

An investigation on Italian researchers' collaborative habits

*RAJANI KANTA SAHU
ASSISTANT PROFESSOR, Mtech,
Department of CSE
Gandhi Institute for Technology, Bhubaneswar.*

Abstract

Traditional approaches to analyzing research cooperation by area begin with categorizing publications based on the context of interest. In this paper, we suggest a new method based on a more nuanced categorization of the authors' respective fields. If the goal is to utilize the suggested approach as a standard against which to measure an individual's tendency to cooperate, then the method is more precise. This study applies the new methodology to all Italian university researchers in the hard sciences, gauging their propensity to collaborate across a spectrum of settings (both within and outside of the university) and between different types of collaborators (intramural peers, domestic partners, and international partners). We use a simulation to demonstrate that the outcomes vary significantly from those achieved by using canonical methods.

Introduction

There has been a significant increase in cooperation for the advancement of science during the last several decades. Analysis of co authorship (Melin and Persson, 1996) has repeatedly shown that the proportion of publications with a single author is steadily decreasing (Abt, 2007; Udine et al., 2012). Contextual variables, beginning with the research discipline, might influence the level of involvement in the various modalities of cooperation (intramural/extramural, domestic/international, interdisciplinary/interdisciplinary) (Gazni et al., 2012; Yossarian and Voyageur, 2004). Because of variables including expensive equipment costs, the need for massive sample sizes, and the method of assigning authorship, publications in the so-called "big scientific" fields tend to have much more authors than those in other fields (Cronin, 2001; Glance and Schubert, 2004). The diverse specialties involved and the varying propensities to cooperate of the individual scientists

might lead to a wide range of activating cooperation strategies, even within a single subject (Pipette & Ross, 1992). (Newman, 2001; Moody, 2004). In order to investigate the processes at the very root of cooperation and define the most relevant policies for its management, it is necessary to have a firm grasp on the many ways in which it manifests across domains and disciplines. This has the potential to boost research output (Wagner and Clydesdale, 2005). This article seeks to examine how researchers from various fields work together. It is common practice for such research to begin with a categorization of relevant papers. Instead, we base our methodology on the traditional divisions between scientific disciplines. This method is available due to a peculiar aspect of the Italian educational system. The Italian Ministry of Universities and Research (MIUR) keeps a database of all national academics, and each one is assigned to exactly one Scientific Disciplinary Sector, which seems

to be unprecedented (SDS). There are 370 of these fields³, which are further organized into 14 Adas at universities (Adas). Assigning authors to works makes it feasible to examine how often researchers from various subjects and disciplines work together and in what ways. It is possible to quantify the disparity between the two sets of findings by using the conventional technique based on the categorization of publications for the same population.

Scientific collaboration and its determinants

In the early stages of a scientific collaboration, when it is required to enhance familiarity and establish a climate of trust among collaborators, the ability to communicate effectively, informally, and flexibly is one of the key components for growth of productive scientific collaboration (Tractor and Landry, 1997). Therefore, it is not surprising that many partnerships are launched via in-person interactions, whether they take place in the workplace, at conferences, or as part of a well orchestrated kickoff event (Lauder, 2001; Wagner and Clydesdale, 2005). In long-distance cooperation, when monitoring is more challenging, face-to-face interactions might assist to alleviate coordination issues during the implementation phase by preventing "free riding" and reducing partner dispute (Hinds and Bailey, 2003). Houseman et al. (2010), Abram et al. (2009), and Larivière et al. (2006) all find that the likelihood of collaboration decreases as the distance between the scientists' respective home organizations increases. This trend may be due to the significance of face-to-face contacts. This would also explain why scientists from different sized universities use different forms of co-authorship (Kate, 2000): those from large universities tend to collaborate primarily with colleagues from the same university or

from foreign organizations, while those from smaller universities, due to the scarcity of their own intramural colleagues and the lower "relational" value of these connections, tend to work with colleagues belonging to other domestic universities. The overall decrease in travel expenditures in recent years is most likely connected to the rise in scientific cooperation, particularly on an international scale (Houseman et al., 2010). However, the single most significant element in the noticeable growth in extramural scientific partnerships is the spread of low-cost new communications technologies that considerably minimize the qualitative distinction between remote and face-to-face contact (Cairn cross, 1997; Olsen and Olsen, 2000).

Methodology

Studying research partnerships typically involves defining the type of partnership (intramural vs. extramural, infra-disciplinary vs. interdisciplinary, public-private vs. international, etc.), the setting (a discipline or a group of universities), and the tool (the co-author ships of the publications). Then, all the articles that may be linked to the given context are sorted according to the studied methods of cooperation. According to Gazing et al. (2012), for instance, the existence or absence of an author affiliated with a foreign organization is used to categorize the articles referable to a field as "international." The percentage of total publications in the field that are categorized as "international" provides a measure of the prevalence of international cooperation within the field. Starting with Panamanian's (1983) "Degree of Collaboration," continuing on to Lawani's (1986) "Collaborative Index," Aquifer et al(1988) 's "Collaborative Coefficient," and finally Egg he's (1991) "Revised Collaborative Coefficient," this methodology underpins

all of the principal indicators of co-authorship developed in the literature. Using a common measurement for all of your publishing data is another method for studying who wrote what. In order to assess the likelihood of scientists collaborating in the form under consideration, the single scientist is used as the primary analytical unit. For the phenomena studied by Gazing et al. (2012) once again, using individual scientists as the base analytical unit would allow for assessment of the tendency to international cooperation for the scientists that are part of a field. At least two other groups have taken a similar tack, and they are Martin-Simpered et al. (2002) and Abram et al (2011). The latter quantified Italian scholars' inclination for international cooperation by field, tallying the proportion of each scholar's total publications that were written in conjunction with foreign organization colleagues. Although limited to 93 Spanish university-based geologists, Martin-Simpered et al. determined each researcher's "degree of collaboration," which they defined as the ratio of coauthored publications to the researcher's total number of publications, and their "degree of national collaboration," which they defined as the ratio of coauthored publications with colleagues from at least one national organization to the researcher's total number of publications.

Methods of obtaining data and scope of study Our investigation relied on data from the aforementioned Ministry of Universities and Research database, which included profiles of Italian academics. The authors then take the data set of these authors' papers and pull it from the Italian Observatory of Public Research (ORP), a database they created and manage using data licensed from the Woos. By starting with the raw data of Italian publications in Woos between 2006 and 2010, and then applying a complex algorithm for disambiguate the true identity of the authors and their institutional affiliations

(for details see Tangelo et al., 2011), we are able to attribute each publication⁴ to the university scientist or scientists (full, associate, and assistant professors) that produced it, with a harmonic average of precision and recall (F-measure) equal to 96 (error of 4%). The biometric data set includes the following information for each publication: a full list of authors; a full list of authors' addresses; a sub-list of solely the academic authors, with their SDS/UDA and university affiliations.

Indicators and methods

We will begin with a single scientist from a well-established field and compare their average propensity to collaborate in four different forms: in general, within their own institution, with researchers from other institutions in their home country, and with scientists from institutions in other countries. The first kind of cooperation is a super set of the others; it is the predisposition to cooperate in general. We build a "author-publication" matrix with m rows and n columns according to the number of active academics and n publications. This matrix has a size of 36,211 by 197,460. Next, we link each scholar with his or her output (p) within that time frame. For each professor, we can determine the total number of collaborative publications (cp), the number of intramural (within the same institution) publications ($clip$), and the number of extramural (within other domestic institutions) publications ($cede$) because we know the total number of authors and the total number of Italian and foreign organizations involved in each publication (extramural international - cap). Using these numbers, we may calculate indications of people's varying propensities to work together, from which we can also get average inclinations by industry and specialty:

- Propensity to collaborate $C = \frac{cp}{p}$
- Propensity to collaborate intramurally $CI = \frac{cip}{p}$
- Propensity to collaborate extramurally at the domestic level $CED = \frac{cedp}{p}$
- Propensity to collaborate extramurally at the international level $CEF = \frac{cefip}{p}$

Results and discussion

The many types of co-authorship may be analyzed, and distinct UDAs and their individual SDSs can be described, using the calculated C, CI, CED, and CEF values based on the registered propensity values for respective member academics. Our results from sections 4.1 and 4.2 detail these analyses. Later, in Section 4.3, we look at how these four metrics are related to one another.

Collaborative Tendency 4.1 Propensity to Work Together in Different Fields

Academics from the different UDAs studied had varying propensities for cooperation in all forms, including intramural, extramural domestic, and international partnerships. We give a table for each kind of partnership, illustrating, on a per UDA basis: I) the proportion of UDA faculty members with zero collaboration propensity; II) the percentage of UDA faculty members with maximum (100%) collaboration propensity; III) the average value of the UDA faculty's collaboration propensity. The Kruskal-Wallis test (Kruskal and Wallis, 1952) is applied to all of the UDAs, and the Mann-Whitney U test (Mann and Whitney, 1947) is used to confirm the differences in the inclinations recorded for the academics who belong to each field. These non-parametric tests allow us to see whether there is a greater or lesser tendency for academics to work together in one UDA compared to another. Using the `kruskal.test` and `Wilcox.test`

functions, we do this analysis, and the findings (available in Supplemental Material - S1) demonstrate that almost all of the comparisons we made had a high level of significance. Conclusions allow for grouping UDAs according to their varying degrees of cooperation. The values of cooperation propensity are shown in Table 2. These seem extraordinarily high, which is consistent with other previous studies using alternative methods demonstrating that the proportion of co-authored articles within the "biometric" fields is currently over 90%. (Abe, 2007; Gazing et al., 2012). Table 2 shows that many UDAs do not vary much from one another in their inclination to cooperate, despite the fact that these differences are typically statistically significant based on the findings of the Mann Whitney U test. The average willingness to work together approaches 100% in the fields of medicine, agriculture and veterinary medicine, biology, and chemistry. All in all, these findings corroborate those of Haiti and Hong (1997) and Gazing et al.

Table 2: Propensity to collaborate, per UDA (percentage values)

UDA	Mean C	% C = 0%	% C = 100%
Medicine (MED)	99.4	0.1	94.8
Chemistry (CHE)	99.2	0.1	94.8
Agricultural and veterinary sciences (AVS)	99.1	0.2	95.7
Biology (BIO)	99.1	0.1	94.4
Earth sciences (EAR)	97.6	0.7	90.6
Industrial and information engineering (IIE)	97.1	0.5	85.5
Pedagogy and psychology (PPS)	96.7	1.4	89.8
Physics (PHY)	96.6	1.0	81.5
Civil engineering (CEN)	94.3	1.7	81.5
Mathematics and computer sciences (MAT)	89.1	3.3	68.6
Economics and statistics (ECS)	84.0	8.3	70.1
Total	97.2	0.9	89.0

When comparing intramural cooperation, the disparities between the different UDAs seem to be considerably more evident. Table 3 demonstrates a disparity of about 40% between the UDA with the highest value (Chemistry) and the one with the lowest value (Physics) (Economics and statistics). Similar to what was shown using the Mann-Whitney U test, the

probability of collaborating with colleagues from the same institution is rather high in the four UDAs of Medicine, Agricultural and Veterinary Sciences, Biology, and Chemistry. This finding makes sense when you take into account the fact that professors in these fields often work in laboratories owned by their own institution, which are sometimes shared for budgetary reasons between other colleagues, encouraging the growth of cooperation. When it comes to industries, Industrial and information engineering have the second greatest tendency for this kind of cooperation. This finding makes sense when you think about the fact that, like many other academic fields, this one relies heavily on shared resources like labs, equipment, and software between faculty members at the same institution, making it easier to foster teamwork. In addition, many engineering studies are the end result of research projects commissioned by businesses and carried out by academics, who typically collaborate with their peers at the same institution rather than those at other institutions to save money on travel, time, and other overhead costs while increasing their reach throughout the region.

Table 3: Propensity to intramural collaboration, per UDA (percentage values)

UDA	Mean CI	% CI = 0%	% CI = 100%
Chemistry (CHE)	83.5	2.4	46.1
Industrial and information engineering (IIE)	82.2	3.9	46.9
Agricultural and veterinary sciences (AVS)	81.2	4.3	51.8
Medicine (MED)	81.1	3.6	45.9
Biology (BIO)	78.8	4.2	45.8
Civil engineering (CEN)	73.4	8.8	46.3
Physics (PHY)	66.7	8.7	29.2
Earth sciences (EAR)	62.0	11.4	31.1
Pedagogy and psychology (PPS)	59.6	18.2	35.8
Mathematics and computer sciences (MAT)	54.1	20.5	25.4
Economics and statistics (ECS)	44.0	36.0	26.7
Total	75.4	7.2	42.3

One way to categorize extramural partnerships is by whether or not the extramural organization is located inside or outside of the same country as the collaborating institution. Average propensities to work

with scientists from different domestic organizations outside of the institution are shown in Table 4. Once again, disparities across fields are quite large, with almost half a century separating the UDAs with the highest (Physics) and lowest (Social Studies) average propensities, respectively (Industrial and information engineering).

Table 4 shows the national average UDA's propensity for extramural cooperation (percentage values)

UDA	Mean CED	% CED = 0%	% CED = 100%
Physics (PHY)	72.5	5.6	24.9
Medicine (MED)	62.4	8.2	20.6
Earth Sciences (EAR)	58.6	13.1	23.3
Biology (BIO)	57.4	9.8	17.5
Chemistry (CHE)	49.8	8.3	9.3
Pedagogy and Psychology (PPS)	48.5	26.3	22.2
Agricultural and veterinary sciences (AVS)	47.1	17.3	13.9
Economics and Statistics (ECS)	38.0	38.9	19.1
Mathematics and Computer Sciences (MAT)	33.6	33.3	10.7
Civil engineering (CEN)	26.0	44.3	8.2
Industrial and information engineering (IIE)	24.8	33.1	5.3
Total	50.3	17.0	15.7

According to the findings about the inclinations to interact in various ways (Section 4.1), in certain UDAs, academics have a tendency to work with scientists from both their own institution and from other organizations. This trend is most pronounced in areas of study where a significant amount of institutional support is needed or when collaboration with other institutions is essential. Sometimes, scientists may prefer one kind of cooperation over another due to the varying degrees of coordination necessary under each. The trend toward publications with a small number of authors may also influence the selection of a single mode of cooperation in particular fields. The Spearman non-parametric correlation between each professor's results on the four indicators C, CI, CED, and CEF is calculated using the R record function (R Development Core Team, 2012) to assess the connections between the various collaboration propensities. Table 8 displays the

findings, allowing one to assess the degree to which the four types of cooperation are associated both globally and for each UDA.

Table 8 shows the association between markers of collaboration propensity using the Spearman test, as reported by UDA.

UDA	C-CI	C-CED	C-CEF	CI-CED	CI-CEF	CED-CEF
AVS	0.20***	0.10***	0.05**	-0.35***	-0.25***	0.00
BIO	0.23***	0.12***	0.02	-0.21***	-0.25***	-0.03
CEN	0.43***	0.14***	0.10***	-0.44***	-0.25***	0.04
CHE	0.25***	0.12***	0.04*	-0.28***	-0.21***	-0.01
EAR	0.21***	0.21***	0.11***	-0.28***	-0.25***	0.08**
ECS	0.41***	0.35***	0.27***	-0.20***	-0.23***	-0.03
IIE	0.38***	0.12***	0.06***	-0.44***	-0.32***	0.15***
MAT	0.43***	0.25***	0.20***	-0.32***	-0.30***	0.03
MED	0.17***	0.10***	0.02	-0.29***	-0.26***	0.07***
PHY	0.33***	0.34***	0.20***	-0.02	-0.04	0.37***
PPS	0.24***	0.21***	0.14***	-0.30***	-0.30***	-0.03
Total	0.36***	0.21***	0.08***	-0.21***	-0.27***	0.12***

Significance level: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

As the statistics reveal, there is a positive relationship between C and every other metric of interest. This finding holds true across all UDAs, albeit to varying degrees, and hints at how scientists who embrace various types of cooperation show a stronger propensity to cooperate in general. With such wide variation across UDAs, it might be difficult to make sense of the correlation between C and CEF, or the tendency to interact at the international level. In Economics and statistics, in particular, there is a favorable and statistically significant association between academic collaboration and the activation of relations with foreign organizations (Table 4, which shows that only 60% of academics in this field communicate).

Conclusions

Several researchers have taken an interest in the study of research collaboration's many forms, hoping to learn if and how patterns of collaboration differ between

disciplines and sectors, and to theorize about the factors that might underlie such differences. Indicators of incidence based on counting articles have been used up till now. Instead, the authors here advocate for a novel methodological approach that takes the individual scientist as its unit of analysis. There are many benefits to using this method, including the fact that it facilitates the study of inter-disciplinary partnerships on a grand scale. In addition, the proposed method allows for a more accurate depiction of researchers' inclination to collaborate in various forms, whether with direct colleagues or with other organizations: in fact, quantifying the phenomenon of collaboration through counting publications implies obvious distortions in the case where productivity, apart from collaboration intensity, is not distributed in a homogeneous fashion (the real-world situation) am. The correct ex-ante definition and ex-post control of policies to create, alter, or maintain the conditions for various forms of collaboration within any reference context depends on the implementation of reliable collaboration measurement systems. Given the mutually beneficial effects that scientific collaboration can have on a country's ability to generate and disseminate new information, it's no surprise that many countries have policies designed to encourage it. The individual scientist's propensity to collaborate can be measured, allowing the policy's impact to be verified on the actors who are ultimately the policy's goal. In addition, beginning with data on individual scientists, one can obtain the measurement of the propensity to collaborate for the research group and organizational unit at increasing levels, which can then be the subject of specific policy. In conclusion, our method provides more appropriate measures than those previously proposed in the literature to support the implementation of policy aimed at influencing scientific collaborations.

By applying our method to the research efforts of Italian academics, we have been able to quantify the extent to which they are inclined to engage in a variety of forms of interdisciplinary cooperation. The findings can inform future policy decisions about how to best encourage collaborations between researchers across disciplines, as well as how to evaluate the efficacy of existing initiatives at individual institutions and across the research enterprise.

References

- Abramo, G., Angelo, C.A., Di Costa, F., Palazzi, M., (2009). *University-industry collaboration in Italy: a biometric examination*. *Innovation*, 29(6-7), 498-507.
- Abram, G., Angelo, C.A., Di Coats, F. (2009a). *Research collaboration and productivity: Is there correlation?* *Higher Education*, 57(2), 155–171.
- Abram, G., Angelo, C.A., Palazzi, M. (2011). *The relationship between scientists' research performance and the degree of internationalization of their research*. *Econometrics*, 86(3), 629–643.
- Abram, G., Cicero, T., Band, E. (2000). Angelo, C.A. (2011a). *The dangers of performance-based research funding in non-competitive higher education systems*. *Econometrics*, 87(3), 641-654.
- Abram, G., Angelo, C.A., Di Coats, F. (2012). *Identifying interdisciplinary through the disciplinary classification of coauthors of scientific publications*. *Journal of the American Society for Information Science and Technology*, 63(11), 2206–2222.
- Abe, H.A. (2007). *The future of single-authored papers*. *Econometrics*, 73(3), 353–358.
- Acedo, F.J., Barbarossa, C., Casanova, C., Galen, J.L. (2006). *Co-authorship in Management and Organizational studies: An empirical and network analysis*. *Journal of Management Studies*, 43(5), 957–983.
- Aquifer, I., Barrel, Q., Ague, J. (1988). *Collaborative coefficient: A single measure of the degree of collaboration in research*. *Econometrics*, 14(3-4), 421–433.
- Archibald, D., Coco, A., (2004). *International partnerships for knowledge in business and academia a comparison between Europe and the USA*. *Innovation*, 24, 517–528.