Long-term correlations in short, non-stationary time series: an application to international R&D collaborations

Lorenzo Righetto* Alessandro Spelta† Emanuele Rabosio‡ Fabio Pammolli§¶

February 18, 2019

Abstract

Within the perimeter of patent collaboration networks, the average distance of collaborations and the number of countries involved per each collaboration have been shown to have increased steadily in time. Less attention, though, has been devoted to assessing whether growth of cross-country collaborations is stable or robust in time. To address this scientific question we focus on the identification of long-term correlations (i.e. persistence in time). Our data sets consists of time series of yearly average collaboration radii and of cross-border links in the Euro-American subsystem of the global collaboration network for the period 1978 - 2014. To detect long-term correlations, we use Detrended Fluctuation Analysis, a method that is used to measure persistence in signals. In addition, we devise a general and original procedure to assess the statistical significance of results for short time series. Our results, showing that long-term correlations do exist in the great majority of our signals, reinforce the hypothesis of a diminishing role of geographical distance in technological collaborations. Results at the level of nations show that a significant degree of heterogeneity in scaling values can be detected within Europe, irrespectively of the substantial efforts towards the set-up of an integrated European Research Area.

R&D international collaborations Detrended Fluctuation Analysis integration of R&D systems

*Center for Analysis, Decisions, and Society (CADS) - Human Technopole, Milano, Italy; lorenzo.righetto@htechnopole.it

†Center for Analysis, Decisions, and Society (CADS) - Human Technopole, Milano, Italy; alessandro.spelta@htechnopole.it

‡Center for Analysis, Decisions, and Society (CADS) - Human Technopole, Milano, Italy; emanuele.rabosio@htechnopole.it

§Politecnico di Milano, Department of Management, Economics and Industrial Engineering; fabio.pammolli@polimi.it

¶Center for Analysis, Decisions, and Society (CADS) - Human Technopole, Milano, Italy

1 Introduction

Networks of collaborative agreements are recognized as an ever-widening organization form, especially in high technology, knowledge-intensive fields (Arora and Gambardella, 1994; Arora et al., 2001; Orsenigo et al., 2001). In particular, networks of collaborations among innovators have been indeed extensively used to represent and analyze the division of innovative labor since the seminal paper by Freeman (Freeman, 1991). The structure of the international network of patent collaboration, in terms of co-inventions or co-assignation of patents, has been evaluated broadly in recent literature and the average distance of collaborations, together with the number of countries involved per each collaboration, has been shown to have increased steadily in time (Chessa et al., 2013; Morescalchi et al., 2015). Several factors play a role in this ongoing phenomenon, including increasing ease of knowledge sharing, physical transportation, diminishing language barriers, and so on. Similar studies concerning academic research collaborations have shown generally coherent results (Frenken et al.; Hoekman et al., 2010; Waltman et al., 2011), with some notable exceptions highlighting the concurrent domestic increase of collaborations (Maisonobe et al., 2016). Despite the growing body of literature on knowledge networks, the analysis of their temporal evolution still remains at its infancy. Less attention has been indeed devoted, for instance, to assessing whether the growth of geographical distance and of the number of cross-country technological collaborations may be considered stable or robust, beyond the common global trend associated with the globalization. In this paper, we make use of Detrended Fluctuation Analysis (DFA) (Peng et al., 1994) to detect the presence of long-term correlations in our time series. In other words, this corresponds to determining whether the observed increases in geopolitical features of technological collaborations can be considered persistent in time. So far, DFA has been used in different fields, e.g. to investigate diseased states in physiological studies (Hardstone et al., 2012), as well as prey search behaviors in ecological studies (Viswanathan et al., 1996).

In the context of R&D collaborations, this analysis can be used to bring further evidence of a decreasing role of distance, in the context of globalization (Disdier and Head, 2008), in case collaboration patterns are shown to display such long-term correlations – e.g. meaning that the increase of the distances of collaborations in the short run makes them more likely to increase also in the future. In fact, DFA represents a useful tool as it purges the signal from the underlying trends and is, as such, particularly indicated for time series which exhibit strong non-stationarity (Hu et al., 2001), as we observe in this case (see Chessa et al., 2013; Morescalchi et al., 2015). We apply DFA to time series of yearly average collaboration radii and of yearly number of cross-border links, recorded for all collaborations (co-inventions and co-assignations) at the OECD TL3 aggregation scale – all regions in the European Union (EU) and in the United States (US) are considered – from OECD RegPat registry of patent applications, which currently spans from 1978 to 2014. We choose to track both these measures, even though they convey similar information, for several reasons - for instance, to control for specific geographical, historical or economic reasons that might make it easier for a region to initiate international collaborations, but in a limited geographical range. We restrict the analysis to the EU+US subsystem to better highlight the role of US nodes in determining the observed EU scaling values, in line with previous literature focusing on these two important subsystems

(Owen-Smith et al., 2002; Chessa et al., 2013). We aggregate results at the national level (and, for the most part, at "continental" level) as in Viswanathan et al. (1996) to obtain a more reliable estimate of scaling behaviours.

In the literature, several papers have highlighted the shortcomings of the DFA method in terms of estimation errors for short time series (see e.g. Delignieres et al., 2006) or even of its real detrending power (Bryce and Sprague, 2012). We take into account these fore-warnings by performing several controls on synthetically generated time series of specified length and known DFA scaling. On one side, we make sure that trend removal is actually performed, following Hu et al. (2001). On the other side, by repeating the synthetic generation of time series, we obtain a measure of the estimation error at varying time series length and scaling factor. We use these measures to complement our empirical observations with confidence intervals, e.g. to detect whether uncorrelation can be excluded with statistical significance.

Results show that long-term correlations do exist in the great majority of our signals. This result shows that the suppression of the global underlying trend does not affect the positive long-term correlation of the selected time series. This general result does not prevent us from running comparisons, e.g. between the EU+US system and each possible subsystem – down to European national level – to detect marginal effects of, say, including US nodes to the magnitude of long-term positive correlations in EU signals. In particular, we assess the degree of heterogeneity of observed scaling values among EU countries, to evaluate the deviation of the system from a completely integrated European R&D area.

The paper is organized as follows. Section 2 describes the dataset we use and the metrics we extract from it. We describe in Section 3 the DFA method in general and in the specific application to our case study. Section 4 is devoted to the presentation of results and their discussion. Conclusions in Section 5 close the paper. In the Appendix, we describe the details of our DFA application to the case under study and the controls on the soundness of our DFA scaling value estimation.

2 Data

We make use of the data provided by the OECD RegPat registry, which contains all patent records from the European Patent Office (EPO) since the first application (1978) until present. We treat 2014 as the latest complete year of application to allow for the review process to be completed for all patents in that year. In addition to that, in RegPat all assignee and inventor addresses are attributed a single OECD TL3 level geographical location for each of the OECD partner countries. We restrict our analysis to EU and US regions, which constitute, together, around 85% of the total number of OECD TL3 regions in RegPat (4700 out of 5552). This spatial segmentation corresponds to NUTS3 regions for EU and to counties for US. We complement this dataset by calculating centroid coordinates for each of these regions. To increase the number of nodes in this network, we use all European regions within the 2013 edition of the NUTS3 map (which also includes not "full" EU countries such as Switzerland and Turkey). What we obtain is a spatial characterization of the patent collaboration network, which we analyze in terms of co-assignments (companies collaborating for R&D regarding a specific patent) and co-inventions (different scientists/researchers collaborating as part of one or more organizations). In the former case, the reported address is related to the location of the company's headquarter, while in the latter case the location may refer to the inventor's residence or the address of her/his working place (which may differ from the related assignee's headquarter address, of course).

We extract all patent numbers which relate to multiple records (i.e. the patent is the result of a co-assignment or of a co-invention) and compute two metrics from each collaboration and for each of the involved nodes (regions):

- the average of the distance (which we call *radius of co-assignment/co-invention*, ρ_A/ρ_I in symbols) from the focal node to all other collaborating nodes (this metric can be zero), computed on a sphere and expressed in kilometres;

- the number of cross-border (i.e. cross-country or cross-state in the US case) links that the focal node collects within the collaboration (again, this metric can be zero). Cross-border links for co-assignments and co-inventions are expressed as CB_A/CB_I in symbols, respectively.

We aggregate these metrics by computing the average yearly radius of collaboration (no finer temporal scale is allowed by the RegPat database) and the yearly sum of cross-border links in each node. We repeat this calculation for 3 different segmentations of our network:

- EU: only European nodes are considered (i.e. US nodes do not contribute to the computations);

- US: only US nodes are considered (cross-border links, in this case, refer to *cross-state* links);

- EU+US: the whole system is considered.

In the end, we obtain time series of both metrics for all nodes covered by the focal segmentation and for the time span 1978-2014. We further exclude the starting year 1978 since the number of applications was extremely low, around 0.1% of the total.

3 Detrended Fluctuation Analysis

We make use of DFA (Peng et al., 1994), a widely used technique for the identification of long-term correlations in time series (see e.g. Hardstone et al. (2012) for a compelling review of this method). The first step of DFA consists in transforming the time series under study in a random walk path, by integrating the time series $x(t)$ as follows:

$$y(t) = \sum_{k=1}^{t} [x(k) - < x >], \tag{1}$$

where $< x >$ is the mean value of $x(t)$. The transformed signal is then segmented in windows of various sizes Δn. In each window, and repeatedly for all values of Δn, a polynomial of order s (hence, the order of the DFA applied to the focal case is fitted to the integrated data, in order to obtain an estimated set of points $y'(k)$. Please note that, as a result of the controls we perform (see the Appendix), we choose to use DFA-2 (i.e. $s = 2$). The value of the *fluctuation function* $F(\Delta n)$ for that particular value of Δn is the average of the standard deviation between the polynomial fit y' and the data, over all data points N, for all windows:

$$F(\Delta n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[y'(k) - y(k)]^2} \tag{2}$$

When $F(\Delta n)$ is plotted against Δn in a double-logarithmic plot, the relationship is expected to be linear and the slope of the regression line defines the DFA scaling α. This scaling is a measure of the presence of self-similarity and, relatedly, of long-term correlations in the signal, as it tracks down the scaling of dispersion around a regressor for increasing window sizes. In particular, the value of α can describe the following signal behaviours:

- $0 < \alpha < 0.5$: the signal has long-term memory and is anti-correlated;
- $0.5 < \alpha < 1$: the signal has long-term memory and is correlated;
- $\alpha = 0.5$: the signal is uncorrelated (has no memory);
- $\alpha > 1$: the signal is non-stationary.

This relates to the intrinsic properties of the signal in the sense that, at least in the range between 0 and 1, it represents three different behaviors of the random walker corresponding to the time series under study (see also Fig. 3 in Hardstone et al., 2012, for a pictorial representation) :

- $0 < \alpha < 0.5$ (anti-correlation): the random walker moves preferentially in the opposite direction with respect to the previous step; the resulting cumulated walk (i.e. cumulated time series) will show very small fluctuations in time and thus the slope of the fluctuation function at increasing window sizes will be relatively low;
- $0.5 < \alpha < 1$ (correlation): the random walker tends to repeat the same moves it has performed previously, inducing wide fluctuations in time; thus, when measuring the error between a regressor and the data at large window sizes, this error will be much larger than the error measured at small window sizes (i.e. a higher slope);
- $\alpha = 0.5$: when the choice of a direction of movement is independent from the previous steps, in analogy with pure diffusion the mean square displacement in time of the random walker will scale as 0.5.

In our case study, following Viswanathan et al. (1996), we extend the method to the case of several short signals grouped together, by calculating $F(\Delta n)$ over all possible windows in the subsystem under study. In brief, each value of the standard deviation of residuals is averaged over N (the size of the individual time series, i.e. 36) times the number of nodes constituting that particular subsystem. We use here 6 values of Δn, corresponding to the exact divisors of 36 (i.e. $\Delta n = 4, 6, 9, 12, 18, 36$). We also standardize values with respect to individual mean and standard deviation values, to further homogenize the sample data in terms of geographical range and country "size".

Subsystem	ρ_A	ρ_I	CB_A	CB_I
EU	0.37	0.30	0.22	0.75
US	0.15	0.12	0.20	0.37
EU+US	0.28	0.19	0.23	0.52
$(\text{EU+US})_{EU}$	0.40	0.31	0.26	0.78
$(\text{EU+US})_{US}$	0.17	0.13	0.18	0.30

Table 1: Fraction of nodes in each subsystem displaying a monotonic trend according to Mann-Kendall test. ρ_A: average yearly radius of co-assignment; ρ_I: average yearly radius of co-invention; CB_A: number of yearly cross-border links according to co-assignments; CB_I: number of yearly cross-border links according to co-inventions.

We are aware, of course, that the series under study might contain strong non-stationarities which might undermine the results of DFA (Bryce and Sprague, 2012). Nevertheless, the effect of trends on DFA has been studied in the literature and workarounds to ascertain the true underlying scaling of the signal have been demonstrated in Hu et al. (2001). In particular, in Hu et al. (2001) it emerges that DFA of order x (termed DFA-x) can only neutralize the effect of $x-1$ order trends. For additional information on the application of DFA to our case study and on the controls we performed (including the ones for determining the actual detrending power of DFA in our case), please see the Appendix.

4 Results and Discussion

As a first assessment of the general evolution of the metrics under study, we calculate the prevalence of nodes for which a Mann-Kendall test at significance level 0.05 detects a monotonic trend. We report the results in Table 1. The values we obtain are substantial, confirming that there is an underlying global trend in this metrics that is affecting a large portion of regions, especially in the EU. As a caveat, we have to specify that higher statistics we obtain for the EU system, with respect to the US system, are likely to be due to the “submission bias” that is intrinsic to using the RegPat database (which comes from EPO data). In this respect, results concerning the US subsystem should be looked at in relative terms.

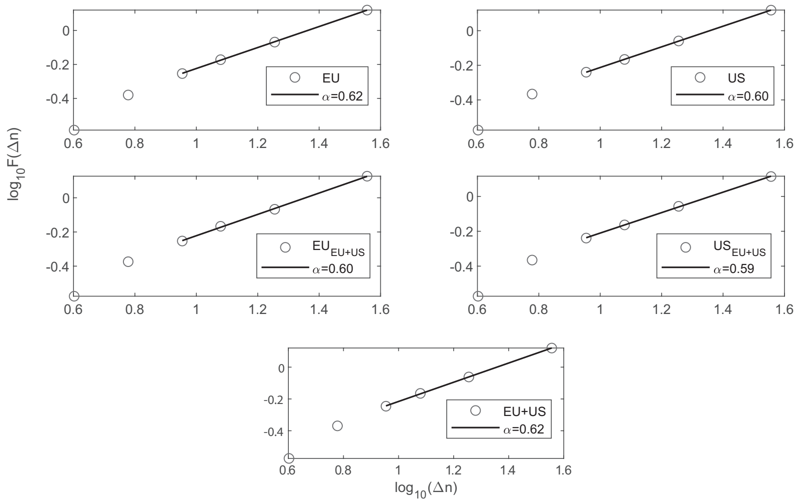
Despite the high prevalence of non-stationary individual signals in our data, our application of DFA-2 can purge the trend effects from the estimated scaling, as shown in the Appendix. Table 2 shows the results of the application of the method to the 3 selected subsystems, plus the scaling obtained when the EU and the US are embedded in the EU+US system. This allows us to detect positive or negative effects of a combined collaboration system on the robustness of these series. In very general terms, scaling values turn out to indicate positive long-term correlation in all the selected subsystems, for all metrics analyzed (see Fig. 1 for an example of fluctuation plots and their relative estimated scaling). We confirm this result by reshuffling our data series, as performed in Castillo et al. (2015). We randomly reassign data points to different nodes for each year and estimate the DFA-2 scaling. We verify that series become uncorrelated in this case (i.e. values lie within the 5-95th percentile region of an uncorrelated synthetic signal).

Co-assignment metrics also show lower scaling values than those regarding co-invention metrics, as a result of co-assignment being a much rarer occurrence. As specified earlier, US scaling values are generally lower than in the EU subsystem, except for that of the

Subsystem	ρ_A	ρ_I	CB_A	CB_I
EU	0.62	0.68	0.62	0.69
US	0.60	0.71	0.58	0.66
EU+US	0.60	0.63	0.62	0.65
$(\text{EU+US})_{EU}$	0.62	0.67	0.63	0.70
$(\text{EU+US})_{US}$	0.59	0.61	0.58	0.62

Table 2: DFA-2 scalings in each subsystem, for each of the measured quantities. Symbols as in Table 1.

Figure 1: Loglog plots of fluctuation function values versus Δn for ρ_A (radius of co-assignment) in the 3 selected subsystems (EU, US, EU+US) and in the EU/US subsytems when embedded into the EU+US subsystem. Black solid lines show the linear fit we use to estimate the DFA-2 scaling.



radius of co-invention, showing a perhaps greater ease of establishing stable connections, even at longer distance, within the US subsystem. It has to be noted that this might also stem, for instance, from a more geographically sparse distribution of company locations. This last hypothesis might also explain the discrepancy between the co-assignment and the higher co-invention scaling values for the US subsystem.

To single out the contribution of "leading nodes", i.e. the nodes showing a significant monotonic trend (see Table 1), we estimate the scaling when all these nodes are removed from each subsystem. Table 3 shows the results of this analysis. Scaling values become generally closer to the "uncorrelated" mark ($\alpha = 0.5$, which would entail that the underlying signal can be associated to white noise and, as such, nothing can be said about its persistence in time), but, still, they lie outside the 95^{th} percentile threshold values for uncorrelated signals. In general, subsytems do not seem to change their qualitative behavior and can be said to be robust to the removal of their most performing nodes. It stems from these results that the increasing patterns that are observed in technological collaboration metrics at continental level can be considered persistent in time, regardless of the underlying trend and of the presence of the most performing nodes.

4.1 Long-term correlation at EU national level

Results, so far, show a general picture that looks remarkably similar, at least qualitatively, for all metrics and subsystems we have taken into account. The structure of the data

Subsystem	ρ_A	ρ_I	CB_A	CB_I
EU	0.61	0.63	0.59	0.58
US	0.58	0.60	0.57	0.61
EU+US	0.59	0.60	0.59	0.58
$(\text{EU+US})_{EU}$	0.61	0.62	0.60	0.58
$(\text{EU+US})_{US}$	0.57	0.60	0.57	0.58

Table 3: DFA-2 scaling values in each subsystem, for each of the measured quantities, when nodes showing monotonic trends (see Table 1) are removed. Symbols as in Table 1.

allows us, though, to perform a much more particular analysis, i.e. at EU national level. In this context, we point out that several political initiatives are currently in place to favor the integration of the European R &D collaboration system (see e.g. Schengell and Lata, 2013; Chessa et al., 2013; Arrieta et al., 2017).

To see whether reducing the geographical scale of the analysis might unravel different properties of local systems and, thus, relevant heterogeneities within the EU system, we single out 18 EU countries having more than 10 "active" regions in at least one of the datasets (i.e. they have at least 10 regions having at least one co-assignment or co-invention in the selected time span), and estimate their correspondent DFA-2 scaling for all metrics. Table 4 shows the list of selected countries and the results of this analysis. We include "partial" or younger EU members to see whether longer or fuller EU membership is associated with DFA-2 scaling results. We remark here that, due to the lower number of data points at national scale, we treat as potentially uncorrelated all scaling values below a 95^{th} percentile threshold (marked with a star in Table 4), that we obtain from generating uncorrelated synthetic signals (see the Appendix for details). In this respect, although there is not enough information to perform a thorough statistical analysis, it seems notable that uncorrelated behavior can arise both "within" (as in the Portugal case) and outside EU-15 borders, and viceversa (see e.g. Hungary and Poland, where relationships between inventors, especially, seem to be maintained in time, whereas it seems much more unlikely for Polish and Hungarian companies to participate in a co-assignment). Switzerland emerges as the best performing country for all metrics, pinpointing the role of geographical (i.e. the country being a natural continental hub in this respect, not to mention the compresence of several languages) but also institutional (i.e. the high number of multinational corporation HQs and of international organizations) factors in maintaining long-distance collaborations in time.

Moreover, it also seems relevant to investigate which countries improve their scaling when US nodes are taken into account while calculating the selected metrics. Figure 2 shows the results (in terms of percentage change) for all countries and metrics. We single out variations exceeding the average estimation error for the relative number of active regions in each country and for each metric (see the Appendix for details). Also, we highlight increments/decrements that bring the scaling above/below the uncorrelation threshold we have set for that country and metric (see Table A1 in the Appendix). We group countries in positively, negatively and contrasted/neutrally affected groups, based on the presence of a significant increase, or decrease, or no significant change/concurrent presence of significant decrease and increase, respectively. Interestingly, this grouping highlights relatively coherent blocks, in terms of economic relationships. As expected, this analysis pinpoints the national systems which have stronger relationships, traditionally,

Country	ρ_A	ρ_I	CB_A	CB_I
Austria: AT	0.68	0.72	0.68	0.78
Belgium: BE	0.62	0.71	0.69	0.78
Bulgaria: BG	0.52*	0.58*	0.57*	0.67
Switzerland: CH	0.77	0.81	0.86	0.92
Germany: DE	0.63	0.70	0.63	0.70
Greece: EL	0.60*	0.60*	0.52*	0.53*
Spain: ES	0.65	0.68	0.61	0.66
France: FR	0.61	0.70	0.65	0.78
Croatia: HR	0.56*	0.61*	0.62*	0.51*
Hungary: HU	0.60*	0.73	0.51*	0.74
Italy: IT	0.65	0.70	0.62	0.68
Netherlands:NL	0.60	0.68	0.65	0.78
Poland: PL	0.52*	0.73	0.54*	0.66
Portugal: PT	0.48*	0.66	0.45*	0.60*
Romania: RO	0.51*	0.61	0.51*	0.55*
Sweden: SE	0.62	0.76	0.53*	0.78
Turkey: TR	0.48*	0.55*	0.46*	0.56*
United Kingdom: UK	0.61	0.59	0.63	0.69

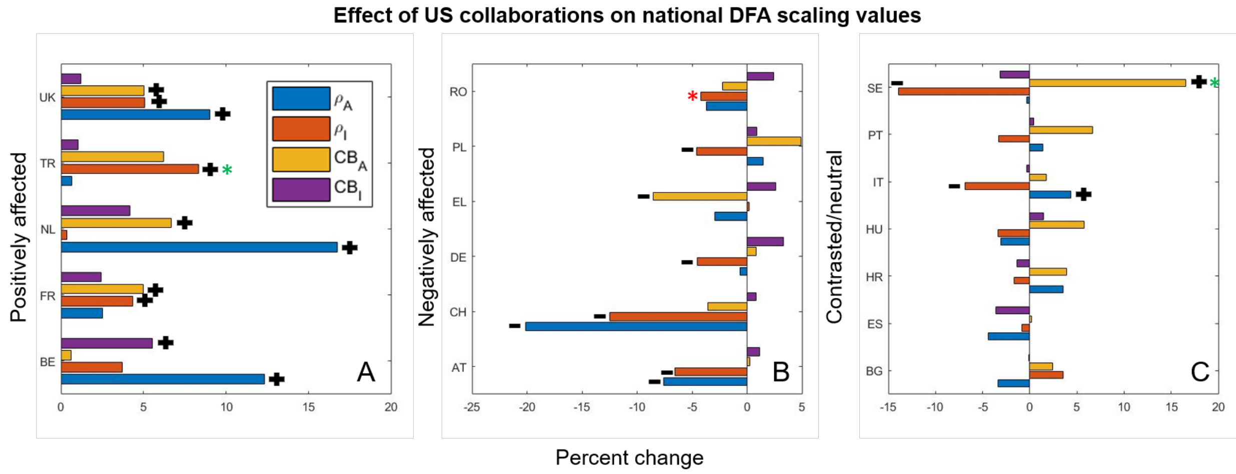
Table 4: DFA-2 scaling values in each of the selected 18 EU countries, when only EU nodes are taken into account, for all measured quantities. Symbols as in Table 1. All country codes refer to the ISO classification, except for Greece (EL). $'*'$: values below the 95^{th} percentile threshold for uncorrelated signals listed in Table A1 in the Appendix.

with the US, such as the UK and Turkey. The significant increase in the DFA metrics of these countries, that is observed when US collaborations are accounted for, implies that the collaborations with US firms and inventors are long-standing and persistent in time. In particular, when you remove US collaborations, the time series for the radius of co-invention in Turkey become uncorrelated in time, even. Interestingly, also countries which entertain solid relationships with the UK display a steady increase in all metrics (i.e. Belgium, Netherlands and France; see panel A in Fig.2). On the contrary, the grouping of negatively affected countries reveals a block of notoriously interconnected countries from the economic point of view (i.e. Austria, Germany, Switzerland, Poland and Romania). Interestingly, Switzerland emerges as the most negatively affected country, as only the scaling of cross-border co-invention links increases, rather slightly. We note that this does not imply necessarily that Switzerland is performing worse than other countries (e.g. has less collaborations with the US; as hinted by data on life sciences industry in Owen-Smith et al., 2002, the opposite is probably true), but that its collaborations with the US are more erratic, which also means that they could be more ubiquitous and dynamic. All other countries show a non-univocal behavior, although many improve their co-assignment scaling value, and, more expectedly, cross-border links become more stable in time.

4.2 Correlation of national DFA scaling values with features of the local collaboration graph

From the previous results, it seems apparent that external factors play a role in determining the observed scaling values in different countries. Correlations between patenting

Figure 2: Percentage changes in DFA-2 scalings for all the selected metrics and in each of the selected EU countries, when the whole EU+US subsystem is considered in metrics calculation. Positive values indicate that metrics increase when US collaborations are accounted for, and viceversa. Light blue bars: radius of co-assignment ρ_A; dark orange: radius of co-invention ρ_I; yellow: cross-border links in co-assignments CB_A; purple: cross-border links in co-inventions CB_I. '+/-': the positive/negative difference is higher in magnitude than the average estimation error for the corresponding length of the time series (see the Appendix and Table A1). '*': the difference is such that the scaling passes from uncorrelated to correlated (green star) or viceversa (red star). Uncorrelation is defined according to the thresholds in Table A1 in the Appendix. Panel A ("Positively affected"): countries whose metrics are positively affected by US collaborations (i.e. at least one shows a significant increase and, concurrently, there are no significant decreases). Panel B ("Negatively affected"): countries whose metrics are negatively affected by US collaborations (i.e. at least one shows a significant decrease and, concurrently, there are no significant increases); Panel C ("Contrasted/neutral"): countries whose metrics show either no significant change, or concurring significant increases and decreases.



rates in general and local economies have been studied in the past (Guellec and de la Potterie, 2001; De Rassenfosse and van Pottelsberghe de la Potterie, 2007) and seem redundant in our case. What might be more interesting, perhaps, is to see whether certain features of the national collaboration networks can be associated with the observed scaling values. This might shed light on the organizational properties of national R&D systems that seemingly promote a greater persistence of technological collaboration time series. In this context, we point out that our previous results on the heterogeneity of DFA scaling values among EU countries and, also, of their dependence on US collaborations, show that a completely integrated European Research Area is far from being fulfilled. This is interesting also from a policy perspective, since it is known that, for instance, EU initiatives towards technological collaboration have been directed to increase certain properties of the EU R&D network (e.g. clustering), even though the impact on actual knowledge diffusion is controversial (Cowan and Jonard, 2004). To this end, we extract the cumulative weighted adjacency matrices of the selected countries for years 2010-2014, in which the weight of each link is defined as the inverse of the number of collaborations between each pair of nodes. We use this cumulative network as a representation of the current state of the collaboration network, in terms of topological properties. This is performed for co-assignment and co-invention networks both. We then calculate a set of graph properties that we deem relevant to describe the topological structure of the network (i.e. density, assortativity, average closeness, network clustering coefficient, central point dominance and diameter). We then correlate these topological features with the observed scaling values, normalized by the corresponding standard deviation to account for the different variances arising from different numbers of active regions and different scaling values (see the Appendix for details). Results are shown in Table 5.

In general, only co-assignment scaling values turn out to exhibit significant correlations with graph properties, hinting at the fact that systemic properties might be more important for co-assignment time series, rather than for co-invention ones. This was also apparent from the actual scaling values (Table 5), where countries with seemingly uncorrelated co-assignment time series had positively correlated co-invention time series.

Among all the graph properties we extracted, a few ones emerge as associated with higher DFA scaling values. In particular, the diameter (i.e. the greatest distance between any pair of vertices) of national co-assignment networks is positively correlated to DFA scaling values at very high significance level, highlighting the role of more developed and complete networks. Assortativity (a measure of the attachment preference to nodes with similar degree) and clustering are the other two important characteristics showing significant, positive correlations with DFA scaling values. Considering that all these three features show themselves significant, positive cross-correlations, it results that persistence in time of the co-assignation metrics is related to systems with specialized – but strongly connected – subnational clusters.

Graph metric	ρ_A	$\rho_{A,US}$	CB_A	$CB_{A,US}$	ρ_I	$\rho_{I,US}$	CB_I	$CB_{I,US}$
Density	0.14	0.16	0.03	0.03	0.35	0.32	0.43	0.43
Assortativity	0.49^*	0.50^*	0.55^*	0.55^*	-0.05	-0.06	0.04	0.04
Closeness	0.24	0.22	0.33	0.32	0.03	0.01	-0.01	0.01
Clustering	0.44	0.43	0.51^*	0.49^*	0.29	0.28	0.23	0.23
CPD	0.27	0.26	0.23	0.21	-0.11	-0.11	-0.22	-0.23
Diameter	0.71^{**}	0.71^{**}	0.71^{**}	0.71^{**}	-0.14	-0.16	-0.03	-0.03

Table 5: Correlation coefficients between normalized DFA-2 scalings in each of the selected 17 EU countries and the corresponding graph metrics (CPD: Central Point Dominance). Asterisks show the p-value associated with the statistically significant correlation coefficients: *: $p < 0.1$; **: $p < 0.05$. Other symbols as in Table 1.

5 Conclusions

We have applied Detrended Fluctuation Analysis to time series of selected metrics concerning international patent collaborations (be them co-assignments or co-inventions), based on the OECD RegPat database. In particular, we have focused on two metrics: the average yearly radius of collaboration and the yearly number of cross-border collaborations. We have extracted these metrics at the level of OECD TL3 regions, the spatial segmentation provided by RegPat, limited to the EU and US systems.

We can summarize the main conclusions of this work as follows:

- when evaluating long-term correlation properties at "continental" scale (that is, EU, or US, or EU+US), the time series of the selected metrics show that the increases that we observe in our data set are persistent in time and, as such, can be expected to protract after our observation period; also, these persistence properties are resistant, to some extent, to the removal of the best performing regions;

- when calculating scaling values at EU national level, different qualitative and quantitative behaviors emerge; the observed changes in scaling values when US collaborations are taken into account shed light on which countries have stronger overseas relationships; in particular, the UK, Turkey and traditionally related countries such as Belgium, France and the Netherlands all increase their persistence metrics, when US collaborations are accounted for; on the contrary, the "German block" (i.e. Germany and German-speaking countries and also German neighbors such as Poland) shows significant decreases;

- even though the number of selected nations is necessarily low, several and significant associations between DFA scaling values and features of the national internal collaboration graph are found; in particular, co-assignment persistence in general emerges as correlated to the diameter (i.e. the maximum distance in the graph), assortativity (i.e. the attachment preference of nodes to nodes of similar degree) and clustering of said graph, highlighting the role of the development of national R&D systems in guaranteeing persistence of international or simply long-distance collaborations.

We maintain that this study can help deepen our understanding of the intrinsic properties of the evolution of international cooperation on R&D, in terms of the persistence in time of the distance between collaborating regions and of the number of cross-border collaborations each region has. In particular, our results confirm that there is a general, persistent increase of relevant metrics of international technological collaborations, especially in the EU, but they also show that the performance of single EU countries and also the dependence on US collaborations are extremely varied. This latter result can shed further light on the difficulties in creating a fully integrated European Research Area.

Despite the shortcomings of the data (i.e. series being short and, often, non-stationary), we have performed several controls on our estimation of the DFA scaling to improve the consistency of our results. Up to now, applications of DFA to short data series have been scarce, to the best of our knowledge, and we believe that our effort represents a relevant methodological advance in this respect. In particular, with respect to previous literature (Viswanathan et al., 1996), we have provided a procedure to associate measures of statistical confidence to the observed scaling values.

Having highlighted the limitations of this study, a more extensive characterization of spatial-temporal patterns of international collaboration can come from the study of individual inventors, especially when a finer time scale can be accounted for. Studies on the exploration of the "space of ideas" are also a potential extension of this method to the study of innovation processes. A finer time scale is available from the USPTO open dataset (US Patent and Trademark Office, 2018), even though such individual-based study has to rely on a thorough disambiguation of patent records. A few methods indeed have been applied with considerable effort and success (Li et al., 2014; Morrison et al., 2017), so we believe this study might set a new path for time series analysis in the field of innovation studies.

References

Arora, A.; Gambardella, A. *Res Policy* **1994**, *23*, 523–532.

Arora, A.; Fosfuri, A.; Gambardella, A. *Ind Corp Change* **2001**, *10*, 419–451.

Orsenigo, L.; Pammolli, F.; ; Riccaboni, M. *Res Pol* **2001**, *30*, 485–508.

Freeman, C. *Res Policy* **1991**, *20*, 499–514.

Chessa, A.; Morescalchi, A.; Pammolli, F.; Penner, O.; Petersen, A. M.; Riccaboni, M. *Science* **2013**, *339*, 650–651.

Morescalchi, A.; Pammolli, F.; Penner, O.; Petersen, A. M.; Riccaboni, M. *Res Policy* **2015**, *44*, 651–668.

Frenken, K.; Hoekman, J.; Kok, S.; Ponds, R.; van Oort, J., F.and van Vliet In *Death of Distance in Science? A Gravity Approach to Research Collaboration.*; Pyka, A., Scharnhorst, A., Eds.

Hoekman, J.; Frenken, K.; Tijssen, R. *Res Pol* **2010**, *39*, 662–673.

Waltman, L.; Tijssen, R.; van Eck, N. *J Inform* **2011**, *5*, 574–582.

Maisonobe, M.; Eckert, D.; Grossetti, M.; Jegou, L.; Milard, B. *J Inform* **2016**, *10*, 1025–1036.

Peng, C.-K.; Buldyrev, S. V.; Havlin, S.; Simons, M.; Stanley, H. E.; Goldberger, A. L. *Phys Rev E* **1994**, *49*, 1685.

Hardstone, R.; Poil, S.-S.; Schiavone, G.; Jansen, R.; Nikulin, V. V.; Mansvelder, H. D.; Linkenkaer-Hansen, K. *Front Physiol* **2012**, *3*, 450.

Viswanathan, G. M.; Afanasyev, V.; Buldyrev, S.; Murphy, E.; Prince, P.; Stanley, H. E. *Nature* **1996**, *381*, 413.

Disdier, A.-C.; Head, K. *Rev Econ Stat* **2008**, *90*, 37–48.

Hu, K.; Ivanov, P. C.; Chen, Z.; Carpena, P.; Stanley, H. E. *Phys Rev E* **2001**, *64*, 011114.

Owen-Smith, J.; Riccaboni, M.; Pammolli, F.; Powell, W. W. *Manag Sci* **2002**, *48*, 24–43.

Delignieres, D.; Ramdani, S.; Lemoine, L.; Torre, K.; Fortes, M.; Ninot, G. *J Math Psychol* **2006**, *50*, 525–544.

Bryce, R.; Sprague, K. *Nat Sci Rep* **2012**, *2*, 315.

Castillo, R. D.; Kloos, H.; Holden, J. G.; Richardson, M. J. *Front Physiol* **2015**, *6*, 138.

Schengell, T.; Lata, R. *Pap Reg Sci* **2013**, *92*, 555–577.

Arrieta, O. A. D.; Pammolli, F.; Petersen, A. *Sci Adv* **2017**, *3*, e1602232.

Guellec, D.; de la Potterie, B. v. P. *Res Policy* **2001**, *30*, 1253–1266.

De Rassenfosse, G.; van Pottelsberghe de la Potterie, B. *OxRep* **2007**, *23*, 588–604.

Cowan, R.; Jonard, N. *JEDC* **2004**, *28*, 1557–1575.

US Patent and Trademark Office, PatentsView database. 2018; http://www.patentsview.org/.

Li, G.-C.; Lai, R.; D'Amour, A.; Doolin, D. M.; Sun, Y.; Torvik, V. I.; Yu, A. Z.; Fleming, L. *Res Policy* **2014**, *43*, 941–955.

Morrison, G.; Riccaboni, M.; Pammolli, F. *Nat Sci Rep* **2017**, *4*, 170064.

A Appendix

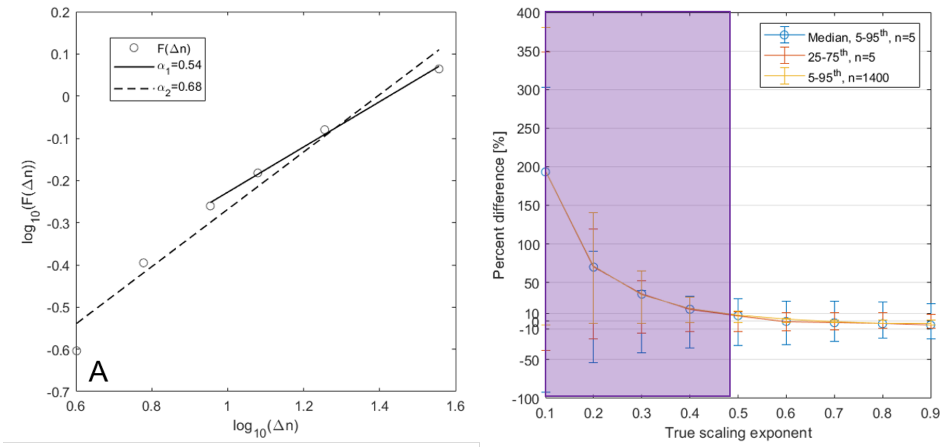
A.1 Specific methodology of DFA application

We devise a specific methodology to recover the DFA scaling value, given the particular conditions of our case study, i.e. short, "composite" time series, possibly characterized by non-stationarity. We generate n synthetic Gaussian noise series of length t (in our case, $t = 36$), characterized by a specific scaling value (i.e. $\alpha = 0.5$). Here, we choose $n = 1000$ to decrease the estimation error and thus to pinpoint the effects of trend. As performed in Hu et al. (2001), we superimpose to an arbitrary fraction of these series a trend of a given order. We estimate the DFA scaling for $1,000$ realizations and for different values of the fraction of nodes with a superimposed trend (spanning from 10% to 100%). In general, we note that DFA can best detect the true scaling of the underlying signal when the first two window lengths (i.e. $\Delta n = \{4, 6\}$), which likely suffer from bad parameterization, are discarded (Fig. A.1A). In the DFA-1 case, scaling turns out to be $\simeq 2$ as in Hu et al. (2001) for any non-negative fraction of nodes with a linear trend ($\alpha_{DFA-1} = 2.00 \pm 6.34 \cdot 10^{-4}$). The distributions of estimated DFA scalings at different values of the fraction of nodes with a trend, instead, turn out to be indistinguishable from the one coming from realizations with no trend superimposition in a two-sided Wilcoxon rank-sum test ($\alpha_{DFA-2} = 0.54 \pm 1.5 \cdot 10^{-3}$; see below for a quantitative analysis of the estimation error). At the same time, the presence of higher order trends (e.g. quadratic) in a fraction of the n series would manifest in a clearly non-stationary scaling (i.e. $\alpha_{DFA-2} = 3.05 \pm 3.16 \cdot 10^{-4}$). Thus we make use of DFA-2 as the best trade-off between detrending power and local overparameterization (our window lengths being highly constrained at $t = 36$).

A.2 Additional controls

We find appropriate, due to the low absolute number of data points and to the shortness of the maximum window length, to assess quantitatively the degree of error associated with our estimate, as revealed e.g. by Fig. A.1A. To this end, we perform 1,000 repetitions of the synthetic series generation, at varying scaling values, and calculate the average error between our estimate and the true scaling. For comparison purposes, we repeat

Figure A.1: Panel A: loglog plot of fluctuation function values versus the corresponding window length of 1,400 synthetically generated time series of total length 36 (gray circles) and fixed scaling (0.5). The black solid line shows the linear fit over the last 4 window lengths, while the dashed solid line shows the linear fit over the whole set of window lengths. Panel B: average and error bars between estimated and true DFA-2 scaling for 1,000 repetitions of time series generation, at varying scaling and number of nodes n. Light blue circles show the median error, while whiskers extend to distribution percentiles (light blue: 5-95th, $n = 5$; dark orange: 25-75th, $n = 5$; yellow: 5-95th, $n = 1400$). The lilac shaded panel excludes the scaling values that are outside the range found in this paper's results.



this simulation for the aforementioned minimum value $n = 5$ and for $n = 1400$, which is approximately the number of "active" EU regions(i.e. appearing in the RegPat database) and, thus, represents a measure of the value of n we use at the aggregate scale. We show in Fig. A.1B that this error converges to zero for growing values of the true underlying scaling. When $n = 5$, only in a few cases the error goes beyond 10%, for extreme values of the 5-95th error bar. Increasing the number of nodes does not change the median error dramatically but narrows the distribution of error values considerably. This seems in accordance with similar results on short data series shown in Delignieres et al. (2006). Much greater deviations are found when the underlying series is anti-correlated (shaded area in Fig. A.1B), but we anticipate that this is almost never found in our results.

Due to the relatively high standard deviation that is found when the number of regions under study is low, we deem necessary to have a measure of confidence when discriminating between uncorrelated and long-term correlated signals. This is particularly relevant for the national scale analysis, when the number of active regions spans from 5 to 402. To this end, we retrieve the 95th percentile value of the scaling value observed in, again, 1,000 repetitions of synthetic series generation of an uncorrelated signal ($\alpha = 0.5$), for each value of the number of regions that we find in the selected countries (see Table A1 for the complete list). This corresponds to performing a 1-tailed t-test and sets a sort of threshold above which we exclude uncorrelation in the observed data, with relative safety. Additionally, we compute the average percent error from the true scaling to have a comparison value for the variations we observe i.e. when considering EU/US collaborations at national scale. Table A1 shows both threshold values for all the values of the number of active regions and their correspondent country. We do not report results regarding the super–national scale, as our results are all well beyond the threshold value in all these cases.

We also deem necessary to account for the different standard deviations we observe when varying number of active regions, especially when we correlate the observed scalings to exogenous factors (see Section 4). To this end, we normalize the observed scalings by their associated standard deviation, which we obtain by observing the distribution of 1,000 synthetic realizations of signals having the same parameters (i.e. DFA true scaling, number of active regions). This is meant to remove the confounding effect of different variances in the data, when performing correlation analysis.

Country	n_{ρ_A}	α_{95}	$\bar{\Delta}_{\%}$	n_{ρ_I}	α_{95}	$\bar{\Delta}_{\%}$	n_{CB_A}	α_{95}	$\bar{\Delta}_{\%}$	n_{CB_I}	α_{95}	$\bar{\Delta}_{\%}$
AT	33	0.59	5.8	35	0.59	5.7	33	0.59	5.8	35	0.59	5.7
BE	44	0.59	5.3	44	0.59	5.3	38	0.59	5.5	44	0.59	5.3
BG	10	0.64	9.6	23	0.61	6.7	5	0.68	13.2	14	0.63	8.3
CH	26	0.60	6.4	26	0.60	6.4	25	0.61	6.6	26	0.60	6.4
DE	401	0.55	3.4	402	0.55	3.4	333	0.56	3.5	402	0.55	3.4
EL	29	0.60	6.1	39[1]	0.59	5.4	25	0.61	6.6	33	0.59	5.8
ES	50	0.58	5.0	56	0.58	4.9	41	0.59	5.5	54	0.58	5.0
FR	95	0.57	4.3	96	0.57	4.3	87	0.57	4.4	96	0.57	4.3
HR	11	0.64	9.2	19	0.61	7.1	7	0.67	11.3	15	0.62	8.0
HU	20	0.61	7.0	20	0.61	7.0	13	0.63	8.5	20	0.61	7.0
IT	109	0.57	4.1	110	0.57	4.0	80	0.57	4.5	107	0.57	4.1
NL	40	0.59	5.5	40	0.59	5.5	39	0.59	5.4	40	0.59	5.5
PL	55[2]	0.58	5.0	72	0.58	4.5	26	0.60	6.4	69	0.58	4.6
PT	19	0.61	7.1	23	0.61	6.7	11	0.64	9.2	20	0.61	7.0
RO	13[3]	0.63	8.5	36	0.59	5.7	8	0.65	10.6	30	0.60	6.2
SE	21	0.61	6.9	21	0.61	6.9	21	0.61	6.9	21	0.61	6.9
TR	27	0.60	6.2	50	0.58	5.0	9	0.65	9.9	31	0.60	6.1
UK	120[4]	0.57	4.0	125	0.57	3.9	101	0.57	4.2	124	0.57	4.1

Table A1: Number of active regions in each country and for each measured metric ($n_{\rho_{A/I}}$, $n_{CB_{A/I}}$) together with corresponding control statistics. α_{95} represents the 95^{th} percentile value of the distribution of the estimated scalings in 1,000 repetitions of a synthetically generated series with true $\alpha = 0.5$. $\Delta_{\%}$ is the percent error between the estimated and the observed scalings, averaged along the 1,000 repetitions of different values of the true scaling, ranging from 0.5 to 0.9, is also shown. Other symbols as in Table 1. Superscripts in the number of active regions indicate a difference between the EU-only and the EU+US cases. [1]: $n = 40$, $\alpha_{95} = 0.59$, $\bar{\Delta}_{\%} = 5.5$; [2]: $n = 56$, $\alpha_{95} = 0.58$, $\Delta_{\%}$=4.9;[3]: $n = 14$, $\alpha_{95} = 0.63$, $\Delta_{\%} = 8.3$; [4]: $n = 121$, $\alpha_{95} = 0.57$, $\Delta_{\%} = 4.1$.