Analysis of the Spatial Dynamics of Intra- v.s. Inter-Research Collaborations across Countries[1]

Lili Wang[1] and Mario Coccia[2]

[1] *wang@merit.unu.edu*
UNU-MERIT, Keizer Karelplein 19, 6211 TC, Maastricht (The Netherlands)

[2] *mario.coccia@ircres.cnr.it*
CNR -- National Research Council of Italy, 10024 Moncalieri, Torino (Italy)

Abstract
The purpose of this paper is to analyse the evolutionary pattern of international research collaborations. Using publication data from 1997 to 2012, this study decomposes international collaborations into two complementary types, intra-collaboration (within the same geographical area) and inter-collaboration (across different geographical areas). Our results show that the geographical concentration of international research collaborations is reducing. The formation of new network structure of international research collaborations is driven by the increase of inter-research collaborations of countries across different geographical areas rather than intra-collaborations of countries within the same geographical area.

Conference Topic
International collaboration

Introduction
Scientific collaborations have been widely acknowledged to be efficient in managing time and labour in research labs (Coccia, 2014; Solla Price & Beaver, 1966), improving research quality (Presser, 1980; Narin et al., 1991; Katz & Hicks, 1997) and spurring the breakthroughs of scientific research for supporting competitiveness (Coccia, 2012). A number of factors have contributed to the continuous increase of international research collaborations and co-authored papers (Beaver & Rosen, 1978; Frame & Carpenter, 1979; Katz & Martin, 1997). Along with the steady rise of international scientific collaborations, a better understanding on the structure of the global research network across geo-economic areas and its evolutionary pattern are needed for scholars and policy makers.

The high heterogeneity across countries – in terms of size, scientific capacity of the national system of innovation, etc. – generates a variety of patterns of the international research collaborations (Melin, 1999; Narin et al., 1991; Ozcan & Islam, 2014). A main issue in economics of science is to determine how and to which extent countries are engaged in international research collaborations so as to understand the behaviour of knowledge flows and to design research policies for improving the scientific research production which will in turn to enhance national competitiveness.

Luukkonen et al. (1992) maintain that the map of collaborative connections between countries corresponds to a geographical map. Frame et al. (1977, p. 502), considering data of 1973, claim that: "the production of mainstream science is more heavily concentrated in the hands of a few countries". Hoekman et al. (2010), using data on co-publications in European countries, show that research collaborations are geographically localized and despite a research heterogeneity in European countries in terms of research collaboration patterns, there

[1] Mario Coccia gratefully acknowledges financial support from United Nations University -The Maastricht Economic and Social Research Institute on Innovation and Technology (Contract ID 606U U-04 76) where this joint research was conducted while he was a visiting researcher.

is “a gradual convergence is taking place toward a more integrated interconnected European science system” (Hoekman et al., 2010, p. 672).

The purpose of this research is to investigate the evolutionary pattern of international research collaborations across countries. Emphasis is placed on two complementary collaboration types, i.e. intra- and inter-collaborations. The former refers to research collaborations conducted by countries within the same geographical area; the latter refers to research collaborations engaged by countries from different geographical areas.[2] Increase of intra-collaborations indicates that cooperation is more and more bounded within certain geographical territories, while increase of inter-collaborations signals the fade of geographical limit.

The main research questions of this paper are:

- How does the distribution of international collaborations across countries evolve over time?
- What type of research collaborations (inter- or intra-) plays a more important role in reshaping the global collaborative scientific network across geo-economic areas?
- How do inter- and intra- connections change in the global collective network?

The analysis of the temporal and spatial evolution of these patterns is of great scientific interest for researchers and policy makers in order to better master knowledge flow and optimize collaborative research output across countries.

Data and methodology

The data of this study are collected from publications in academic journals covered by the Science Citation Index (SCI) and Social Sciences Citation Index (SSCI). In particular, this study refers to dataset by National Science Foundation (2014)-National Center for Science and Engineering Statistics, special tabulations from Thomson Reuters (2013), SCI and SSCI. Collaboration data cover two years 1997 and 2012 and 40 countries (see the list in Appendix A). These 40 countries produce about 97% of the global total articles over 1997-2012. The 40 countries are classified into eight geographical areas: North America, South America, Europe Union, Other Europe, Middle East, Africa, Asia and Australia/Oceania (see Appendix A).

The analysis consists of the following steps:

- Firstly, to analyse the worldwide distribution of international collaborations, this study uses Lorenz curves and Gini coefficient. Lorenz curve is indicated by $L(X)$, then Gini coefficient can be derived as follows:

$$Gini\ coefficient\ (G) = 1 - 2\int_0^1 L(X)dX \quad (1)$$

  G is main indicator of concentration of the distribution of data.
- Secondly, to map the research connections between countries, both absolute collaborative output (number of articles) and collaboration intensity are considered. The former data set demonstrates the major players in the global collaboration research network while the latter puts all countries into one comparable framework. Although the matrix of co-authored papers between countries provides us main information concerning the output co-occurrence, the number of collaborated output might have different meanings for the collaborating country pair due to their different research capacity. For instance, suppose that a research collaborative pair is formed by Country A (of which the number of total publications is 1000) and Country B (of which the number of total publications in 10,000). Collaboration intensity (the ratio of collaborative output to national total publications) presents a stronger collaboration

[2] The under studied geographical areas are: North America, South America, Europe Union, Other Europe, Middle East, Africa, Asia and Australia/Oceania.

link for country A than B. Therefore, extra caution should be exercised when analysing the collaborative connections between research partners.

Based on eight geographical groups, this study disentangles intra-collaborations (between countries located in the same geographical area) from inter-collaborations (between countries of different geographical areas).[3]

Salton and Jaccard indexes are both valuable in measuring relative collaboration intensity (cf. Luukkonen et al., 1993). The collaboration index by Salton's measure (CSI) is

$$CSI = \frac{CO_{ij}}{\sqrt{P_i * P_j}} \quad (2)$$

whereas, the Jaccard's measure (CJI) is given by:

$$CJI = \frac{CO_{ij}}{P_i + P_j - CO_{ij}} \quad (3)$$

Where $\boldsymbol{CO_{ij}}$ is the number of co-authored papers between country *i* and country *j*

$\boldsymbol{P_i}$ is the total publication number by country *i*

$\boldsymbol{P_j}$ is the total publication number by country *j*

In addition, to understand the intra- and inter- collaborations by Salton and Jaccard indices (equations (2) and (3)), the adapted intra- and inter- collaboration intensities are

- $\mathrm{CSI}_{\mathrm{intra}} = \frac{\mathrm{CO_{ij}}}{\sqrt{\mathrm{P_i * P_j}}}$ (*i* & *j* ∈ same geographical area) (4)
- $CSI_{inter} = \frac{CO_{ij}}{\sqrt{P_i * P_j}}$ (*i* & *j* ∈ different geographical areas) (5)
- $CJI_{intra} = \frac{CO_{ij}}{P_i + P_j - CO_{ij}}$ (*i* & *j* ∈ same geographical area) (6)
- $CJI_{inter} = \frac{CO_{ij}}{P_i + P_j - CO_{ij}}$ (*i* & *j* ∈ different geographical areas) (7)

Coefficient of variation is also applied to assess the dispersion of data.

- Thirdly, from a dynamic perspective, this study applies network analysis to explore the structure of international collaborations and its changes from 1997 to 2012. In particular, intra- and inter- scientific ties across countries are distinguished from each other in the networks.

Empirical analysis

Global distribution of scientific research and collaborations

It has been well recognized that research capability and resources are unevenly distributed in the world, and hence scientific research output is concentrated in certain countries which are scientifically strong (Frame et al., 1977). By measuring the statistical dispersion of total publications and international collaborations, Table 1 shows that the Gini coefficient of internationally co-authored papers is lower than that of total publications, which means the former is distributed more evenly across countries than the latter. Most importantly, the Gini coefficients for both types of scientific outputs are decreasing over years. This means that the distributions of total publications and internationally co-authored papers both became less geographically concentrated in the later years.

[3] Refer to Appendix A for detailed group information.

Table 1. Gini Coefficient over years

	1997	*2002*	*2007*	*2012*
Total publications	0.67	0.63	0.61	0.59
Internationally co-authored papers	0.60	0.58	0.56	0.54

Dynamics of international collaborations

Salton and Jaccard measures are considered for estimating the collaboration intensity (Figure B1 and B2, in the Appendix B). The arithmetic mean of Salton measure is as twice as that of Jaccard measure, which is in line with Hamers, et al. (1989). However, the coefficient of variation in Jaccard is somewhat higher than that of Salton (see Fig. B1 and B2), indicating a greater dispersion of collaboration intensities is measured by Jaccard index. As the aim of this study is to analyse collaborative research variability between countries, intensities derived from Jaccard index seem to be more suitable.[4]

At the level of geographical groups, Figure 1 shows the relationship of the intra- and inter-collaboration intensities between 1997 and 2012. Red dots represent the inter-collaboration intensity and green ones represent intra-collaboration intensities. A dot being above diagonal line indicates that the collaboration intensity of this observed unit has increased in 2012 in contrast to that of 1997. Likewise, a dot underneath the diagonal indicates that the international collaboration intensity has decreased in 2012 compared to that of 1997. The fact that all the dots lying above the diagonal line suggests that both intra- and inter- collaboration intensities in all geographical areas have improved over years. On the other hand, by comparing the red and green dots, it is of great interest to observe that inter-collaborations in all geographical areas have increased dramatically while intra-collaborations stay mostly low and close to the diagonal line. The intra-collaboration intensity in the European Union (EU) is the only exception with high level of intra-collaborations in both 1997 and 2012, which is a phenomenon of "Europeanisation" as discussed by Mattsson et al. (2008). In general, this figure shows that intra-collaborations tend to be static while inter-collaborations exhibit high dynamics of growth.

[4] In the rest of the paper, we present only results calculated based on Jaccard measure. Similar results using Salton measure are available upon request.

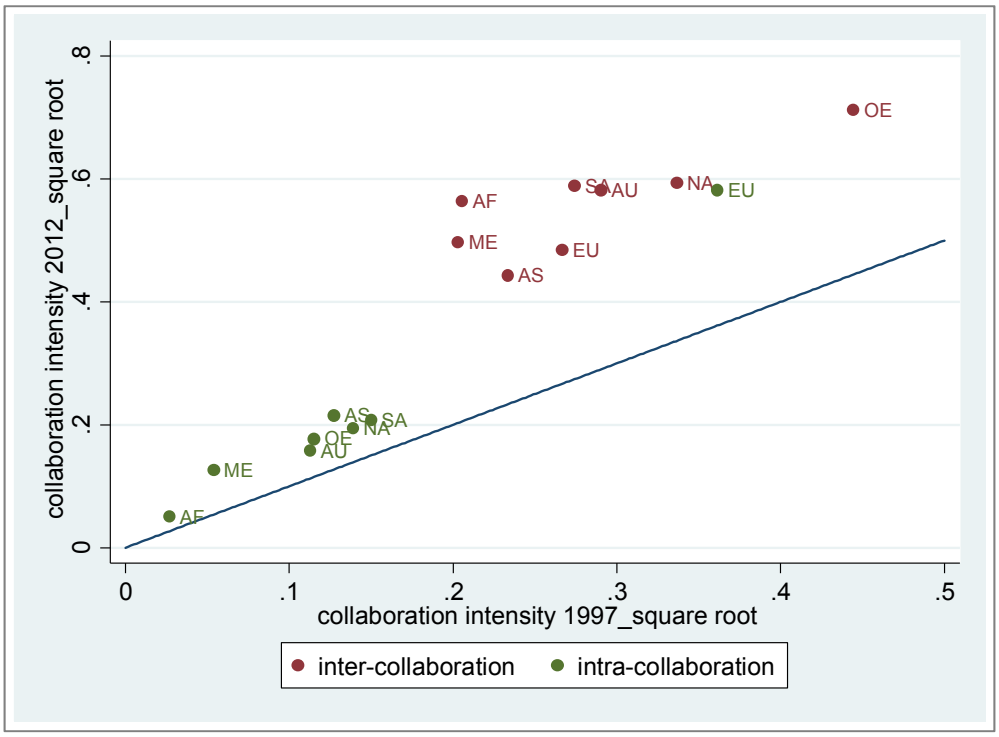

Figure 1. Comparison of international collaboration intensity (inter vs. intra)

Note: 1) The eight geographical areas are: North America (NA), South America (SA), European Union (EU), Other Europe (OE), Middle East (ME), Africa (AF), Asia (AS) and Australia/Oceania (AU). 2) Collaboration intensity is measured by Jaccard index.

To further understand the changes of collaborative performance in individual countries, Figure 2 presents the intra- and inter-collaboration intensity in the 40 under studied countries. Countries in European Union are the only ones showing growth of both intra- and inter-collaborations. This can be the result of European Commission's policy which stimulates cooperation between European countries. In the rest countries, the intra-collaboration performance looks all static, while inter-collaborations have risen obviously. Among all the countries, a group of Asian countries (China, India, Japan, Singapore, and South Korea) show relatively slow growth in inter-collaborations.

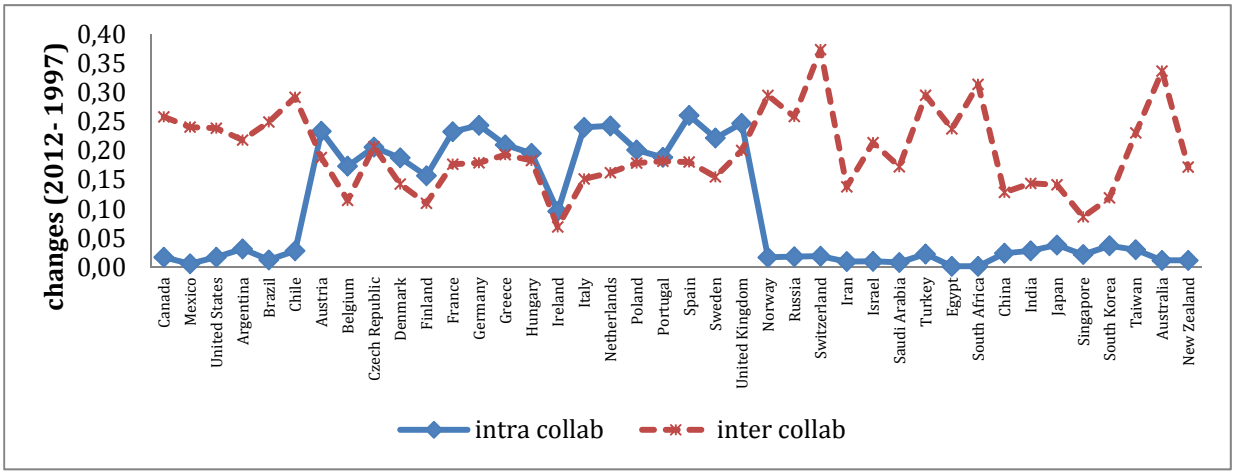

Figure 2. Changes of international collaboration intensity by country (inter vs. intra)

Note: 1) Collaboration intensity is measured by Jaccard index. 2) The value of y-axis is calculated by the collaboration intensity in 2012 minus that in 1997.

Networks of research collaborations

Based on Jaccard collaboration intensity, collaborative networks across 40 countries in 1997 and 2012 are provided in Figure 3 and 4. The thickness of each edge between two nodes reflects the strength of their collaborative relationship. The higher collaboration intensity one country pair has, the thicker their connection line is. In order to distinguish between intra- and inter-collaborations, geographical areas are presented in different colours.[5] Lines connecting nodes in different colours represent inter-collaborations, while those between nodes in same colours represent intra-collaborations. The size of each node embodies its aggregated collaboration intensity (including both intra- and inter- collaborations).

Figure 3 shows that scientific collaboration networks have been, to some degree, formed by geographic ties. Apart from the intensive connections between European countries (intra-collaborations), there are a few geographically biased small clusters are of great interest. The rectangular cluster in Nordic countries (formed by Denmark, Sweden, Norway and Finland) and the triangular cluster in South America (formed by Chile, Brazil and Argentina) both indicate that scientific collaborations are geographically localized. Besides these small clusters, in North America, a strong tie is observed between United States and Canada. In Asia, China is mainly connected with Japan. In Australia/Oceania, New Zealand has a strong connection only with Australia.

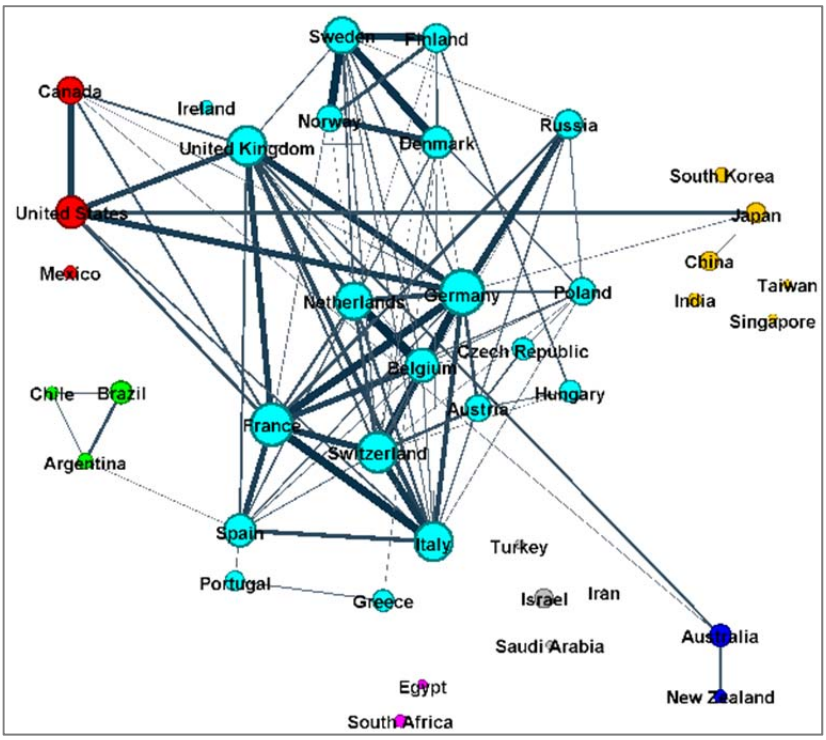

Figure 3. Network of global research connections in 1997.

Note: 1) A filter of 0.0083 is applied in this figure, which means that edges with collaboration intensity less than 0.0083 are omitted. 2) The thickness of each edge between two nodes reflects the strength of their collaborative relationship. 3) The size of each node embodies its aggregated collaboration intensity.

[5] To emphasize the effect of geographical locations, *European Union* and *Other Europe* are regarded as one group in the network figures (Fig. 3 and 4).

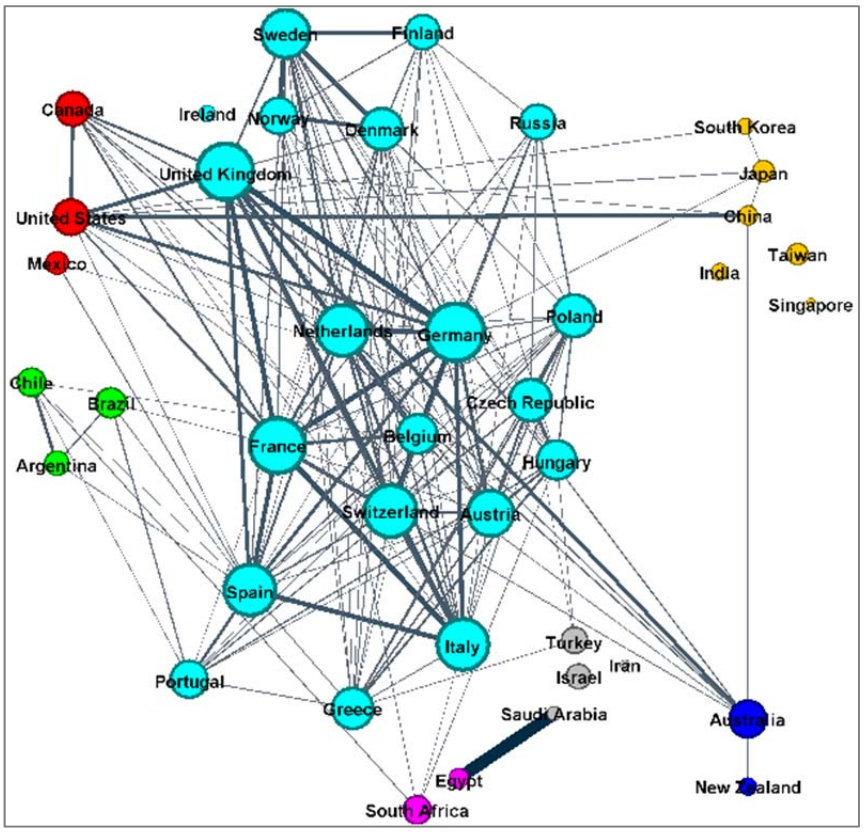

Figure 4. Network of global research connections in 2012.

Note: 1) The network in 2012 is much denser than that of 1997. In order to keep the visualization compact and readable, filter applied in this figure is as twice high as the 1997 figure. Edges with collaboration intensity less than 0.016 are omitted. 2) The thickness of each edge between two nodes reflects the strength of their collaborative relationship. 3) The size of each node embodies its aggregated collaboration intensity.

In order to understand the dynamics of international collaborations, it is necessary to compare the structure of networks in the earlier year 1997 (Fig. 3) with that of the later year 2012 (Fig. 4). In contrast with 1997, the aggregated collaboration intensity (embodied by the circle size of each node) for most countries has increased in 2012. In particular, an important observation is that, the variety of inter-collaborations (lines between different coloured nodes) has grown significantly in 2012, while the connection strength between major intra-collaborative partners (nodes with the same colours) stayed roughly at original level of 1997.

In contrast with the structure in 1997 (Fig. 3), the rectangular Nordic cluster and triangular South American cluster in 2012 have both increased their inter-connections with countries beyond their geographic neighbours (see Fig. 4). The strong tie between Chile and Brazil (i.e. intra-collaboration) has been weakened while both Chile and Brazil developed new inter-collaborative partnerships with countries from other geographical areas. Similarly, the tie between Finland and Denmark became relatively weaker, whereas both of them established more connections with various countries. Due to the effect of "Europeanisation" of this geo-economic area, the new major collaboration partners are still within Europe, but far beyond the old Nordic limit in the later year.

Asian countries, though still with relatively low collaboration intensity, have increased scientific cooperation with the United States (i.e. known as type of inter-collaborations). In particular, China has developed a very strong collaborative tie with the United States and a reasonable partnership with Australia, which are both inter-collaborations. Yet as the second

largest producer of scientific publications, China did not develop any new strong collaborative ties (i.e. intra-collaborations) within its own geographical area.

Located in North America, Mexico seemed to have developed new collaborative research partners only beyond its own geographical area (i.e. inter-collaborations). As one of the most dynamic countries regarding international research collaborations, South Africa seemed to have built inter-collaborative relationships mainly in Europe and South America. Different from the isolated situation in the earlier stage (1997), Egypt and Saudi Arabia developed an extremely strong research partnership in 2012.[6] Their connection with each other was so strong that they hardly had any cooperation with any third countries.

Conclusions

The main lessons learned of this research can synthetized as follows:

1) The Gini coefficients for total publications and collaborations were both smaller in 2012 than 1997, indicating that the distribution among the under studied 40 countries became more and more balanced. Nevertheless, it is worthwhile to note that the distribution of total publications was more divergent than that of internationally co-authored papers.
2) In the process of evolution of international collaborations, evidence shows significant difference between intra- and inter- collaborations. In all geographical areas, except European Union, the intra collaboration performances exhibited a steady-state pattern, whereas inter-collaborations in the global network research structure have risen dramatically.
3) From a dynamic point of view, the comparison of 1997 and 2012 networks shows that inter-collaborations (between countries from different geographical areas) have grown significantly in the later stage, while the connection strength between major intra-collaborative partners stayed mostly unchanged. This finding indicates that recent research network across countries has a higher global inter-connection beyond geographical territorials, which is likely driven by advances of ICT and transportation new technologies and improvement of socio-economic systems.

In short, the increase of research collaborations between countries from different geographical areas has reshaped the global structure of international scientific collaborations. In the modern process of knowledge production, countries seem to be looking for more diverse collaborative partners worldwide.

References

Beaver, de B.D. & Rosen, R. (1978). Studies in scientific collaboration – Part. I. The professional origins of scientific co-authorship, *Scientometrics, 1*, 65-84.

Coccia M. (2010). Democratization is the Driving Force for Technological and Economic Change, *Technological Forecasting & Social Change*, *77*, 248-264.

Coccia, M. (2012). Political economy of R&D to support the modern competitiveness of nations and determinants of economic optimization and inertia, *Technovation*, *32*, 370-379

Coccia M. (2014). Driving forces of technological change: The relation between population growth and technological innovation-Analysis of the optimal interaction across countries, *Technological Forecasting & Social Change, 82*, 52-65

de Solla Price D. & de B. Beaver D. (1966). Collaboration in an invisible college, *American Psychologist*, *21*, 1011-1018.

Frame J . D. & Carpenter M. P. (1979). International research collaboration, *Social Studies of Science*, *9*, 481-497.

Frame, J.D., Narin, F. & Carpenter, M.P. (1977). The distribution of world science, Social *Studies of Science*, *7*, 501-516.

[6] Although Egypt and Saudi Arabia are classified into different groups, they are in geographically adjacent. Therefore their collaborative relationship can be still regarded as a result of geographical localization.

Hamers, L., Hemeryck, Y., Herweyers, G., Janssen, M., Keters, H., Rousseau, R., & Vanhoutte, A. (1989). Similarity measures in scientometric research: the Jaccard index versus Salton's cosine formula. *Information Processing & Management*, *25*(3), 315-318.

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

Katz, J. S. & Martin, B. R. (1997). What is research collaboration? *Research Policy*, *26*, 1-18.

Katz, J.S. & Hicks, D. (1997). How much is a collaboration worth? A calibrated bibliometric model. *Scientometrics*, *40*, 541-554.

Luukkonen, T., Persson, O., & Sivertsen, G. (1992). Understanding patterns of international scientific collaboration, *Science, Technology & Human Values*, *17*, 101-126.

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

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

Melin, G. (1999). Impact of national size on research collaboration. *Scientometrics*, *46*, 161-170.

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

Ozcan, S. & Islam, N. (2014). Collaborative networks and technology clusters –The case of nanowire. *Technological Forecasting and Social Change*, *82*, 115-31.

Presser, S. (1980). Collaboration and the quality of research. *Social Studies of Science*, *10*, 95-101.

Appendix A. Country/economy of the sample

nr	country	Geo-Economic Area
1	Canada	North America
2	Mexico	
3	United States	
4	Argentina	South America
5	Brazil	
6	Chile	
7	Austria	European Union
8	Belgium	
9	Czech Republic	
10	Denmark	
11	Finland	
12	France	
13	Germany	
14	Greece	
15	Hungary	
16	Ireland	
17	Italy	
18	Netherlands	
19	Poland	
20	Portugal	
21	Spain	
22	Sweden	
23	United Kingdom	
24	Norway	Other Europe
25	Russia	
26	Switzerland	
27	Iran	Middle East
28	Israel	
29	Saudi Arabia	
30	Turkey	
31	Egypt	Africa
32	South Africa	
33	China	Asia
34	India	
35	Japan	
36	Singapore	
37	South Korea	
38	Taiwan	
39	Australia	Australia/Oceania
40	New Zealand	

Appendix B:

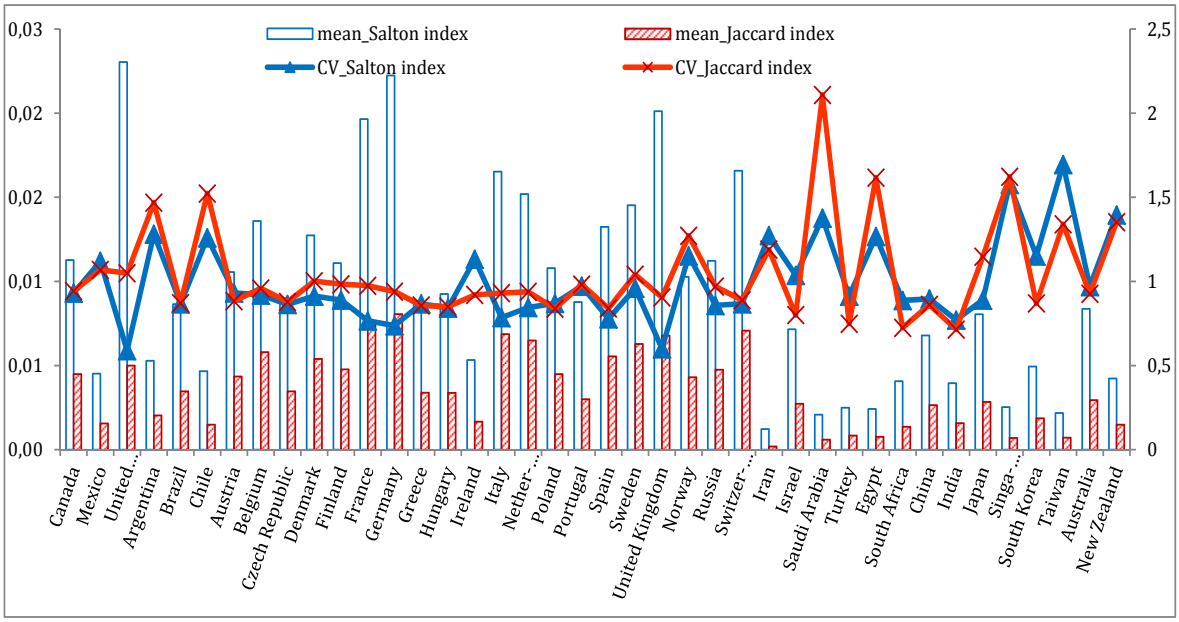

Figure B1. Mean and coefficient variation for collaboration indices (Salton vs. Jaccard) 1997

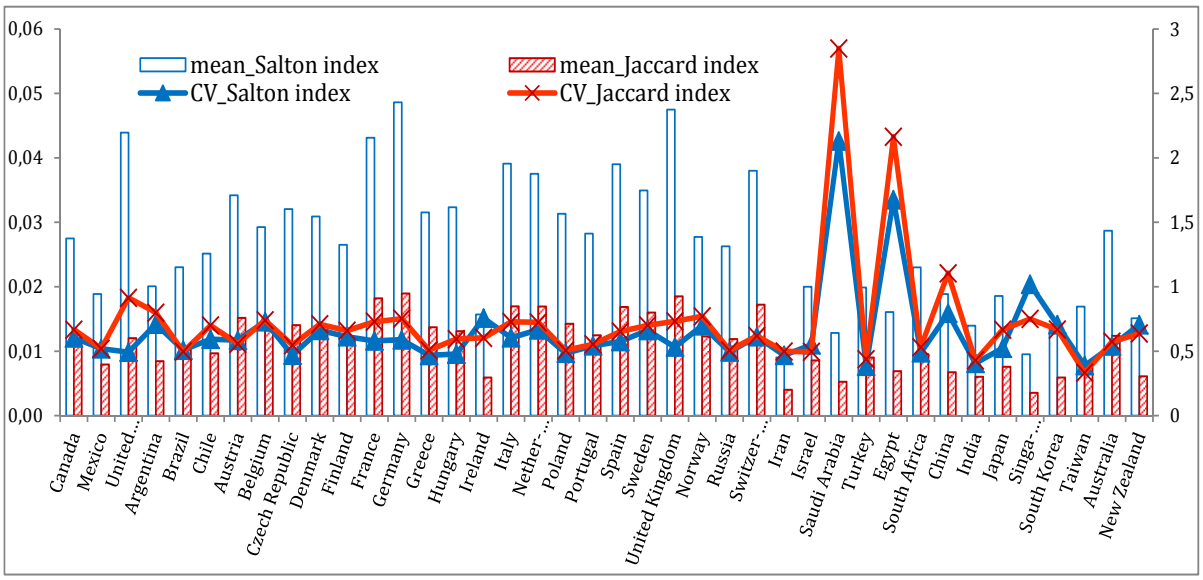

Figure B2. Mean and coefficient variation for collaboration indices (Salton vs. Jaccard) 2012