Team Production, Scientific Competition and Interdisciplinary Research

Carayol, N.; Maublanc, F.

STI 2018 Leiden

23rd International Conference on Science and Technology Indicators
"Science, Technology and Innovation Indicators in Transition"

STI 2018 Conference Proceedings

Proceedings of the 23rd International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

Chair of the Conference

Paul Wouters

Scientific Editors

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

Layout

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

The articles of this collection can be accessed at https://hdl.handle.net/1887/64521

ISBN: 978-90-9031204-0

Team Production, Scientific Competition and Interdisciplinary Research

Nicolas Carayol[*], François Maublanc[**]

[*] *nicolas.carayol@u-bordeaux.fr*
GREThA, University of Bordeaux, Avenue Léon Duguit, PESSAC, 33608 (FRANCE)

[**] *francois.maublanc@u-bordeaux.fr*
GREThA, University of Bordeaux, Avenue Léon Duguit, PESSAC, 33608 (FRANCE)

Introduction

As scientific enquiry complexifies and as communities of researchers expand, Science becomes increasingly specialized, fragmented into consistent bodies of knowledge and their associated research communities, called disciplines. A discipline can be defined as a "branch of knowledge, instruction, learning, teaching, or education" (Alvargonzález, 2011). As such, scientific disciplines also define socio-professional communities for research production, evaluation and teaching. However, real world problems are "rarely confined to the artificial boundaries of academic disciplines" and "cannot be adequately addressed by single disciplines alone" (Choi and Pak, 2006; p. 357-358). To solve them it is thus often necessary to combine skills and efforts from different disciplines.

In this article, we focus on interdisciplinarity as a team production problem that we embed in a Tullock competition framework. Team leaders may undertake research projects that are characterized by a singular CES production function. Production factors are sub-team efforts, each one in a distinct discipline. Each project is defined as a given vector of production coefficients and an elasticity of substitution between factors. Formally, our model is close to Kolmar and Rommeswinkel (2013). We however leave aside the question of free-riding in teams and consider that a leader maximizes a group utility function to focus on the impact of the production coefficients and of the elasticity of substitution on the variables at equilibrium.

Solving the model, we show that outcome depends negatively on a diversity index of the discipline contributions, which is also function of the elasticity of substitution. It turns out that under certain sufficient conditions, our index and measures of interdisciplinarity point in the same direction, meaning that project interdisciplinarity goes in the opposite direction than all performance indicators. These results are consistent with most scientific research being performed within a single discipline, and articles impact to be decreasing with interdisciplinarity. Our model can also explain that some interdisciplinary research actually have huge impact as a result of a selection bias, not as an outcome of interdisciplinarity. Our results support the idea that the outcome is less sensitive to interdisciplinarity when the team is of higher efficiency.

We test our predictions on a database of more than four hundred thousant scientific articles authored by a set of nearly thirty thousand professors employed by French research institutions.

The model

Consider a research project i involving m_i disciplines. To realize it, researchers must exert efforts in each discipline. Denote $x_i = \left(x_{i1}, \dots, x_{im_i}\right)$ the vector of efforts of dimension m_i, where each entry of coordinate $s = 1, \dots, m_i$ is x_{is}, the effort of team i in discipline s. It is equivalent to see the participation of each discipline in the project as corresponding to a subteam effort. X_i represents the aggregate effort :

$$X_i = \sum_{s=1}^{m_i} x_{is}.$$

Let ω_{is} denote the efficiency of effort in discipline s in this very project i, and $\omega_i = (\omega_{i1}, \dots, \omega_{im_i})$ the vector of efficiencies. Without loss of generality, assume that disciplines in each project i are sorted in descending order of efficiency :

$$\omega_{i1} \geq \omega_{i2} \geq \cdots \geq \omega_{im_i},$$

and we normalize those coefficients such that $\sum_{s=1}^{m_i} \omega_{is} = 1$.

In our model we assume the following hypotheses :

A1. The output of research project i, q_i, is obtained via a constant elasticity of substitution (CES) production function as follows:

$$q_i = \Omega_i \left(\sum_{s=1}^{m_i} \omega_{is}\, x_{is}^{\sigma_i} \right)^{\frac{1}{\sigma_i}},$$

with Ω_i the group efficiency parameter and $\sigma_i \in]-\infty; 0[\cup]0; 1]$ the elasticity of substitution between the efforts spent in each discipline.

A2. Competition in science is modelled by a Tullock contest where the probability p_i that project i obtains the prize V is equal to

$$p_i = \frac{q_i}{Q},$$

with Q the total research production:

$$Q = \sum_{l=1}^{G} q_l,$$

and G the number of research projects in competition to obtain a prize of value V.

A3. A leader makes the decisions on efforts for the whole project, meaning that we suppose no free-riding in research teams.

A4. The cost function of efforts is linear and unitary, *i.e.* $c(x) = x$.

Let u_i define the utility function of project i :

$$u_i = p_i V - X_i.$$

The leader of each project $i = 1, \dots, G$ chooses the vector x_i of efforts to maximize the payoff function:

$$\max_{x_{i1}, \dots, x_{im_i}} u_i$$

We define

$$\Gamma_i = \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} \right)^{1-\frac{1}{\sigma_i}}, \forall i = 1, \dots, G$$

to be a synthetic indicator of the production function of team i. Without loss of generality, let's suppose that research projects are sorted in ascending order so that:

$$\frac{\Gamma_1}{\Omega_1} \leq \frac{\Gamma_2}{\Omega_2} \leq \cdots \leq \frac{\Gamma_G}{\Omega_G}.$$

Let define $g \leq G$ the number of active research projects in the competition.

After computing equilibrium effort spent by team i in discipline j (x_{ij}^*), we obtain the other variables at equilibrium : Q^*, p_i^*, u_i^* and X_i^*. All of them depend on the synthetic indexes Γ_j and Ω_j , $\forall j = 1, \ldots, g$.

To appreciate Γ in concrete terms, we rewrite it as a function of the contributions of the disciplines to the production. Note that the CES function allows us to determine this last variable. Denote λ_{ij} the contribution of each discipline j to the production of team i :

$$\lambda_{ij} = \frac{\omega_{ij} x_{ij}^{\sigma}}{\sum_{s=1}^{m_i} \omega_{is} x_{is}^{\sigma}}.$$

At equilibrium, λ_{ij}^* is given by:

$$\lambda_{ij}^* = \frac{\omega_{ij}^{\frac{1}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}}}$$

Further, it is possible to rewrite Γ_i as:

$$\Gamma_i = \left(\sum_{s=1}^{m_i} (\lambda_{is}^*)^{1-\sigma_i} \right)^{\frac{1}{\sigma_i}}$$

Γ_i is in fact an index of diversity (Hill index) of the contributions of the disciplines to the production of team i :

$$\Gamma_i = H_i^{1-\sigma_i}(\lambda_i^*)$$

In our model, we obtain the following results:

R1. Teams that are characterized by a larger Γ_i are disadvantaged in the contest as their production, their probability to win as well as their participation decrease.

R2. A higher team efficiency decreases the penalty to produce diversity.

R3. If other teams have a higher Γ_j, then the penalty of team i to engage a project with a higher Γ_i is lower.

R4. Raising competition has a positive effect on the production of more disciplinary and efficient teams and a negative effect on the production of less efficient and more diverse teams.

In this paper we first intend to test empirically **R1**.

Testing empirically our model

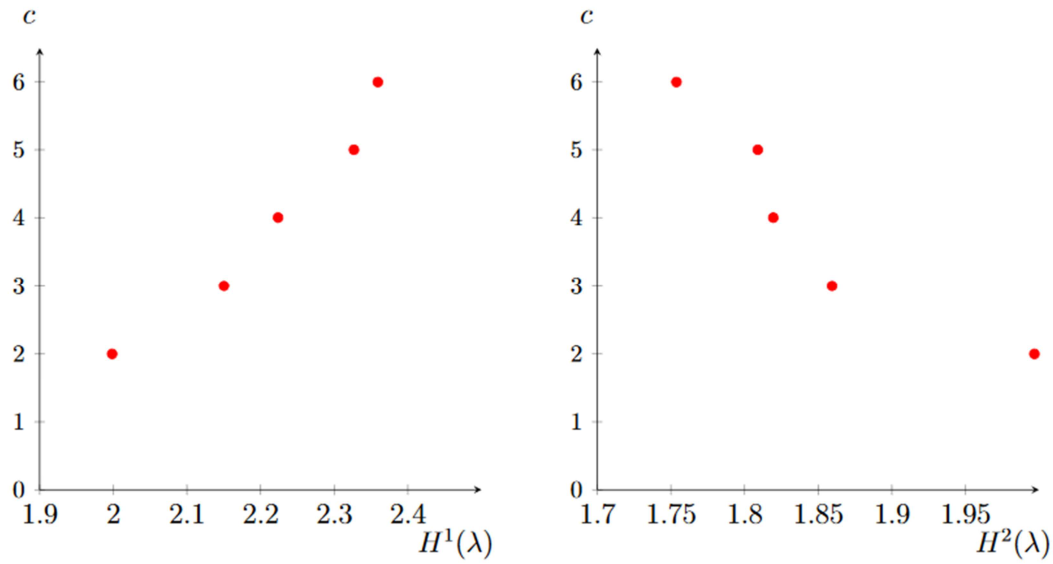
In our theoretical model, the comparative statics results obtained essentially depend on a diversity index of power $1 - \sigma_i$. Remark that σ_i is specific to each research project, and this is an unknown parameter. An econometric estimation of this parameter requires to make the hypothesis that research projects have similar technologies, meaning $\sigma_i = \sigma, \forall i$. However such estimations are difficult to conduct. In the literature, the Hill Index of power β is often employed to measure interdisciplinarity. Therefore, one way could be to understand the circumstances under which our index of diversity $H_i^{1-\sigma}$ and the measure of interdisciplinarity H_i^{β} vary in the same direction. In other words, we want to find the conditions such as the ranking of projects according to Γ is close to the ranking of projects according to H^{β}. This is not an easy question that we illustrate in the following by an example in Table 1. We consider five research projects where the first four columns present the contribution of each discipline to a given research paper i while c_i in the fifth column is the number of citations. We compute the Hill index for $\beta = 1$ (exponential of the Shannon-Wiener index) and $\beta = 2$ (Simpson index) and we plot the number of citations versus interdisciplinarity measured by these two indicators in Figure 1.

Table 1 : Contributions of each discipline λ_{ij} and impact of the paper c_i.

i	λ_{i1}	λ_{i2}	λ_{i3}	λ_{i4}	c_i
1	0.52	0.48	0	0	2
2	0.68	0.27	0.05	0	3
3	0.7	0.24	0.04	0.02	4
4	0.72	0.14	0.12	0.02	5
5	0.74	0.09	0.09	0.08	6

Although we use exactly the same data, we can conclude that interdisciplinarity has a positive effect on the number of citations if we employ the Shannon-Wiener index (Figure 1(a)) while it has a negative effect if we consider the Simpson index (Figure 1(b)). In our example, projects are ranked by ascending order of the contribution of the first discipline. The Shannon-Wiener index, more sensitive to the weight of rare disciplines than the Simpson one, is increasing as the weight of discipline 1 is similar for projects 2 to 5 but contributions are becoming more and more egalitarian between disciplines 2, 3 and 4. On the contrary, the Simpson index is more sensitive to the weight of the first discipline and therefore decreases. In this example the ranking of projects according to their interdisciplinarity is completely different depending on the choice of the index. Thus indicators of interdisciplinarity are not necessarily highly and positively correlated.

Figure 1: Number of citations as a function of Interdisciplinarity



It turns out that the rankings based on all versions of the Hill index coincide (and thus coincide with the Γ_i) when the number of disciplines is equal to two. Under this condition, we can replace Γ_i by any interdisciplinarity measurement nested in the Hill index in the above propositions (R1-R4).

Data

We test those predictions on about four hundred thousand articles authored by thirty thousand tenured researchers and faculty member employed by academic institutions in France, from 1999 to 2014 and published by journals referenced in the WoS.

Our main dependant variable is the number of citations obtained by a paper in a fixed window of three years after its date of publication. Descriptive statistics show that this variable is overdispersed, with a mean of 6,91 and a standard deviation of 16,68. A remarkable fact is the proportion of papers obtaining 0 citation, which is close to 22%. We observe huge differences between disciplines with for instance a mean number of citations equal to 8.57 in Basic Biology versus 3.87 in Human Sciences.

The independent variable of interest is interdisciplinarity. With a 11 disciplines nomenclature, 73% of the papers from the database span over at most two disciplines. For each paper we look at its reference list and attribute the disciplines of the Journal in which they were published. Thus we obtain the distribution of disciplines with their associated weight for the reference list of a given paper. To compute interdisciplinarity, we consider the Hill-index for some special values of β: $\beta = +\infty$ (Berger-Parker index), $\beta = 2$ (Simpson index) and $\beta = 1$ (exponential of the Shannon-Wiener index). These indicators take into account variety (number of disciplines involved in a paper) and balance (evenness of the distribution of disciplines). Since Porter (2009), a new dimension has been added to the measure of interdisciplinarity (disparity) to take into account the cognitive distance between the disciplines. Because this variable doesn't appear in our theoretical model, we ignore this dimension in our empirical study. Collecting the references of a paper is not easy, particularly for Social and Human Sciences. To avoid the calculation of interdisciplinary indexes with a very little number of references, we drop papers with less than four references linked to the WoS categories like in the paper of Yegros-Yegros and al (2015). We obtain a dataset of 177,525 papers published between 1999 and 2015.

We include in our regressions a list of controls that previous studies considered as influencing the number of citations. To refine the list, we first analyze the correlations between the variables given in Figure 2. We can remark that the number of authors and the number of different organisations participating to a paper are highly correlated (0.91). Even if these two variables are of interest for our study, putting them simultaneously in the regressions gives opposite signs for the coefficients. Therefore we choose to consider the number of authors rather than the number of different organisations. To control for the "quality" of researchers, we include a variable equal to the maximum of the h-index of the French researchers participating to a paper. We also control for the nature of the collaboration with a dummy equal to 1 if the paper is an international collaboration, 0 otherwise.

We include variables such as dummies for the year of publication, for the type of document (93% of the papers are research articles, the remaining being conference proceedings and letters), for the disciplines and for the researchers affiliations to French institution(s).

Figure 2 : Correlation matrix

		1	2	3	4	5	6	7	8
1	Number of citations	1							
2	Berger-Parker	-0.020***	1						
3	Simpson	-0.022***	0.97***	1					
4	Shannon	-0.023***	0.92***	0.98***	1				
5	Number of authors	0.055***	-0.027***	-0.026***	-0.020***	1			
6	Number of different organisations	0.10***	-0.045***	-0.046***	-0.042***	0.91***	1		
7	International Collaboration	0.099***	-0.023***	-0.027***	-0.029***	0.090***	0.21***	1	
8	maximum h-index	0.030***	-0.031***	-0.030***	-0.027***	0.079***	0.16***	0.11***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Empirical results

In this section, we report the results of our estimations dealing with the effect of interdisciplinarity on impact. Because the number of citations per article is over dispersed, we estimate negative binomial models rather than Poisson models. Figure 3 displays the results of these estimations in three blocks. Each of one tests the studied relationship by measuring interdisciplinarity successively with the Berger-Parker (regression "a"), the Simpson (regression "b") and the Shannon-Wiener (regression "c") indexes. In the first block we directly introduce interdisciplinarity indexes as an independent variable. The second one integrates a square term of each index to test the possible existence of a U or U-inverted relationship between interdisciplinarity and impact. In the last block, we introduce the logarithm of interdisciplinary indexes to interpret our relationship in terms of semi-elasticities.

The first block shows that the coefficient associated with interdisciplinarity is strongly significant and negative, whatever the index considered. This seems to validate the first proposition of our model, therefore the idea that ceteris paribus, increasing interdisciplinarity implies a decrease in the number of citations. After the introduction of a square term (second block), we observe a U-relationship between interdisciplinarity and the number of citations. We obtain that the value of the Berger-Parker index which minimizes the number of citations is around 1,682. 91,5% of the papers with less than two disciplines have an interdisciplinary index lower than this threshold. For the Simpson index, we obtain a threshold of 1.752 (82.9% of the papers below this value) and for the Shannon-Wiener, a threshold of 1.829 (80.3% of the papers below this value). For the last block, we can conclude that increasing by 10% the index of interdisciplinarity decreases by 4-5% the number of citations.

Looking at the other variables, the number of authors increases significantly and positively with the impact of the papers. However the effect is limited: adding one author only raises by 0.1% the number of citations. Papers which are internationally co-authored have a greater impact compared to the others as it increases by 46% the number of citations.

Robustness checks and extensions

Considering the complete database (therefore including papers with more than two disciplines) marginally changes the results compared to the previous section. We still observe the negative effect of interdisciplinarity on the number of citations. If the coefficient associated with the square term of each index is still significant, the values of the interdisciplinary indicators minimizing the number of citations are such that more than 99% of the papers are below this threshold, meaning that the turning point is only generated by extreme values. Increasing by 10% the index of interdisciplinarity decreases by 5-6% the number of citations, which is consistent with the previous results. Instead of dropping papers with less than four references, we conduct the same regressions by deleting papers with less than six references or considering the complete database. These new estimations change very slightly our results.

As an alternative to the nomenclature of OST, we employ the WoS categories. Considering successively papers with less than two WoS subject categories then all the database, we find results with the same order of magnitude. We conduct logit regressions to see the effect of interdisciplinarity on the probability that the paper belongs to the 10% highest cited papers. This doesn't change our conclusions.

In the next months, we will conduct other types of robustness checks.

Figure 3: Negative binomial estimates for the effect of interdisciplinarity on the number of citations.

	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)	(3a)	(3b)	(3c)
Berger-Parker	-0.280			-2.097					
	(-15.38)			(-10.57)					
Berger-Parker^2				0.623					
				(9.26)					
ln(Berger-Parker)							-0.420		
							(-16.17)		
Simpson		-0.276			-2.194				
		(-17.09)			(-11.06)				
Simpson^2					0.626				
					(9.78)				
ln(Simpson)								-0.432	
								(-17.78)	
Shannon			-0.340			-2.634			
			(-18.44)			(-10.68)			
Shannon^2						0.720			
						(9.40)			
ln(Shannon)									-0.551
									(-19.02)
Number of authors	0.00106	0.00105	0.00105	0.00105	0.00104	0.00104	0.00106	0.00105	0.00104
	(10.20)	(10.19)	(10.17)	(10.19)	(10.16)	(10.11)	(10.19)	(10.18)	(10.16)
International Collaboration	0.468	0.468	0.468	0.468	0.468	0.468	0.468	0.468	0.468
	(54.58)	(54.65)	(54.69)	(54.66)	(54.70)	(54.68)	(54.61)	(54.67)	(54.70)
maximum h-index	0.0309	0.0307	0.0306	0.0306	0.0305	0.0303	0.0308	0.0306	0.0305
	(43.66)	(43.53)	(43.41)	(43.46)	(43.28)	(43.17)	(43.61)	(43.48)	(43.35)
Observations	177525	177525	177525	177525	177525	177525	177525	177525	177525
Log-Likelihood	-484880.2	-484822.5	-484768.5	-484805.7	-484738.7	-484687.7	-484855.3	-484796.4	-484742.7
AIC	970044.4	969929.0	969822.9	969897.4	969763.5	969661.5	969994.6	969876.7	969771.5
BIC	971476.7	971361.3	971265.4	971339.8	971205.9	971103.9	971426.9	971309.1	971213.9

t statistics in parentheses

References and Citations

Adams, J. D., Black, G. C., Clemmons, J. R. & Stephan, P. E. (2005), 'Scientific teams and institutional collaborations: Evidence from us universities, 1981–1999', Research policy 34 (3), 259–285.

Alvargonzález, D. (2011), 'Multidisciplinarity, interdisciplinarity, transdisciplinarity, and the sciences', International Studies in the Philosophy of Science 25 (4), 387–403.

Choi, B. C. & Pak, A. W. (2006), 'Multidisciplinarity, interdisciplinarity and transdisciplinarity in health research, services, education and policy: 1. definitions, objectives, and evidence of effectiveness', Clinical and investigative medicine 29(6), 351

Kolmar, M. & Rommeswinkel, H. (2013), 'Contests with group-specific public goods and complementarities in efforts', Journal of Economic Behavior & Organization 89, 9–22.

Leahey, E., Beckman, C. M. & Stanko, T. L. (2017), 'Prominent but less productive: The impact of interdisciplinarity on scientists' research', Administrative Science Quarterly 62 (1), 105–139.

Porter, A. & Rafols, I. (2009), 'Is science becoming more interdisciplinary? measuring and mapping six research fields over time', Scientometrics 81 (3), 719–745.

Rhoten, D. & Parker, A. (2004), 'Risks and rewards of an interdisciplinary research path', Science 306 (5704), 2046–2046.

Sá, C. M. (2008), '"interdisciplinary strategies" in us research universities', Higher Education 55 (5), 537–552.

Schuitema, G. & Sintov, N. D. (2017), 'Should we quit our jobs? challenges, barriers and recommendations for interdisciplinary energy research', Energy Policy 101, 246–250.

Wang, J., Thijs, B. & Glänzel, W. (2015), 'Interdisciplinarity and impact: Distinct effects of variety, balance, and disparity', PloS one 10 (5), e0127298.

Yegros-Yegros, A., Rafols, I. & D'Este, P. (2015), 'Does interdisciplinary research lead to higher citation impact? the different effect of proximal and distal interdisciplinarity, PloS one 10 (8), e0135095.