of other more performing national schemes aimed at simply boosting research regardless the multiple ends ERA declared to have. Also co-authoring with other Western European countries (West_AC), though positive and significant, is not particularly high (0.15). On top of these considerations, it is worth remembering that this small increase in terms of influence is gained at parity of number of co-authoring countries (no_coll), which implies that although small, membership to EU is a positive contribution gained at parity of the advantage found in testing hypothesis 1.

With this second hypothesis, we try to test if the influence of research might be caused by simply joining the EU. Results show a positive effect, though much less prominent if compared to the capacity that EU membership yields in terms of expanding members' presence in international co-authored publications (first hypothesis). This is consistent also once checked by the covariates in analysis, meaning that although networks of co-authored publications might increase due to, for instance, shared infrastructures at regional level, this may not necessarily have a remarkable proportional reflection in the way scientists appreciate these publications.

At the bottom of both Tables 4 and 5, margins of treatment by time can be observed. Both hypotheses are confirmed as statistically significant, as Table 3 also displays without considering any covariates. Figure 2 to this regard plots the results of the difference-in-difference tests scored

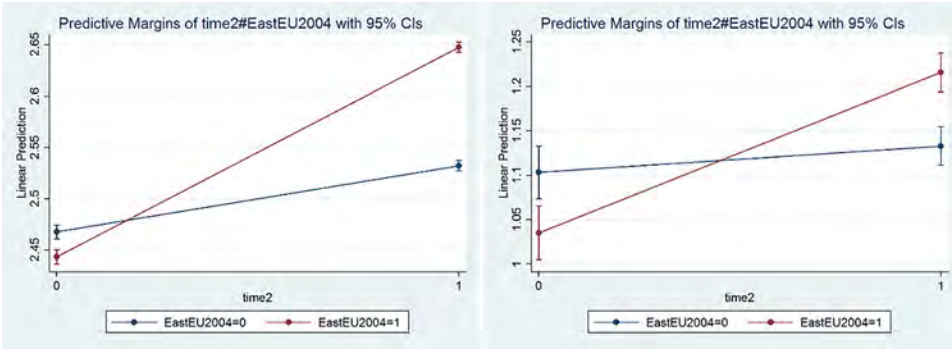

Fig. 2. Predictions of the two independent variables for the two groups of observations for years 1995–2003 and 2004–2015, considering covariates indicated in respectively Tables 4 and 5

out of Tables 4 and 5. For both figures one may appreciate the extent to which respective hypothesis are confirmed, with a much larger effect for the first hypothesis.

DISCUSSION AND CONCLUSIONS

Following Wagner and Leydesdorff (2005a), the main pattern in international collaborations among scholars based in different countries is that international co-authored publications is a self-performing network, which results from free choices (autopoiesis). Nevertheless, ERA, EU, and its funding schemes, due to some specific features and declared aims, ought to shape networks integrating the European space. ERA is also an attempt to broaden the borders of EU with its external Associate Countries and other countries geographically located in Europe and beyond. This paper tries to assess if and to what extent EU membership has been able to set an advantage in comparison to other non-EU member States by means of expanding the number of countries co-authoring international publications (hypothesis 1), and also raising the respective influence of research (hypothesis 2). This study uses an increasingly popular quasi-experimental test (difference-in-differences) to answer this question, taking into account pre and post entry in the EU by the ten countries that joined in 2004, comparing them with other European countries that, at least until now, did not join the EU. Checking also by several covariates, which are arguably the most relevant according to literature, the paper brings to the conclusion that joining the EU brought advantages in term of expanding countries' networks, allowing more dense webs of co-authored publications. Moreover, there is a significant, tough less prominent, gain in terms of influence (CNCI – normalised citations by discipline and years after publication). These findings are relevant considering that the openness of the labour market of academics within ERA has potential detrimental consequences just for those countries which perform less well than Western ones (e.g., best researchers based in some countries poached by some Western universities). As Chessa et al. demonstrate (2013), some countries (EU04 among them) witness impoverishment, rather than boosting science, because of the naturally asymmetrical fluxes of scientists' mobility, especially if this sort of mobility is incentivised. In such a way, this paper shows that some possible negative consequences in terms of brain drain are completely counterbalanced. Any analysis of ERA, on the other hand, is necessarily multidimensional, and this study focuses on only a specific component of its goals. Any assessment of other specific aspects of goals promoted by ERA would require further and different research designs.

This study gives original contributions to the topic of co-authored publications from a Central and Eastern European perspective. Although other studies take into account the construction of ERA with a focus on international research, this paper gives novel insight, as ittakes into account more possible covariates and it analyzes microdata instead of aggregates. It is also a different sort of analysis in comparison to recent studies (Arrieta et al., 2017) because the percentage of international papers over the total publications is a different variable from the number of countries in internationally co-authored publications. Although higher participation in saturation is more likely to be a sign of marginality rather than of strength and centrality, the construction of ERA has international co-authoring publications as one of its political aims, which does not necessarily impair other goals – that of influence (or "excellence" in nowadays common political jargon). At parity of countries in terms of relevance and investment, it is

interesting to observe if accessing countries (EU04) are improving more their density in global publications or not. This means that a country in Central Eastern Europe is not better off if it is able to have less often access to multiple co-authoring opportunities. This latter condition may be provided by funding schemes, projects, physical and regulative infrastructures which ERA tries to offer also beyond EU borders. Nevertheless, membership as such to EU appears to guarantee this in a more effective way, at least for the years analysed in this paper and provided other unobserved conditions (i.e. social, economic, demographic, political or other factors) might have played a role in engendering these results.

Ultimately, the paper finds that accessing EU has been a positive change, other factors kept constant. In these terms, EU membership has brought some advantages for those EU04 countries, although othEastERA countries equally could access participation in projects. These advantages, namely the capacity to increase the ray of collaborations among scientists, is a relevant advantage that will require further research in the future to unpack if, for instance, EU membership happened along with other changes.

These findings parallel a recent work about patents (De Noni, Orsia, & Belussi, 2018). In terms of de-fragmentation, this study gives some elements in favour of other empirical works (Leydesdorff & Wagner, 2008), highlighting the way the global network of science is increasing, and giving insights about the role ERA might have in counterbalancing some of the non-desirable consequences of this global growth.

In terms of policy recommendations, this paper can draw attention to what would happen if EU opted for a two speeds process of integration (EC 2017). It is possible in fact to imagine that desirable and undesirable consequences are the final outcomes of multiple initiatives whose respective evaluations are more complicated than a simple quasi-experimental design can offer. Also, any debate about extra-European Countries (i.e. BRICS Countries) wishing to strengthen the ties with ERA might find insights about how and the extent to which they might benefit from strengthening links with ERA.

This paper has a list of limitations we need to account. It does not disentangle by every single country of the world, keeping, for instance, Australasian countries aggregated. Further analyses may try to understand the contribution of Chinese co-authored publications for instance. Moreover, this study has dropped Romania, Bulgaria and Croatia, opting for simplifying the research design and arguing that a census of more than a half million observations is reliable and valid enough. For the sake of future analyses, it would be worth understanding whether having joined in different moments the EU produces similar results. Another limitation is the following: although ERA established co-authored publications as a goal, co-authored publications are likely to depend also by the same state-of-the-art of the integration process that ERA itself is monitoring. Some indicators about the degree of accomplishment of the ERA's goals at country level would shed further light, if combined with advanced bibliometric research design. Yet, this paper is based on the fact that entering the EU establishes a set of opportunities that go beyond the participation in single EU funding schemes. Nevertheless, it is also true that single countries might pursue policies of more openness in general, and in science in particular, regardless of their option to join the EU. Also, the same preparation to join the EU may be considered an approximation of the process of sharing resources, opportunities, and more broadly values. In this sense the year 2004, though a watershed for those countries and for the EU as a whole, does not grasp entirely the complexity of the phenomenon, which implied changes before and also after the EU

accession. Overall, further analyses should take into account more indicators, different measures, and also respective different statistical approaches. This paper also overlooks some asymmetrical advantages possibly derived from different patterns (Glänzel, 2001). In terms of triggered initiatives of collaborations, co-authorships in this dataset do not indicate who the first author is, which is a proxy of triggering initiative (Wang & Wang, 2017). Moreover, tailored studies ought to test if each European funding scheme is serving effectively its declared scopes – i.e. ERC publications (EC 2015), which is here overlooked. Although schemes have different rationales and aims, any analysis by a funding agency should preferably look also at productivity. Also, the institutional level, rather than the country level, may expand knowledge in the topic. Last, the paper overlooks the possible role of different domestic funding agencies, which might have been established with overt international aims.

ACKNOWLEDGEMENT

The support of the Economic and Social Research Council, the Office for Students and Research England (grant reference ES/M010082/1) is gratefully acknowledged along with support from the Centre for Global Higher Education (CGHE), UCL Institute of Education, London.

ABOUT THE AUTHOR

Dr Giulio Marini is from October 2020 Lecturer (Teaching) at the Social Research Institute – Faculty of Education and Society UCL. Previously he was Research Associate at the Centre for Global Higher Education, EPS Department, IoE UCL since 2016. He has previously worked in post-doctoral positions at Scuola Normale Superiore, Pisa (Italy), Centro de Investigação de Políticas do Ensino Superior (CIPES), Porto (Portugal), The National Research Council (Italy) and Sapienza University (Italy), where he got his PhD in Methodology for Social Sciences.

REFERENCES

Allik, J. (2013). Factors affecting bibliometric indicators of scientific quality. *Trames*, *17*(67/62), 3, 199–214.

Arrieta, et al. (2017). Quantifying the negative impact of brain drain on the integration of European science. *Science Advances*, *3*(4), e1602232.

Braun, T., & Glänzel, W. (1996). International collaboration: Will it be keeping alive East European research? *Scientometrics*, *36*(2), 247–254.

Breschi, S., & Cusmano, L. (2004). Unveiling the texture of a European Research Area: Emergence of oligarchic networks under EU framework Programmes. *International Journal of Technology Management*, *27*(8), 747–772.

Cavallaro, M., & Lepori, B. (2021). Institutional barriers to participation in EU framework programs: Contrasting the Swiss and UK cases. *Scientometrics*, *126*, 1311–1328. https://doi.org/10.1007/s11192-020-03810-0.

Chessa, A., Morescalchi, A., Pammolli, F., Penner, O., Petersen, A. M., & Riccaboni, M. (2013). Is Europe evolving toward an integrated research area? *Science*, *339*, 650–651.

Commission of the European Communities (2000). *Presidency conclusions of the Lisbon European Council*. Brussels: Commission of the European Communities.

Davydchyk, M., Mehlhausen, T., & Priesmeyer-Tkocz, W. (2017). The price of success, the benefit of setbacks: Alternative futures of EU-Ukraine relations. *Futures*, *95*, 35–46.

De Noni, I., Orsia, L., & Belussi, F. (2018). The role of collaborative networks in supporting the innovation performances of lagging-behind European regions. *Research Policy*, *47*(1), 1–13.

European Commission (2000). *Towards a European Research Area*. Bruxelles.

European Commission (2007). *Green paper. The European Research Area: New perspectives*. COM. 2007. 161 final, 4th April.

European Commission (2012). *A reinforced European Research Area partnership for excellence and growth*. COM 394 COM, 392 final, 17th July.

European Commission (2015). *Comparative scientometric assessment of the results of ERC funded projects (D5 report)*. Luxemburg.

European Commission (2016). *ERA progress report 2016. Science metrix study. Data gathering and information for the 2016 ERA monitoring*. Technical Report. Bruxelles, https://doi.org/10.2777/35265.

European Commission (2017). *Reflections and scenarios for the EU27 by 2025*. Bruxelles: White Paper.

European Commission (2020). *Towards a 2030 vision on the future of universities in Europe*. Luxembourg: Policy Report.

Frenken, K. (2002). A new indicator of European integration and an application to collaboration in scientific research. *Economic Systems Research*, *14*, 345–361.

Georghiou, L. (2001). Evolving frameworks for European collaboration in research and technology. *Research Policy*, *30*, 891–903.

Glänzel, W. (2001). National characteristics of international scientific co-authorship. *Scientometrics*, *66*, 231–240.

Glänzel, W., & Schubert, A. (2001). Double effort = double impact? A critical view at international co-authorship in chemistry. *Scientometrics*, *50*(2), 199–214.

Glänzel, W., Schubert, A., & Czerwon, H. J. (1999). A bibliometric analysis of international scientific cooperation of the European union, 1985–1995. *Scientometrics*, *45*(2), 185–202.

Gopalan, M., Rosinger, K., & Ahn, J. B. (2020). Use of quasi-experimental research designs in education research: Growth, promise, and challenges. *Review of Research in Education*, *44*(1), 218–243. https://doi.org/10.3102/0091732X20903302.

Hoekman, J., Frenken, K., & Tijssen, R. J. W. (2010). Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy*, *39*(5), 662–673.

Inzelt, A., Schubert, A., & Schubert, M. (2009). Incremental citation impact due to international co-authorship in Hungarian higher education institutions. *Scientometrics*, *78*(1), 37–43.

Kahn, M. (2018). Co-Authorship as a proxy for collaboration: A cautionary tale. *Science and Public Policy*, *45*(1), 117–123.

Khor, K. A., & Yu, L.-G. (2016). Influence of international co-authorship on the research citation impact of young universities. *Scientometrics*, *107*(3), 1095–1110.

Kozak, M., Bornmann, L., & Leydesdorff, L. (2015). How have the eastern European countries of the former warsaw pact developed since 1990? A bibliometric study. *Scientometrics*, *102*, 1101–1117.

Kozlowski, J., Radosevic, S., & Ircha, D. (1999). History matters: The inherited disciplinary structure of the post-communist science in countries of central and Eastern Europe and its restructuring. *Scientometrics*, *45*(1), 137–166.

Lemarchand, G. (2012). The long-term dynamics of co-authorship scientific networks: Iberoamerican Countries (1973–2010). *Research Policy*, *41*, 291–305.

Lepori, B., Reale, E., & Larédo, P. (2014). Logics of integration and actors' strategies in European joint programs. *Research Policy*, *43*(2), 391–402.

Lewison, G. (1993). The contribution of European community less favoured region research outputs to economic and social cohesion. *Scientometrics*, *28*(2), 217–229.

Leydesdorff, L., & Wagner, C. (2008). International collaboration in science and the formation of a core group. *Journal of Informetrics*, *2*(4), 317–325.

Luukkonen, T., Tijssen, R., Persson, O., & Sivertsen, G. (1993). The measurement of international scientific collaboration. *Scientometrics*, *28*(1), 15–36.

Makkonen, T., & Mitze, T. (2016). Scientific collaboration between 'old' and 'new' member states: Did joining the European Union make a difference? *Scientometrics*, *106*, 1193–1215.

Mali, F., Pustovrh, T., Platinovšek, R., Kronegger, L., & Ferligoj, A. (2017). The effects of funding and co-authorship on research performance in a small scientific community. *Science and Public Policy*, *44*(4), 486–496.

Marini, G. (2018). International co-authorships and the role of the European union as a funder: An eastern European perspective. *CGHE Working Paper*, *39*, 1–24.

Mattsson, P., Laget, P., Nillson, A., & Sundberg, C. J. (2008). Intra-EU vs. extra-EU scientific co-publication patterns in EU. *Scientometrics*, *75*(3), 555–574.

Moed, H. F., De Bruin, R. E., Nederhof, A. J., & Tijssen, R. J. W. (1991). International scientific cooperation and awareness within the European community: Problems and perspectives. *Scientometrics*, *21*, 291–311.

Must, U. (2006). New" countries in Europe—research, development and innovation strategies vs bibliometric data. *Scientometrics*, *6*(2), 241–248.

Narin, F., Stevens, K., & Whitlow, E. S. (1991). Scientific cooperation in Europe and the citation of multinationally authored papers. *Scientometrics*, *21*(3), 313–323.

Nedeva, M. (2012). Between the global and the national: Organising European science. *Research Policy*, *42*(1), 220–230.

Ovalle-Perandones, M. A., Gorraiz, J., Wieland, M., Gumpenberger, C., & Olmeda-Gómez, C. (2013). The influence of European Framework Programmes on scientific collaboration in nanotechnology. *Scientometrics*, *97*(1), 59–74.

Pajić, D. (2015). Globalization of the social sciences in eastern Europe: Genuine breakthrough or a slippery slope of the research evaluation practice? *Scientometrics*, *102*(3), 2131–2150.

Persson, O., Glänzel, W., & Danell, R. (2004). Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, *60*(3), 421–432.

Reger, G., Bührer, S., Balthasar, A., & Bättig, C. (1998). Influence of non-membership of the European Union on collaboration in European R&D networks: The case of Switzerland. *Science and Public Policy*, *25*(3), 171–183.

Rodriguez-Navarro, A., & Narin, F. (2018). European paradox or delusion—are European science and economy outdated? *Science and Public Policy*, *45*(1), 14–23.

Scherngell, T., & Lata, R. (2011). Towards an integrated European Research Area? Findings from eigenvector spatially filtered spatial interaction models using European framework programme data. *Papers in Regional Science*, *92*(3), 555–577.

Smith, E., Franke, J., Jarvis, A., Sureka, P., Chirico, S., Peter, V., Simmonds, P., & Kolar, P. (2015). *Assessment of progress in achieving ERA in member states and associated countries. Final report to DG research and innovation*. London: ICF & Technopolis.

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*, 711–722.

Teodorescu, D., & Tudorel, A. (2014). An examination of "citation circles" for social sciences journals in Eastern European countries. *Scientometrics*, *99*(2), 209–231.

Tijssen, R. J. W. (2008). Are we moving towards an integrated European Research Area? *COLLNET Journal of Scientometrics and Information Management*, *2*(1), 19–25.

Ulnicane, I. (2015). Broadening aims and building support in science, technology and innovation policy: The case of the European Research Area. *Journal of Contemporary European Research*, *11*(1), 31–49.

Vinkler, P. (2008). Correlation between the structure of scientific research, scientometric indicators and GDP in EU and non-EU countries. *Scientometrics*, *74*(2), 237–254.

Wagner, C., & Leydesdorff, L. (2005a). Network structure, self-organization, and the growth of international collaboration in science. *Research Policy*, *34*, 1608–1618.

Wagner, C., & Leydesdorff, L. (2005b). Mapping the network of global science: Comparing international co-authorships from 1990 to 2000. *International Journal of Technology and Globalisation*, *1*(2), 185–208.

Wagner, C., Whetsell, T. A., & Leydesdorff, L. (2017). Growth of international collaboration in science: Revisiting six specialties. *Scientometrics*, *110*(3), 1601–1614.

Wang, X., Liu, D., Ding, K., & Wang, X. (2012). Science funding and research output: A study on 10 countries. *Scientometrics*, *91*, 591–599.

Wang, J., & Shapira, P. (2011). Funding acknowledgement analysis: An enhanced tool to investigate research sponsorship impacts: The case of nanotechnology. *Scientometrics*, *87*(3), 563–586.

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.

Yegorov, I. (2009). Post-Soviet science: Difficulties in the transformation of the R&D systems in Russia and Ukraine. *Research Policy*, *38*, 600–609.

Zgaga, P. (2014). The role of higher education centres in research and policy: A case from a European periphery. *Studies in Higher Education*, *39*(8), 1393–1404.

Zimmerman, E., Glänzel, W., & Bar-Ilan, J. (2009). Scholarly collaboration between Europe and Israel: A scientometric examination of a changing landscape. *Scientometrics*, *78*(3), 427–446.

Appendix

Additionals notes on Dataset construction

The extraction of publications depends by how searches are launched over the repository selected (InCite Clarivate). Since it is possible to search any publication authored by people affiliated in Country "A" and Country "B", it is possible to replicate this search multiple times

keeping of changing other features (e.g., years of publication). This search permits to have outputs co-authored by two given Countries – {country_A ∩ country_B} in set theory notation. For simplicity we considered 49 Countries or Regions for comprehensive worldwide publications. This operation allows to have a relatively low number of dyads [$(N^2-N)/2 = 2{,}352$]. All these dyads have been extracted. For a handful of these extractions, some further launches were necessary due to the limit of 50.000 outputs per time allowed by InCites. Last, for Eastern European Countries, all publications, domestic included, have been extracted in order to be able to compute separately *dom_avs* variable as already described (see Variables section).

The first 15 Countries to join the EU have been grouped in the analysis with the label "EU15". For the case of EU04 also a further binary variable is used. *EU04_within* is computed in order to disentangle those co-authored publications which happened *only* within the 10 Countries comprising EU04, ackowledging literature about this possible type of networking (Teodorescu & Tudorel, 2014). This latter distinction is consistent also with other literature (Glänzel & Schubert, 2001) contemplating the possibility of asymmetrical patterns in co-authorships, and also with the more general assumption that co-authored publications may happen more likely when geographical distance is shorter (Hoekman et al., 2010).

The list of dummy variables are kept in the dataset to identify with which country each publication is co-authored by. This list was also used to elimite duplications,[2] by means of collapsing publications by their univocal Web of Science identifier. The counting variable out of this process equals the generation of the variable indicating the number of Countries present in each publication (no_coll; see below).

Further details about variables.

No_coll. This variable is computed collapsing observations by unique identifier (a univocal string code released by Web of Science) and consequently generating a variable that counts the number of duplications.

RDGDP_pub. This variable represents the public expenditure in R&D as a percentage of GDP by countries – World Bank source. Averages between 1995 and 2003, and between 2004 and 2015 are computed separately. Since co-authored publications are based on more countries by definition, the value by each observation is an average among any country present in the publication.

Dom_avs. This variable is computed from the average of CNCI among domestic publications of each country under analysis (EU04 and "control group"). This figure is computed from a different dataset. The average of non-internationally co-authored publications is used as covariate for both hypotheses. For the number of co-authoring countries (no_coll – hypothesis 1), the higher the average of influence of domestic publications, the higher the probability that scholars in that country may find the opportunity to co-author a paper with scholars in other countries. Consequently, the lower the average of influence in domestic publications, the higher the probabilities that publications from that country are present only when co-authorships are close to saturation (Breschi & Cusmano, 2004). The argument followed by these authors is that a co-authored publication in a certain country happens more likely when scholars in that country are on average more influential than other countries in domestic publications. For the

[2]Duplications are expected on the ground that within all publications co-authored by County A and Country B there could be also some publications co-authored by Country A and Country C. In *set theory* notation: {country_A ∩ country_B}/{country_A ∩ country_C} ≠ ∅.

hypothesis regarding influence of research (CNCI as dependent variable), the higher the domestic influence, the higher also the influence of the publications co-authored internationally. This assumption is also consistent with the main literature concerning the autopoietic nature of international scientific collaborations, as recalled in the literature review.

More information about the dataset are available upon request.