



HAL
open science

Competition, Interdisciplinarity and Teams in Science

François Maublanc

► **To cite this version:**

François Maublanc. Competition, Interdisciplinarity and Teams in Science. Economics and Finance. Université de Bordeaux, 2018. English. NNT : 2018BORD0354 . tel-02076695

HAL Id: tel-02076695

<https://theses.hal.science/tel-02076695>

Submitted on 22 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE PRÉSENTÉE
POUR OBTENIR LE GRADE DE
**DOCTEUR DE
L'UNIVERSITÉ DE BORDEAUX**

ÉCOLE DOCTORALE ENTREPRISE, ÉCONOMIE, SOCIÉTÉ (ED 42)
SPÉCIALITÉ SCIENCES ÉCONOMIQUES

Par **François MAUBLANC**

**COMPETITION, INTERDISCIPLINARITY AND
TEAMS IN SCIENCE**

Sous la direction de
M. Nicolas CARAYOL et M. Sébastien ROUILLON

Soutenue le 13 décembre 2018

Membres du jury :

M. Francis BLOCH

Professeur des Universités, Paris School of Economics, *Président*

M. Nicolas CARAYOL

Professeur des Universités, Université de Bordeaux, *Co-directeur de thèse*

M. Luis CORCHÓN

Professeur des Universités, Universidad Carlos III de Madrid, *Rapporteur*

M. Pierre MOHNEN

Professeur des Universités, Maastricht University, *Rapporteur*

M. Sébastien ROUILLON

Professeur des Universités, Université de Bordeaux, *Co-directeur de thèse*

The University of Bordeaux is not to provide any approval or disapproval regarding the opinions this PhD dissertation includes. These opinions must be considered as being solely those of their author.

L'Université de Bordeaux n'entend ni approuver, ni désapprouver les opinions particulières émises dans cette thèse. Ces opinions sont considérées comme propres à leur auteur.

Remerciements

Mes premiers remerciements vont tout naturellement à mes deux co-directeurs de thèse, Nicolas Carayol et Sébastien Rouillon, pour leur professionnalisme, leurs conseils, leur soutien, leur disponibilité et leur patience. Vous m'avez offert l'opportunité de réaliser mon mémoire de master puis cette thèse dans les meilleures conditions. Le métier de chercheur nécessite d'être passionné, mais aussi d'être formé. J'ai énormément appris auprès de vous pendant ces cinq années.

Je remercie infiniment l'ensemble des membres du jury, Francis Bloch, Luis Corchón et Pierre Mohnen, d'avoir accepté de rapporter et d'examiner cette thèse, et de se déplacer à Bordeaux pour cette soutenance.

Je tiens à remercier l'Université de Bordeaux et plus particulièrement le GREThA pour tout le soutien matériel, administratif et financier nécessaire à la réalisation de cette thèse. Je remercie plus particulièrement son directeur, Marc-Alexandre Sénégas, qui a toujours veillé à ce que je puisse préparer ma thèse dans les meilleures conditions et permis le financement de mes participations à des conférences en France ou à l'étranger. Merci à l'ensemble de l'équipe administrative, en particulier Camille Embarek, Julie Vissaguet et Anne Horain, pour leur aide précieuse dans toutes les démarches. J'exprime également ma gratitude à Hélène Jacquet et Marie Campain pour m'avoir fait confiance sur la réalisation d'une étude pour l>IDEX.

Ce travail n'aurait jamais pu voir le jour sans les ingénieurs du GREThA Karine Onfroy et Edouard Kleinpeter. Vous m'avez considérablement facilité la tâche en rendant les bases de données rapidement exploitables. Vous avez été très disponibles et vous avez apporté des réponses compréhensibles sur des questions pourtant très techniques. Un grand merci à vous deux !

Tout au long de mon parcours à l'Université de Bordeaux, j'ai assisté à des cours dispensés par d'excellents professeurs. Je souhaite tous les remercier ici, et plus particulièrement Pascale Roux. J'ai eu la chance de t'avoir d'abord en tant que chargée de travaux dirigés. C'était en 2008, c'était même mon tout premier TD à l'université. Je me rappelle encore du sujet de l'exposé, sur la crise du pétrole (déjà à l'époque !). Je me souviens que dans le cadre du plan Licence et de l'enseignant référent, tu m'avais demandé ce que je voulais faire plus tard. Je t'avais répondu: une thèse ! Et bien nous y voilà... 10 ans plus tard ! Entre temps, j'ai assisté à ton cours de microéconomie de licence 2 et

de licence 3, puis celui d'Économie de l'innovation en master. Et puis j'ai eu la chance de faire partie de ton équipe pédagogique en microéconomie pendant quatre ans !

Je tenais aussi à remercier l'ensemble des responsables des équipes pédagogiques auxquelles j'ai pu appartenir pendant ces cinq années: Hervé Hocquard, Cristina Bada-
rau, Marie Lebreton et Sébastien Rouillon. Cela a été un vrai plaisir de travailler à vos côtés, et il n'y a rien de plus agréable que de donner des travaux dirigés lorsqu'en amont, tout est préparé avec grand professionnalisme.

Mes remerciements vont également aux étudiants pour leur écoute, leur motivation et pour ne pas avoir râlé les quelques fois où j'ai pu amputer sur votre pause pour finir un exercice. Vous avez aussi contribué à améliorer mes enseignements. Mention spéciale à la promotion de Licence 2 Economie 2016-2017 ! Je garderai quelques souvenirs mémorables de ces années. Les copies d'histoire des faits économiques sont toujours un bonheur à corriger ! Et puis il y a eu aussi les cours dans les anciens préfabriqués, aujourd'hui détruits : la salle D et sa piscine au pied du tableau les jours de pluie, ou encore la salle B et une partie du faux-plafond qui s'effondre en plein cours !

Merci aussi à tous les autres doctorants et jeunes docteurs, pour les moments de partage, les pauses de 16h30 et toutes les discussions passionnées que j'ai pu avoir avec vous. Bien sûr les anciens collègues économistes du développement du bureau D323 : Sébastien Michiels, Azyadé Nematollahi-Gillet, Pierre Levasseur, Riana Razafimandimby et Claudia Noumedem Temgoua, ainsi que ceux du bureau E337: Marianne Lanoé, Emmanuel Lorenzon, Jérémy Rastouil, Elodie Carpentier et Adil El Aichouchi. Je les remercie pour leur aide parfois technique mais très précieuse, surtout Elodie et Marianne pour les commandes Stata ! Merci également à Fadoua Chiba, Jeanne Métivier et Samuel Klebaner pour l'organisation impeccable des séminaires doctorants, et aussi pour les afterworks bien sympathiques. Je ne peux pas oublier Nicolas Yol et Thibaud Deguilhem, et les supers souvenirs du week-end à Saint-Lary Soulan (ou plutôt Grailhen, 22 habitants !). Je remercie aussi Suneha Seetahul qui finit et partage les mêmes galères de fin de thèse que moi, Romain Jourdeuilh et Nicolas Bedu. Je salue plus particulièrement Viola Lamani et Dan Tran, que j'ai connus sur les bancs de l'université en première année de licence, et avec qui j'ai fait toute ma scolarité jusqu'en doctorat ! Je souhaite du courage et de la réussite à tous les doctorants en cours de thèse: Lionel Cosnard, Bao Nguyen, Léo Malherbe, Louis Olie, Lucile Marchand, Thomas Eekhout, Valentina di Lasio, Julien Seaux, Martin Paquette, Elodie Rossi, Lucie Piasser, Coralie Kersulec, Léo Delpy, Damien

Girollet et Robin Vos.

Cette thèse n'aurait pas pu voir le jour non plus sans les week-ends de repos salvateur dans les Landes et sans mes amis musiciens. Que ce soit le Sanguinet Jazz Band, le Brass Born Quintet, la Banda Lous Pegailouns, l'Orchestre Symphonique du Bassin d'Arcachon et le Brass Landes, merci à vous tous, et plus particulièrement à Yoann, David, Louis, Théo et Thomas.

Je remercie enfin chaleureusement toute ma famille. J'ai une pensée toute particulière pour mes deux grands-mères qui n'ont pas pu voir l'aboutissement de ce travail.

Contents

1	Tullock contests with an uncertain number of prizes	9
1	Introduction.	9
2	The model.	11
3	Prize allocation mechanisms and equilibrium outcomes	13
3.1	Winner Selection Mechanism (WSM)	13
3.2	Looser Elimination Mechanism (LEM)	15
4	Comparative statics	18
4.1	Distribution of the number of prizes	18
4.2	Number of contestants	25
5	Conclusion	28
6	Appendix	29
6.1	Winner Selection Mechanism (Clark and Riis (1996), subsection 3.1)	29
6.2	Looser Selection Mechanism (Fu et al. (2014), subsection 3.2)	31
6.3	First Order and Second Order Stochastic Dominance	33
6.4	Proofs associated to the Winner Selection Mechanism (WSM)	35
6.5	Proofs associated to the Loser Elimination Mechanism (LEM)	46
2	Team Production Function in Science	53
1	Introduction	53
2	The model	56
3	Complexity of the production function	58
4	Equilibrium outcomes	59
5	Comparative statics	60

5.1	Project i characteristics	60
5.2	Competition intensity	62
6	Extensions	66
7	Conclusion	69
8	Appendix	71
8.1	Proof of Proposition 8	71
8.2	Proof of Proposition 9	75
8.3	Proof of Proposition 10	77
8.4	Proof of Proposition 11	78
8.5	Proof of Proposition 12	79
8.6	Proof of Proposition 13	80
8.7	Proof of Proposition 14	82
8.8	Properties of Υ_i	84
3	Do complex projects lead to higher citation impact?	87
1	Introduction	87
2	From the model to the data	88
3	Literature review	93
4	Data and Methods	95
4.1	Data set	95
4.2	Variables	96
5	Results	111
5.1	1-2 disciplines	111
5.2	1-10 disciplines	115
6	Robustness checks	129
7	Is interdisciplinarity increasing over time?	132
8	Discussion	138

9	Conclusion	141
10	Appendix	143
10.1	Proof of Proposition 16	143
10.2	Proof of Proposition 17	144
10.3	More than 6 references	145
10.4	CWTS nomenclature	161
10.5	WoS subject categories	177
10.6	Tobit estimates	191
10.7	Individual Fixed Effects	204
	List of tables	231
	Liste of figures	241
	Table of contents	243

General Introduction

Are ideas getting harder to find? This question formulated by [Bloom et al. \(2017\)](#) stems from the observation of a substantial diminishing of research productivity coupled with a considerable rise in research efforts. Between 1930 and 2014, they estimate that the effective number of researchers has been multiplied by 23 while the total factor productivity growth rate of the United States has been divided by 2 (Figure 1). This observation seems confirmed in a wide range of cases. One of the most striking facts that emerge from their paper concerns the Moore's Law ([Moore, 1965](#)), which refers to the doubling of the number of transistors in microprocessors every two years, a remarkably stable relationship over time (Figure 2). To achieve the same annual growth rate of the number of transistors ($\approx 35\%$), the number of researchers required in 2014 is 18 times that in the early 1970s. Other applications in agriculture and medical technologies show similar patterns with a decrease in the research productivity at an annual rate around 3.5%.

The decline of research productivity is not a new question in the innovation economics literature. [Bloom et al. \(2017\)](#) are resurrecting an older debate first introduced by [Evenson \(1993\)](#) and later by [Lanjouw and Schankerman \(2004\)](#) who observe that the ratio of the number of patented innovations to the research and development expenditures (or the number of scientists and engineers) has been approximatively divided by two between the 1960's and 1990 in the United States, United Kingdom, Germany and France.

is indisputable, a significant part is also due to increasing transaction costs and other market inefficiencies. Nevertheless, a consensus is emerging among scientists on the fact that research productivity is sharply declining.

This is a major concern since ideas and their adoption are at the core of economic progress. The first economic growth models (Romer, 1990) suppose constant research productivity implying that the growth rate of ideas is proportional to the number of people working in the research sector. It results a scale effect: if the share of workers devoted to research is constant over time, then the total factor productivity and the growth rate of the economy are proportional to the size of the labor force. However there is no empirical evidence of such a relation. Jones (1995) argues that the growth rate of ideas does not only depend on the number of researchers, but also on the stock of knowledge in the economy in two different ways. The first one implicitly assumes that knowledge is infinite: the number of new ideas increases with the stock of knowledge in the economy, signifying that previous discoveries make easier to find new ideas. The second one infers that knowledge is finite: finding new ideas decreases the stock of those not yet discovered, making the new ones harder to find. In both cases the growth rate of ideas decreases with the stock of knowledge available in the economy. These models assume that research productivity is constant over time. However the previous empirical observations seem to support that research productivity is declining, meaning that these models potentially overestimate the growth rate of ideas. Segerstrom (1998) is the first to explore the consequences of R&D becoming more and more difficult over time in a R&D endogenous growth model. He assumes that research productivity is declining given that the most “obvious” ideas are discovered first, making the new ones harder to find. However the nature of ideas (*obvious* or *complex*) remains a fuzzy concept.

Naturally comes the question of why research productivity is declining. In the pharmaceutical sector, Pammolli et al. (2011) observe that the R&D expenditures have substantially increased while the number of new drugs being approved remains stable. According to the authors, R&D expenditures have been concentrated on high risk of failure research areas. The productivity crisis in the pharmaceutical sector may be temporary since firms have invested high amounts of money on risky research areas with the hope to innovate disruptively rather than incrementally. This might momentarily lead to a decline in the research productivity that can be compensated in the future by a greater number of patents, drugs and molecular entities on the market. According to Akcigit et al. (2013),

the underinvestment of governments and private sector in basic research is detrimental since basic research generates greater positive spillovers for the economy than applied research. Given that collaboration in Science is becoming more the rule than the exception (Adams et al., 2005), coordination costs and free-riding could also explain the decline in the research productivity.

Perhaps the more natural reason to explain this phenomenon is that research is becoming more and more *complex*. According to the *Oxford Dictionary of English*, *complex* has several meanings. The first one is “*Not easy to analyse or understand; complicated or intricate*”. As knowledge accumulates, it becomes increasingly difficult to “reach the frontier”. Jones (2009) argues that researchers need to spend more and more time in education to the detriment of time to innovate as the knowledge frontier moves further. There is an increasing educational burden so that age of first invention increases over time, individual knowledge narrows leading to a greater specialization, and team size grows. Consequently, the labor consumption increases and thus the associated costs, leading to a decline in research productivity.

The second meaning of *complexity* is “*Consisting of many different and connected parts*.” Doing research is intrinsically a productive problem: it requires to combine previous academic works in a particular way to produce new ideas. Here increasing the *complexity* of research projects means that the *knowledge production function* has changed over time, in a way that leads to an increase of the labor consumption and a decline of research productivity. Such explanation does not exist in the literature yet, and is the main subject of this thesis. To do so, we propose in Chapter 2 a theoretical model to understand how variations in the knowledge production function that we suppose of the constant elasticity of substitution form affects scientific output and productivity. Among many results, we find that, at equilibrium, all the variables depend on a synthetic index of the knowledge production function that we define to be its *complexity*. We show how changes in this index modify the knowledge production and productivity at equilibrium.

To empirically test our model (Chapter 3), we consider one particular form of complexity, called *interdisciplinarity*¹, which requires combining skills and efforts from different

¹One important challenge in the Philosophy of Science has been to conceptualise different forms of combining disciplines. These academic works led to different expressions such as multidisciplinary, interdisciplinarity or transdisciplinarity, which have been employed indifferently for a long time although a clear classification exists since the paper of Choi and Pak (2006).

research areas or academic disciplines. More precisely, the National Academy of Science defines interdisciplinarity as follows:

Interdisciplinary research is a mode of research by teams or individuals that integrates information, data, techniques, tools, perspectives, concepts, and/or theories from two or more disciplines or bodies of specialized knowledge to advance fundamental understanding or to solve problems whose solutions are beyond the scope of a single discipline or area of research practice.

If interdisciplinarity is a form of complexity, this is not the only form of complexity. Our choice to study interdisciplinarity is motivated by the fact that this topic has received a growing interest from the scientific community in recent years. Various stakeholders see interdisciplinary research as a mean to increase the social returns to public funding of academic research. This question is all the more important given that many studies point out the barriers to interdisciplinary research raised by the intrinsic organization of scientific communities. Journal editors and referees encounter difficulties to evaluate such work since they are often specialised in a single discipline. While standards of evaluation often exist for monodisciplinary papers, they are extremely rare or even non-existent for multidisciplinary research ([Pautasso and Pautasso, 2010](#)). Reviewers often disagree concerning the evaluation of the quality of those papers ([Lamont et al., 2006](#)) which may lead to higher rejection rates or at least complicates the process of publication. Multidisciplinary researchers seem to be penalised for academic promotion and tenures ([Rhoten and Parker, 2004](#)). Even though universities increasingly try to value interdisciplinary research, their organisational structure is often based on monodisciplinary departements, which limits interactions between researchers from different fields of science ([Sá, 2008](#)).

A remarkable fact is that this form of complexity is increasing over time. To the best of our knowledge, such an empirical fact has not been established yet. We deal with two databases to support this assertion. The first one contains more than 400,000 research articles in academic journals published by at least a French researcher between 1999 and 2013. The second one is the full Web of Science between 2013 and 2017 which contains around 8,000,000 publications over the period. In both cases, we observe that the average complexity of academic publications increases over time at a constant rate, between 1 and 3.5% per year (Figures 3 and 4), whatever the index considered. Although many factors affect research productivity, its decline rate calculated by [Bloom et al. \(2017\)](#)

approximatively corresponds to our complexity growth rate. This observation goes in the same direction than our theoretical conclusions on the relation between the complexity of the knowledge production function and the research productivity. Nevertheless comes the question of why researchers engage in more and more *complex* problems, if it depresses their productivity. We open the debate in the last part of this document.

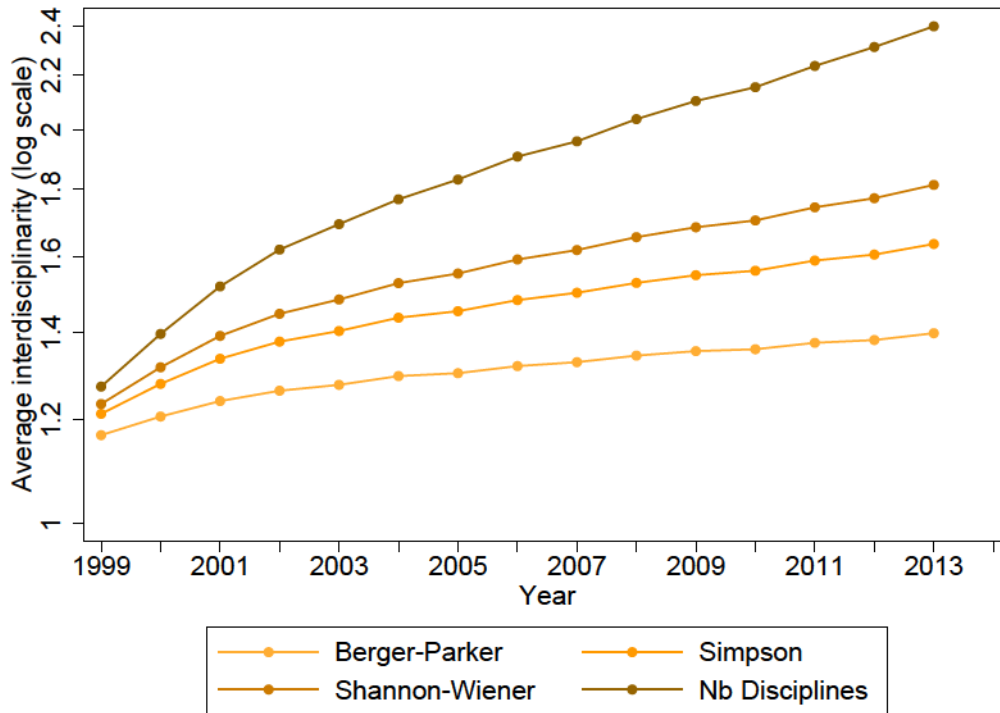


Figure 3: Evolution of the interdisciplinarity of research articles published by at least a French researcher over the period 1999-2013, Web of Science database (400,000 observations), OST disciplinary nomenclature.

In addition to the growing complexity, Science has evolved towards an incredibly increase of the competition over the last decades. The scientific labor force has substantially grown. But above all, the number of selective procedures has considerably increased to fund projects (creation of research funding agencies), to promote researchers or to grant scientific awards for instance. Decision-makers aim to increase the efficiency of the allocation of resources. However researchers often have to face uncertainty concerning the number of competitors but also the prize structure, which can change their incentives to exert effort. In this thesis, we theoretically address these questions. In Chapter 1, we are interested in how competition and uncertainty concerning the number of prizes change players incentives to exert effort. In Chapter 2, we aim to focus on the effect of competition on efforts and output, both in the qualitative (greater efficiency of competitors) and quantitative (number of projects) dimensions. We obtain interesting results that we

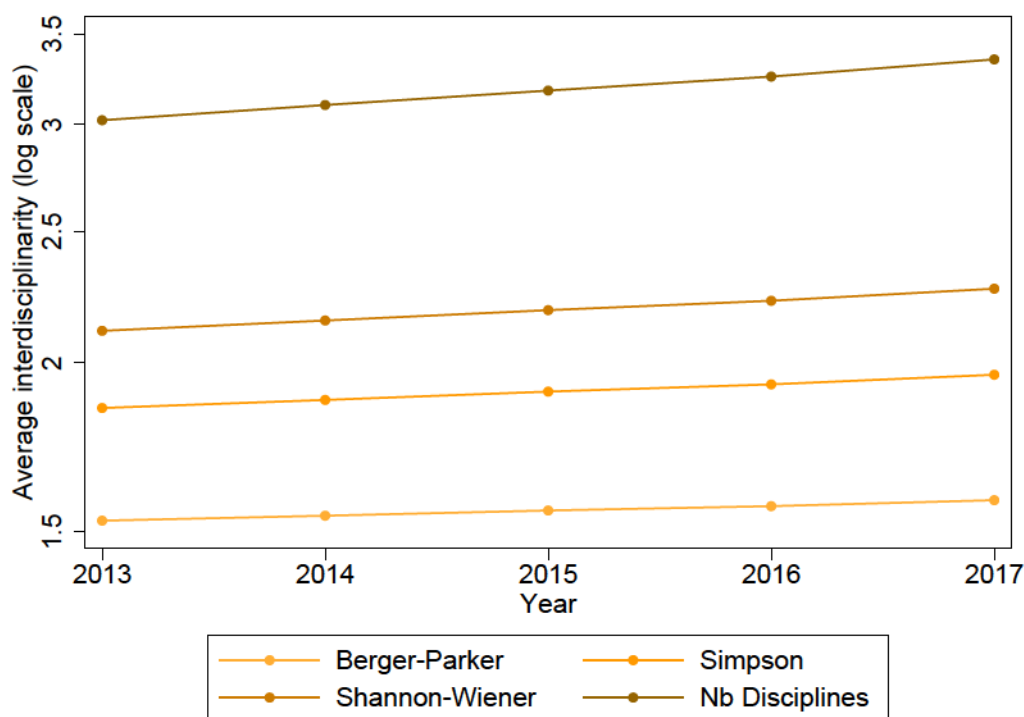


Figure 4: Evolution of the interdisciplinarity of research articles over the period 2013-2017, Web of Science database (8,000,000 observations), OST disciplinary nomenclature.

develop all along these two chapters.

This thesis contains three chapters.

We propose in **Chapter 1** an original contribution on the Tullock contest theory. We consider multiple-prize contests where the number of prizes to be awarded is uncertain. We determine the symmetric Nash equilibrium of the contest game. Under a specification of the model commonly used in the literature, we show that the aggregate contest expenditure decreases with the expectation on the number of prizes (first-order stochastic dominance), with the risk in the number of prizes (second-order stochastic dominance), and increases with the number of contestants regardless of the prize allocation mechanism. We give sufficient conditions such that the same holds under a general specification. Accordingly, complete information concerning the number of prizes maximizes the aggregate contest expenditure.

In **Chapter 2**, we propose a theoretical model to deal with the effect of the complexity of the production function on the output and productivity at equilibrium. We model competition in Science as a Tullock contest between research projects for the obtention of a prize (a publication slot in a journal for instance). We conceptualize a research

project as an idea and a team of researchers. Each idea is associated to a given knowledge production function whose factors are subteam efforts, each one in a distinct field of expertise. We abstract from free-riding issues and intra-team coordination to focus on the characteristics of the knowledge production function, that we suppose of the CES form. We characterize the Nash equilibrium of the corresponding contest game. We find that all the variables are function of a synthetic index of the knowledge production function, that we call *complexity* of the project given its properties. We derive comparative statics with respect to this indicator and analyze the consequences of the number of projects in the contest. Among the results, we show that teams treating more complex problems are penalized since their participation, their impact, their productivity and their probability to obtain the prize decrease. We find that competition exerts a differentiated effect on project impact depending on their own *complexity*.

In **Chapter 3**, we analyze the effect of one form of *complexity* on the impact of research projects, called interdisciplinarity. We first rewrite our index of complexity as a function of the contributions of each field of expertise to the outcome and show that it belongs to the family of Hill diversity indexes, a standard interdisciplinary measure. We confirm our theoretical predictions on an original dataset of nearly four hundred thousand authorship participations of approximately thirty thousand French professors and researchers over the period 1999-2013 for which we have individual data. We show that interdisciplinarity is increasing over time before discussing how this may affect the research productivity.

We finally conclude and give perspectives for further research dealing with the relation between complexity and research productivity.

Tullock contests with an uncertain number of prizes¹

1 Introduction.

[Sisak \(2009\)](#) surveyed the literature on multiple-prize contests². She shows that this framework can be relevant in many situations, taking examples from rent-seeking activities, patents and R&D races, licenses, labor markets, sports and so on. [Sisak \(2009\)](#) classifies the literature along two main dimensions, based on the specification of the contest success function (Tullock versus fully discriminating contest success function) and on the adoption of single versus multiple efforts (the contestants exert an overall effort for all prizes or can allocate it more specifically to a sub-group of prizes). The central finding is that with risk-neutral and symmetric contestants, a contest designer aiming at maximizing the aggregate effort should always prefer to allocate a single prize rather than splitting it into several smaller prizes. However, dividing the prize can be optimal in situations with asymmetric players for instance ([Szymanski and Valletti, 2005](#)).

More recently, the literature on multiple-prize contests has been extended in several ways. [Fu et al. \(2014\)](#) propose a prize allocation mechanism which selects-out losers while the literature had so far considered a winner selection mechanism ([Clark and Riis, 1996](#)). [Minchuk and Sela \(2014\)](#) study an all-pay auction with multiple prizes, where the players have a common value for all the certain prizes and a private valuation for the uncertain one. [Stracke et al. \(2014\)](#) analyze the effect of the prize structure on effort in a dynamic two-stage elimination contest. If a single prize maximizes the aggregate effort when players are risk-neutral, a two prizes structure may dominate a winner-takes-all contest if players are risk-averse. In a working paper, [Balafoutas et al. \(2017\)](#) propose

¹This paper is co-authored with Sébastien Rouillon.

²The very first contributions on multiple-prize contests are [Glazer and Hassin \(1988\)](#) and [Berry \(1993\)](#).

to test experimentally an all-pay auction with multiple endogenous and uncertain prizes. Keeping constant the expected number of prizes, they observe that more uncertainty leads to a decrease in the effort spent by players.

Surprisingly, the case of multiple-prize contests with an uncertain number of prizes has never been investigated theoretically, although this is a natural assumption and an immediate extension of the literature just surveyed. Many real-life examples correspond to this situation. In firms, workers may compete without knowing exactly the number of promotions. In Science, researchers compete to obtain funds without knowing precisely how funding agencies allocate their budget between projects. To publish their paper, researchers have limited informations concerning the acceptance rate of journals. In the academic job market, PhD students are not aware of the exact number of jobs of assistant professors or researchers that will be available after their PhD defense. The purpose of the present paper is to provide a first attempt to fill the gap concerning uncertainty on the number of prizes in contests.

We consider a Tullock contest success function with risk-neutral and symmetric players, assuming that the number of (identical) prizes to be awarded is a random variable, with a common knowledge probability distribution. We characterize the Nash equilibrium of the corresponding contest game and identify some of its properties. Under a specification commonly used in the literature on multiple-prize contests (Berry, 1993; Clark and Riis, 1996), we find that the aggregate effort of the contestants increases when the uncertain number of prizes gets smaller in the sense of the first-order stochastic dominance or when the risk in the number of prizes decreases in the sense of the second-order stochastic dominance. We discuss how the valuation function and the specification of the probability function can modify the properties previously determined. We also provide a numerical illustration to display less expected behaviors.

This paper also contributes to the literature from a technical point of view. Our results follow from the properties (monotonicity and concavity) of a *discrete* function first appeared in Clark and Riis (1996) then in Fu et al. (2014). We show here how to extend it to a *continuous and twice differentiable* function, by using of the *psi* (or *digamma*) function (Abramowitz and Stegun, 1964). This greatly eases the analysis of its properties, which are obtained from first-order and second-order derivatives. We also enhance the conditions of existence of a symmetric equilibrium of the contest games given

in the papers of [Clark and Riis \(1996\)](#) and [Fu et al. \(2014\)](#).

Finally, this paper is related to [Münster \(2006\)](#), [Lim and Matros \(2009\)](#), [Myerson and Wärneryd \(2006\)](#), and [Kahana and Klunover \(2015\)](#), who extend the contest literature to situations where the number of contestants is uncertain. They show that the (*ex-ante*) aggregate effort in a contest with population uncertainty is smaller than its counterpart in a contest with population certainty and the same expected number of contestants. Under some conditions, our paper gives the analog finding for contests with prize uncertainty.

The rest of the paper is organized as follows. Section 2 sets out the model. Section 3 presents the prize allocation mechanisms and characterizes the Nash equilibrium of the contest games. Section 4 deals with the properties of the equilibrium outcome. Section 5 concludes. Several proofs are given in the Appendix.

2 The model.

We consider n (risk neutral) players competing in a nested contest awarding k prizes, with $1 \leq k < n$. All prizes have the same value, given by the valuation function $V(k)$ ³. The contestants simultaneously exert an overall effort for all prizes in order to get a chance to win one prize and no more. The individual effort of a player i is denoted by x_i and the aggregated effort is denoted by $X = \sum_{j=1}^n x_j$. Let $f(x_i)$ be the impact function of player i ([Myerson and Wärneryd, 2006](#)), and assume that this is a strictly increasing and concave function, with $f(0) = 0$ ⁴. To ease the notation, it will be useful below to define $g(x_i) \equiv f(x_i)/f'(x_i)$ ⁵.

In the literature on multiple-prize contests ([Berry, 1993](#); [Clark and Riis, 1996](#)), it is standard to use the specification where $f(x) = x^r$, with $0 < r \leq 1$, and $V(k) = V/k$. To ease the comparison, below we will sometimes refer to it as a benchmark case. However, it is worth noting that most of the analysis and results will be derived within the general model.

Let $P_i(x_1, \dots, x_n; k)$ denote the (*ex ante*) probability that player i wins one prize in the certain case. The originality of our paper is that we assume that the players are unaware

³For technical reasons, it will be convenient to assume below that $V(k)$ is defined and twice differentiable for all $k \in [1, n]$.

⁴These assumptions are commonly used in the literature.

⁵Remark that $g(x_i)$ is increasing. Indeed, $f(x_i) > 0$ and $f''(x_i) \leq 0$ imply that $g'(x_i) = 1 - f(x_i)f''(x_i)/(f'(x_i))^2 > 0$.

of the exact number of prizes to be awarded and only know that this number is distributed between 1 and K , according to a probability distribution $\pi(k)$. In this setting with an uncertain number of prizes, each player i expects to win one prize with probability

$$\mathbb{E}[P_i(x_1, \dots, x_n; k)] = \sum_{k=1}^K \pi(k) P_i(x_1, \dots, x_n; k).$$

We consider first the case where the players observe the number of prizes k to be allocated before they choose their level of effort. We then deal with the case where the players only know that the number of prizes k is distributed between 1 and K , according to the probability distribution $\pi(k)$.

When the number of prizes is known with certainty, each player i observes k and chooses x_i to maximize

$$P_i(x_1, \dots, x_n; k) V(k) - x_i. \quad (1.1)$$

For an interior solution, the following first and second-order conditions must be satisfied

$$\frac{\partial}{\partial x_i} P_i(x_1, \dots, x_n; k) V(k) - 1 = 0, \text{ for all } i,$$

$$\frac{\partial^2}{\partial x_i^2} P_i(x_1, \dots, x_n; k) V(k) < 0, \text{ for all } i,$$

as well as the participation constraint

$$P_i(x_1, \dots, x_n; k) V(k) - x_i \geq 0.$$

Consider now the case where the players choose their effort under uncertainty. Knowing that k is distributed between 1 and K , according to a probability distribution $\pi(k)$, each player maximizes his expected utility with respect to his own effort

$$\sum_{k=1}^K \pi(k) P_i(x_1, \dots, x_n; k) V(k) - x_i.$$

For an interior solution, the following first and second-order conditions must be fulfilled

$$\sum_{k=1}^K \pi(k) \frac{\partial}{\partial x_i} P_i(x_1, \dots, x_n; k) V(k) - 1 = 0, \text{ for all } i,$$

$$\sum_{k=1}^K \pi(k) \frac{\partial^2}{\partial x_i^2} P_i(x_1, \dots, x_n; k) V(k) < 0, \text{ for all } i,$$

as well as the participation constraint

$$\sum_{k=1}^K \pi(k) P_i(x_1, \dots, x_n; k) V(k) - x_i \geq 0.$$

3 Prize allocation mechanisms and equilibrium outcomes

We propose two specifications for the probability function $P_i(x_1, \dots, x_n; k)$ here, each of them associated to a different prize allocation process. We deal with the [Clark and Riis \(1996\)](#) winner selection mechanism (WSM) and the [Fu et al. \(2014\)](#) loser elimination mechanism (LEM). We also characterize the Nash equilibrium of the corresponding contest games.

3.1 Winner Selection Mechanism (WSM)

The prizes are awarded in k rounds, according to the iterative process initially described in [Clark and Riis \(1996\)](#)⁶. Let $N(\kappa)$ denote the set of players still remaining in the contest at round κ ⁷. The conditional probability⁸ that any player i in $N(\kappa)$ wins the prize in the κ -th round is equal to

$$p_i^\kappa(x_1, \dots, x_n) = \begin{cases} \frac{f(x_i)}{\sum_{j \in N(\kappa)} f(x_j)} & \text{if } \sum_{j \in N(\kappa)} f(x_j) > 0 \\ \frac{1}{n+1-\kappa} & \text{if } x_1 = \dots = x_n \end{cases}.$$

If player i is drawn at random during round κ , the set of remaining players then evolves according to

$$N(\kappa + 1) = N(\kappa) - \{i\}.$$

⁶More details are given in Appendix 6.1.

⁷Clearly, $N(1) = \{1, \dots, n\}$.

⁸The contest success function has been axiomatized by [Skaperdas \(1996\)](#).

The process is repeated until all prizes are allocated. Assuming that $x_j = x$, for all $j \neq i$, Clark and Riis (1996) show that

$$P_i(x_i, x, \dots, x; k) = \frac{f(x_i)}{f(x_i) + (n-1)f(x)} + \sum_{\kappa=2}^k \left[\prod_{\lambda=1}^{\kappa-1} \left(1 - \frac{f(x_i)}{f(x_i) + (n-\lambda)f(x)} \right) \right] \frac{f(x_i)}{f(x_i) + (n-\kappa)f(x)}.$$

Considering the symmetric equilibrium (letting $x_i = x$ for all i), we can calculate that⁹

$$\frac{\partial}{\partial x_i} P_i(x, \dots, x; k) = \frac{n-k}{n} \left(\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^{n-k} \frac{1}{r} \right) \frac{f'(x)}{f(x)}. \quad (1.2)$$

Proposition 1 gives the conditions of existence and the expression of the Nash equilibrium of the game when the number of prizes is known with certainty.

Proposition 1 *There exists a symmetric pure strategy equilibrium $x_i^* = x^*$ for all i if the following sufficient condition is satisfied*

$$\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^{n-k} \frac{1}{r} \leq 1, \quad (1.3)$$

where all the players exert an effort satisfying

$$g(x^*) = A_{WSM}(k, n), \quad (1.4)$$

where we denote¹⁰

$$A_{WSM}(k, n) = V(k) \frac{n-k}{n} \left(\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^{n-k} \frac{1}{r} \right). \quad (1.5)$$

Proof is given in Appendix 6.4.3. We draw attention to the fact that our proof improves the precedent ones concerning the conditions of existence of the Nash equilibrium. Clark and Riis (1998) give a sufficient condition on the number of prizes k and the parameter r of the impact function $f(x) = x^r$ to obtain a local maximum, so that the utility function

⁹For reasons that will become clear below, the expression of the first-order condition used here, though equivalent, differs from that in (Clark and Riis, 1996). The proof is given in Appendix 6.4.1.

¹⁰Clark and Riis (1996) write $A_{WSM}(k, n) = V(k) \left(k - \sum_{j=0}^{k-1} (k-j)/(n-j) \right) / n$, which is equivalent.

is concave around $x_i = x^*$. Here we give the proof for any impact function $f(x)$ and show that player i 's utility function is concave with x_i over the interval $[0; x]$.

When the players choose their effort under uncertainty, we establish in Proposition 2 the symmetric Nash equilibrium of the game and its conditions of existence.

Proposition 2 *There exists a symmetric pure strategy equilibrium $x_i^* = x^*$ for all i if the following sufficient condition is fulfilled*

$$\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^{n-K} \frac{1}{r} \leq 1,$$

where all players exert an effort satisfying

$$g(x^*) = \sum_{k=1}^K \pi(k) A_{WSM}(k, n), \quad (1.6)$$

with

$$A_{WSM}(k, n) = V(k) \frac{n-k}{n} \left(\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^{n-k} \frac{1}{r} \right).$$

Note also that effort at equilibrium in the uncertain case is a linear combination of equilibrium efforts in the certain case weighted by the probabilities.

3.2 Looser Elimination Mechanism (LEM)

Instead of selecting-in the winners in each nested round like in the model of Clark and Riis (1996), Fu et al. (2014) propose a mechanism which selects-out losers¹¹. In each round κ , one player and no more is eliminated and doesn't receive any prize¹². The elimination process continues until k players survive, or equivalently $n - k$ players are removed from the initial set of contestants. At the end of round $\kappa = n - k$, each survivor receives a prize of value $V(k)$.

Let $S(\kappa)$ denote the set of survivors remaining in the contest at round κ ¹³. The conditional probability¹⁴ q_i^κ that a player i in $S(\kappa)$ is eliminated in the κ -th round is

¹¹More details are given in Appendix 6.2.

¹²Fu et al. (2014) allows eliminated players to receive a prize of positive value. Here we make the assumption that eliminated players get no prize to have a perfect match with the Clark and Riis (1996) prize structure.

¹³Clearly $S(1) = 1, \dots, n$

¹⁴The contest elimination function has been axiomatized by Lu and Wang (2016).

equal to

$$q_i^\kappa(x_1, \dots, x_n) = \begin{cases} \frac{\frac{1}{f(x_i)}}{\sum_{j \in S(\kappa)} \frac{1}{f(x_j)}} & \text{if } \sum_{j \in S(\kappa)} f(x_j) > 0 \\ \frac{1}{n+1-\kappa} & \text{if } x_1 = \dots = x_n \end{cases}$$

If player i is eliminated at round κ , then the set of survivors in next period evolves according to

$$S(\kappa + 1) = S(\kappa) - \{i\}$$

Since only one player is eliminated in each round, the process ends at round $\kappa = n - k$. Assuming $x_i = x$ for all $j \neq i$, the probability that player i wins one prize with the loser elimination mechanism simplifies to

$$P_i(x_i, x, \dots, x, k) = \prod_{\kappa=1}^{n-k} \frac{(n - \kappa) f(x_i)}{(n - \kappa) f(x_i) + f(x)}.$$

We obtain that¹⁵

$$\frac{\partial}{\partial x_i} P_i(x, \dots, x, k) = \frac{k}{n} \left(\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^k \frac{1}{r} \right) \frac{f'(x)}{f(x)}. \quad (1.7)$$

Proposition 3 gives the conditions of existence and the expression of the Nash equilibrium of the game when the number of prizes is known with certainty.

Proposition 3 *There exists a symmetric pure strategy equilibrium $x_i^* = x^*$ if the following sufficient condition is satisfied*

$$\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^k \frac{1}{r} \leq 1, \quad (1.8)$$

where all the players exert an effort satisfying

$$g(x^*) = A_{LEM}(k, n), \quad (1.9)$$

with

$$A_{LEM}(k, n) = V(k) \frac{k}{n} \left(\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^k \frac{1}{r} \right). \quad (1.10)$$

¹⁵The proof is given in Appendix 6.5.1.

Proof is given in Appendix 6.5.2. Since Fu et al. (2014) have used the same specification as Clark and Riis (1996) ($f(x) = x^r$), our proof enhances the Fu et al. (2014) one. Indeed we consider a more general impact function $f(x)$ and we show that player i 's utility function is concave with x_i over the interval $[x; +\infty[$, not just around $x_i = x$. Here the sufficient condition in (1.8) involves that we are able to prove the existence of a symmetric pure strategy equilibrium when $f(x) = x$ as long as the total effort spent by players doesn't exceed the total value distributed in the contest.

Now we turn to the case where the players choose their effort under uncertainty. Proposition 4 gives the conditions of existence and the expression of the symmetric Nash equilibrium of the game.

Proposition 4 *There exists a symmetric pure strategy equilibrium $x_i^* = x^*$ if the following sufficient condition is satisfied*

$$\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^{\underline{k}} \frac{1}{r} \leq 1, \quad (1.11)$$

with $\underline{k} \in [1; K]$ the minimal integer such that (1.11) is true, and

$$\pi(k) = 0 \quad \forall k < \underline{k},$$

where all players exert an effort satisfying

$$g(x^*) = \sum_{k=1}^K \pi(k) A_{LEM}(k, n), \quad (1.12)$$

with

$$A_{LEM}(k, n) = V(k) \frac{k}{n} \left(\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^k \frac{1}{r} \right).$$

Proof is given in Appendix 6.5.2. The sufficient condition in (1.11) means that when $f(x) = x$, we are able to prove the existence of a symmetric pure strategy equilibrium as long as the aggregate effort exerted by players is lower than the total expected value distributed in the contest. Again, effort at equilibrium in the uncertain case is a linear combination of equilibrium efforts in the certain case weighted by the probabilities.

4 Comparative statics

Here we derive some comparative statics of the equilibrium outcome under uncertainty. We show how the probability distribution of the number of prizes and the number of participants affect the aggregate contest expenditure. For convenience, we sometimes drop the subscript relative to the prize allocation mechanism (WSM or LEM). Thus, this means that our assertions are valid regardless of the prize allocation mechanism.

4.1 Distribution of the number of prizes

We consider here the effect of the probability distribution of the number of prizes. Formally, we compare the equilibrium outcome under two distributions, denoted by $\underline{\pi}(k)$ and $\bar{\pi}(k)$. We let \underline{x} be the equilibrium effort in the contest with $\underline{\pi}(k)$, and \bar{x} the equilibrium effort in the contest with $\bar{\pi}(k)$. We search for conditions on the valuation function $V(k)$ such that $\underline{x} > \bar{x}$, considering in turn first-order and second-order stochastic dominance¹⁶.

From our previous analysis, we know that the equilibrium efforts respectively satisfy

$$g(\underline{x}) = \sum_{k=1}^K \underline{\pi}(k) A(k, n) \quad \text{and} \quad g(\bar{x}) = \sum_{k=1}^K \bar{\pi}(k) A(k, n).$$

As $g(x)$ is increasing, the condition for $\underline{x} > \bar{x}$ writes

$$\sum_{k=1}^K \underline{\pi}(k) A(k, n) > \sum_{k=1}^K \bar{\pi}(k) A(k, n), \quad (1.13)$$

where we recall that

$$A_{WSM}(k, n) = V(k) \frac{n-k}{n} \left(\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^{n-k} \frac{1}{r} \right)$$

and

$$A_{LEM}(k, n) = V(k) \frac{k}{n} \left(\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^k \frac{1}{r} \right).$$

Below, we investigate properties of the valuation function $V(k)$ such that inequality (1.13) will hold true, dealing in turn with first-order and second-order stochastic domi-

¹⁶See Appendix 6.3 for definitions of these concepts.

nance. Assuming that $\bar{\pi}(k)$ first-order stochastically dominates $\underline{\pi}(k)$, it will be sufficient to find shapes of the valuation function $V(k)$ generating a function $A(k, n)$ decreasing in k (Courtault et al., 2006). Assuming that $\underline{\pi}(k)$ second-order stochastically dominates $\bar{\pi}(k)$, it will be sufficient to identify shapes of the valuation function $V(k)$ inducing a function $A(k, n)$ concave in k (Rothschild and Stiglitz, 1971).

Remarking that $A(k, n)$ is a discrete function of k ¹⁷, our strategy for investigating its properties is as follows. Using the *psi* (or *digamma*) function (Abramowitz and Stegun, 1964), for all *real* numbers $k \in [1, n]$, we construct a continuous and twice differentiable function $F(k)$, coinciding with $A(k, n)$ for every integer k ¹⁸. We then calculate the first-order and second-order derivatives of $F(k)$. We finally use the derivatives to give sufficient conditions on $V(k)$ ensuring the monotonicity and the concavity of $A(k, n)$ with respect to k .

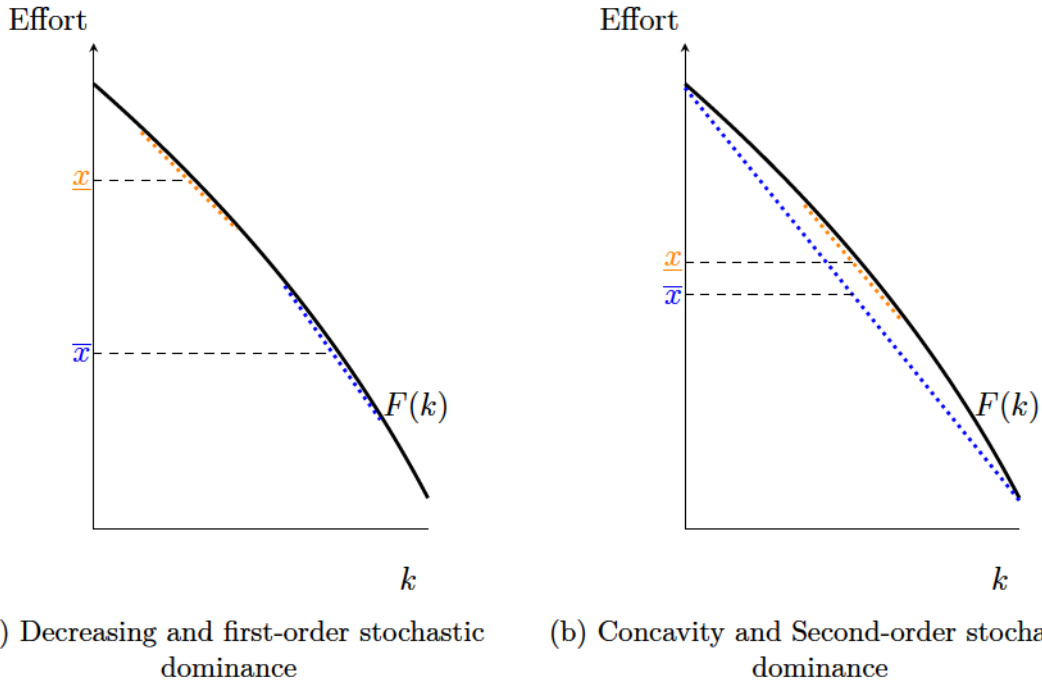


Figure 1.1: Influence of the probability distribution on the number of prizes (first-order and second-order stochastic dominance) on effort

Example 1 To illustrate our methodology, we provide an example in Figure 1.1. Panel (a) considers two contests where the number of prizes is uncertain and distributed according to two probability distributions: $\underline{\pi}(3) = \underline{\pi}(7) = 0.5$ and $\bar{\pi}(9) = \bar{\pi}(13) = 0.5$. Obviously, $\bar{\pi}(k)$ first-order stochastically dominates $\underline{\pi}(k)$. We represent the function $F(k)$ which coincides

¹⁷This is due to the presence of the term $\sum_{r=1}^{n-k} 1/r$ or $\sum_{r=1}^k 1/r$.

¹⁸Formally, $F(k) = A(k, n)$ when $k = 1, \dots, n$.

with $A(k, n)$ (the equilibrium effort when the number of prizes is certain) for every integer k . Since equilibrium effort in the uncertain case is a linear combination of equilibrium efforts in the certain case weighted by the probabilities, we can easily determine on the graph the efforts \underline{x} and \bar{x} (respectively in orange and blue). Here the decrease of $F(k)$ leads to $\underline{x} > \bar{x}$. Panel (b) considers two contests with the same expected number of prizes but with a different variance. The probability distributions are given by $\underline{\pi}(6) = \underline{\pi}(10) = 0.5$ and $\bar{\pi}(1) = \bar{\pi}(15) = 0.5$, meaning that $\underline{\pi}(k)$ second-order stochastically dominates $\bar{\pi}(k)$. Given that $F(k)$ is concave with k , we observe $\underline{x} > \bar{x}$.

For all $k \in [1, n]$, let us define

$$F_{WSM}(k) = V(k) \frac{n-k}{n} (\psi(n+1) - \psi(n-k+1)), \quad (1.14)$$

and

$$F_{LEM}(k) = V(k) \frac{k}{n} (\psi(n+1) - \psi(k+1)), \quad (1.15)$$

where ψ is the *psi* (or *digamma*) function (Abramowitz and Stegun, 1964), defined for any positive real χ by

$$\psi(\chi) = -\gamma + \sum_{j=0}^{\infty} \frac{\chi-1}{(j+1)(j+\chi)}, \quad (1.16)$$

where γ is the *Euler constant*. Knowing that

$$\psi(\eta+1) = -\gamma + \sum_{j=1}^{\eta} \frac{1}{j},$$

where η denotes any positive integer (Abramowitz and Stegun, 1964), we can verify that

$$F_{WSM}(k) = V(k) \frac{n-k}{n} \left(\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^{n-k} \frac{1}{r} \right) = A_{WSM}(k, n)$$

and

$$F_{LEM}(k) = V(k) \frac{k}{n} \left(\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^k \frac{1}{r} \right) = A_{LEM}(k, n)$$

for all integer values $k = 1, \dots, n$. Moreover, knowing that the first-order and second-order derivatives of ψ are (Abramowitz and Stegun, 1964)

$$\psi'(k) = \sum_{j=0}^{\infty} \frac{1}{(j+k)^2} \quad \text{and} \quad \psi''(k) = -\sum_{j=0}^{\infty} \frac{2}{(j+k)^3}, \quad (1.17)$$

we know that $F(k)$ is twice differentiable.

We can calculate that

$$F'_{WSM}(k) = -\frac{1}{n} \left(\begin{aligned} &-(V'(k)k + V(k)) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+n-k+1)} \\ &+ V(k)k \sum_{j=0}^{\infty} \frac{j+1}{(j+n+1)(j+n-k+1)^2} \end{aligned} \right) \quad (1.18)$$

$$F'_{LEM}(k) = -\frac{1}{n} \left(\begin{aligned} &-(V'(k)k + V(k)) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \\ &+ V(k)k \sum_{j=0}^{\infty} \frac{1}{(j+k+1)^2} \end{aligned} \right) \quad (1.19)$$

and

$$F''_{WSM}(k) = -\frac{2}{n} \left(\begin{aligned} &-\frac{1}{2} (V''(k)k + 2V'(k)) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+n-k+1)} \\ &+ (V'(k)k + V(k)) \sum_{j=0}^{\infty} \frac{j+1}{(j+n+1)(j+n-k+1)^2} \\ &+ V(k)k \sum_{j=0}^{\infty} \frac{j+1}{(j+n+1)(j+n-k+1)^3} \end{aligned} \right) \quad (1.20)$$

$$F''_{LEM}(k) = -\frac{2}{n} \left(\begin{aligned} &-\frac{1}{2} (V''(k)k + 2V'(k)) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \\ &+ (V'(k)k + V(k)) \sum_{j=0}^{\infty} \frac{1}{(j+k+1)^2} \\ &+ V(k)k \left(\sum_{j=0}^{\infty} \frac{1}{(j+k+1)^3} \right) \end{aligned} \right). \quad (1.21)$$

Before we derive general results, it is worth considering the benchmark specification, where $V(k) = V/k$, which is the one used in the literature (Berry, 1993; Clark and Riis, 1996). For the first-order derivatives, we obtain clearcut results, as the two conditions respectively simplify to

$$F'_{WSM}(k) = -\frac{V}{n} \sum_{j=0}^{\infty} \frac{j+1}{(j+n+1)(j+n-k+1)^2} < 0$$

and

$$F'_{LEM}(k) = -\frac{V}{n} \sum_{j=0}^{\infty} \frac{1}{(j+k+1)^2} < 0$$

Proposition 5 *With the valuation function $V(k) = V/k$, equilibrium effort \underline{x} is larger than \bar{x} when $\bar{\pi}(k)$ first-order stochastically dominates $\underline{\pi}(k)$.*

Proposition 5 parallels the results of (Clark and Riis, 1996; Berry, 1993), showing that when the number of prizes is certain, the contestants expend more in a contest with a single prize than in a contest with several prizes. Roughly speaking, our result means that

when the number of prizes is uncertain, the contestants exert more effort in the contest that awards the smallest expected number of prizes (in the sense of first-order stochastic dominance). With a valuation function $V(k) = V/k$, each contestant has more chances to win a prize but of lower value. As a consequence, this creates less incentives for players to exert effort in the contest.

Looking at the second-order derivatives, we also obtain clearcut results with the benchmark specification $V(k) = V/k$:

$$F''_{WSM}(k) = -2\frac{V}{n} \sum_{j=0}^{\infty} \frac{j+1}{(j+n+1)(j+n-k+1)^3} < 0.$$

and

$$F''_{LEM}(k) = -2\frac{V}{n} \sum_{j=0}^{+\infty} \frac{1}{(j+k+1)^3} < 0.$$

Proposition 6 *With the valuation function $V(k) = V/k$, equilibrium effort \underline{x} will be larger than \bar{x} when $\underline{\pi}(k)$ second-order stochastically dominates $\bar{\pi}(k)$.*

This result is new in the literature. This means that when the number of prizes is uncertain, the contestants expend more in the contest with less risk in the number of prizes (in the sense of second-order stochastic dominance). Note first that Proposition 6 doesn't depend on the prize allocation mechanism. Remark that the first-order and second-order derivatives of $F(k)$ (expressions (1.18) to (1.21)) are functions of $V(k)$. With a general specification for the valuation function $V(k)$, we can extend our results in several directions.

Using condition (1.18) and (1.19), we see that $F'(k) < 0$ whenever the *aggregate* value at stake (i.e., $V(k)k$) is non-increasing in the number of prizes to be allocated (i.e., $V'(k)k + V(k) \leq 0$). Then, assuming that $\bar{\pi}(k)$ first-order stochastically dominates $\underline{\pi}(k)$, condition (1.13) is true and $\underline{x} > \bar{x}$. In other words, the contestants exert more effort in the contest with $\underline{\pi}(k)$ than in the contest with $\bar{\pi}(k)$.

Likewise, using (1.20) and (1.21), we see that $F''(k) < 0$ whenever the *aggregate* value at stake (i.e., $V(k)k$) is non-decreasing and weakly concave in the number of prizes to be allocated (i.e., $V'(k)k + V(k) \geq 0$ and $V''(k)k + 2V'(k) \leq 0$). Then, assuming that $\underline{\pi}(k)$ second-order stochastically dominates $\bar{\pi}(k)$, condition (1.13) is true and $\underline{x} > \bar{x}$, meaning that the contestants expend more in the contest with $\underline{\pi}(k)$ than in the contest with $\bar{\pi}(k)$.

Finally, it should also be noted that the set of functions $V(k)$ such that $F'(k) < 0$ and $F''(k) < 0$, is stable to positive combinations of its elements. Indeed, it should be clear that if $V_1(k)$ and $V_2(k)$ are two valuation functions satisfying $F'(k) < 0$ and $F''(k) < 0$, so is the valuation function $V(k) = \lambda V_1(k) + \mu V_2(k)$, with $\lambda > 0$ and $\mu > 0$.

We have discussed some conditions on the valuation function $V(k)$ such that Propositions 5 and 6 still hold. In the following, we propose an example to illustrate the case where these conditions are no more satisfied.

Example 2 We use the specification where $f(x) = x^r$, with $r = 1$, $V(k) = V/k^\mu$, $V = 1$ and $\mu \geq 0$. We consider the (Clark and Riis, 1996) prize allocation mechanism. The number of contestants n is set equal to 100. We consider a class of probability distributions $\pi(k)$, characterized by two parameters $1 < \bar{k} < 8$ and $0 < v < 1$, as represented in Figure 1.2.¹⁹ Accordingly, the number of prizes has an expected value equal to \bar{k} and a variance equal to v .

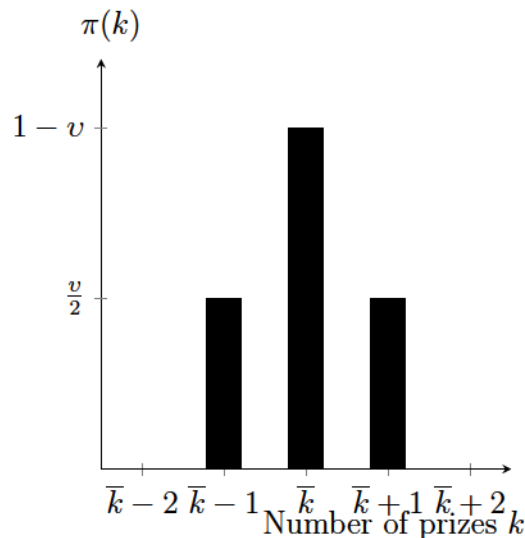


Figure 1.2: Probability distribution on the number of prizes

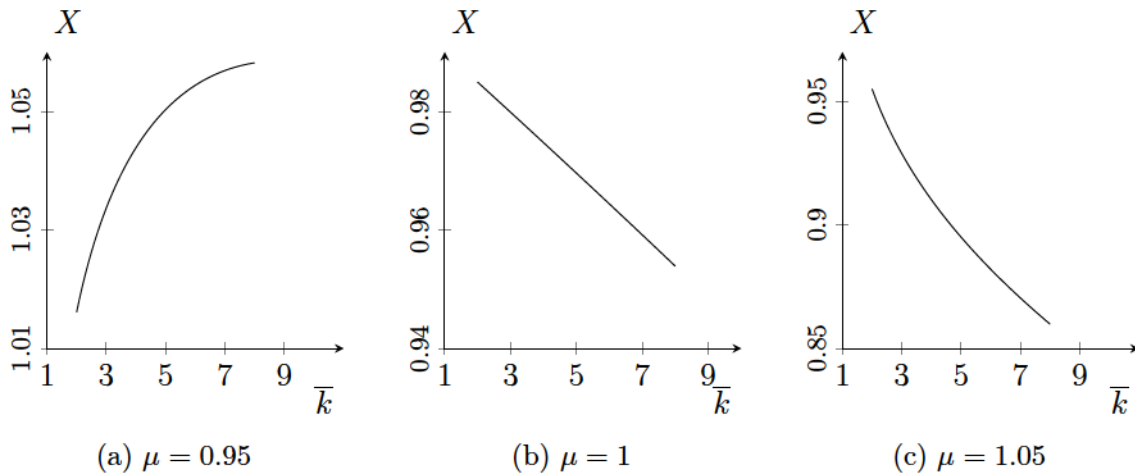
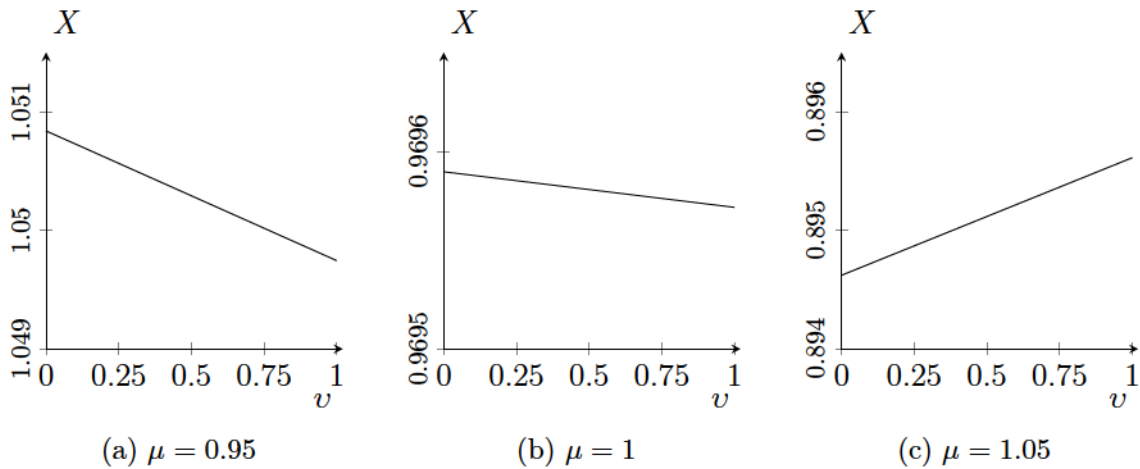
This class of probability distributions is convenient to deal separately with the first- and second-order stochastic dominance effects that we are interested in. An increase of \bar{k} alone gives a new probability distribution that first-order stochastically dominates the initial one. A decrease of v alone yields a new probability distribution that second-order stochastically dominates the initial one.

¹⁹Formally, $\pi(\bar{k}) = 1 - v$, $\pi(\bar{k} - 1) = \pi(\bar{k} + 1) = v/2$ and $\pi(k) = 0$ otherwise.

Let us first illustrate the effect of a larger expected number of prizes. Figure 1.3 represents the total effort X as a function of the expected number of prizes \bar{k} , the variance v being set to 0.5. The two panels 1.3(b) and 1.3(c), dealing with the cases where $\mu = 1$ and $\mu = 1.05$, are meant to illustrate our previous findings in this section. We observe that the aggregate effort is decreasing as the expected number of prizes increases. In fact, we have shown theoretically that this will be the case whenever $V(k)k$ is non increasing (i.e., here, for all $\mu \geq 1$). Panel 1.3(a) supplies an example, dealing with the case where $\mu = 0.95$, such that the aggregate effort is increasing with the expected number of prizes \bar{k} .

Let us now illustrate the effect of a larger variance of the number of prizes. Figure 1.4 plots the total effort X as a function of the variance v , the expected number of prizes \bar{k} being set to 5. The two panels 1.4(a) and 1.4(b), dealing with the cases where $\mu = 0.95$ and $\mu = 1$, are given to confirm our previous results in this section. We observe that the aggregate effort is decreasing as the variance of the number of prizes increases. In fact, we have proved theoretically that this will be the case whenever $V(k)k$ is non decreasing and weakly concave (i.e., here, for all $\mu \leq 1$). Panel 1.4(c) provides an example, dealing with the case where $\mu = 1.05$, such that the aggregate effort is increasing with the variance of the number of prizes v .

Another factor that can modify the sign of the second derivative of $F(k)$ is the specification of the conditional probability to be selected or eliminated. In this paper, we have only considered those developed by Clark and Riis (1996) and Fu et al. (2014). However in the literature, there exists other forms for the conditional probability to be selected or eliminated (Berry, 1993; Chowdhury and Kim, 2014), the former being largely criticized by Clark and Riis (1996) for instance. Given that equilibria in the certain and uncertain case are easy to obtain with these specifications, and that effort at equilibrium can be considered as a continuous function of the number of prizes, we do not report the calculations and the proofs. Results are unchanged concerning the first-order stochastic dominance but may change concerning the second-order stochastic dominance. In particular, we obtain that \underline{x} is larger than \bar{x} when $\bar{\pi}(k)$ second-order stochastically dominates $\pi(k)$. In other words, increasing the risk in the number of prizes in the sense of the second-order stochastic dominance increases effort. Naturally comes the question to find general conditions on the probability function $P_i(x_1, \dots, x_n; k)$ such that the latter generates a function $A(k, n)$ concave or convex with k . The main difficulty is that nothing

Figure 1.3: Total effort as a function of \bar{k} Figure 1.4: Total effort as a function of v

ensures that $A(k, n)$ is a continuous function of the number of prizes k , and that we can find a continuous and twice differentiable function to extend $A(k, n)$ like for these two prize allocation mechanisms.

4.2 Number of contestants

We now consider the effect of the number of participants in the contest. Formally, we first derive the comparative statics of the *aggregate* contest expenditure $X = nx$ with respect to the number of contestants n . We then provide conditions on the impact function such that X is increasing in n .

From our previous analysis, we know that the equilibrium effort satisfies

$$g(x^*) = \sum_{k=1}^K \pi(k) A(k, n).$$

Let us use the following expression of $A_{WSM}(k, n)$ found in [Clark and Riis \(1996\)](#),

$$A_{WSM}(k, n) = \frac{V(k)}{n} \left(k - \sum_{j=0}^{k-1} \frac{k-j}{n-j} \right), \quad (1.22)$$

and of $A_{LEM}(k, n)$, that we can deduce from [Fu et al. \(2014\)](#),

$$A_{LEM}(k, n) = \frac{V(k)}{n} \left(n - k - \sum_{j=0}^{n-k-1} \frac{n-k-j}{n-j} \right). \quad (1.23)$$

Letting

$$\begin{aligned} B_{WSM}(n) &\equiv \sum_{k=1}^K \pi(k) \frac{V(k)}{n} \left(k - \sum_{j=0}^{k-1} \frac{k-j}{n-j} \right) \\ B_{LEM}(n) &\equiv \sum_{k=1}^K \pi(k) \frac{V(k)}{n} \left(n - k - \sum_{j=0}^{n-k-1} \frac{n-k-j}{n-j} \right), \end{aligned}$$

the equilibrium effort is implicitly defined by

$$g(x^*) = B(n).$$

Remarking that $A_{LEM}(k, n)$ is a discrete function of n , we use again the *digamma* function to construct a continuous and once differentiable function $F(n)$ coinciding with $A_{LEM}(k, n)$ for every integer n . Using [\(1.15\)](#), we obtain immediately

$$F_{LEM}(n) = V(k) \frac{k}{n} (\psi(n+1) - \psi(k+1))$$

and we can verify that

$$B_{LEM}(n) = \sum_{k=1}^K \pi(k) V(k) \frac{k}{n} (\psi(n+1) - \psi(k+1))$$

for all integer values n .

Using the implicit function theorem, we can show that

$$\frac{dx}{dn} = \frac{B'(n)}{g'(x)}.$$

By differentiation of $X = nx$, we then find that

$$\frac{dX}{dn} = x + n \frac{B'(n)}{g'(x)}.$$

From this, knowing that $g(x) = B(n)$ and $g'(x) > 0$, we can show that

$$\frac{dX}{dn} > 0 \Leftrightarrow \frac{g'(x)x}{g(x)} > -\frac{B'(n)n}{B(n)}.$$

Now, knowing that

$$B_{WSM}(n)n = \sum_{k=1}^K \pi(k)V(k) \left(k - \sum_{j=1}^{k-1} \frac{k-j}{n-j} \right)$$

$$B_{LEM}(n)n = \sum_{k=1}^K \pi(k)V(k)k(\psi(n+1) - \psi(k+1)),$$

and since

$$\frac{\partial}{\partial n} \left(-\sum_{j=1}^{k-1} \frac{k-j}{n-j} \right) = \sum_{j=1}^{k-1} \frac{k-j}{(n-j)^2} > 0$$

$$\psi'(n+1) = \sum_{j=0}^{+\infty} \frac{1}{(j+n+1)^2} > 0,$$

$B(n)n$ is increasing in n , implying that

$$B(n) + B'(n)n > 0$$

and

$$-\frac{B'(n)n}{B(n)} < 1.$$

Therefore, a sufficient condition for $dX/dn > 0$ is

$$\frac{g'(x)x}{g(x)} \geq 1.$$

Proposition 7 *Aggregate effort at equilibrium increases with the number of contestants whenever the elasticity of $g(x)$ is at least equal to one.*

In particular, the benchmark specification $f(x) = x^r$ satisfies Proposition 7, for all r^{20} .

5 Conclusion

This paper is the first one to analyze the case of multiple-prize contests with an uncertain number of prizes. We extend the multiple-prize contest initiated in [Berry \(1993\)](#) and in [Clark and Riis \(1996\)](#), assuming that the number of (identical) prizes to be awarded is a random variable.

Using the same specification as [Clark and Riis \(1996\)](#) (contest success function and prize valuation function) then [Fu et al. \(2014\)](#) (contest elimination function), we find that the aggregate contest expenditure increases when the number of prizes to be awarded is expected to be smaller and less risky, in the sense of first-order and second-order stochastic dominance, and when the number of contestants increases. We discuss the robustness of these properties (valuation function, specification of the conditional probability), both by providing sufficient conditions such that they still hold true and by proposing numerical illustrations showing opposite results.

This paper also illustrates how to use the *psi* (or *digamma*) function ([Abramowitz and Stegun, 1964](#)), in order to extend a discrete function appearing in [Clark and Riis \(1996\)](#) and [Fu et al. \(2014\)](#) to a continuous and twice differentiable function, which helps deriving the comparative statics of the equilibrium outcome. Our belief is that this trick may prove to be useful in other settings in the contest literature (uncertainty on the number of players for instance).

Future work includes risk-aversion and different prize valuations in a multiple-prize contest with an uncertain number of prizes. Indeed the literature exposed in introduction has shown that these assumptions may clearly modify the results. However this leads to complex calculations and we are not sure to obtain a tractable form.

²⁰Indeed, when $f(x) = x^r$, we have $g(x) = f(x)/f'(x) = x/r$ and $g'(x)x/g(x) = 1$.

6 Appendix

6.1 Winner Selection Mechanism (Clark and Riis (1996), subsection 3.1)

The prize allocation mechanism described by Clark and Riis is the following : “*The players make one rent-seeking contribution which is valid for all k rounds of a nested contest. The winner of the first prize is decided by using the probability distribution which arises after the rent-seeking outlays of all n players are collected. The winner of this round is then eliminated and the second prize is distributed by using the probability distribution which arises when we exclude the outlay of the first winner. This process continues until all k prizes have been distributed. Thus the probability distribution is updated after each round to reflect the fact that previous winners are eliminated from future rounds.*”

Let’s give an example with $k = 3$ prizes and $n = 10$ players. The contest designer organises three rounds to allocate the prizes. Consider that player i is the blue player. The other contestants that we suppose symmetric are represented in black. In the first round (Figure 1.5(a)), player i has a probability p_i^1 to win the prize (issue W). If he really wins, he is excluded from the other rounds of the contest as he can’t win no more than one prize.

Otherwise, player i loses (issue L) with the complementary probability $1 - p_i^1$ and a black player wins the first round. The latter is excluded from the future rounds while player i participates to the next round. The contest designer organises the second round (Figure 1.5(b)) to allocate the second prize where 9 players still remain. If player i wins (with probability p_i^2) then he is excluded from the future rounds. Otherwise, he loses with probability $1 - p_i^2$ and a black player wins the second round.

A final round is organised to allocate the third and last prize (Figure 1.5(c)). 8 players still remain : player i and seven symmetric contestants. Player i wins with probability p_i^3 or loses with probability $1 - p_i^3$. As there are no more prizes to award, the process of allocation stops.

Finally, the probability that player i wins a prize in a contest where three prizes are allocated is :

$$p_i^1 + (1 - p_i^1) p_i^2 + (1 - p_i^1) (1 - p_i^2) p_i^3$$

which corresponds to $P_i(x_i, x, 3)$.

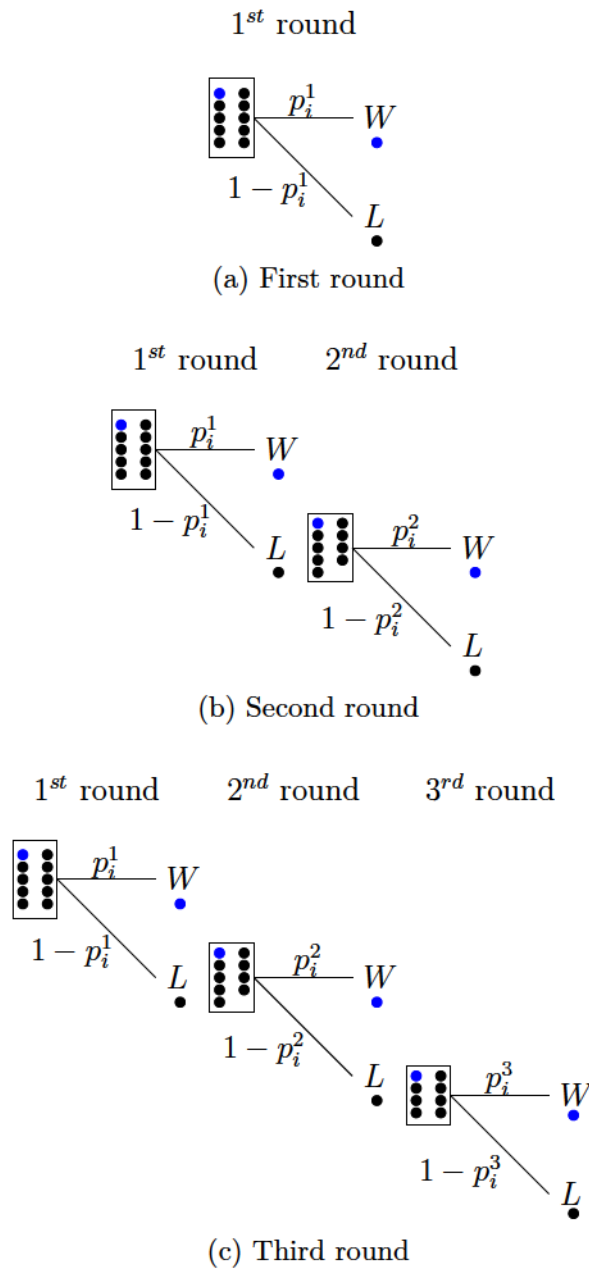


Figure 1.5: Mechanism of Clark and Riis (1996)

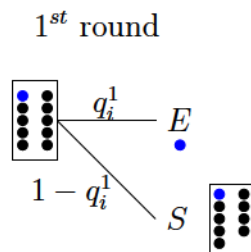
6.2 Looser Selection Mechanism (Fu et al. (2014), subsection 3.2)

The prize allocation mechanism proposed by Fu et al. (2014) is the *mirror image* of the mechanism of Clark and Riis (1996). Here, losers are sequentially eliminated to obtain the set of winners. Consider a contest among $n \geq 2$ players with $1 \leq k < n$ prizes. Each player exert effort only once. The loser elimination process lasts several rounds through which $n - k$ among the n players are eliminated. In each period, one player is eliminated, and when k players are left, the identical prizes of value $V(k)$ are distributed to the survivors.

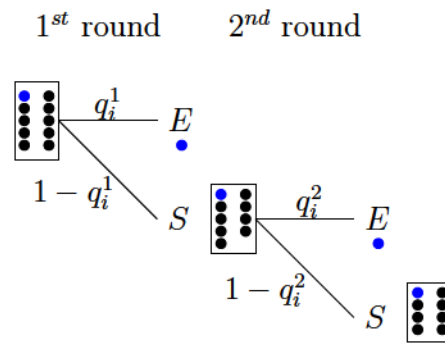
To illustrate this mechanism, let's take an example with $n = 10$ players and $k = 7$ prizes. This means that three players need to be eliminated to obtain the set of winners ; thus three rounds are organised. In the following, we only describe the edges of the game tree in which player i appears.

Consider that player i is the blue player. We suppose that the other contestants are symmetric and represented in black. For each round, note E the event “*the player is eliminated*” and S the event “*the player is a survivor*”.

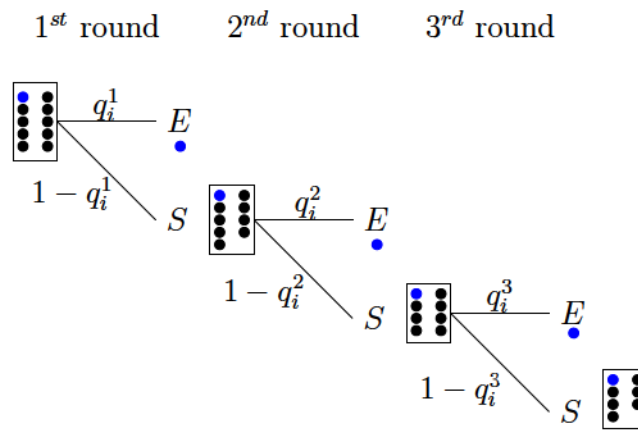
In the first round, player i is eliminated with probability q_i^1 . Otherwise, he survives with the complementary probability $1 - q_i^1$, which means that a black player has been removed from the contest. In this case, the set of survivors after the first round is composed of player i and eight symmetric black players.



Then a new round is organised to eliminate one of the nine remaining players. Player i is removed from the contest with probability q_i^2 . Otherwise, a black player is eliminated and contestant i survives in the second round with probability $1 - q_i^2$.



A final round is organised to exclude one of the eight remaining players. Concerning player i and conditional to be a survivor in the first and the second round, the event E occurs in the third round with probability q_i^3 and the event S occurs with probability $1 - q_i^3$. If player i has survived, the final set of winners is composed of player i and six other symmetric players.



The probability that player i wins a prize in a contest where seven prizes are allocated is :

$$P_i(x_i, x, 7) = (1 - q_i^1) (1 - q_i^2) (1 - q_i^3)$$

This expression can be generalized for any $k < n$:

$$P_i(x_i, x; k) = \prod_{\kappa=1}^{n-k} (1 - q_i^\kappa)$$

6.3 First Order and Second Order Stochastic Dominance

In economics of uncertainty, it is convenient to use the first-order and the second-order stochastic dominance to compare different lotteries according to the wealth they distribute. In our model, we compare contests according to the number of prizes they allocate. Using the first-order and second-order stochastic dominance on the number of prizes seems more relevant than on the total value distributed in the contest. As justification, consider the valuation function $V(k) = V/k$. The total value distributed in the contest is V , whatever the number of prizes and the probability distribution. However the way V is split in different prizes changes the players incentives to exert effort. In the following we present the criteria of first-order and second-order stochastic dominance on the number of prizes.

6.3.1 First-Order Stochastic Dominance

Denote by $\underline{\Pi}(k)$ and $\overline{\Pi}(k)$ the cumulative distribution functions corresponding to the probability density functions of the number of prizes $\underline{\pi}(k)$ and $\overline{\pi}(k)$. We say that a contest with $\overline{\Pi}(k)$ dominates a contest with $\underline{\Pi}(k)$ in the sense of first-order stochastic dominance if for all $k \in [1; +\infty[$:

$$\overline{\Pi}(k) \leq \underline{\Pi}(k)$$

6.3.2 Second-Order Stochastic Dominance

We say that $\underline{\pi}(k)$ is less risky than $\overline{\pi}(k)$ in the sense of the second order stochastic dominance if

$$\forall t \in [1; +\infty[, \int_1^t \underline{\Pi}(k) dk \leq \int_1^t \overline{\Pi}(k) dk,$$

Or equivalently

$$\forall t \in [1; +\infty[, \int_1^t (\underline{\Pi}(k) - \overline{\Pi}(k)) dk \leq 0. \quad (1.24)$$

It is convenient to consider a special case of the second order stochastic dominance where the expected number of prizes is the same for both probability distributions (mean-preserving spread). We give an illustration in Figure 1.6. Consider the two following probability density functions such that $\overline{\pi}(k)$ is a mean-preserving spread of $\underline{\pi}(k)$.

Note that the expected number of prizes is equal to 5 for both probability distributions. The cumulative distribution functions are given in the following table:

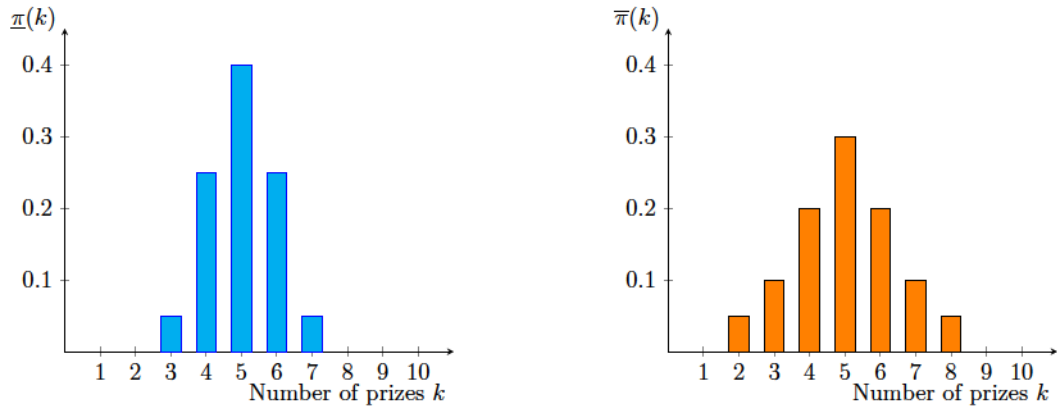


Figure 1.6: Mean preserving spread

k	[1; 2[[2; 3[[3; 4[[4; 5[[5; 6[[6; 7[[7; 8[[8; +∞[
$\underline{\Pi}(k)$	0	0	0.05	0.3	0.7	0.95	1	1
$\overline{\Pi}(k)$	0	0.05	0.15	0.35	0.65	0.85	0.95	1
$\underline{\Pi} - \overline{\Pi}$	0	-0.05	-0.1	-0.05	0.05	0.1	0.05	0

In Figure 1.7, we observe that condition (1.24) is satisfied.

t	$\int_1^t (\underline{\Pi}(k) - \overline{\Pi}(k)) dk$
[1; 2[0
[2; 3[$-0.05t + 0.1$
[3; 4[$-0.1t + 0.25$
[4; 5[$-0.05t + 0.05$
[5; 6[$0.05t - 0.45$
[6; 7[$0.1t - 0.75$
[7; 8[$0.05t - 0.4$
[8; +∞[0

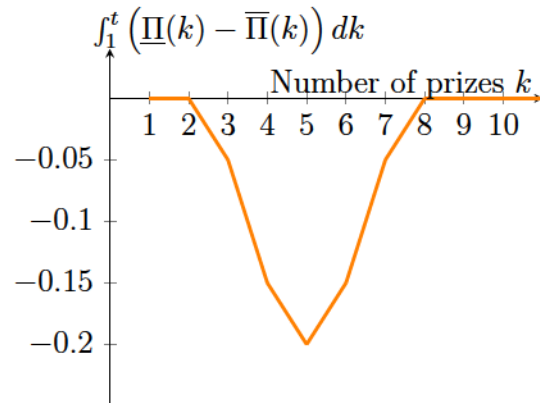


Figure 1.7: Illustration of the Second-Order Stochastic Dominance on the number of prizes

6.4 Proofs associated to the Winner Selection Mechanism (WSM)

6.4.1 Proof of expression (1.2)

Considering player i 's point of view, we assume that $x_j = x$, for all $j \neq i$. Then, for all round κ , we can write

$$p_i^\kappa(x_i, x, \dots, x) = \frac{f(x_i)}{f(x_i) + (n - \kappa) f(x)}.$$

If $k = 1$, player i chooses x_i to maximize

$$\frac{f(x_i)}{f(x_i) + (n - 1) f(x)} V(1) - x_i.$$

Therefore, the first-order condition for x_i to maximize player i 's payoff is

$$\frac{(n - 1) f(x) f'(x_i)}{(f(x_i) + (n - 1) f(x))^2} V(1) - 1 = 0.$$

If $x_i = x$, this simplifies to

$$V(1) \frac{n - 1}{n^2} \frac{f'(x)}{f(x)} - 1 = 0.$$

If $k > 1$, player i chooses x_i to maximize

$$P_i(x_i, x, \dots, x; k) V(k) - x_i,$$

where

$$P_i(x_i, x, \dots, x; k) = \frac{f(x_i)}{f(x_i) + (n - 1) f(x)} + \sum_{\kappa=2}^k \left[\prod_{\lambda=1}^{\kappa-1} \left(1 - \frac{f(x_i)}{f(x_i) + (n - \lambda) f(x)} \right) \right] \frac{f(x_i)}{f(x_i) + (n - \kappa) f(x)}$$

An equivalent expression is

$$P_i(x_i, x, \dots, x; k) = \sum_{\kappa=1}^k \frac{f(x_i)}{(n - \kappa) f(x)} \left(\prod_{\lambda=1}^{\kappa} \frac{(n - \lambda) f(x)}{f(x_i) + (n - \lambda) f(x)} \right).$$

By differentiation, we can get

$$\frac{\partial}{\partial x_i} P_i(x_i, x, \dots, x; k) = \sum_{\kappa=1}^k \left[\frac{f'(x_i)}{(n-\kappa)f(x)} \left(\prod_{\lambda=1}^{\kappa} \frac{(n-\lambda)f(x)}{f(x_i)+(n-\lambda)f(x)} \right) + \frac{f(x_i)}{(n-\kappa)f(x)} \left(- \sum_{s=1}^{\kappa} \frac{(n-\lambda)f(x)f'(x_i)}{(f(x_i)+(n-\lambda)f(x))^2} \prod_{\mu \neq \lambda} \frac{(n-\mu)f(x)}{f(x_i)+(n-\mu)f(x)} \right) \right].$$

which can be rewritten after some arrangements

$$\frac{\partial}{\partial x_i} P_i(x_i, x, \dots, x; k) = \sum_{\kappa=1}^k \left[\left(\frac{f'(x_i)}{(n-\kappa)f(x)} \right) \left(\prod_{\lambda=1}^{\kappa} \frac{(n-\lambda)f(x)}{f(x_i)+(n-\lambda)f(x)} \right) \left(1 - \sum_{\lambda=1}^{\kappa} \frac{f(x_i)}{f(x_i)+(n-\lambda)f(x)} \right) \right].$$

If $x_i = x$, this simplifies to²¹

$$\frac{\partial}{\partial x_i} P_i(x, \dots, x; k) = \frac{1}{n} \sum_{\kappa=1}^k \left(1 - \sum_{\lambda=1}^{\kappa} \frac{1}{n+1-\lambda} \right) \frac{f'(x)}{f(x)}.$$

We can show by induction that

$$\sum_{\kappa=1}^k \left(1 - \sum_{\lambda=1}^{\kappa} \frac{1}{n-\lambda+1} \right) = (n-k) \left(\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^{n-k} \frac{1}{r} \right),$$

which implies that

$$\frac{\partial}{\partial x_i} P_i(x, \dots, x; k) = \frac{n-k}{n} \left(\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^{n-k} \frac{1}{r} \right) \frac{f'(x)}{f(x)}$$

²¹Remark here that $\prod_{\lambda=1}^{\kappa} \frac{n-\lambda}{n+1-\lambda} = \frac{n-\kappa}{n}$.

6.4.2 Proof of Proposition 1

Second-order condition A sufficient condition for a global maximum of player i 's payoff is obtained if

$$\frac{\partial}{\partial x_i} P_i(x_i, x, \dots, x; k) = \sum_{\kappa=1}^k \left[\left(\frac{f'(x_i)}{(n-\kappa)f(x)} \right) \left(\prod_{\lambda=1}^{\kappa} \frac{(n-\lambda)f(x)}{f(x_i) + (n-\lambda)f(x)} \right) \left(1 - \sum_{\lambda=1}^{\kappa} \frac{f(x_i)}{f(x_i) + (n-\lambda)f(x)} \right) \right]$$

is decreasing with x_i . If we consider the following condition

$$1 - \sum_{\lambda=1}^{\kappa} \frac{f(x_i)}{f(x_i) + (n-\lambda)f(x)} > 0, \text{ for all } 1 \leq \kappa \leq k,$$

then, the expression under square brackets above is the product of three terms, each positive and decreasing in x_i (given that $f'(x_i) > 0$ and $f''(x_i) \leq 0$). Clearly, this implies that

$$\frac{\partial^2}{\partial (x_i)^2} P_i(x_i, x, \dots, x; k) < 0,$$

which gives a sufficient second-order condition for a global maximum ²².

Nevertheless the proposed inequality is not necessarily satisfied for all x_i ²³. However no player will invest more than $V(k)$ in the contest (otherwise his utility is negative) so we need to verify this condition for all $x_i \in [0; V(k)]$.

Moreover remark that our condition holds true for all $1 \leq \kappa \leq k$ if it holds true for $\kappa = k$. Thus we obtain a sufficient condition for a global maximum:

$$1 - \sum_{\lambda=1}^k \frac{f(x_i)}{f(x_i) + (n-\lambda)f(x)} > 0, \text{ for all } x_i \in [0; V(k)], \text{ for all } k.$$

Given that this condition may not be necessarily fulfilled, we can at least show that player i 's utility function is concave over the interval $[0; x]$ if the number of prizes is such that

$$\sum_{\lambda=1}^k \frac{1}{n-\lambda+1} \leq 1,$$

which can be rewritten as

$$\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^{n-k} \frac{1}{r} \leq 1,$$

²²(Clark and Riis, 1998) provide a sufficient second-order condition for a local maximum only. It writes $1 - \sum_{\lambda=1}^{\kappa} \frac{1}{n+1-\lambda} > 0$, for all $1 \leq \kappa \leq k$. Clearly, it can be derived from our analysis in the case where $x_i = x$.

²³For instance, if $x_i \rightarrow +\infty$, then $\frac{f(x_i)}{f(x_i) + (n-\lambda)f(x)}$ tends to 1 and the left hand side of the inequality tends to $1 - \kappa \leq 0$.

and approximated by $k \leq 0.632n$ for large values of n according to [Clark and Riis \(1998\)](#). Indeed, since $f(x_i) = 0$ when $x_i = 0$ and given that

$$\sum_{\lambda=1}^k \frac{f(x_i)}{f(x_i) + (n - \lambda)f(x)}$$

is increasing with x_i , our condition

$$1 - \sum_{\lambda=1}^k \frac{f(x_i)}{f(x_i) + (n - \lambda)f(x)} > 0$$

is true for all $x_i \in [0; x]$.

Participation constraint At equilibrium each player spends effort such that

$$g(x^*) = A(k, n)$$

if his participation constraint is fulfilled

$$P_i(x^*; k)V(k) - x^* > 0.$$

Consider first $g(x) = x^{24}$. If we suppose

$$1 - \sum_{\lambda=1}^k \frac{1}{n - \lambda + 1} > 0,$$

that we can rewrite

$$\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^{n-k} \frac{1}{r} \leq 1,$$

then we obtain

$$P_i(x^*; k)V(k) - x^* = \frac{V(k)}{n} \left[n \sum_{\lambda=1}^k \frac{1}{n - \lambda + 1} + k \left(1 - \sum_{\lambda=1}^k \frac{1}{n - \lambda + 1} \right) \right] > 0.$$

Thus each player participates to the contest.

For a general $g(x)$ function, we proceed as follows. Given that $g(x)$ is a continuous and increasing function of x on $[0; +\infty[$, with $g(0) = 0$ and $\lim_{x \rightarrow +\infty} g(x) = +\infty$, then g is a bijection from $[0; +\infty[$ to $[0; +\infty[$ and admits an inverse function $g^{-1} : [0; +\infty[\rightarrow [0; +\infty[$.

²⁴This corresponds to $f(x) = x$ for instance.

It follows that the participation constraint can be expressed as

$$P_i(x^*; k)V(k) - g^{-1}(A(k, n)) \geq 0.$$

If we find a function $g_1(x)$ such that the participation constraint is fulfilled

$$P_i(x^*; k)V(k) - g_1^{-1}(A(k, n)) \geq 0,$$

then this is also the case for any function $g_2(x) \geq g_1(x)$ ²⁵. Or we have found that this is the case for $g_1(x) = x$. Thus we can conclude that for any function $g(x) \geq x$, the participation constraint is fulfilled. Given that

$$g(x) = \frac{f(x)}{f'(x)},$$

we can rewrite $g(x) \geq x$ as

$$\frac{f'(x)x}{f(x)} \leq 1.$$

In other words, all the impact functions $f(x)$ with an elasticity lower or equal to one satisfy the participation constraint.

²⁵ Consider two functions $g_1(x)$ and $g_2(x)$ such that $g_1(x) \leq g_2(x) \forall x$. Then

$$g_1^{-1}(g_1(x)) \leq g_1^{-1}(g_2(x))$$

since g_1 is a continuous and increasing function of x . This leads to

$$x \leq g_1^{-1}(g_2(x)).$$

Since $x = g_2^{-1}(g_2(x))$ we deduce that

$$g_2^{-1}(g_2(x)) \leq g_1^{-1}(g_2(x)).$$

It follows that if $g_1(x) \leq g_2(x)$ then

$$g_2^{-1}(A(k, n)) \leq g_1^{-1}(A(k, n)).$$

6.4.3 Proof of Proposition 2

First-order condition An interior solution satisfies the following first-order condition if

$$\sum_{k=1}^K \pi(k) \frac{\partial}{\partial x_i} P_i(x_i, x, \dots, x; k) V(k) - 1 = 0, \text{ for all } i,$$

Given that we have calculated the derivative of $P_i(x_i, x; k)$ in subsection 6.4.1, we easily obtain equilibrium effort in the uncertain case

$$g(x^*) = \sum_{k=1}^K \pi(k) V(k) \frac{n-k}{n} \left(\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^{n-k} \frac{1}{r} \right).$$

Second-order condition An interior solution satisfies the following second-order condition if

$$\sum_{k=1}^K \pi(k) \frac{\partial^2}{\partial x_i^2} P_i(x_i, x, \dots, x; k) V(k) < 0, \text{ for all } i.$$

A sufficient condition is to satisfy

$$\frac{\partial^2}{\partial x_i^2} P_i(x_i, x, \dots, x; k) < 0$$

$\forall k = 1, \dots, K$. Note that no player will invest more than the highest prize value distributed in the contest (otherwise his utility is negative). Given the discussion in the certain case, the condition to obtain a global maximum is

$$1 - \sum_{\lambda=1}^K \frac{f(x_i)}{f(x_i) + (n-\lambda)f(x)} > 0, \text{ for all } x_i \in \left[0; \max_{k=1, \dots, K} V(k)\right].$$

If this condition is not verified, we can show, as previously, that player i 's utility function is concave over the interval $[0; x]$ if the number of prizes is such that

$$\sum_{\lambda=1}^K \frac{1}{n-\lambda+1} \leq 1,$$

that can be rewritten as

$$\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^{n-k} \frac{1}{r} \leq 1,$$

and approximated by $K \leq 0.632n$ for high values of n .

Participation constraint At equilibrium each player spends effort such that

$$g(x^*) = \sum_{k=1}^K \pi(k)A(k, n)$$

if his participation constraint is fulfilled

$$\sum_{k=1}^K \pi(k)P_i(x^*; k)V(k) - x^* > 0.$$

Consider first $g(x) = x$ ²⁶. If we suppose

$$1 - \sum_{\lambda=1}^k \frac{1}{n - \lambda + 1} > 0,$$

then we obtain

$$\sum_{k=1}^K \pi(k)P_i(x^*; k)V(k) - x^* = \frac{1}{n} \sum_{k=1}^K \pi(k)V(k) \left[n \sum_{\lambda=1}^k \frac{1}{n - \lambda + 1} + k \left(1 - \sum_{\lambda=1}^k \frac{1}{n - \lambda + 1} \right) \right] > 0.$$

Thus each player participates to the contest.

For a general $g(x)$ function, we proceed as follows. Given that $g(x)$ is a continuous and increasing function of x on $[0; +\infty[$, with $g(0) = 0$ and $\lim_{x \rightarrow +\infty} g(x) = +\infty$, then g is a bijection from $[0; +\infty[$ to $[0; +\infty[$ and admits an inverse function $g^{-1} : [0; +\infty[\rightarrow [0; +\infty[$.

It follows that the participation constraint can be expressed as

$$\sum_{k=1}^K \pi(k)P_i(x^*; k)V(k) - g^{-1} \left(\sum_{k=1}^K \pi(k)A(k, n) \right) \geq 0.$$

If we find a function $g_1(x)$ such that the participation constraint is fulfilled

$$\sum_{k=1}^K \pi(k)P_i(x^*; k)V(k) - g_1^{-1} \left(\sum_{k=1}^K \pi(k)A(k, n) \right) \geq 0,$$

then this is also the case for any function $g_2(x) \geq g_1(x)$ ²⁷. Or we have found that this is the case for $g_1(x) = x$. Thus we can conclude that for any function $g(x) \geq x$, the

²⁶This corresponds to $f(x) = x$ for instance.

²⁷ Consider two functions $g_1(x)$ and $g_2(x)$ such that $g_1(x) \leq g_2(x) \forall x$. Then

$$g_1^{-1}(g_1(x)) \leq g_1^{-1}(g_2(x))$$

since g_1 is a continuous and increasing function of x . This leads to

$$x \leq g_1^{-1}(g_2(x)).$$

participation constraint is fulfilled. Given that

$$g(x) = \frac{f(x)}{f'(x)},$$

we can rewrite $g(x) \geq x$ as

$$\frac{f'(x)x}{f(x)} \leq 1.$$

In other words, all the impact functions $f(x)$ with an elasticity lower or equal to one satisfy the participation constraint.

Since $x = g_2^{-1}(g_2(x))$ we deduce that

$$g_2^{-1}(g_2(x)) \leq g_1^{-1}(g_2(x)).$$

It follows that if $g_1(x) \leq g_2(x)$ then

$$g_2^{-1} \left(\sum_{k=1}^K \pi(k)A(k, n) \right) \leq g_1^{-1} \left(\sum_{k=1}^K \pi(k)A(k, n) \right).$$

6.4.4 Proof of equations (1.18) and (1.20)

$F_{WSM}(k)$ is twice differentiable and is equal to

$$F_{WSM}(k) = V(k) \frac{n-k}{n} (\psi(n+1) - \psi(n-k+1))$$

We calculate $F'_{WSM}(k)$ and $F''_{WSM}(k)$:

$$\begin{aligned} F'_{WSM}(k) &= \frac{V'(k)}{n} (n-k) (\psi(n+1) - \psi(n-k+1)) \\ &\quad - \frac{V(k)}{n} (\psi(n+1) - \psi(n-k+1) - (n-k)\psi'(n-k+1)) \end{aligned}$$

$$\begin{aligned} F''_{WSM}(k) &= \frac{V''(k)}{n} (n-k) (\psi(n+1) - \psi(n-k+1)) \\ &\quad - 2 \frac{V'(k)}{n} (\psi(n+1) - \psi(n-k+1) - (n-k)\psi'(n-k+1)) \\ &\quad - \frac{V(k)}{n} (2\psi'(n-k+1) + (n-k)\psi''(n-k+1)). \end{aligned}$$

We report some intermediate calculations using (1.17)

$$\begin{aligned} \psi(n+1) - \psi(n-k+1) &= \sum_{j=0}^{\infty} \frac{k}{(j+n+1)(j+n-k+1)} \\ \psi'(n-k+1) &= \sum_{j=0}^{\infty} \frac{1}{(j+n-k+1)^2} \\ \psi''(n-k+1) &= - \sum_{j=0}^{\infty} \frac{2}{(j+n-k+1)^3} \end{aligned}$$

which allow us to rewrite $F'_{WSM}(k)$ and $F''_{WSM}(k)$ as

$$\begin{aligned} F'_{WSM}(k) &= \frac{V'(k)}{n} \sum_{j=0}^{\infty} \frac{(n-k)k}{(j+n+1)(j+n-k+1)} \\ &\quad - \frac{V(k)}{n} \left(\sum_{j=0}^{\infty} \frac{k}{(j+n+1)(j+n-k+1)} - \sum_{j=0}^{\infty} \frac{n-k}{(j+n-k+1)^2} \right) \end{aligned}$$

$$\begin{aligned}
F''_{WSM}(k) &= \frac{V''(k)}{n} \sum_{j=0}^{\infty} \frac{(n-k)k}{(j+n+1)(j+n-k+1)} \\
&\quad - 2 \frac{V'(k)}{n} \left(\sum_{j=0}^{\infty} \frac{k}{(j+n+1)(j+n-k+1)} - \sum_{j=0}^{\infty} \frac{n-k}{(j+n-k+1)^2} \right) \\
&\quad - 2 \frac{V(k)}{n} \left(\sum_{j=0}^{\infty} \frac{1}{(j+n-k+1)^2} - \sum_{j=0}^{\infty} \frac{n-k}{(j+n-k+1)^3} \right).
\end{aligned}$$

First we focus on $F'_{WSM}(k)$:

$$F'_{WSM}(k) = -\frac{1}{n} \left(\begin{array}{c} -V'(k)k \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+n-k+1)} \\ +V(k) \left(\sum_{j=0}^{\infty} \frac{k}{(j+n+1)(j+n-k+1)} - \sum_{j=0}^{\infty} \frac{n-k}{(j+n-k+1)^2} \right) \end{array} \right)$$

We add and remove the following expression

$$V(k) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+n-k+1)},$$

which leads to

$$\begin{aligned}
F'_{WSM}(k) &= -\frac{1}{n} \left(\begin{array}{c} -(V'(k)k + V(k)) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+n-k+1)} \\ +V(k) \left(\sum_{j=0}^{\infty} \frac{k}{(j+n+1)(j+n-k+1)} - \sum_{j=0}^{\infty} \frac{n-k}{(j+n-k+1)^2} + \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+n-k+1)} \right) \end{array} \right) \\
&= -\frac{1}{n} \left(\begin{array}{c} -(V'(k)k + V(k)) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+n-k+1)} \\ +V(k) \left(\sum_{j=0}^{\infty} \frac{n}{(j+n+1)(j+n-k+1)} - \sum_{j=0}^{\infty} \frac{n-k}{(j+n-k+1)^2} \right) \end{array} \right)
\end{aligned}$$

We finally obtain expression (1.18):

$$F'_{WSM}(k) = -\frac{1}{n} \left(\begin{array}{c} -(V'(k)k + V(k)) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+n-k+1)} \\ +kV(k) \sum_{j=0}^{\infty} \frac{j+1}{(j+n+1)(j+n-k+1)^2} \end{array} \right)$$

Looking at $F''(k)$:

$$F''_{WSM}(k) = -\frac{2}{n} \left(\begin{array}{c} -\frac{1}{2}V''(k)k \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+n-k+1)} \\ +V'(k) \left(\sum_{j=0}^{\infty} \frac{k}{(j+n+1)(j+n-k+1)} - \sum_{j=0}^{\infty} \frac{n-k}{(j+n-k+1)^2} \right) \\ +V(k) \left(\sum_{j=0}^{\infty} \frac{1}{(j+n-k+1)^2} - \sum_{j=0}^{\infty} \frac{n-k}{(j+n-k+1)^3} \right) \end{array} \right)$$

We add and remove the following expression in the first two terms of the sum

$$V'(k) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+n-k+1)},$$

which leads to

$$F''_{WSM}(k) = -\frac{2}{n} \left(\begin{array}{c} -\frac{1}{2}(V''(k)k + 2V'(k)) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+n-k+1)} \\ +V'(k) \left(\sum_{j=0}^{\infty} \frac{k}{(j+n+1)(j+n-k+1)} - \sum_{j=0}^{\infty} \frac{n-k}{(j+n-k+1)^2} + \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+n-k+1)} \right) \\ +V(k) \left(\sum_{j=0}^{\infty} \frac{1}{(j+n-k+1)^2} - \sum_{j=0}^{\infty} \frac{n-k}{(j+n-k+1)^3} \right) \end{array} \right)$$

Rearranging the terms, it is possible to rewrite $F''_{WSM}(k)$ as:

$$F''_{WSM}(k) = -\frac{2}{n} \left(\begin{array}{c} -\frac{1}{2}(V''(k)k + 2V'(k)) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+n-k+1)} \\ +V'(k) \left(-\sum_{j=0}^{\infty} \frac{n-k}{(j+n-k+1)^2} + \sum_{j=0}^{\infty} \frac{n}{(j+n+1)(j+n-k+1)} \right) \\ +V(k) \sum_{j=0}^{\infty} \frac{j+1}{(j+n-k+1)^3} \end{array} \right)$$

We finally obtain expression (1.20):

$$F''_{WSM}(k) = -\frac{2}{n} \left(\begin{array}{c} -\frac{1}{2}(V''(k)k + 2V'(k)) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+n-k+1)} \\ + (V'(k)k + V(k)) \sum_{j=0}^{\infty} \frac{j+1}{(j+n+1)(j+n-k+1)^2} \\ +V(k)k \sum_{j=0}^{\infty} \frac{j+1}{(j+n+1)(j+n-k+1)^3} \end{array} \right)$$

6.5 Proofs associated to the Loser Elimination Mechanism (LEM)

6.5.1 Proof of expression (1.7)

Player i chooses to maximize

$$P_i(x_i, x, \dots, x; k) V(k) - x_i,$$

where

$$P_i(x_i, x, \dots, x, k) = \prod_{\kappa=1}^{n-k} \frac{(n-\kappa) f(x_i)}{(n-\kappa) f(x_i) + f(x)}.$$

Since

$$\frac{\partial}{\partial x_i} \left(\frac{(n-\kappa) f(x_i)}{(n-\kappa) f(x_i) + f(x)} \right) = f'(x_i)(n-\kappa) \frac{f(x)}{((n-\kappa) f(x_i) + f(x))^2},$$

we obtain by differentiating $P_i(x_i, x, \dots, x; k)$ with respect to x_i

$$\frac{\partial}{\partial x_i} P_i(x_i, x, \dots, x; k) = \frac{f'(x_i)}{f(x_i)} \left(\prod_{\kappa=1}^{n-k} \frac{(n-\kappa) f(x_i)}{(n-\kappa) f(x_i) + f(x)} \right) \left(\sum_{\kappa=1}^{n-k} \frac{f(x)}{(n-\kappa) f(x_i) + f(x)} \right)$$

If $x_i = x$, this simplifies to

$$\frac{\partial}{\partial x_i} P_i(x, \dots, x; k) = \frac{f'(x)}{f(x)} \left(\prod_{\kappa=1}^{n-k} \frac{n-\kappa}{n-\kappa+1} \right) \left(\sum_{\kappa=1}^{n-k} \frac{1}{n-\kappa+1} \right)$$

and after some intermediate calculations, we can show that

$$\frac{\partial}{\partial x_i} P_i(x, \dots, x; k) = \frac{k}{n} \left(\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^k \frac{1}{r} \right) \frac{f'(x)}{f(x)}.$$

6.5.2 Proof of Proposition 4

First-order condition An interior solution satisfies the following first-order condition if

$$\sum_{k=1}^K \pi(k) \frac{\partial}{\partial x_i} P_i(x_i, x, \dots, x; k) V(k) - 1 = 0, \text{ for all } i,$$

Since we have calculated the derivative of $P_i(x_i, x, \dots, x; k)$ in subsection 6.5.1, we deduce the expression of equilibrium effort in the uncertain case

$$g(x^*) = \sum_{k=1}^K \pi(k) V(k) \frac{k}{n} \left(\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^k \frac{1}{r} \right).$$

Second-order condition We first consider the case where the number of prizes is known with certainty and equal to k . The interior solution is a global maximum for player i if

$$\frac{\partial}{\partial x_i} P_i(x_i, x, \dots, x; k) = \frac{f'(x_i)}{f(x_i)} \left(\prod_{\kappa=1}^{n-k} \frac{(n-\kappa)f(x_i)}{(n-\kappa)f(x_i) + f(x)} \right) \left(\sum_{\kappa=1}^{n-k} \frac{f(x)}{(n-\kappa)f(x_i) + f(x)} \right)$$

is decreasing with x_i .

An equivalent expression is

$$\frac{\partial}{\partial x_i} P_i(x_i, x, \dots, x; k) = f'(x_i) \left(\frac{1}{f(x_i)} \prod_{\kappa=1}^{n-k} \frac{(n-\kappa)f(x_i)}{(n-\kappa)f(x_i) + f(x)} \right) \left(\sum_{\kappa=1}^{n-k} \frac{f(x)}{(n-\kappa)f(x_i) + f(x)} \right)$$

If we consider the following condition

$$1 - \sum_{\kappa=1}^{n-k} \frac{f(x)}{(n-\kappa)f(x_i) + f(x)} > 0,$$

then the three terms of the first derivative of $P_i(x_i, x, \dots, x; k)$ with respect to x_i are positive and decreasing with x_i .

The problem is that this condition is not satisfied for all x_i . Simply, when $x_i = 0$, the condition becomes $1 - (n - k) \leq 0$. Nevertheless we can show that player i 's utility function is concave over the interval $[x; +\infty[$ if the number of prizes is such that

$$\sum_{\kappa=1}^{n-k} \frac{1}{n - \kappa + 1} \leq 1, \tag{1.25}$$

or equivalently

$$\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^k \frac{1}{r} \leq 1,$$

which can be approximated by $k \geq 0.368n$ for large values of n . Indeed

$$\sum_{\kappa=1}^{n-k} \frac{f(x)}{(n-\kappa)f(x_i) + f(x)}$$

is decreasing with x_i so that if our condition is true for $x_i = x$, it holds true for $x_i \geq x$, if the number of prizes is such that (1.25) is fulfilled.

When the number of prizes is uncertain, an interior solution satisfies the following second-order condition if

$$\sum_{k=1}^K \pi(k) \frac{\partial^2}{\partial x_i^2} P_i(x_i, x, \dots, x; k) V(k) < 0, \text{ for all } i.$$

A sufficient condition is to satisfy

$$\pi(k) \frac{\partial^2}{\partial x_i^2} P_i(x_i, x, \dots, x; k) \leq 0$$

$\forall k = 1, \dots, K$, with a strict inequality for at least a value of k .

Denote by \underline{k} the minimum number of prizes such that

$$\sum_{\kappa=1}^{n-\underline{k}} \frac{1}{n-\kappa+1} \leq 1. \tag{1.26}$$

Then the condition above is satisfied for all $k \in [\underline{k}; K]$.

Thus under condition (1.26) and if $\pi(k) = 0 \forall k \in [1; \underline{k} - 1]$, player i 's utility function is concave over the interval $[x; +\infty]$.

Participation constraint The proof is close to the one given in Appendix 6.4.3. Each player participates to the contest if

$$\sum_{k=1}^K \pi(k) P_i(x^*; k) V(k) - x^* \geq 0.$$

With $g(x) = x$ and if we suppose

$$1 - \sum_{\kappa=1}^{n-k} \frac{1}{n - \kappa + 1} > 0,$$

or equivalently

$$\sum_{r=1}^n \frac{1}{r} - \sum_{r=1}^k \frac{1}{r} \leq 1,$$

we obtain

$$\sum_{k=1}^K \pi(k) P_i(x^*; k) V(k) - x^* = \frac{1}{n} \sum_{k=1}^K \pi(k) k V(k) \left(1 - \sum_{\kappa=1}^{n-k} \frac{1}{n - \kappa + 1} \right) > 0$$

Thus each player participates to the contest. We can show like in Appendix 6.4.3 that for a function $g(x) \geq x$, the participation constraint is fulfilled. As a consequence, any impact function $f(x)$ with an elasticity lower or equal to one satisfies the participation constraint.

6.5.3 Proof of equations (1.19) and (1.21)

$F_{LEM}(k)$ is twice differentiable is equal to

$$F_{LEM}(k) = \frac{V(k)}{n} k [\psi(n+1) - \psi(k+1)].$$

We calculate $F'_{LEM}(k)$ and $F''_{LEM}(k)$:

$$\begin{aligned} F'_{LEM}(k) &= \frac{V'(k)}{n} k (\psi(n+1) - \psi(k+1)) \\ &\quad + \frac{V(k)}{n} (\psi(n+1) - \psi(k+1) - k\psi'(k+1)) \end{aligned}$$

$$\begin{aligned} F''_{LEM}(k) &= \frac{V''(k)}{n} k (\psi(n+1) - \psi(k+1)) \\ &\quad + 2 \frac{V'(k)}{n} (\psi(n+1) - \psi(k+1) - k\psi'(k+1)) \\ &\quad + \frac{V(k)}{n} (-2\psi'(k+1) - k\psi''(k+1)) \end{aligned}$$

We report some intermediate calculations using (1.17)

$$\begin{aligned} \psi(n+1) - \psi(k+1) &= \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \\ \psi'(k+1) &= \sum_{j=0}^{\infty} \frac{1}{(j+k+1)^2} \\ \psi''(k+1) &= - \sum_{j=0}^{\infty} \frac{2}{(j+k+1)^3} \end{aligned}$$

which allow us to rewrite $F'_{LEM}(k)$ and $F''_{LEM}(k)$ as

$$F'_{LEM}(k) = -\frac{1}{n} \left(\begin{array}{c} -V'(k)k \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \\ +V(k) \left(\sum_{j=0}^{\infty} \frac{k}{(j+k+1)^2} - \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \right) \end{array} \right)$$

$$F''_{LEM}(k) = -\frac{2}{n} \left(\begin{array}{c} -\frac{1}{2}V''(k)k \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \\ -V'(k) \left(\sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} - \sum_{j=0}^{\infty} \frac{k}{(j+k+1)^2} \right) \\ +V(k) \left(\sum_{j=0}^{\infty} \frac{1}{(j+k+1)^2} + \sum_{j=0}^{\infty} \frac{k}{(j+k+1)^3} \right) \end{array} \right)$$

First we focus on $F'_{LEM}(k)$:

$$F'_{LEM}(k) = -\frac{1}{n} \left(\begin{array}{c} -V'(k)k \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \\ +V(k) \left(\sum_{j=0}^{\infty} \frac{k}{(j+k+1)^2} - \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \right) \end{array} \right)$$

We add and remove the following expression

$$V(k) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)},$$

which leads to

$$\begin{aligned} F'_{LEM}(k) &= -\frac{1}{n} \left(\begin{array}{c} -(V'(k)k + V(k)) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \\ +V(k) \left(\sum_{j=0}^{\infty} \frac{k}{(j+k+1)^2} - \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} + \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \right) \end{array} \right) \\ &= -\frac{1}{n} \left(\begin{array}{c} -(V'(k)k + V(k)) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \\ +V(k) \left(\sum_{j=0}^{\infty} \frac{k}{(j+k+1)^2} \right) \end{array} \right) \end{aligned}$$

We finally obtain expression (1.19)

$$F'_{LEM}(k) = -\frac{1}{n} \left(\begin{array}{c} -(V'(k)k + V(k)) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \\ +V(k)k \sum_{j=0}^{\infty} \frac{1}{(j+k+1)^2} \end{array} \right).$$

Looking at $F''(k)$

$$F''_{LEM}(k) = -\frac{2}{n} \left(\begin{array}{c} -\frac{1}{2}V''(k)k \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \\ -V'(k) \left(\sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} - \sum_{j=0}^{\infty} \frac{k}{(j+k+1)^2} \right) \\ +V(k) \left(\sum_{j=0}^{\infty} \frac{1}{(j+k+1)^2} + \sum_{j=0}^{\infty} \frac{k}{(j+k+1)^3} \right) \end{array} \right)$$

We add and remove the following expression in the first two terms of the sum

$$V'(k) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)},$$

which leads to

$$F''_{LEM}(k) = -\frac{2}{n} \left(\begin{array}{c} -\frac{1}{2}(V''(k)k + 2V'(k)) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \\ -V'(k) \left(\sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} - \sum_{j=0}^{\infty} \frac{k}{(j+k+1)^2} - \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \right) \\ +V(k) \left(\sum_{j=0}^{\infty} \frac{1}{(j+k+1)^2} + \sum_{j=0}^{\infty} \frac{k}{(j+k+1)^3} \right) \end{array} \right)$$

Rearranging the terms, it is possible to rewrite $F''_{LEM}(k)$ as:

$$F''_{LEM}(k) = -\frac{2}{n} \left(\begin{aligned} &-\frac{1}{2} (V''(k)k + 2V'(k)) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \\ &+ V'(k) \sum_{j=0}^{\infty} \frac{k}{(j+k+1)^2} \\ &+ V(k) \left(\sum_{j=0}^{\infty} \frac{1}{(j+k+1)^2} + \sum_{j=0}^{\infty} \frac{k}{(j+k+1)^3} \right) \end{aligned} \right)$$

We finally obtain expression (1.21):

$$F''_{LEM}(k) = -\frac{2}{n} \left(\begin{aligned} &-\frac{1}{2} (V''(k)k + 2V'(k)) \sum_{j=0}^{\infty} \frac{n-k}{(j+n+1)(j+k+1)} \\ &+ (V'(k)k + V(k)) \sum_{j=0}^{\infty} \frac{1}{(j+k+1)^2} \\ &+ V(k)k \left(\sum_{j=0}^{\infty} \frac{1}{(j+k+1)^3} \right) \end{aligned} \right).$$

Team Production Function in Science¹

1 Introduction

“If I have seen further it is by standing on the shoulders of Giants” wrote Newton to Hooke in 1676 in a letter that became famous (Newton and Turnbull, 1961). Newton’s quote doesn’t only highlight the importance of great authors’ contributions, but also the role of prior knowledge in the discovery of new ideas. Doing research implies to combine previous academic works, a form of input, to create research output: this is intrinsically a productive problem. Naturally each idea differs in the nature of inputs (by requiring different sets of prior knowledge), but mainly in their assembly, in other words the type of *knowledge production function*.

This is of interest since a large literature in innovation economics tries to understand how existing knowledge produces new knowledge. Interdisciplinarity and novelty are currently the two main topics dealing with this question. The former studies the combination of knowledge and expertise from different disciplines, the latter deals with the *“recombination of pre-existing knowledge components in an unprecedented fashion”* (Wang et al., 2017). The question of the production of new knowledge is also approached from the angle of geography of innovation at a more macroeconomic level (Jaffe, 1989). Nevertheless to the best of our knowledge, no theoretical contribution deals with knowledge production at the project level, and mainly the way to empirically test it. This paper contributes in that sense.

In addition to technological considerations, a wide variety of factors influences scientific production. Members composing the team may differ in their intellectual background, their ability, their working methods or their interest in the project for instance. Larger

¹This paper is co-authored with Nicolas Carayol.

teams may engender bureaucratic structures (Walsh and Lee, 2015), coordination issues and free-riding behaviors (Albanese and Van Fleet, 1985). The competition environment also plays a central role. Researchers engage in projects to develop new ideas and diffuse their results by writing papers. In the well-established “*publish or perish*” culture, academic scientists need to obtain publications in prestigious journals to give visibility to their work, gain reputation and consequently be promoted and funded. However all submitted papers are not published: journals also compete to be the most prominent and only award the articles they consider to be the best. An important feature of academic competition is the imperfection of the selection mechanism of papers by journals. On the one hand, some articles are published then withdrawn some years later after the detection of scientific fraud (Shewan and Coats, 2010). On the other hand, there exists many examples of papers initially rejected by some journals which are considered nowadays by the scientific community as major contributions of their discipline, for instance in economics (Gans and Shepherd, 1994).

In this paper, we model competition in Science as a Tullock contest between research projects for the obtention of a prize (a publication slot in a journal for instance). We conceptualize a research project as an idea and a team of researchers. Each idea is associated to a unique knowledge production function whose factors are subteam efforts, each one in a distinct field of expertise. Depending on the areas of expertise involved and the level of efforts required, researchers form a team to complete the project. We abstract from free-riding issues and intra-team coordination to focus on changes in the knowledge production function.

Technically, our model is inspired by the literature of group contests, whose first contribution is from Katz et al. (1990). Assuming symmetric players, they define the team aggregate effort as the sum of individual spendings. Baik (1993) extends their work to allow for asymmetric contestants who differ according to their prize valuations. Lee (2012) investigates a contest model where each group’s probability of winning is only determined by the minimum effort spent within each team. Chowdhury et al. (2013) treat the case where each group’s probability of winning depends exclusively on the highest effort expended within each team. Kolmar and Rommeswinkel (2013) adopt a team contest success function with a constant elasticity of substitution technology.

In those models, each individual produces an effort constituting a factor of the aggregate group output also called team *impact function* (Myerson and Wärneryd, 2006). In

our approach, each factor is the effort exerted by a subteam of researchers in a field of expertise. Here the knowledge production function reflects the way efforts are combined between research areas to obtain output. For the knowledge production function, we adopt a constant elasticity of substitution technology. Contrary to [Kolmar and Rommeswinkel \(2013\)](#), we deal separately with the effects of technical changes in the production function and the valuations of the prize by the subteams. This distinction reveals interesting indexes and properties that we discuss in this paper.

Since researchers are becoming more and more specialized, research projects often involve teamwork. Only for the sake of simplicity of our model, we suppose that a team leader makes the decisions on efforts for the whole project. In essence, we abstract from free-riding issues. [Adams et al. \(2005\)](#) argue that “*the system of reward for priority in discovery severely punishes shirking by team members and rewards good work with publication, reputation, and income*”. In this paper, we are interested in the consequences of technical changes in the knowledge production function on output and productivity. In fact, the choice of a centralized versus decentralized decision-making doesn’t change qualitatively our results, what we show in the last part of this paper.

We characterize the Nash equilibrium of the corresponding contest game. We find that all the variables are function of a synthetic index of the knowledge production function, that we call *complexity* of the production function given its properties. We derive comparative statics with respect to this indicator and analyze the effect of the number of projects in the contest. Among the results, we show that teams treating more complex problems are penalized since their participation, their probability to obtain the prize, their production and their productivity decrease. We also show that competition exerts a differentiated effect on project output depending on its own complexity.

This paper is organized as follows. Section 2 sets out the model. Section 3 defines the complexity of the production function. Section 4 characterizes the Nash equilibrium of the game. Section 5 deals with the properties of the equilibrium outcome. Section 6 extends the model by introducing different valuations for the prize between and within teams. Section 7 concludes. Proofs are given in Appendix.

2 The model

We consider a contest where G research projects compete for the obtention of a unique prize of value V (a publication in a scientific journal for instance). We assume that the contest is not perfectly discriminating so that competition in Science is a probabilistic contest. Formally, denote p_i the probability that project i obtains the prize and assume that p_i is of the Tullock form

$$p_i = \frac{q_i}{Q} \quad \forall i = 1, \dots, G, \quad (2.1)$$

with q_i the production function of team i and Q the total research production

$$Q = \sum_{l=1}^G q_l. \quad (2.2)$$

The project with the highest q has the greatest probability to obtain the prize, but the latter is not necessarily equal to one like in a perfectly discriminating contest.

Denote X_i the aggregate effort and α_i the effort productivity of project i

$$\alpha_i = \frac{q_i}{X_i}. \quad (2.3)$$

We assume that the cost function of efforts is linear and unitary, i.e. $c(x) = x$. Let u_i designate the utility function of project i

$$u_i = p_i V - X_i. \quad (2.4)$$

In this model, we assume that the knowledge production function q_i is of the constant elasticity of substitution form $\forall i = 1, \dots, G$. Formally, let define $m_i \geq 1$ the number of fields of expertise involved in project i . Denote $x_i = (x_{i1}, \dots, x_{im_i})$ the vector of efforts of dimension m_i , where each entry of coordinate $s = 1, \dots, m_i$ is x_{is} , the effort spent in research area s of project i . The aggregate effort X_i is equal to

$$X_i = \sum_{s=1}^{m_i} x_{is}. \quad (2.5)$$

We suppose that to each field of expertise $s = 1, \dots, m_i$ involved in project i corresponds a subteam of researchers.

Let ω_{is} be the effort intensity coefficient in research area s of project i , and $\omega_i = (\omega_{i1}, \dots, \omega_{im_i})$ the vector of those coefficients in this same project. Without loss of generality, assume that the coordinates of ω_i are sorted in descending order:

$$\omega_{i1} \geq \omega_{i2} \geq \dots \geq \omega_{im_i}, \quad (2.6)$$

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

Let define Ω_i a positive constant that can reflect the intrinsic quality of idea i or the efficiency of team i (organizational capacity, researchers ability...).

The knowledge production function of project i is defined by

$$q_i = \Omega_i \left(\sum_{s=1}^{m_i} \omega_{is} x_{is}^{\sigma_i} \right)^{\frac{1}{\sigma_i}}, \quad (2.7)$$

with $\sigma_i \in]-\infty; 0[\cup]0; 1]$ and $\epsilon_i = \frac{1}{1-\sigma_i}$ the elasticity of substitution between the efforts spent in each field of expertise. If σ_i tends to $-\infty$, efforts are perfect complements; $\sigma_i = 1$ corresponds to the situation where efforts are perfect substitutes. The Cobb-Douglas function is obtained if σ_i approaches 0 in the limit.

When efforts are perfect substitutes, the knowledge production function is the sum of efforts in the different fields of expertise. If some coordinates of vector x_i are null (at least 1 and at most $m_i - 1$ values), then $(\sum_{s=1}^{m_i} \omega_{is} x_{is}^{\sigma_i})^{\frac{1}{\sigma_i}} > 0$ if efforts are perfect substitutes whereas $(\sum_{s=1}^{m_i} \omega_{is} x_{is}^{\sigma_i})^{\frac{1}{\sigma_i}} = 0$ once efforts become more complementary. In substance, the elasticity of substitution indicates the nature of the project. Spending efforts in all the fields of expertise initially involved in the project is not necessary to get knowledge production when efforts are perfect substitutes. In other words, the interaction between research areas is low since knowledge production can be obtained with only a unique field of expertise. On the contrary, researchers must work in all the fields of expertise implicated in project i when efforts are complementary, otherwise knowledge production is null. Thus interaction between fields of expertise is more important when $\sigma_i \rightarrow -\infty$ than when $\sigma_i = 1$.

From the constant elasticity of substitution production function, we derive a synthetic indicator that we call *complexity* of the production function. Next section presents this index.

3 Complexity of the production function

Definition 1 *Let define*

$$\Gamma_i = \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} \right)^{1-\frac{1}{\sigma_i}}, \forall i = 1, \dots, G \quad (2.8)$$

to be the complexity of the constant elasticity of substitution production function.

Here, we need to be precise with the word “*complexity*” that can mean several things. According to the first sense given in the Oxford Dictionary of English, something complex *consists of many different and connected parts*. Our index Γ_i coincides with this definition since the different parts correspond to the fields of expertise and the degree of connection is given by σ_i . The second sense given in the Oxford Dictionary of English is something *not easy to analyse or understand; complicated or intricate*. Applied to research, this definition suggests that a project is complex since it requires a lot of knowledge or skills to be completed. This is definitively not the sense of the *complexity* of a project that we express in this paper.

We justify the use of *complexity* given the properties of this index that we state in the following proposition.

Proposition 8 Γ_i *increases with:*

- (1) *the number of fields of expertise m_i ,*
- (2) *the evenness of the distribution of ω_{is} ,*
- (3) *the complementarity between efforts in the different fields of expertise ($\sigma_i \rightarrow -\infty$).*

Γ_i *belongs to* $[1; m_i]$.

A detailed proof is given in Appendix 8.1.

Proposition 8 means that the more a process of production requires different inputs, the more these inputs are involved in a balanced way and the more it is difficult to switch inputs to obtain output, the more the production function is complex.

In the next section, we identify the Nash equilibrium of the model.

4 Equilibrium outcomes

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

$$\max_{x_{i1}, \dots, x_{im_i}} u_i. \quad (2.9)$$

Without loss of generality, let's suppose that research projects are sorted in ascending order so that:

$$\frac{\Gamma_1}{\Omega_1} \leq \frac{\Gamma_2}{\Omega_2} \leq \dots \leq \frac{\Gamma_G}{\Omega_G}. \quad (2.10)$$

We say that research project i is inactive if x_i is the null vector. Project i participates to the contest if at least one entry of vector x_i is strictly positive.

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

$$\bar{z} = \frac{1}{g-1} \sum_{j=1}^g \frac{\Gamma_j}{\Omega_j}.$$

The resolution of program (2.9) is very close to [Kolmar and Rommeswinkel \(2013\)](#). The following proposition characterizes the Nash equilibrium of the game.

Proposition 9 *Given the ranking of projects in (2.10), the active projects are those with the g lowest values of Γ/Ω , with g the smallest integer such that*

$$\frac{\Gamma_{g+1}}{\Omega_{g+1}} \geq \bar{z}, \quad (2.11)$$

or G if this value doesn't exist.

At equilibrium, the demand of labor in research area s of project i is

$$x_{is}^* = \begin{cases} \frac{\omega_{is}^{\frac{1}{1-\sigma}}}{\Omega_i \Gamma_i^{\frac{1}{\sigma-1}}} \frac{1}{\bar{z}} \left(1 - \frac{1}{\bar{z}} \frac{\Gamma_i}{\Omega_i}\right) V & \text{if } i = 1, \dots, g ; \forall s = 1, \dots, m_i \\ 0 & \text{if } i = g + 1, \dots, G \end{cases}$$

Proof is given in Appendix 8.2.

Substituting equilibrium spendings from (2.1) to (2.5) and in (2.7), we finally obtain:

$$\begin{aligned} p_i^* &= 1 - \frac{1}{\bar{z}} \frac{\Gamma_i}{\Omega_i} \\ Q^* &= \frac{V}{\bar{z}} \\ q_i^* &= Q^* p_i^* \\ u_i^* &= (p_i^*)^2 V \\ X_i^* &= \frac{\Gamma_i}{\Omega_i} p_i^* Q^* \\ \alpha_i^* &= \frac{\Omega_i}{\Gamma_i} \end{aligned}$$

$\forall i = 1, \dots, g, 0$ otherwise.

Note that if only two projects compete, then they are necessarily active so $g = G = 2$. If $G \geq 3$, then selection effects can occur ($g < G$): projects with high values of Γ/Ω relatively to other teams may not participate to the contest.

Remark 1 *All the variables at equilibrium ($\alpha_i^*, p_i^*, u_i^*, q_i^*, Q^*$) depend on Γ_i/Ω_i , \bar{z} and V .*

\bar{z} characterizes the smoothness of the effective competition at equilibrium, given that \bar{z} is close to the mean of the Γ/Ω of projects participating to the contest.

5 Comparative statics

In this section, we derive some comparative statics of the equilibrium outcome which are new compared to the paper of [Kolmar and Rommeswinkel \(2013\)](#). We show successively how the own characteristics of project i then those of competitors $j \neq i$ affect player i 's probability, outcome, productivity, effort and output at equilibrium. We mainly focus on q_i^* given that we have proxies for research project's impact (see chapter 3). We also analyze the consequences of the entry of new projects in the contest.

5.1 Project i characteristics

We have already discussed the effect of Γ_i/Ω_i on project i participation to the contest in Proposition 9. As a reminder, teams with a higher ratio Γ/Ω (those lower ranked in (2.10))

may not participate to the contest. Thereby increasing *ceteris paribus* the complexity Γ_i reduces the possibility for project i to take part in the competition. In the following proposition, we summarize the effect of Γ_i on the main variables at equilibrium.

Proposition 10 *Teams that are characterized by a larger Γ_i are disadvantaged in the contest since their participation, their probability to win (p_i^*), their output (q_i^*) as well as their productivity (α_i^*) decrease. This relation is convex for q_i^* .*

Proof is given in Appendix 8.3. To illustrate the three first assertions of Proposition 10, let's consider the following example.

i	ω_{i1}	ω_{i2}	ω_{i3}	Ω_i	Γ_i	p_i^*	q_i^*
1	1			1	1	0.293	0.207
2	0.97	0.03		1.1	1.144	0.264	0.187
3	0.96	0.04		1.1	1.183	0.240	0.17
4	0.93	0.04	0.03	1.2	1.352	0.203	0.144
5	0.5	0.5		1.33	2	0	0
6	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	2	3	0	0

Table 2.1: Example of a contest with 6 projects in competition

Example 3 *Imagine that six projects compete for the obtention of a prize of value $V = 1$ (Table 2.1). Project 1 only implies a unique field of expertise. Other projects are complex since they involve at least two research areas. To ease the example, we suppose that all the research projects have the same elasticity of substitution $\epsilon = 1$ ($\sigma_i = \sigma = 0$). In column 6 we compute the complexity of each research project. Columns 7 and 8 give respectively the probability to obtain the prize and the output of each research project at equilibrium. Let's consider projects 2 and 3 characterized by the same $\Omega_i = 1.1$ but a greater complexity for project 3. Indeed, $\Gamma_3 > \Gamma_2$ since these two research projects only differ from the distribution of ω_i (ω_2 being more uneven than ω_3), while the number of fields of expertise and the elasticity of substitution remain the same. We observe that the probability to obtain the prize at equilibrium and the output are greater for project 2 than for project 3. Concerning projects 5 and 6, even though they have high values of Ω , their complexity is so important that they don't participate to the contest.*

The convexity of q_i^* with respect to Γ_i can be interpreted in the following way. From Proposition 8, Γ_i belongs to $[1; m_i]$. When $\Gamma_i = 1$ the project is not complex since it

involves a unique field of expertise. The convexity of q_i^* with respect to Γ_i indicates that the highest penalty on output appears when the project begins to be complex, that is to say when Γ_i raises from 1 to $1 + \delta$ with $\delta \rightarrow 0^+$. Such an increase necessarily implies that the project goes from 1 to at least a second field of expertise.

Consider now the productivity α_i^* . Proposition 10 stems that α_i^* is decreasing with Γ_i . Remember that α_i^* is the ratio of outcome q_i^* to the total effort X_i^* . If the former is decreasing with Γ_i , the latter is not monotonous as established in the following proposition.

Proposition 11 *An increase of the complexity of project i has a positive effect on the total effort exerted if*

$$\frac{\Gamma_i}{\Omega_i} < \frac{\bar{z}}{2}, \quad (2.12)$$

and a negative one if

$$\frac{\Gamma_i}{\Omega_i} > \frac{\bar{z}}{2}.$$

Proof is given in Appendix 8.4.

If the ratio Γ_i/Ω_i is such that (2.12) is fulfilled, then raising Γ_i leads to a higher effort X_i^* . However, the increase of X_i^* is insufficient to compensate the negative effect of the production function technological change. As a consequence, the greater use of inputs doesn't translate into outcome q_i^* . Since the numerator of α_i^* decreases while its denominator increases, project i 's productivity declines with Γ_i . If condition (2.12) is not fulfilled, project i doesn't already belong to the most competitive projects. Increasing Γ_i discourages project i researchers to exert effort. As a consequence, q_i^* decreases. For the productivity α_i^* , the numerator and the denominator decrease; however the net effect is negative. We can also notice that α_i^* doesn't depend on the characteristics of other teams $j \neq i$. In (2.10), we have ranked projects in ascending order of Γ/Ω . In fact, (2.10) is also the ranking of projects by descending order of productivity at equilibrium.

5.2 Competition intensity

Here we analyze the effect of the competitive environment on project i main variables at equilibrium. We successively consider two forms of competition intensity: more efficient

projects (lower ratio Γ_k/Ω_k , $k \neq i$) and more competitors. In both cases, the effect is not monotonous and depends on the value of Γ_i/Ω_i . Next proposition gives the threshold under which decreasing the complexity of other teams $k \neq i$ has a positive (respectively negative) effect on project i output.

Proposition 12 *A decrease of the complexity of project $k \neq i$ has a positive effect on project i output if*

$$\frac{\Gamma_i}{\Omega_i} < \frac{\bar{z}}{2}, \quad (2.13)$$

and a negative one if

$$\frac{\Gamma_i}{\Omega_i} > \frac{\bar{z}}{2}.$$

Proof is given in Appendix 8.5.

Since $X_i^* = \frac{\Gamma_i}{\Omega_i} q_i^*$, the condition such as X_i^* raises after a decrease of Γ_k is exactly the same as for q_i^* . Decreasing the complexity of other projects $k \neq i$ increases the strenght of the competition for project i since other teams have more chances to win the contest. For projects with high Γ_i/Ω_i (the less efficient ones), the effect is negative on output since the increase of the competition gives disincentives to spend effort X_i^* at equilibrium. On the contrary, for projects with low Γ_i/Ω_i , the increase of the strenght of competition creates incentives to exert efforts.

Now we deal with an increase of the competition intensity by the number of projects. Consider the situation where g projects out of G are active as the initial contest. Assume that a new project characterized by Γ_{new} and Ω_{new} enters the contest. Denote r the number of active projects out of $G + 1$ after this entry. Let

$$\hat{z} = \frac{1}{r-1} \sum_{j=1}^r \frac{\Gamma_j}{\Omega_j}.$$

Remark that the new project stays inactive in the contest if

$$\frac{\Gamma_{new}}{\Omega_{new}} \geq \hat{z}.$$

Obviously in this situation the previous equilibrium remains unchanged.

Otherwise if there exists at least an integer $h \in [1, g]$ such that

$$\frac{\Gamma_{new}}{\Omega_{new}} \leq \frac{\Gamma_h}{\Omega_h},$$

then the new project is necessarily active.

Note that the number of active projects r after the entry can be greater or smaller than g , but lower or equal to $g + 1$. To show that $r \leq g + 1$, consider project $g + 1$. Since project $g + 1$ does not participate in the initial contest, we have

$$\frac{\Gamma_{g+1}}{\Omega_{g+1}} \geq \bar{z}. \quad (2.14)$$

A new project enters the contest if

$$\frac{\Gamma_{new}}{\Omega_{new}} \leq \bar{z}.$$

Project $g + 1$ participates to the contest after the entry if

$$\frac{\Gamma_{g+1}}{\Omega_{g+1}} \leq \frac{1}{g} \left(\sum_{j=1}^g \frac{\Gamma_j}{\Omega_j} + \frac{\Gamma_{new}}{\Omega_{new}} \right)$$

Using the constraint of participation of the new project, this leads to the following inequality

$$\frac{\Gamma_{g+1}}{\Omega_{g+1}} \leq \bar{z},$$

which contradicts (2.14). If the new project enters the contest, then necessarily $r \leq g + 1$.

In particular, if the new project is such that:

$$\frac{\Gamma_g}{\Omega_g} \leq \frac{\Gamma_{new}}{\Omega_{new}} < \bar{z},$$

then $r = g + 1$.

Next proposition describes the effect of the entry of a new project in the contest on q_i^* .

Proposition 13 *Consider a contest where g projects out of G are active. If a new project enters the contest such that:*

$$\frac{\Gamma_{new}}{\Omega_{new}} \leq \bar{z},$$

then at equilibrium q_i^* is greater in the new contest if

$$\frac{\Gamma_i}{\Omega_i} < \frac{\widehat{z}\bar{z}}{\widehat{z} + \bar{z}}$$

and lower in the new contest if

$$\frac{\Gamma_i}{\Omega_i} > \frac{\widehat{z}\bar{z}}{\widehat{z} + \bar{z}}.$$

Proof is given in Appendix 8.6. Note that if \widehat{z} is very close to \bar{z} , so that the smoothness of the effective competition at equilibrium remains unchanged, then we obtain the same condition as (2.13). Increasing the competition intensity by the number of projects has a positive effect on output for less complex projects while it can lead to a decrease of output or even to a non-participation for more complex projects. We obtain a differentiated effect of competition on output like in contests with different valuations for the prize (Matros, 2006).

To illustrate Propositions 12 and 13, we give the following example.

Example 4 *Let's consider the initial situation where two projects compete for a prize of value $V = 1$. We give their characteristics in Table 2.1a. Imagine that another research project joins the contest. Depending on the value of $\Gamma_{new}/\Omega_{new}$, we show in Figure 2.1b the effect of this entry on projects 1 and 2 output at equilibrium. Obviously, if the new project stays inactive ($\Gamma_{new}/\Omega_{new} \geq 3$, zone A), nothing happens and the new equilibrium is identical to the initial situation. In zone B, all the projects are active. Since condition (2.13) is checked for project 1, q_1^* increases when the complexity of the new project decreases (thus a greater intensity of competition). For project 2, we observe the opposite result. The competition becomes too fierce and discourages researchers to spend efforts. In particular, if the new project is characterized by $\Gamma_{new}/\Omega_{new} \leq 1$, project 2 no longer participates to the contest. In zone C, project 2 is inactive and condition (2.13) is now unchecked. The competition becomes too strong for project 1 against the new project and as a consequence, effort and output at equilibrium are decreasing when $\Gamma_{new}/\Omega_{new}$ diminishes.*

i	ω_{i1}	ω_{i2}	ω_{i3}	Ω_i	Γ_i	p_i^*	q_i^*
1	1			1	1	0.666	0.222
2	0.5	0.5		1	2	0.333	0.111

(a) Initial Contest

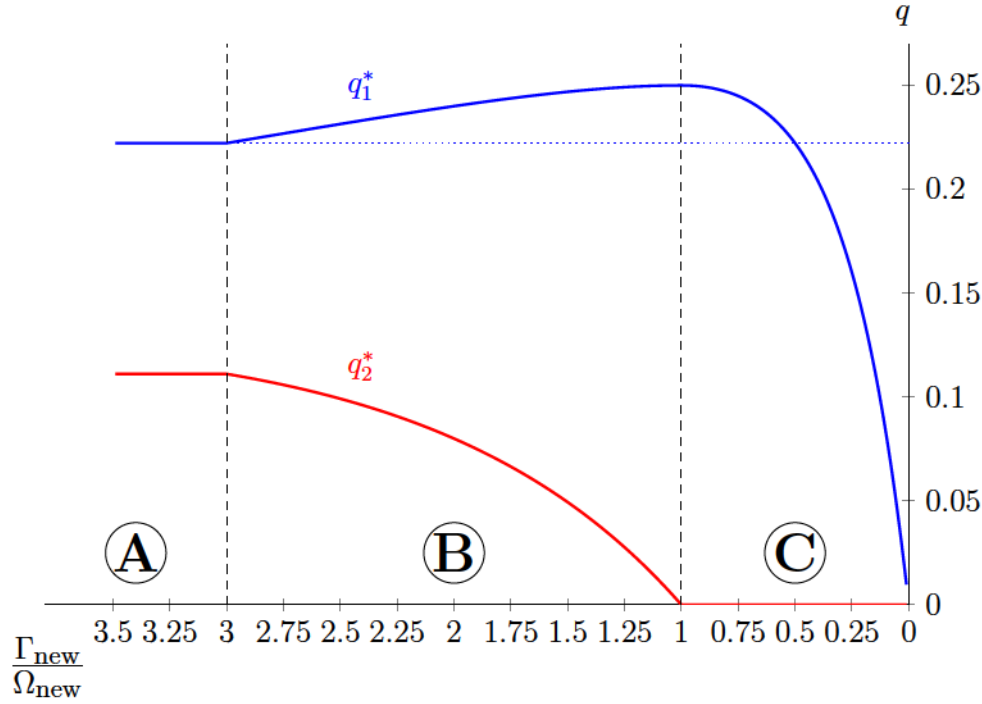
(b) Effect of the entry of a new project in the contest on q_1^* and q_2^*

Figure 2.1: Example to illustrate Propositions 12 and 13

6 Extensions

In the above, we have restricted the model of [Kolmar and Rommeswinkel \(2013\)](#) to the situation where each team has the same valuation for the prize. We have shown that all the variables at equilibrium depend on a synthetic indicator (Γ_i) called the *complexity* of the production function. This index only depends on the characteristics of project i knowledge production function.

In the following, we enrich the model by incorporating different valuations for the prize between and within teams. We also assume that each subteam chooses his own level of effort. In fact, we return to the framework of [Kolmar and Rommeswinkel \(2013\)](#). We introduce v_{ij} to be the value of project i for the subteam working in field of expertise j and $v_i = (v_{i1}, \dots, v_{im_i})$ the vector of those valuations. We keep the notations of sections 2 and 4 for the other variables. We also suppose that the decision-making is decentralized,

i.e. each subteam chooses his effort to maximize his own utility function

$$u_{ij} = p_i v_{ij} - x_{ij}, \quad (2.15)$$

with u_{ij} the utility of subteam j involved in project i .

Let define

$$\Upsilon_i = \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{\sigma_i}{1-\sigma_i}} \right)^{1-\frac{1}{\sigma_i}}, \forall i = 1, \dots, G \quad (2.16)$$

to be a synthetic index of the characteristics of project i . Suppose that research projects are sorted in ascending order so that:

$$\frac{\Upsilon_1}{\Omega_1} \leq \dots \leq \frac{\Upsilon_G}{\Omega_G} \quad (2.17)$$

We keep the notation $g \leq G$ for the number of active projects. Let

$$\bar{z} = \frac{1}{g-1} \sum_{j=1}^g \frac{\Upsilon_j}{\Omega_j}.$$

Next proposition from [Kolmar and Rommeswinkel \(2013\)](#) characterizes the Nash equilibrium of this game².

Proposition 14 *Given the ranking of projects in (2.17), the active projects are those with the g lowest values of Υ/Ω , with g the smallest integer such that*

$$\frac{\Upsilon_{g+1}}{\Omega_{g+1}} \geq \bar{z},$$

or G if this value doesn't exist.

At equilibrium, the demand of labor in research area s of project i is

$$x_{is}^* = \begin{cases} \frac{(\omega_{is} v_{is})^{\frac{1}{1-\sigma}}}{\Omega_i \Upsilon_i^{\frac{1}{\sigma-1}}} \frac{1}{\bar{z}} \left(1 - \frac{1}{\bar{z}} \frac{\Upsilon_i}{\Omega_i} \right) & \text{if } i = 1, \dots, g ; \forall s = 1, \dots, m_i \\ 0 & \text{if } i = g + 1, \dots, G \end{cases}$$

Proof is given in Appendix 8.7.

²We adapt their proposition to our notations.

Substituting equilibrium spendings in the main variables of the model, we get

$$\begin{aligned}
 p_i^* &= 1 - \frac{1}{\bar{z}} \frac{\Upsilon_i}{\Omega_i} \\
 Q^* &= \frac{1}{\bar{z}} \\
 q_i^* &= Q^* p_i^* \\
 X_i^* &= \frac{\left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{1}{1-\sigma_i}} \right) \Upsilon_i^{\frac{1}{1-\sigma_i}}}{\Omega_i} p_i^* Q^* \\
 \alpha_i^* &= \frac{\Omega_i}{\left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{1}{1-\sigma_i}} \right) \Upsilon_i^{\frac{1}{1-\sigma_i}}}.
 \end{aligned}$$

$\forall i = 1, \dots, g$, 0 otherwise.

Note that the comparative statics of p_i^* and q_i^* with respect to Υ_i are the same than those that we derived with respect to Γ_i in section 5. However Υ_i has a more complex shape, as well as X_i^* and α_i^* .

Here we search for conditions on the valuation vector v_i such that the properties of Γ_i given in Proposition 8 still hold for Υ_i . In other words, we want to see if *all other things being equal*, increasing the number of fields of expertise, the evenness of the distribution of ω_{is} or the complementarity between efforts also raises Υ_i .

In Appendix 8.8, we prove that decreasing σ_i increases Υ_i whatever the vector v_i . Thus Υ_i behaves similarly to Γ_i after an increase of the complementarity between efforts, regardless of the vector of valuations. The situation is different for the two other dimensions of the complexity, since the effect of a change in the number of fields of expertise or the evenness of the distribution of ω_{is} depends on v_i . Consider two fields of expertise j and t with $\omega_{ij} > \omega_{it}$. A sufficient condition to increase Υ_i after any transfer from ω_{ij} to ω_{it} is to verify $v_{ij} < v_{it}$ if $\sigma_i > 0$ and $v_{ij} > v_{it}$ if $\sigma_i < 0$. Given a vector v_i , this means that we can find situations such that increasing the evenness of the distribution of the ω_{is} decreases Υ_i . In other words, an increase of our index of complexity Γ_i can lead to a higher probability and output at equilibrium because of the structure of the vector of valuations v_i .

Example 5 *Imagine a journal publishing papers involving two academic disciplines. The reputation of this journal differs between the research areas. Suppose that the vector of valuations is $v_i = (10, 20)$. With $\omega_i = (2/3, 1/3)$ and $\sigma_i = 0.5$, we get $\Upsilon_i = 0.15$. With*

$\omega_i = (0.6, 0.4)$ and $\sigma_i = 0.5$, we have $\Upsilon_i \approx 0.147$. Thus the latter project is more complex than the former one but has a lower Υ_i , involving a greater probability and a higher output at equilibrium.

Looking at (2.16), Υ_i is a combination of two indexes. If all the coordinates of the ω_i -vector are identical ($\omega_{is} = 1/m_i \forall s = 1, \dots, m_i$), then Υ_i is proportional to the inverse of the generalized mean of the valuations denoted by $M_{\frac{\sigma_i}{1-\sigma_i}}$:

$$M_{\frac{\sigma_i}{1-\sigma_i}} = \left(\sum_{s=1}^{m_i} v_{is}^{\frac{\sigma_i}{1-\sigma_i}} \right)^{\frac{1-\sigma_i}{\sigma_i}}$$

When all the valuations of the prize within the team are identical, then Υ_i is proportional to Γ_i , our index of complexity. In fact, the results presented in section 5 still hold when there exists heterogeneity concerning the valuations of the prize between teams but not within teams.

7 Conclusion

This paper analyzes the effects of technical changes in the knowledge production function in a Tullock contest framework. We adopt a constant elasticity of substitution technology following [Kolmar and Rommeswinkel \(2013\)](#), considering first that the valuations between and within teams are identical. We show that at equilibrium, the main variables of the model are function of a synthetic index that we call *complexity* of the production function given its properties.

We find that the more the production function is complex, the more teams are disadvantaged in the contest given that their participation, their probability to win, their output and their productivity decrease. We also obtain that competition intensity exerts a positive effect on effort and output of most efficient teams, while it discourages those already treating more complex problems.

Our result is robust to different valuations between teams, but does not hold anymore when we introduce heterogeneity within teams. We need to conduct other extensions. In particular, we want to see if our results still prevail in a game with incomplete information, by supposing that each team knows his knowledge production function but not those of

competitors. Moreover, since we have supposed linear costs, our index may be modified if we use a strictly convex cost function. For further work, we want to see if other types of production function can engender other indexes at equilibrium with similar properties. For the moment, we want to empirically test our relation between complexity and output by using a dataset on scientific publications, which is the subject of the next chapter.

8 Appendix

8.1 Proof of Proposition 8

Parts (1) and (2): Let's analyze the impact of the number of fields of expertise m_i and the vector ω_i on Γ_i .

The objective is to find the ω_i -vector maximizing Γ_i .

Let's consider a vector $\underline{\omega}_i$. We say that $\underline{\omega}_i$ is improvable if there exists a vector $\overline{\omega}_i$ providing a higher Γ_i .

As it is a tedious work to compare all the possible vectors, one way to proceed is to see if any Lorenz-improvement in ω_i -vector increases Γ_i .

To compare ω_i -vectors, we use the concept of Lorenz dominance. In a research project involving m_i fields of expertise sorted like in (2.6), we say that $\widetilde{\omega}_i$ is Lorenz dominated by $\widehat{\omega}_i$ if and only if we have :

$$\sum_{s=0}^k \widetilde{\omega}_{i,m_i-s} \leq \sum_{s=0}^k \widehat{\omega}_{i,m_i-s}$$

$\forall k$ with $0 \leq k < m$ and the inequality is strict in at least one case.

A nice property of the Lorenz dominance concerns transfers. Indeed, consider an inegalitarian $\widetilde{\omega}_i$ -vector and operate a transfer which reduces the difference between two of its coordinates ω_{ij} and ω_{it} ($j < t$) to obtain the vector $\widehat{\omega}_i$. Then $\widehat{\omega}_i$ Lorenz dominates $\widetilde{\omega}_i$.

The two extreme cases are the egalitarian vector $\omega_i = (1/m_i, \dots, 1/m_i)$ and the perfect inegalitarian vector $\omega_i = (1, 0, \dots, 0)$. Example 6 with Figure 2.2 illustrates this concept.

Example 6 $\widehat{\omega}_i = (0.3, 0.3, 0.25, 0.15)$ and $\widetilde{\omega}_i = (0.5, 0.2, 0.2, 0.1)$.

In this example, note that the vector $\widehat{\omega}_i$ is obtained from $\widetilde{\omega}_i$ by transferring a weight of 0.2 from $\widetilde{\omega}_{i1}$ to $\widetilde{\omega}_{i2}$ (0.1), $\widetilde{\omega}_{i3}$ (0.05) and $\widetilde{\omega}_{i4}$ (0.05). Then we deduce that $\widehat{\omega}_i$ Lorenz dominates $\widetilde{\omega}_i$.

From (3.1), let's compute the partial derivative of Γ_i with respect to ω_{ij} :

$$\frac{\partial \Gamma_i}{\partial \omega_{ij}} = -\frac{1}{\sigma_i} \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} \right)^{-\frac{1}{\sigma_i}} \omega_{ij}^{\frac{1}{1-\sigma_i}-1} \quad (2.18)$$

For another field of expertise $t > j$, we obtain similarly:

$$\frac{\partial \Gamma_i}{\partial \omega_{it}} = -\frac{1}{\sigma_i} \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} \right)^{-\frac{1}{\sigma_i}} \omega_{it}^{\frac{1}{1-\sigma_i}-1} \quad (2.19)$$

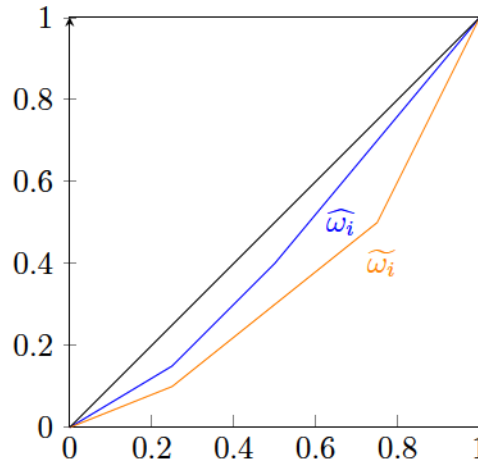


Figure 2.2: Illustration of Lorenz-dominance

Let's compare (2.18) and (2.19). Since $t > j$ and according to (2.6), we have $\omega_{ij} > \omega_{it}$. The impact of an infinitesimal transfer from ω_{ij} to ω_{it} (Lorenz domination of the new vector) depends on the comparison between $-\frac{1}{\sigma_i} \omega_{ij}^{\frac{1}{1-\sigma_i}-1}$ and $-\frac{1}{\sigma_i} \omega_{it}^{\frac{1}{1-\sigma_i}-1}$. For ω_{ij} and ω_{it} , a more egalitarian vector in the sense of Lorenz increases Γ_i if:

$$\frac{\partial \Gamma_i}{\partial \omega_{it}} > \frac{\partial \Gamma_i}{\partial \omega_{ij}}$$

In other words, it is enough to show that the function

$$f(x) = -\frac{1}{\sigma_i} x^{\frac{1}{1-\sigma_i}-1}$$

is decreasing in x when $x \in [0, 1]$.

Since we have:

$$f'(x) = -\frac{1}{1-\sigma_i} x^{\frac{1}{1-\sigma_i}-2} < 0 \quad \forall \sigma_i \in]-\infty, 0[\cup]0, 1],$$

we conclude that:

$$\frac{\partial \Gamma_i}{\partial \omega_{it}} > \frac{\partial \Gamma_i}{\partial \omega_{ij}} \quad \forall \omega_{it} < \omega_{ij}$$

and thus any Lorenz-improvement in the ω_i -vector increases Γ_i . We deduce that the vector ω_i which minimizes Γ_i is $(1, 0, \dots, 0)$, leading to $\Gamma_i = 1$. For a given m_i , the vector ω_i which maximizes Γ_i is $(\frac{1}{m_i}, \dots, \frac{1}{m_i})$, leading to $\Gamma_i = m_i$. Thus $\Gamma_i \in [1; m_i]$.

Part (3) : Our proof is close to Beck and Schögl (1995) dealing with the RÄnyi information (pages 50-53).

Derivating Γ_i with respect to σ_i , we obtain

$$\begin{aligned}
\frac{\partial \Gamma_i}{\partial \sigma_i} &= \frac{\Gamma_i}{\sigma_i^2} \left[\ln \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} \right) - \frac{\sigma_i}{1-\sigma_i} \frac{\sum_{s=1}^{m_i} (\ln \omega_{is}) \omega_{is}^{\frac{1}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}}} \right] \\
&= \frac{\Gamma_i}{\sigma_i^2} \left[\sum_{k=1}^{m_i} \frac{\omega_{ik}^{\frac{1}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}}} \ln \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} \right) - \sum_{k=1}^{m_i} \frac{\omega_{ik}^{\frac{1}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}}} \frac{\sigma_i}{1-\sigma_i} \ln \omega_{ik} \right] \\
&= \frac{\Gamma_i}{\sigma_i^2} \sum_{k=1}^{m_i} \left[\frac{\omega_{ik}^{\frac{1}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}}} \left(\ln \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} \right) - \ln \omega_{ik}^{\frac{1}{1-\sigma_i}-1} \right) \right] \\
&= -\frac{\Gamma_i}{\sigma_i^2} \sum_{k=1}^{m_i} \left[\frac{\omega_{ik}^{\frac{1}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}}} \left(\ln \omega_{ik}^{\frac{1}{1-\sigma_i}-1} - \ln \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} \right) \right) \right] \\
&= -\frac{\Gamma_i}{\sigma_i^2} \sum_{k=1}^{m_i} \frac{\omega_{ik}^{\frac{1}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}}} \ln \left(\frac{\omega_{ik}^{\frac{1}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}}} \right).
\end{aligned}$$

Let

$$P_{ik} = \frac{\omega_{ik}^{\frac{1}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}}}.$$

Finally, we obtain

$$\frac{\partial \Gamma_i}{\partial \sigma_i} = -\frac{\Gamma_i}{\sigma_i^2} \sum_{k=1}^{m_i} P_{ik} \ln \left(\frac{P_{ik}}{\omega_{ik}} \right).$$

This sum is always positive since it is the Kullback-Leibler divergence between P_{ik} and ω_{ik} . Indeed

$$\sum_{k=1}^{m_i} P_{ik} \ln \left(\frac{P_{ik}}{\omega_{ik}} \right) = -\sum_{k=1}^{m_i} P_{ik} \ln \left(\frac{\omega_{ik}}{P_{ik}} \right).$$

Since \ln is a concave function, by the Jensen inequality we have

$$\sum_{k=1}^{m_i} P_{ik} \ln \left(\frac{\omega_{ik}}{P_{ik}} \right) \leq \sum_{k=1}^{m_i} \ln \left(P_{ik} \frac{\omega_{ik}}{P_{ik}} \right) = \sum_{k=1}^{m_i} \ln (\omega_{ik}) \leq 0,$$

therefore

$$\sum_{k=1}^{m_i} P_{ik} \ln \left(\frac{P_{ik}}{\omega_{ik}} \right) \geq 0,$$

and finally

$$\frac{\partial \Gamma_i}{\partial \sigma_i} \leq 0.$$

Thus more complementarity between efforts (σ_i decreasing) increases Γ_i .

8.2 Proof of Proposition 9

The leader of research project i chooses the vector of efforts x_i to maximize u_i . We obtain the first order conditions $\forall j = 1, \dots, m_i$:

$$\Omega_i \left(\sum_{s=1}^{m_i} \omega_{is} x_{is}^{\sigma_i} \right)^{\frac{1}{\sigma_i} - 1} \omega_{ij} x_{ij}^{\sigma_i - 1} \frac{\sum_{l \neq i}^G q_l}{Q^2} V = 1. \quad (2.20)$$

Remark that the second order condition is always satisfied since

$$(1 - \sigma_i) \left(\omega_{ij} x_{ij}^{\sigma_i} - \sum_{s=1}^{m_i} \omega_{is} x_{is}^{\sigma_i} \right) Q - 2x_{ij}^{\sigma_i} \left(\sum_{s=1}^{m_i} \omega_{is} x_{is}^{\sigma_i} \right)^{\frac{1}{\sigma_i}} < 0.$$

Thus the effort at equilibrium x_{is}^* in field of expertise s is given by the solution to the first-order condition in (2.20) or by the corner solution $x_{is}^* = 0$.

Using (2.20), we deduce the following relation between efforts spent in research project i for two fields of expertise j and t :

$$x_{it} = \left(\frac{\omega_{it}}{\omega_{ij}} \right)^{\frac{1}{1-\sigma_i}} x_{ij}.$$

We get

$$\sum_{s=1}^{m_i} \omega_{is} x_{is}^{\sigma_i} = \frac{x_{ij}^{\sigma_i}}{\omega_{ij}^{\frac{1}{1-\sigma_i}}} \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} \right).$$

Taking the exponent $\frac{1}{\sigma_i} - 1$ on both sides, we obtain

$$\left(\sum_{s=1}^{m_i} \omega_{is} x_{is}^{\sigma_i} \right)^{\frac{1}{\sigma_i} - 1} = \frac{x_{ij}^{1-\sigma_i}}{\omega_{ij}} \frac{1}{\left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} \right)^{1-\frac{1}{\sigma_i}}}.$$

Replacing in the first order condition (2.20) leads to

$$\frac{\sum_{l \neq i}^G q_l}{\left(\sum_{l=1}^G q_l \right)^2} V = \frac{\Gamma_i}{\Omega_i}, \quad (2.21)$$

with

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

Thus project i best response given the strategies of the others is:

$$q_i = \max \left(0, \sqrt{Q_{/i} V \frac{\Omega_i}{\Gamma_i}} - Q_{/i} \right) \quad (2.22)$$

By denoting g the number of projects participating to the contest and summing (2.21)

$\forall i = 1, \dots, g$ we obtain:

$$Q^* = \frac{V}{\bar{z}}. \quad (2.23)$$

with

$$\bar{z} = \frac{1}{g-1} \sum_{j=1}^g \frac{\Gamma_j}{\Omega_j}.$$

From (2.22), project i participates to the contest if

$$V \frac{\Omega_i}{\Gamma_i} > Q_{/i}.$$

Using (2.23) and (2.21), this leads to

$$\frac{\Gamma_i}{\Omega_i} < \bar{z}.$$

Since we denote g the number of active projects, the first inactive project $g+1$ has a value $\Gamma_{g+1}/\Omega_{g+1}$ such that

$$\frac{\Gamma_{g+1}}{\Omega_{g+1}} \geq \bar{z}.$$

Given the ranking of projects in (2.10), g is the smallest integer such that condition (2.11) is satisfied.

8.3 Proof of Proposition 10

We report some intermediate calculations

$$\begin{aligned}\frac{\partial Q^*}{\partial \Gamma_i} &= -\frac{g-1}{\Omega_i} \frac{V}{\left(\sum_{j=1}^g \frac{\Gamma_j}{\Omega_j}\right)^2} < 0 \\ \frac{\partial p_i^*}{\partial \Gamma_i} &= -\frac{g-1}{\Omega_i} \frac{\sum_{j \neq i}^g \frac{\Gamma_j}{\Omega_j}}{\left(\sum_{j=1}^g \frac{\Gamma_j}{\Omega_j}\right)^2} < 0 \\ \frac{\partial^2 Q^*}{\partial \Gamma_i^2} &= \frac{g-1}{\Omega_i^2} \frac{2V}{\left(\sum_{j=1}^g \frac{\Gamma_j}{\Omega_j}\right)^3} > 0 \\ \frac{\partial^2 p_i^*}{\partial \Gamma_i^2} &= \frac{g-1}{\Omega_i^2} \frac{2 \sum_{j \neq i}^g \frac{\Gamma_j}{\Omega_j}}{\left(\sum_{j=1}^g \frac{\Gamma_j}{\Omega_j}\right)^3} > 0.\end{aligned}$$

Since $q_i^* = Q^* p_i^*$, we immediately obtain

$$\frac{\partial q_i^*}{\partial \Gamma_i} = \frac{\partial Q^*}{\partial \Gamma_i} p_i^* + Q^* \frac{\partial p_i^*}{\partial \Gamma_i} < 0.$$

Derivating again $\partial q_i^* / \partial \Gamma_i$ with respect to Γ_i , we get

$$\frac{\partial^2 q_i^*}{\partial \Gamma_i^2} = \frac{\partial^2 Q^*}{\partial \Gamma_i^2} p_i^* + 2 \frac{\partial Q^*}{\partial \Gamma_i} \frac{\partial p_i^*}{\partial \Gamma_i} + Q^* \frac{\partial^2 p_i^*}{\partial \Gamma_i^2} > 0$$

q_i^* is convex with respect to Γ_i .

8.4 Proof of Proposition 11

Remember that

$$X_i^* = \frac{\Gamma_i}{\Omega_i} p_i^* Q^*.$$

Given that

$$\begin{aligned} \frac{\partial p_i^*}{\partial \Gamma_i} &= -\frac{1}{\Omega_i} \frac{1}{\bar{z}} \frac{\sum_{j \neq i}^g \frac{\Gamma_j}{\Omega_j}}{\left(\sum_{j=1}^g \frac{\Gamma_j}{\Omega_j}\right)^2} \\ \frac{\partial Q^*}{\partial \Gamma_i} &= -\frac{1}{\Omega_i} \frac{Q^*}{\sum_{j=1}^g \frac{\Gamma_j}{\Omega_j}}, \end{aligned}$$

we have

$$\begin{aligned} \frac{\partial X_i^*}{\partial \Gamma_i} &= \frac{1}{\Omega_i} p_i^* Q^* + \frac{\Gamma_i}{\Omega_i} \left(\frac{\partial p_i^*}{\partial \Gamma_i} Q^* + p_i^* \frac{\partial Q^*}{\partial \Gamma_i} \right) \\ &= \frac{1}{\Omega_i} p_i^* Q^* + \frac{\Gamma_i}{\Omega_i} \left(-\frac{1}{\Omega_i} \frac{1}{\bar{z}} \frac{\sum_{j \neq i}^g \frac{\Gamma_j}{\Omega_j}}{\sum_{j=1}^g \frac{\Gamma_j}{\Omega_j}} Q^* - \frac{p_i^*}{\Omega_i} Q^* \frac{1}{\sum_{j=1}^g \frac{\Gamma_j}{\Omega_j}} \right) \\ &= \frac{Q^*}{\Omega_i} \left(p_i^* - \frac{\Gamma_i}{\Omega_i} \frac{1}{\bar{z}} \frac{\sum_{j \neq i}^g \frac{\Gamma_j}{\Omega_j}}{\sum_{j=1}^g \frac{\Gamma_j}{\Omega_j}} - \frac{\Gamma_i}{\Omega_i} \frac{p_i^*}{\sum_{j=1}^g \frac{\Gamma_j}{\Omega_j}} \right) \\ &= \frac{Q^*}{\Omega_i} \left(\sum_{j \neq i}^g \frac{\Gamma_j}{\Omega_j} \right) \left(p_i^* - \frac{\Gamma_i}{\Omega_i} \frac{1}{\bar{z}} \right) \\ &= \frac{Q^*}{\Omega_i} \left(\sum_{j \neq i}^g \frac{\Gamma_j}{\Omega_j} \right) \left(1 - 2 \frac{\Gamma_i}{\Omega_i} \frac{1}{\bar{z}} \right) \end{aligned}$$

Finally the derivative of X_i^* with respect to Γ_i is positive if

$$\frac{\Gamma_i}{\Omega_i} < \frac{\bar{z}}{2},$$

and negative if

$$\frac{\Gamma_i}{\Omega_i} > \frac{\bar{z}}{2}.$$

8.5 Proof of Proposition 12

As a reminder,

$$q_i^* = Q^* p_i^*$$

Given that

$$\begin{aligned} \frac{\partial Q^*}{\partial \Gamma_k} &= -\frac{1}{\Omega_k} \frac{Q^*}{\sum_{j=1}^g \frac{\Gamma_j}{\Omega_j}} \\ \frac{\partial p_i^*}{\partial \Gamma_k} &= \frac{1}{\bar{z}} \frac{1}{\Omega_k} \frac{\frac{\Gamma_i}{\Omega_i}}{\sum_{j=1}^g \frac{\Gamma_j}{\Omega_j}}, \end{aligned}$$

we obtain

$$\begin{aligned} \frac{\partial q_i^*}{\partial \Gamma_k} &= \frac{\partial Q^*}{\partial \Gamma_k} p_i^* + Q^* \frac{\partial p_i^*}{\partial \Gamma_k} \\ &= \frac{Q^*}{\Omega_k} \frac{1}{\sum_{j=1}^g \frac{\Gamma_j}{\Omega_j}} \left(-p_i^* + \frac{1}{\bar{z}} \frac{\Gamma_i}{\Omega_i} \right) \\ &= \frac{Q^*}{\Omega_k} \frac{1}{\sum_{j=1}^g \frac{\Gamma_j}{\Omega_j}} \left(\frac{2}{\bar{z}} \frac{\Gamma_i}{\Omega_i} - 1 \right). \end{aligned}$$

Thus the derivative of X_i^* with respect to Γ_k is positive if

$$\frac{\Gamma_i}{\Omega_i} > \frac{\bar{z}}{2},$$

and negative if

$$\frac{\Gamma_i}{\Omega_i} < \frac{\bar{z}}{2}.$$

8.6 Proof of Proposition 13

In the initial contest, project i output at equilibrium q_i^* is given by:

$$q_i^* = \frac{V}{\bar{z}} \left(1 - \frac{1}{\bar{z}} \frac{\Gamma_i}{\Omega_i} \right)$$

After the entry of a new project, project i output at equilibrium q_i^{*AE} is equal to:

$$q_i^{*AE} = \frac{V}{\hat{z}} \left(1 - \frac{1}{\hat{z}} \frac{\Gamma_i}{\Omega_i} \right)$$

The entry exerts a positive effect on project i output if and only if:

$$\begin{aligned} q_i^{*AE} > q_i^* &\Leftrightarrow \frac{V}{\hat{z}} \left(1 - \frac{1}{\hat{z}} \frac{\Gamma_i}{\Omega_i} \right) > \frac{V}{\bar{z}} \left(1 - \frac{1}{\bar{z}} \frac{\Gamma_i}{\Omega_i} \right) \\ &\Leftrightarrow \frac{1}{\hat{z}} - \frac{1}{\bar{z}} > \left(\frac{1}{\hat{z}^2} - \frac{1}{\bar{z}^2} \right) \frac{\Gamma_i}{\Omega_i} \end{aligned}$$

Since $\frac{1}{\hat{z}} > \frac{1}{\bar{z}}$ (see below) we finally obtain

$$q_i^{*AE} > q_i^* \Leftrightarrow \frac{\Gamma_i}{\Omega_i} < \frac{\hat{z}\bar{z}}{\hat{z} + \bar{z}}.$$

To show that $\frac{1}{\hat{z}} > \frac{1}{\bar{z}}$, let's consider separately the case $r = g + 1$ and $r \leq g$.

If $r = g + 1$:

$$\begin{aligned} \frac{1}{\hat{z}} > \frac{1}{\bar{z}} &\Leftrightarrow \frac{r-1}{\sum_{l=1}^r \frac{\Gamma_l}{\Omega_l}} > \frac{g-1}{\sum_{l=1}^g \frac{\Gamma_l}{\Omega_l}} \\ &\Leftrightarrow \frac{g}{\sum_{l=1}^g \frac{\Gamma_l}{\Omega_l} + \frac{\Gamma_{new}}{\Omega_{new}}} > \frac{g-1}{\sum_{l=1}^g \frac{\Gamma_l}{\Omega_l}} \\ &\Leftrightarrow \frac{\Gamma_{new}}{\Omega_{new}} < \bar{z} \end{aligned}$$

which corresponds to the condition such as the new project is active. Therefore $\frac{1}{\hat{z}} > \frac{1}{\bar{z}}$ if $r = g + 1$.

If $r \leq g$

r projects compete in the contest: the new project and the $r - 1$ with the lowest values

of Γ_i/Ω_i .

$$\begin{aligned} \frac{1}{\bar{z}} > \frac{1}{z} &\Leftrightarrow \frac{r-1}{\sum_{l=1}^{r-1} \frac{\Gamma_l}{\Omega_l} + \frac{\Gamma_{new}}{\Omega_{new}}} > \frac{g-1}{\sum_{l=1}^g \frac{\Gamma_l}{\Omega_l}} \\ &\Leftrightarrow (r-1) \left(\sum_{l=1}^{r-1} \frac{\Gamma_l}{\Omega_l} + \sum_{l=r}^g \frac{\Gamma_l}{\Omega_l} \right) > (g-1) \left(\sum_{l=1}^{r-1} \frac{\Gamma_l}{\Omega_l} + \frac{\Gamma_{new}}{\Omega_{new}} \right) \end{aligned} \quad (2.24)$$

Since projects r, \dots, g are inactive in the new contest, we have:

$$\frac{\Gamma_k}{\Omega_k} \geq \frac{1}{r-1} \left(\sum_{l=1}^{r-1} \frac{\Gamma_l}{\Omega_l} + \frac{\Gamma_{new}}{\Omega_{new}} \right) \quad \forall k = r, \dots, g$$

Given that $g-r+1$ projects become inactive in the new contest and using the condition of non-participation just above we deduce that:

$$\begin{aligned} (r-1) \left(\sum_{l=1}^{r-1} \frac{\Gamma_l}{\Omega_l} + \sum_{l=r}^g \frac{\Gamma_l}{\Omega_l} \right) &> (r-1) \left[\sum_{l=1}^{r-1} \frac{\Gamma_l}{\Omega_l} + (g-r+1) \frac{1}{r-1} \left(\sum_{l=1}^{r-1} \frac{\Gamma_l}{\Omega_l} + \frac{\Gamma_{new}}{\Omega_{new}} \right) \right] \\ &> (r-1) \left(\sum_{l=1}^{r-1} \frac{\Gamma_l}{\Omega_l} \right) + (g-r+1) \left(\sum_{l=1}^{r-1} \frac{\Gamma_l}{\Omega_l} + \frac{\Gamma_{new}}{\Omega_{new}} \right) \\ &> g \left(\sum_{l=1}^{r-1} \frac{\Gamma_l}{\Omega_l} \right) + (g-r+1) \frac{\Gamma_{new}}{\Omega_{new}} \\ &> (g-1) \left(\sum_{l=1}^{r-1} \frac{\Gamma_l}{\Omega_l} + \frac{\Gamma_{new}}{\Omega_{new}} \right) + \left(\sum_{l=1}^{r-1} \frac{\Gamma_l}{\Omega_l} - (r-2) \frac{\Gamma_{new}}{\Omega_{new}} \right) \end{aligned}$$

Remark that the first term of the last expression corresponds to the right hand side of (2.24). We can show that the second term is necessarily positive:

$$\sum_{l=1}^{r-1} \frac{\Gamma_l}{\Omega_l} - (r-2) \frac{\Gamma_{new}}{\Omega_{new}} > 0 \Leftrightarrow \frac{\Gamma_{new}}{\Omega_{new}} < \frac{1}{r-2} \left(\sum_{l=1}^{r-1} \frac{\Gamma_l}{\Omega_l} \right)$$

which corresponds to the condition of participation of the new project when $r-1$ already compete. Therefore condition (2.24) is necessarily realized. We conclude that $\frac{1}{z} > \frac{1}{\bar{z}}$.

8.7 Proof of Proposition 14

The utility function of subteam j in project i is:

$$u_{ij} = \frac{q_i}{Q} v_{ij} - x_{ij} \quad (2.25)$$

Each subteam j maximizes his own utility with respect to x_{ij} . We obtain the first order conditions $\forall j = 1, \dots, m_i$:

$$\Omega_i \left(\sum_{s=1}^{m_i} \omega_{is} x_{is}^{\sigma_i} \right)^{\frac{1}{\sigma_i} - 1} \omega_{ij} x_{ij}^{\sigma_i - 1} \frac{\sum_{l \neq i}^G q_l}{Q^2} v_{ij} = 1 \quad (2.26)$$

We deduce that for two fields of expertise j and t involved in the same research project i :

$$x_{it} = \left(\frac{\omega_{it} v_{it}}{\omega_{ij} v_{ij}} \right)^{\frac{1}{1 - \sigma_i}} x_{ij}$$

We obtain

$$\sum_{s=1}^{m_i} \omega_{is} x_{is}^{\sigma_i} = \frac{x_{ij}^{\sigma_i}}{(\omega_{ij} v_{ij})^{\frac{\sigma_i}{1 - \sigma_i}}} \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1 - \sigma_i}} v_{is}^{\frac{\sigma_i}{1 - \sigma_i}} \right) \quad (2.27)$$

Taking the exponent $\frac{1}{\sigma_i} - 1$ on both sides, we get:

$$\left(\sum_{s=1}^{m_i} \omega_{is} x_{is}^{\sigma_i} \right)^{\frac{1}{\sigma_i} - 1} = \frac{x_{ij}^{1 - \sigma_i}}{\omega_{ij} v_{ij}} \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1 - \sigma_i}} v_{is}^{\frac{\sigma_i}{1 - \sigma_i}} \right)^{\frac{1}{\sigma_i} - 1}$$

Replacing in the first order condition (2.26) leads to

$$\frac{\sum_{l \neq i}^G q_l}{\left(\sum_{l=1}^G q_l \right)^2} = \frac{\Upsilon_i}{\Omega_i}, \quad (2.28)$$

with

$$\Upsilon_i = \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1 - \sigma_i}} v_{is}^{\frac{\sigma_i}{1 - \sigma_i}} \right)^{1 - \frac{1}{\sigma_i}}.$$

Project i best response given the strategies of other groups is

$$q_i = \max \left(0, \sqrt{Q_{/i} \frac{\Omega_i}{\Upsilon_i}} - Q_{/i} \right). \quad (2.29)$$

Summing (2.28) $\forall i = 1, \dots, g$, with g the number of participating projects, we get

$$Q^* = \frac{1}{Z}, \quad (2.30)$$

with

$$\bar{z} = \frac{1}{g-1} \sum_{j=1}^g \frac{\Upsilon_j}{\Omega_j}.$$

From (2.29), project i participates to the contest if

$$\frac{\Upsilon_i}{\Omega_i} > Q_{/i}.$$

Using (2.30) then (2.28), we obtain

$$\frac{\Upsilon_i}{\Omega_i} < \bar{z}.$$

Given that g is the number of active projects, and according to the ranking of projects in (2.17), the first inactive project $g+1$ has a value of $\Upsilon_{g+1}/\Omega_{g+1}$ such that

$$\frac{\Upsilon_{g+1}}{\Omega_{g+1}} \geq \bar{z}.$$

Using (2.28) then (2.27), we finally obtain q_i^* and x_{is}^* .

8.8 Properties of Υ_i

Derivative of Υ_i with respect to σ_i .

$$\begin{aligned}
\frac{\partial \Upsilon_i}{\partial \sigma_i} &= \frac{\Upsilon_i}{\sigma_i^2} \left[\ln \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{\sigma_i}{1-\sigma_i}} \right) - \frac{\sigma_i}{1-\sigma_i} \frac{\sum_{s=1}^{m_i} (\ln(\omega_{is} v_{is})) \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{\sigma_i}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{\sigma_i}{1-\sigma_i}}} \right] \\
&= \frac{\Upsilon_i}{\sigma_i^2} \left[\sum_{k=1}^{m_i} \frac{\omega_{ik}^{\frac{1}{1-\sigma_i}} v_{ik}^{\frac{\sigma_i}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{\sigma_i}{1-\sigma_i}}} \ln \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{\sigma_i}{1-\sigma_i}} \right) - \frac{\sigma_i}{1-\sigma_i} \sum_{k=1}^{m_i} \frac{\omega_{ik}^{\frac{1}{1-\sigma_i}} v_{ik}^{\frac{\sigma_i}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{\sigma_i}{1-\sigma_i}}} (\ln(\omega_{ik} v_{ik})) \right] \\
&= \frac{\Upsilon_i}{\sigma_i^2} \sum_{k=1}^{m_i} \left[\frac{\omega_{ik}^{\frac{1}{1-\sigma_i}} v_{ik}^{\frac{\sigma_i}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{\sigma_i}{1-\sigma_i}}} \left(\ln \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{\sigma_i}{1-\sigma_i}} \right) - \frac{\sigma_i}{1-\sigma_i} \ln(\omega_{ik} v_{ik}) \right) \right] \\
&= -\frac{\Upsilon_i}{\sigma_i^2} \sum_{k=1}^{m_i} \left[\frac{\omega_{ik}^{\frac{1}{1-\sigma_i}} v_{ik}^{\frac{\sigma_i}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{\sigma_i}{1-\sigma_i}}} \left(\ln \left(\omega_{ik}^{\frac{1}{1-\sigma_i}-1} v_{ik}^{\frac{\sigma_i}{1-\sigma_i}} \right) - \ln \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{\sigma_i}{1-\sigma_i}} \right) \right) \right] \\
&= -\frac{\Upsilon_i}{\sigma_i^2} \sum_{k=1}^{m_i} \left[\frac{\omega_{ik}^{\frac{1}{1-\sigma_i}} v_{ik}^{\frac{\sigma_i}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{\sigma_i}{1-\sigma_i}}} \ln \left(\frac{\omega_{ik}^{\frac{1}{1-\sigma_i}} v_{ik}^{\frac{\sigma_i}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{\sigma_i}{1-\sigma_i}}} \right) \right]
\end{aligned}$$

Let:

$$P_{ik} = \frac{\omega_{ik}^{\frac{1}{1-\sigma_i}} v_{ik}^{\frac{\sigma_i}{1-\sigma_i}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{\sigma_i}{1-\sigma_i}}}$$

Rewriting the derivative, we finally obtain:

$$\frac{\partial \Upsilon_i}{\partial \sigma_i} = -\frac{\Upsilon_i}{\sigma_i^2} \sum_{k=1}^{m_i} P_{ik} \ln \left(\frac{P_{ik}}{\omega_{ik}} \right) < 0$$

for the same reasons than those given in Appendix 8.1.

Derivative of Υ_i with respect to ω_{ij} .

Derivating Υ_i with respect to ω_{ij} , we get:

$$\frac{\partial \Upsilon_i}{\partial \omega_{ij}} = -\frac{1}{\sigma_i} \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{\sigma_i}{1-\sigma_i}} \right)^{-\frac{1}{\sigma_i}} \omega_{ij}^{\frac{1}{1-\sigma_i}-1} v_{ij}^{\frac{\sigma_i}{1-\sigma_i}}$$

Similarly for another field of expertise $t > j$ we have:

$$\frac{\partial \Upsilon_i}{\partial \omega_{it}} = -\frac{1}{\sigma_i} \left(\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} v_{is}^{\frac{\sigma_i}{1-\sigma_i}} \right)^{-\frac{1}{\sigma_i}} \omega_{it}^{\frac{1}{1-\sigma_i}-1} v_{it}^{\frac{\sigma_i}{1-\sigma_i}}$$

It follows that

$$\frac{\partial \Upsilon_i}{\partial \omega_{it}} > \frac{\partial \Upsilon_i}{\partial \omega_{ij}}$$

if and only if

$$-\frac{1}{\sigma_i} \omega_{it}^{\frac{1}{1-\sigma_i}-1} v_{it}^{\frac{\sigma_i}{1-\sigma_i}} > -\frac{1}{\sigma_i} \omega_{ij}^{\frac{1}{1-\sigma_i}-1} v_{ij}^{\frac{\sigma_i}{1-\sigma_i}}$$

In Appendix 8.1, we have already proved that for $\omega_{ij} > \omega_{it}$:

$$-\frac{1}{\sigma_i} \omega_{it}^{\frac{1}{1-\sigma_i}-1} > -\frac{1}{\sigma_i} \omega_{ij}^{\frac{1}{1-\sigma_i}-1}$$

Thus if

$$v_{it}^{\frac{\sigma_i}{1-\sigma_i}} > v_{ij}^{\frac{\sigma_i}{1-\sigma_i}},$$

then

$$\frac{\partial \Upsilon_i}{\partial \omega_{it}} > \frac{\partial \Upsilon_i}{\partial \omega_{ij}}.$$

and any Lorenz-improvement in the ω_i -vector increases Υ_i . This occurs if:

$$v_{it} > v_{ij} \text{ if } \sigma_i > 0$$

$$v_{ij} > v_{it} \text{ if } \sigma_i < 0$$

Do complex projects lead to higher citation impact? An empirical application to interdisciplinary research¹.

1 Introduction

In the general introduction, we have presented one form of complexity of research project, called interdisciplinarity, which requires to combine knowledge from different research areas to produce scientific output. In the second chapter, we have defined a complexity indicator of the production function. The aim of this chapter is to test this index on a database of scientific publications, and to link the results with the main question of this thesis: the decline of the research productivity over time. To do so, we use an original dataset of nearly four hundred thousand authorship participations of approximately thirty thousand French professors and researchers over the period 1999-2013.

Nevertheless several problems arise. The first one is that efforts exerted by teams are unobservable so we can't calculate the productivity at the project level. However we have found in chapter 2 that the complexity index behaves similarly on the scientific output for which we have proxies. The second problem is that many aspects of our complexity index are unobservable: the effort intensity coefficients (vector ω) and the elasticity of substitution σ . We rewrite this indicator as a function of the contributions of each discipline to the output, and show that it belongs to the family of Hill diversity indexes. This is of interest since this reduces the number of unobservable dimensions. Moreover the obtained rewriting is a standard measure of interdisciplinarity.

¹This paper is co-authored with Nicolas Carayol.

One major finding of this paper is that interdisciplinarity of research projects is increasing over time. We conduct several tests to confirm this assertion. We link this observation to the decline of research productivity and open the debate in the last part of this chapter.

The rest of the paper is organized as follows. Section 2 explains the transition from the model to data. Section 3 reviews the literature on the relationship between interdisciplinarity and impact. Section 4 describes the database and the variables. Section 5 contains the regression results. Section 6 presents robustness checks. Section 7 aims at showing that interdisciplinarity is increasing over time. Section 8 discusses the results. Finally section 9 concludes. Proofs and tables of regression results are given in Appendix.

2 From the model to the data

The main problem to empirically test our model comes from the impossibility to estimate for every project the parameter σ_i of the knowledge production function

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

since efforts spent in each discipline x_{is} and the parameters $\omega_{is} \forall s$ are unobservable. Subsequently, we are unable to compute our complexity index Γ_i

$$\Gamma_i = \left(\sum_{s=1}^m \omega_{is}^{\frac{1}{1-\sigma_i}} \right)^{1-\frac{1}{\sigma_i}}. \quad (3.1)$$

In the following, we aim at finding a rewriting of Γ_i as a function of variables for which we have proxies in our dataset.

Definition 2 *Let define*

$$H^\beta(b) = \left(\sum_{s=1}^m (b_s)^\beta \right)^{\frac{1}{1-\beta}}$$

the Hill index of order $\beta \in [0; +\infty[$ of vector $b = (b_1, \dots, b_m)$, with $b_s \geq 0 \forall s = 1, \dots, m$ and $\sum_{s=1}^m b_s = 1$.

This general index (Hill, 1973) encompasses series of diversity indexes, each one corresponding to a particular value of β :

(1) the Berger-Parker index when $\beta \rightarrow +\infty$:

$$H^\infty(b) = \frac{1}{\max_{s=1, \dots, m} (b_s)};$$

(2) the Simpson index when $\beta = 2$:

$$H^2(b) = \frac{1}{\sum_{s=1}^m (b_s^2)};$$

(3) the exponential of the Shannon-Wiener index when $\beta = 1$:

$$H^1(b) = e^{-\sum_{s=1}^m b_s \ln b_s};$$

(4) the number of fields of expertise when $\beta = 0$:

$$H^0(b) = m.$$

Looking at (3.1), it follows immediately that:

Proposition 15 Γ_i is the Hill index of order $\epsilon_i = \frac{1}{1-\sigma_i}$ of vector ω_i .

Note that β and the elasticity of substitution $\epsilon_i = \frac{1}{1-\sigma_i}$ are defined on the same interval $[0; +\infty[$. Table 3.1 gives the correspondance between σ_i and the associated diversity index of ω_i .

σ_i	Diversity index
1	Berger-Parker index
0.5	Simpson index
0	Exponential of the Shannon-Wiener index
$-\infty$	Number of fields of expertise

Table 3.1: Table of correspondance between σ_i and the diversity indexes of ω_i

Given that Γ_i is a function of technical parameters ω_{is} , $s = 1, \dots, m_i$, Γ_i is difficult to appreciate in concrete terms. We rewrite it as a function of the contributions of each field of expertise to the knowledge production since we have proxies for this variable.

Definition 3 Let define

$$\lambda_{ij} = \frac{\omega_{ij} x_{ij}^\sigma}{\sum_{s=1}^{m_i} \omega_{is} x_{is}^\sigma}. \quad (3.2)$$

the contribution of each research area j to the knowledge production of research project i , and $\lambda_i = (\lambda_{i1}, \dots, \lambda_{im_i})$ the vector of those contributions.

At equilibrium, λ_{ij}^* is given by

$$\lambda_{ij}^* = \frac{\omega_{ij}^{\frac{1}{1-\sigma}}}{\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma}}}. \quad (3.3)$$

First remark that λ_{ij}^* only depends on the characteristics of the knowledge production function of project i . Second, λ_{ij}^* is also equal to the share of effort in the field of expertise j over the total effort spent in project i at equilibrium (x_{ij}^*/X_i^*).

Further, we show in Appendix 10.1 that it is possible to write Γ_i as:

$$\Gamma_i = \left(\sum_{s=1}^{m_i} \lambda_{is}^{*1-\sigma_i} \right)^{\frac{1}{\sigma_i}}. \quad (3.4)$$

It follows immediately that:

Proposition 16 Γ_i is the Hill index of order $1 - \sigma_i$ of vector λ_i .

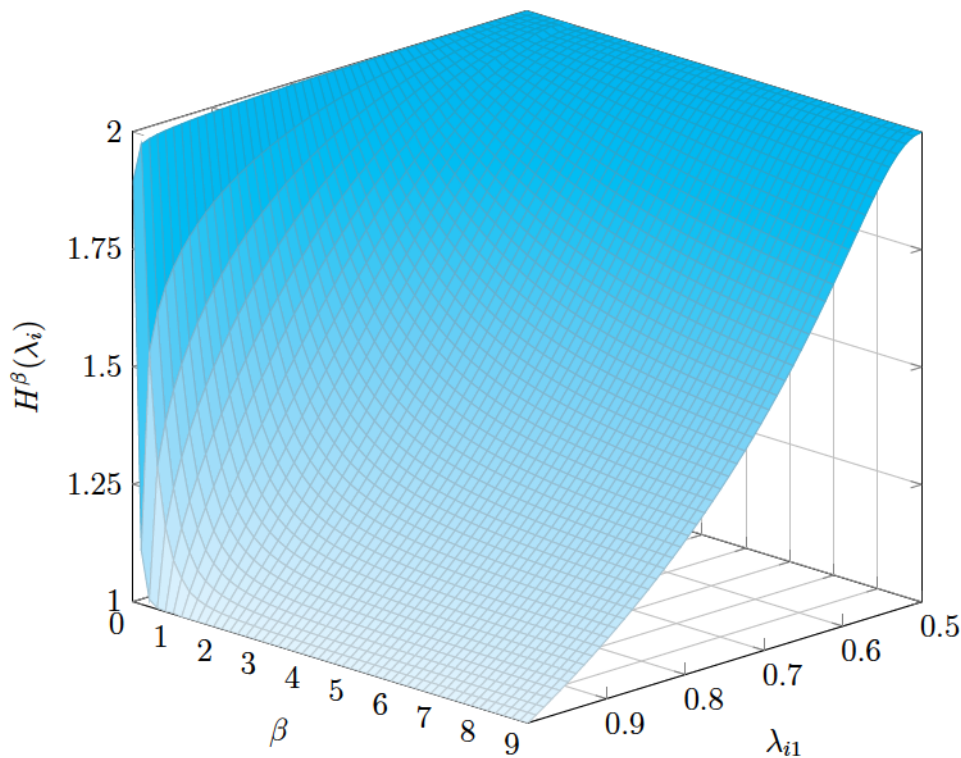
Our index of complexity Γ_i coincides at equilibrium with the diversity indicator of the contributions of the fields of expertise to the production. Since we have proxies for λ_{is} thus only σ_i remains unknown. In other words, given that we observe vector λ_i^* , the problem is reduced to the order of the Hill index.

Proposition 17 If $m_i = m \leq 2$, then the ordering of any two vectors λ_1 and λ_2 is the same according to H_i^β , $\forall \beta \geq 0$.

Proof is given in Appendix 10.2, and we illustrate Proposition 17 in Figure 3.1. If $m_i = 2$ then $\lambda_{i2} = 1 - \lambda_{i1}$, thus $H^\beta(\lambda_i)$ is a function of λ_{i1} and β . It is enough to show that $H^\beta(\lambda_i)$ is decreasing with λ_{i1} when $\lambda_{i1} \in [0.5, 1]$ whatever the value of β .

Corollary 1 If the elasticity of substitution is identical for all the projects ($\sigma_i = \sigma$) and the number of fields $m_i \leq 2$ for all i , then the ranking of projects according to Γ_i is the same as the one obtained with H_i^β for any $\beta \geq 0$.

In other words, considering the subsample of projects involving less than two disciplines, and assuming that all of them have the same elasticity of substitution, no matter

Figure 3.1: Hill index when $m_i = 2$

the choice of the order of the Hill index β , the ranking of projects according to β is exactly the same as this obtained with $1 - \sigma$.

Nevertheless this condition may considerably restrict the sample of papers in our database and only allows us to draw limited conclusions in the relation between complexity and impact. To deal with the entire database, we calculate Spearman's rank correlation coefficients between different disciplinary diversity indexes. Positive values close to 1 indicate that the ranking of projects according to their disciplinary diversity doesn't substantially change whatever the value of β . However Hill indexes may be negatively, non- or weakly correlated with each other, like in the numerical illustration given in Example 7. This may lead to opposite signs for the coefficients associated to the diversity indexes in our regressions. If such a situation were to occur, since we are unable to select the correct index ($\beta = 1 - \sigma$), we can't verify the validity of our model with the entire database. In this situation, we are only able to partially conclude with the subset of projects involving at most two disciplines.

Example 7 Consider the set of projects given in Table 3.2 for which we report the contributions of each discipline to the outcome in columns 2 to 5.

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

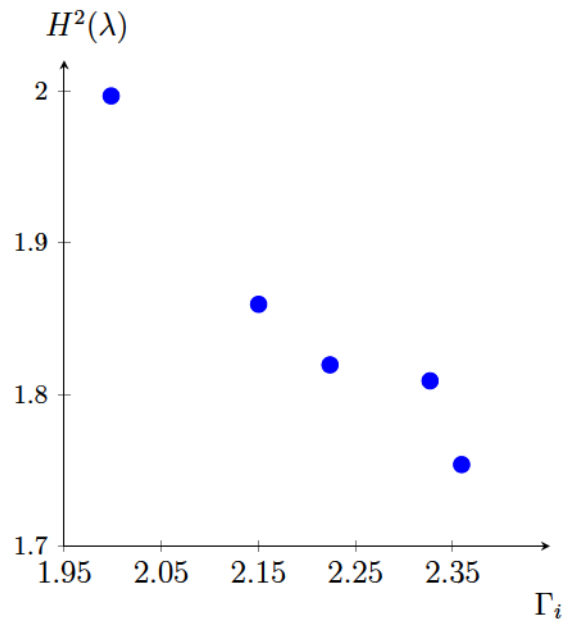
Table 3.2: Contributions of each discipline λ_{ij} to project i 

Figure 3.2: Negative correlation between the Simpson and our complexity indexes.

Suppose that all the projects have the same elasticity of substitution between efforts such that $\sigma_i = 0$, implying that the complexity index of project i (Γ_i) is the exponential of the Shannon-Wiener index ($H^1(\lambda_i)$). If we compute the Simpson index ($H^2(\lambda_i)$) for each project i , we observe in Figure 3.2 that this measure is negatively correlated with the complexity index. This comes from the fact that the lower β is, the more underrepresented disciplines in a project are taken into account by the measure. In our example, projects are ranked by ascending order of the contribution of the first discipline. The Shannon-Wiener index, more sensitive to lower weights than the Simpson one, is increasing with i since there is greater evenness in the contributions of disciplines 2 to 4 while that of discipline 1 is similar for projects 2 to 5. On the contrary, the Simpson index is more sensitive to

the highest weight. Since λ_{i1} is increasing with i , projects are ranked by descending order of the Simpson index. In this example, choosing a random order for the Hill index can lead to opposite results when testing the relation between interdisciplinarity and impact.

3 Literature review

Previous papers have already studied the relation between interdisciplinarity and impact without finding clearcut results. One of the first contributions to this literature comes from [Rinia et al. \(2001\)](#). They propose a descriptive statistical study of interdisciplinarity concerning 200 academic physics research programs in the Netherlands. They assign each paper the research subfields of the journal in order to obtain for a given research program the distribution of publications over research subfields. They define the level of interdisciplinarity of each physics research program as the percentage of non-physics papers, and calculate Spearman rank-correlation coefficients between interdisciplinarity and variables of impact. They highlight a negative or an absence of significant correlation between interdisciplinarity and several measures of impact (citations, number of publications). [Levitt and Thelwall \(2008\)](#) consider multidisciplinary papers as those published in journals referenced in more than one discipline. Conducting statistical tests (correlations, mean comparisons), they observe in social sciences no significant differences between monodisciplinary and multidisciplinary research articles concerning the number of citations. They even note that in life sciences, health sciences and physical sciences, papers involving a unique academic discipline have twice as many citations as multidisciplinary ones.

Instead of using a unique indicator to measure interdisciplinarity, [Wang et al. \(2015\)](#) combine three variables (variety, balance, disparity) and estimate a Poisson model with journal fixed effects to analyse the effect of multidisciplinary on the impact of academic papers (citation and citation speed). Following [Porter and Rafols \(2009\)](#), they operationalise variety as the number of disciplines, balance as the evenness of the distribution of disciplines and disparity as the degree to which the disciplines are different. Using a 3-year citation window, they obtain that the number of citations decreases with variety and disparity whereas balance has no significant effect. Using a 13-year citation window, variety and disparity have a positive effect while balance impacts negatively the number of citations. The authors argue that the best strategy to increase impact in the long-term

is to have one disciplinary core and simultaneously borrow knowledge from some other disciplines, rather than drawing knowledge evenly from multiple disciplines. In a very similar study, [Yegros-Yegros et al. \(2015\)](#) estimate a Tobit regression model to investigate the relation between interdisciplinarity and impact. Using a 5-year citation window, they note a positive effect of variety on the number of citations whereas a higher balance and disparity decreases the impact of academic papers. They emphasize a U-inverted relationship between these independent variables and the number of citations, but some of them are questionable. Looking at Figure 2b (page 14), no paper has a value of balance between 0 and 0.4, which does not prevent the authors from concluding to an artificial increasing relation in this interval between balance and impact. The authors employ the Rao-Stirling diversity index to test simultaneously the effects of variety, balance and disparity. However they found no evidence of a significant relationship between interdisciplinarity and impact.

[Leahey et al. \(2017\)](#) adopt path analytic techniques to study the impact of interdisciplinarity on productivity (number of papers published over one researcher's career) and visibility (number of citations per paper). They point out that increasing by 10% the indicator of interdisciplinarity decreases the number of publications by 7.7% over a researcher's career while it increases the number of citations of a paper by 6.9%. The authors confirm these results by estimating fixed-effect regression model at the person-year level (for productivity) and at the paper level (for visibility). Among other results, they mention that producing interdisciplinary research is associated with a higher variability (measured as the standard deviation) in the number of citations than monodisciplinary research, meaning that scientists who publish more interdisciplinary research are more likely to produce both frequently-cited and rarely-cited works.

This profusion of different results leads us to question the existence of a possible omitted variable that would blur the negative relationship between interdisciplinarity and impact. Our belief is that previous estimations insufficiently control for the characteristics of the teams producing cross-disciplinary research. In particular, the literature indicates that senior researchers are more likely to engage in such a challenge ([Rhoten and Parker, 2004](#); [Carayol and Thi, 2005](#)), as they are more experienced and benefit from a greater professional freedom. An overrepresentation of this category of scientists in the sample of cross-disciplinary researchers could explain why some studies observe a positive effect of interdisciplinarity on the number of citations. Indeed senior researchers often benefit from

a great visibility, a high number of past publications and a dense network of co-authors. The presumed negative effect of cross-disciplinary research could be counterbalanced by the positive effect of these individual characteristics, introducing noise in the estimation of the relation. In this paper, we propose to partially control for the characteristics of the team members (French authors) to avoid such an ambiguity.

A second possible explanation is the choice of the order of the Hill index. In the previous section, we have shown that this may lead to opposite results in the regressions. In this paper we systematically test our relation between interdisciplinarity and impact with four indexes: the Berger-Parker, the Simpson, the exponential of the Shannon-Wiener indexes, and the number of disciplines. Many previous studies only test this relation with a unique index, which could explain the profusion of different results.

In the next section, we present our database.

4 Data and Methods

4.1 Data set

Data collection starts similarly than [Carayol and Lanoë \(2017\)](#) with a list of all researchers and professors associated with one laboratory accredited by the French Ministry of Higher Education and Research around the year 2010. This list contains informations on nearly fifty thousand tenured persons who are full or assistant professors, assistant researchers or research directors. Tenured researchers and professors who are not associated to a laboratory or research unit labeled by the ministry are excluded from the data set². However we estimate that the resulting sample represents at least two thirds of the nation's active academic research tenured workforce.

The collection of publication data of this paper actualizes [Carayol and Lanoë \(2017\)](#). We have collected the scientific articles, letters and reviews published by at least a French researcher in journals listed in the Web of Science (WoS), but for a larger fifteen years period (1999-2013). The same disambiguation technique is also applied based on a \hat{A} seed + expand \hat{A} methodology ([Reijnhoudt et al., 2014](#)). Some researchers could not be properly disambiguated since they have too common surnames. Some others have no publication record that can be attached to them, particularly in Social and Human Sciences.

²The French research system is organized in research laboratories, see [Carayol and Matt \(2004\)](#)

In both cases we exclude them from the database. At the end of this process, our sample contains 27,515 researchers and their 777,278 participations in 401,563 distinct articles published from 1999 to 2013. We drop 5,795 papers which belongs to a tote category, 532 papers for which the disciplinary profile is not correctly specified ($\sum_{s=1}^{m_i} \lambda_{is} \neq 1$), and 43 articles having no collected reference. Our final dataset comprises 395,193 papers.

4.2 Variables

4.2.1 Dependent variables

Number of citations. Our main dependent variable is the number of citations obtained by a paper in a fixed window of three years after its date of publication. This measure is widely used in scientometrics to assess research impact of journal articles. To justify their claims, researchers cite previous works on which their current paper builds. Given a paper, the received number of citations is a form of recognition by the scientific community often interpreted as the “*quality*” of an idea. Nevertheless this approximation must be done with caution. The number of citations depends on the disciplines involved in the paper. Furthermore researchers can negatively cite a journal article to refute an idea. The distinction between citations to acknowledge and those criticizing previous work is not made; both being positively considered in the measure except in some papers (Catalini et al., 2015). The practice of self-citation is widespread and influences the number of citations: if referring to author’s previous work can be relevant, researchers can self-cite their own papers for strategic reasons, like increasing their h-index (Bartneck and Kokkelmans, 2010). The citation coverage differs from one database to another (Web of Science, Scopus, Google Scholar) according to the number of indexed journals, the research areas and the beginning date of data collection. Journals limit the length of submissions which prevents authors to cite their sources (Seglen, 1997). The number of citations is not only affected by measurement problems, but also by the behavior of the scientific community. Wang et al. (2017) argue that “*bibliometric measures are biased against novel research*”. Pasterkamp et al. (2007) show that the number of citations is influenced by the geographic origin of the research publication.

Table 3.3 presents some descriptive statistics on our sample of papers according to their main discipline. Articles published in journals of hard sciences constitute the bulk of our database contrary to those in Social and Human Sciences. The average number

of citations per paper three years after the date of publication varies a lot across fields: more than 9 in Basic Biology versus 2 in Mathematics for instance. Besides the number of citations per paper is overdispersed whatever the discipline, with a mean of 6.74 and a standard deviation of 16.8. In Table 3.4, we provide the cumulative frequency distribution of the number of citations. A remarkable fact is the proportion of papers obtaining 0 citation, which is close to 18.4% in the full database but the double in some research areas as in Social Sciences for instance.

Top 10% and Top 5%. We also consider a dummy equal to 1 if the paper belongs to the top 10% (respectively top 5%) most cited articles in its main discipline in a 3-year window, 0 otherwise.

	Observations	Min	Median	Max	Mean	S.D.
Applied Biology - Ecology	16290	0	4	259	6.45	9.18
Basic Biology	65042	0	5	647	9.09	14.1
Chemistry	64292	0	4	855	6.38	11.1
Engineering Sciences	32914	0	1	182	2.48	4.37
Human Sciences	3520	0	1	72	3.02	4.69
Mathematics	15182	0	1	108	2.00	3.01
Medical Research	93923	0	4	1108	7.90	19.1
Physical Sciences	61196	0	3	3264	6.26	27.1
Sciences of the Universe	40274	0	4	917	7.52	14.0
Social Sciences	2560	0	1	44	1.96	3.11
Total	395193	0	3	3264	6.74	16.8

Table 3.3: Number of citations per paper according to the main discipline (all the database)

Number of citations	Applied Biology Ecology	Basic Biology	Chemistry	Engineering Sciences	Human Sciences	Mathematics	Medical Research	Physical Sciences	Sciences of the Universe	Social Sciences	Total
0	11.74	8.830	13.85	37.51	33.12	35.03	18.63	20.89	14.70	38.20	18.36
1	23.88	17.56	26.27	55.95	51.42	58.15	31.19	34.44	25.81	60.35	31.27
2	35.60	26.53	37.95	68.21	62.50	72.66	41.17	45.56	35.99	74.41	42.03
3	45.86	35.15	47.73	76.85	72.27	81.96	49.43	54.85	44.92	82.46	50.99
4	54.71	42.98	56.13	82.74	78.44	87.60	56.33	62.09	52.66	87.89	58.36
5	62.01	50.01	63.27	87.06	82.84	91.40	62.12	68.08	59.31	91.99	64.53
6	67.88	56.19	68.90	90.13	86.19	94.03	66.86	73.18	64.91	94.22	69.59
7	72.64	61.59	73.75	92.36	88.98	95.57	70.84	77.05	69.53	95.62	73.76
8	76.67	66.24	77.58	93.94	91.25	96.62	74.13	80.32	73.51	96.48	77.20
9	80.03	70.21	80.86	95.26	92.90	97.42	77.07	83.02	76.85	97.30	80.15
10	82.81	73.63	83.48	96.10	94.09	98.09	79.61	85.12	79.68	97.58	82.58

Table 3.4: Cumulative frequency distribution (%) of the number of citations per paper according to the main discipline (all the database)

Number of disciplines	Observations	Frequency	Cumulative Frequency
1	149403	37.81	37.81
2	143748	36.37	74.18
3	72633	18.38	92.56
4	22189	5.61	98.17
5	5619	1.42	99.60
6	1320	0.33	99.93
7	230	0.06	99.99
8	42	0.01	100.00
9	7	0.00	100.00
10	1	0.00	100.00
Total	395192	100.00	100.00

Table 3.5: Number of disciplines per paper (OST nomenclature)

4.2.2 Independent variables

Following the literature, we compute interdisciplinary indexes of a given paper according to the disciplinary profile of its reference list. Such an operation first requires to identify the disciplines. To this end, we mainly consider the disciplinary classification from the OST³ which is characterized by a high level of aggregation (10 disciplines). This nomenclature is suitable to test the predictions of our model since 74% of the papers from the database span over at most two research areas (Table 3.5). Besides the 10 disciplines of this classification remain unchanged over time allowing for temporal comparisons of interdisciplinary indexes. In the robustness checks, we also test our model with two other nomenclatures differing by their granularity : the NOWT⁴ classification with 35 research areas and the WoS subject categories with 252 disciplines. We give in Table 3.98 the correspondance between these nomenclatures.

To estimate the disciplinary profile of a paper, we attribute to each cited reference the disciplines of the journal in which it has been published. Thus for the reference list of a given paper, we obtain the distribution of disciplines with their associated weight (Example 8). To correctly determine this disciplinary background, we need to have a sufficiently long reference list. Many factors can affect this variable : citations practices are different between the disciplines and the type of document (article, conference proceeding, review). Working papers and book chapters are not taken into account in the reference list since they are not publications in journals. The reference coverage in the database is not perfect, particularly in Social and Human Sciences.

³Observatoire des Sciences et Techniques (Paris)

⁴Nederlands Observatorium voor Wetenschap en Technologie (Leiden)

Table 3.6 presents some summary statistics concerning the size of the reference list of papers according to their main discipline. In our database we are able to identify 7.23 cited references per paper on average, and at least 50% of the papers have more than 5 references. However in some academic disciplines (Social Sciences for instance), the mean number of references is too low to estimate properly the disciplinary profile of each paper. Since the type of document may be partly the cause, we give descriptive statistics in Table 3.7 to deal with this question. In particular, conference proceedings have less references than other types of documents. Following Yegros-Yegros et al. (2015), we consider papers with at least 4 references linked to a WoS subject category. This leads to a subsample of 251,479 articles. We present in Tables 3.8 and 3.9 the same descriptive statistics as in Tables 3.6 and 3.7 after dropping papers with less than 4 references. As expected, this reduces the weight of conference proceedings in our database. In our subsample, we are able to identify 10.3 cited references per article on average, with a median at 8 which seems to be sufficient to correctly determine the disciplinary profile of each paper. Then we compute four indicators of interdisciplinarity belonging to the Hill index family: the Berger-Parker ($\beta \rightarrow +\infty$), the Simpson ($\beta = 2$) and the exponential of the Shannon-Wiener ($\beta = 1$) indexes, and the number of disciplines ($\beta = 0$). Naturally it is not possible to test all the Hill indexes since their number is infinite. However the chosen values of β makes sense economically (see chapter 2) and are distributed on the whole interval $[0; +\infty[$. We need to take care when we test the number of disciplines since this variable can only take 10 values (every integer between 1 and 10). In our regressions, we successively integrate the number of disciplines as a continuous then a categorical variable. In the latter, we choose monodisciplinary papers (with 1 discipline) as the reference group.

For better readability, we will sometimes shorten “exponential of the Shannon-Wiener index” to “Shannon index”. In Figure 3.3, we present the distribution of interdisciplinarity according to the Berger-Parker, the Simpson and the Shannon indexes. For the number of disciplines, the information is available in Table 3.5. Whatever the index, the distribution is positively skewed.

	Observations	Min	Median	Max	Mean	S.D.
Applied Biology - Ecology	16290	1	7	133	8.68	7.94
Basic Biology	65042	1	7	249	9.13	8.48
Chemistry	64292	1	5	335	7.31	7.79
Engineering Sciences	32914	1	3	76	3.76	3.63
Human Sciences	3520	1	3	52	4.83	4.65
Mathematics	15182	1	3	55	3.60	3.05
Medical Research	93923	1	6	311	7.67	8.05
Physical Sciences	61196	1	4	847	5.82	7.59
Sciences of the Universe	40274	1	7	246	9.21	9.18
Social Sciences	2560	1	2	75	2.91	2.89
Total	395193	1	5	847	7.23	7.88

Table 3.6: Number of references per paper (all the database)

	Conference Proceeding	Journal Article	Letter	Total
Applied Biology - Ecology	0.60	93.3	6.10	100
Basic Biology	1.12	91.2	7.71	100
Chemistry	0.22	96.4	3.40	100
Engineering Sciences	0.10	99.0	0.94	100
Human Sciences	0.57	94.8	4.60	100
Mathematics	0.020	99.7	0.25	100
Medical Research	5.76	84.7	9.49	100
Physical Sciences	0.19	98.3	1.55	100
Sciences of the Universe	0.17	95.9	3.91	100
Social Sciences	0.078	98.5	1.41	100
Total	1.68	93.2	5.10	100

Table 3.7: Type of research document (all the database)

	Observations	Min	Median	Max	Mean	S.D.
Applied Biology - Ecology	11819	4	9	133	11.2	7.94
Basic Biology	49027	4	9	249	11.5	8.53
Chemistry	42403	4	8	335	10.1	8.31
Engineering Sciences	12439	4	6	76	7.00	4.13
Human Sciences	1709	4	7	52	8.05	4.86
Mathematics	5828	4	5	55	6.47	3.15
Medical Research	62717	4	8	311	10.5	8.49
Physical Sciences	35671	4	7	847	8.59	8.95
Sciences of the Universe	29182	4	9	246	12.0	9.40
Social Sciences	684	4	5	75	6.07	3.98
Total	251479	4	8	847	10.3	8.47

Table 3.8: Number of references per paper (papers with at least 4 references)

	Conference Proceeding	Journal Article	Letter	Total
Applied Biology - Ecology	0.47	91.9	7.67	100
Basic Biology	0.46	90.2	9.38	100
Chemistry	0.12	95.2	4.68	100
Engineering Sciences	0.032	98.1	1.85	100
Human Sciences	0.18	93.9	5.97	100
Mathematics	0	99.6	0.45	100
Medical Research	1.84	85.7	12.5	100
Physical Sciences	0.10	97.6	2.33	100
Sciences of the Universe	0.072	95.1	4.80	100
Social Sciences	0.15	96.6	3.22	100
Total	0.62	92.2	7.13	100

Table 3.9: Type of research document (papers with at least 4 references)

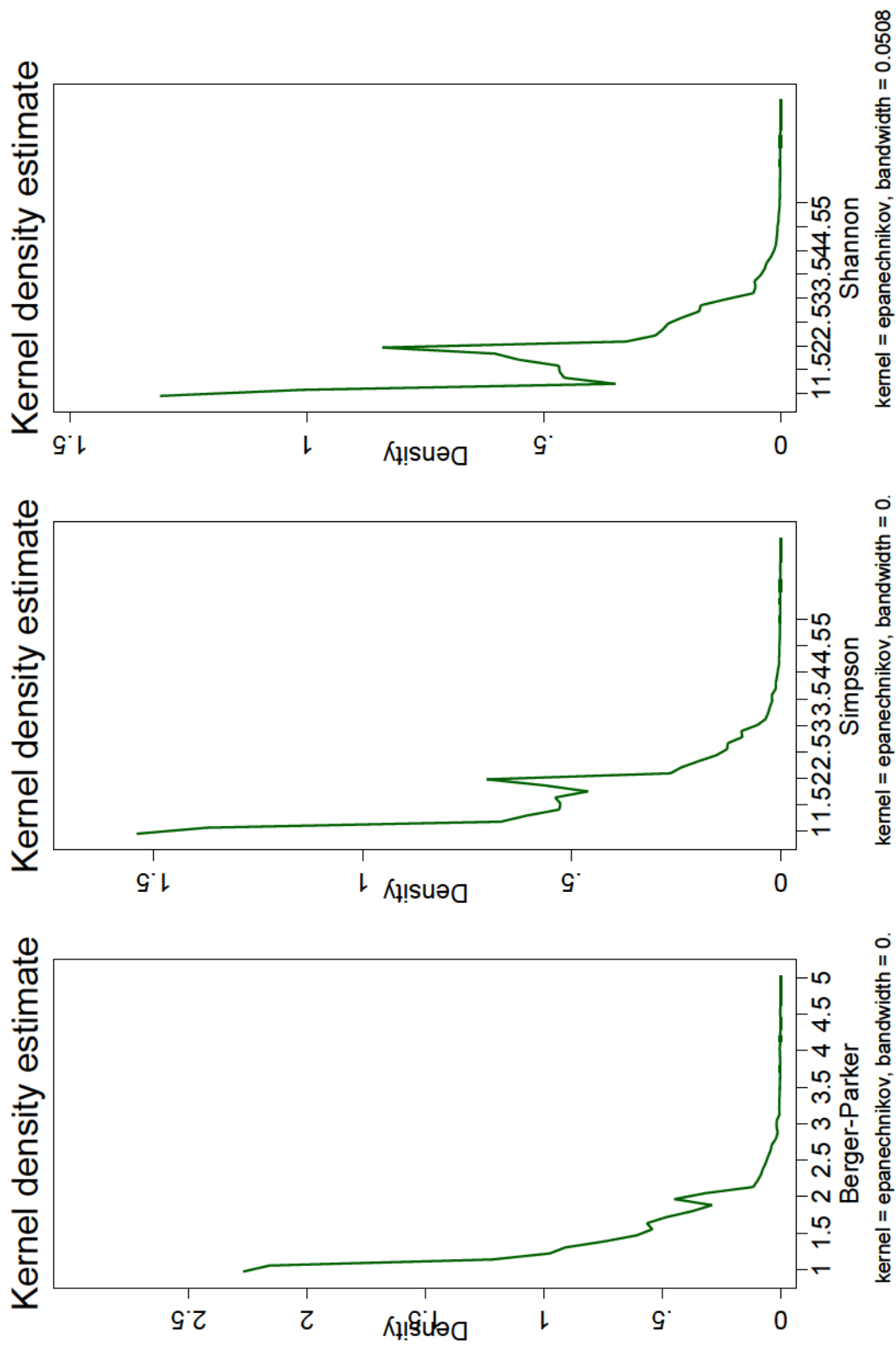


Figure 3.3: Distribution of Interdisciplinarity according to the Berger-Parker, the Simpson and the Shannon indexes (OST nomenclature, papers with at least 4 references)

Example 8 Consider a paper with a reference list of five articles, all of them published in academic reviews. We assign to each reference the disciplines of the corresponding journal like in Figure 3.4. In this example, the first reference is published in a journal dealing both with Economics and Mathematics, the second only in Economics and so forth. We sum these weights for each discipline (2.5 in Economics, 1.5 in Mathematics and 1 in Sociology) before dividing by the total sum (5) to obtain the contribution of each discipline to the paper: $\lambda_{i1} = 0.5$ in Economics, $\lambda_{i2} = 0.3$ in Mathematics and $\lambda_{i3} = 0.2$ in Sociology.

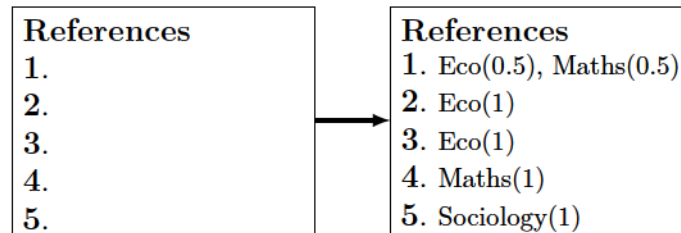


Figure 3.4: Disciplinary profile of a paper

4.2.3 Control variables

We include in our regressions a list of controls that previous studies considered as influencing the number of citations.

H-index. We partially control for individual characteristics like the past academic performance of French researchers at the time of publication. We consider the h -index to measure both the productivity and the impact of each scholar. Proposed by [Hirsch \(2005\)](#), a researcher has an h -index equal to h if h of his or her N papers have at least h citations. This measure improves the previous ones, especially the number of publications and the total number of citations. Indeed the h -index is less sensitive to extreme values: researchers with huge number of uncited publications or high number of citations caused by a single coauthored paper have both a low h -index. This measure is also calculated from the papers with the greatest impact, *i.e.* the most prominent publications of an author. However this index presents disadvantages summarized in the paper of [Bornmann and Daniel \(2007\)](#). Among others, comparing the h -index of authors belonging to different disciplines makes little sense. This measure can also be inflated by self-citations. The h -index depends on the scientific age of researchers since their number of publications and citations increases during their career. Furthermore the h -index doesn't permit to differentiate between an inactive researcher with high past publication performance but outdated and a young scholar working in fashionable fields. Over the 395,193 papers, we

are unable to obtain the h -index at the date of publication for at least a French researcher in 5,183 papers. We create a variable equal to the maximum of the h -index of French researchers in each project to partially deal with the characteristics of team members.

International collaboration. We also control for the nature of the collaboration with a dummy equal to 1 if the paper is internationally co-authored, 0 otherwise. According to [Adams \(2013\)](#), the scientific community is entering in the *fourth age of research, driven by international collaborations between elite research groups*. Internationally co-authored papers are more likely to be cited because teams are composed of high-impact researchers. Even if we don't have informations concerning foreign authors in our database, our international collaboration dummy captures similar informations than the h -index of foreign scholars.

Academic institutions. In our regressions, we partially take into account institutional effects with dummies equal to 1 if an academic institution is part of a given project, 0 otherwise. Research teams can belong to many academic institutions at the same time. In our database, we have 113 French academic institutions. We had to harmonize their names since some of them merged between 1999 and 2013: for instance Aix-Marseille in 2007, Strasbourg in 2009 and Bordeaux in 2013. We also include the number of academic institutions as a control of our regressions.

Number of authors. Previous studies have shown that the number of authors increases impact. [Fox and Faver \(1984\)](#) surveyed the literature on the advantages and disadvantages of academic collaboration. Among others, the separation of tasks resulting from the division of labor coupled with the greater specialization of researchers can enhance productivity. [Goldfinch et al. \(2003\)](#) argue that co-authored papers may benefit from a larger diffusion because of the greater social network of team members. [Herbertz \(1995\)](#) underlines the fact that co-authored papers may lead to a higher number of self-citations. We present some descriptive statistics concerning the number of authors per paper and per year in [Table 3.10](#). We observe that the number of authors is increasing over time. We remark that the average number of authors explodes between 2011 and 2013. However this is due to extreme values given that the median, the 3rd quartile and the 9th decile increase at a regular weight between 1999 and 2013.

Journal Impact Factor. We also control for the journal impact factor. For an

Year	Mean	Median	3rd Quartile	9th Decile	Max	S.D.
1999	9.55	5	7	9	525	38.3
2000	9.91	5	7	10	630	40.4
2001	8.39	5	7	9	743	32.6
2002	8.13	5	7	10	823	30.3
2003	7.98	5	7	10	817	30.5
2004	9.33	5	7	10	939	39.1
2005	9.86	5	7	10	859	44.4
2006	9.93	5	7	10	2512	47.2
2007	9.74	5	7	10	2010	43.7
2008	9.70	5	7	11	3099	47.5
2009	9.86	5	8	11	624	40.0
2010	11.99	5	8	12	3220	87.8
2011	18.47	5	8	12	3178	158.9
2012	27.49	6	8	13	3130	215.9
2013	23.94	6	9	13	3066	186.7
Total	13.12	5	8	11	3220	104.1

Table 3.10: Number of authors per paper

academic review r , its 3-years impact factor at date t is:

$$IF_{r,t} = \frac{\sum_{i=1}^{N_t^r} \sum_{j=t}^{t+3} \text{Citations}_{i,j}}{N_t^r},$$

with N_t^r the number of papers published in t in journal r .

This formula slightly differs from that available in the Web of Science, calculated as follows:

$$IF_{r,t}^{WoS} = \frac{\sum_{i=1}^N \sum_{j=t-2}^{t-1} \text{Citations}_{i,j}}{N_{t-1}^r + N_{t-2}^r}.$$

Our measure is close to the true definition of impact factor: the expected number of citations a paper published in t can obtain between t and $t + 3$. However the last value of impact factor that we can calculate is 3 years before the current year, contrary to the Web of Science (current year).

Whatever the measure, the journal impact factor is a poor predictor of the number of citations a paper can obtain. The citations distribution within a journal is very skewed: a high percentage of articles are never cited while a very small fraction of them get huge impact. However publications in journals with higher impact factor are more prestigious and have better visibility for the scientific community, which may influence the number of citations.

Disciplinary distance. We include a measure of distance between the disciplines.

Rafols and Meyer (2009) first constructed a matrix of citation flows between the 252 WoS subject categories. Denote by D the number of disciplines in this classification. Let c_{ij} be the number of citations obtained by discipline j from discipline i , meaning that i is citing and j is cited. Denote $c_i = (c_{i1}, \dots, c_{iD})$ the vector of citations from discipline i . For two disciplines $i = a, b$:

$$c_a = \begin{bmatrix} c_{a1} \\ \vdots \\ c_{aD} \end{bmatrix} \quad c_b = \begin{bmatrix} c_{b1} \\ \vdots \\ c_{bD} \end{bmatrix}.$$

The matrix of citation flows between disciplines is denoted by C :

$$C = \begin{bmatrix} c_1 & \dots & c_D \end{bmatrix}.$$

We can calculate the measure of distance between disciplines from the citing or the cited dimension. The former is related to knowledge integration while the latter is associated to knowledge diffusion. Thus measuring the distances between disciplines from the citing dimension seems to be more relevant for our application.

Rafols and Meyer (2009) calculates the similarity s_{ab} between disciplines a and b in the citing pattern using the Salton's cosine:

$$s_{ab} = \frac{\sum_{l=1}^D c_{al}c_{bl}}{\sqrt{\sum_{l=1}^D c_{al}^2} \sqrt{\sum_{l=1}^D c_{bl}^2}}$$

s_{ab} is comprised between 0 (perfect dissimilarity) and 1 (perfect similarity). The distance d_{ab} between disciplines a and b is obtained according to:

$$d_{ab} = 1 - s_{ab}$$

Example 9 illustrates the calculation of the distances between disciplines.

Example 9 Consider the matrix of citations between disciplines given in Table 3.11. For instance 114 papers published in discipline 4 have at least 1 reference in a journal classified in discipline 3. The similarity between disciplines 2 and 3 is computed as follows:

$$s_{23} = \frac{136 \times 0 + 734 \times 12 + 65 \times 542 + 12 \times 28}{\sqrt{136^2 + 734^2 + 65^2 + 12^2} \times \sqrt{0^2 + 12^2 + 542^2 + 28^2}} \approx 0.109$$

We deduce the distance between disciplines 2 and 3:

$$d_{23} = 1 - s_{23} = 0.891.$$

Computing the distances for all the disciplines, we obtain the matrix in Table 3.12.

Cited \ Citing	1	2	3	4
1	225	136	0	208
2	36	734	12	32
3	98	65	542	114
4	289	12	28	305

Table 3.11: Example of a matrix of citation flows between disciplines

s_{ij}	1	2	3	4
1	0	0.766	0.702	0.003
2	0.766	0	0.891	0.784
3	0.702	0.891	0	0.664
4	0.003	0.784	0.664	0

Table 3.12: Example of a matrix of distance between disciplines

Rafols et al. (2010) proposes the matrix of citation flows between approximately 225 WoS categories (224 in 2010 and 227 in 2015). Since we mainly consider the classification from the OST with 10 disciplines, we need to aggregate citation flows. Some subject categories are associated with two disciplines in the OST nomenclature. In this situation, we divide by two the citation flows both in the citing and the cited dimensions. *Rafols et al. (2010)* make the matrices of citation flows available to users (2007, 2008, 2009, 2010 and 2015). We compare the distances between disciplines with the 2010 and 2015 matrices. Since the results are very similar, we will consider the 2015 matrix for the whole period (1999-2013). We probably underestimate the distance between the disciplines at the beginning of the study period, but the 2015 matrix has more WoS subject categories.

For each paper, we calculate the mean, the minimum and the maximum distance between disciplines.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 Number of citations	1													
2 Papers in the top 10%	0.54	1												
3 Papers in the top 5%	0.58	0.70	1											
4 Impact Factor	0.44	0.30	0.28	1										
5 Number of references	0.21	0.15	0.13	0.19	1									
6 Berger-Parker	-0.030	-0.038	-0.036	0.015	-0.031	1								
7 Simpson	-0.031	-0.041	-0.038	0.018	-0.030	0.97	1							
8 Shannon	-0.031	-0.041	-0.038	0.020	-0.022	0.91	0.98	1						
9 Number of disciplines	-0.0065	-0.023	-0.020	0.037	0.083	0.61	0.71	0.81	1					
10 Number of authors	0.085	0.059	0.063	0.053	-0.0019	-0.028	-0.027	-0.021	0.016	1				
11 Number of diff. organisations	0.17	0.11	0.12	0.11	0.062	-0.046	-0.047	-0.043	0.0014	0.90	1			
12 International Collaboration	0.14	0.12	0.11	0.15	0.070	-0.025	-0.028	-0.031	-0.024	0.092	0.22	1		
13 Maximum h-index	0.14	0.084	0.083	0.14	0.29	-0.029	-0.029	-0.025	0.016	0.073	0.15	0.10	1	
14 Maximal distance	-0.023	-0.017	-0.014	-0.038	0.038	0.41	0.50	0.59	0.80	0.044	0.040	0.0010	-0.039	1

Table 3.13: Correlation matrix⁵(papers with at least four references, 1-10 disciplines, OST nomenclature)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 Number of citations	1															
2 Papers in the top 10%	0.44	1														
3 Papers in the top 5%	0.47	0.70	1													
4 Impact Factor	0.36	0.29	0.28	1												
5 Number of references	0.35	0.18	0.18	0.21	1											
6 Berger-Parker	-0.060	-0.053	-0.049	-0.081	-0.071	1										
7 Simpson	-0.064	-0.056	-0.052	-0.091	-0.068	0.96	1									
8 Shannon	-0.060	-0.053	-0.049	-0.092	-0.041	0.88	0.97	1								
9 Number of disciplines	-0.0025	-0.0072	-0.0037	-0.035	0.18	0.52	0.65	0.78	1							
10 Number of authors	0.089	0.055	0.055	0.046	-0.0053	-0.045	-0.045	-0.039	-0.010	1						
11 Number of diff. organisations	0.17	0.11	0.11	0.10	0.057	-0.068	-0.069	-0.063	-0.019	0.91	1					
12 International Collaboration	0.11	0.12	0.098	0.14	0.054	-0.041	-0.044	-0.045	-0.030	0.088	0.20	1				
13 Maximum h-index	0.13	0.096	0.094	0.15	0.27	-0.069	-0.068	-0.056	0.033	0.088	0.16	0.12	1			
14 Minimal distance	-0.028	-0.0086	-0.0050	-0.082	-0.064	-0.049	-0.060	-0.062	-0.036	0.064	0.063	0.0022	-0.093	1		
15 Mean distance	-0.027	-0.014	-0.010	-0.092	0.053	0.28	0.36	0.44	0.60	0.040	0.034	-0.014	-0.050	0.71	1	
16 Maximal distance	-0.020	-0.014	-0.0098	-0.077	0.10	0.38	0.47	0.58	0.76	0.023	0.018	-0.016	-0.018	0.48	0.95	1

Table 3.14: Correlation matrix⁶(papers with at least four references, 1-10 disciplines)

Table 3.13⁷ and 3.14 provide the correlation matrix for the variables employed in our analysis. For better readability, we have dropped the stars in the tables. All the coefficients are significant at 1%, except in Table 3.14 where the correlations between the number of disciplines and the number of citations, and the correlation between the number of disciplines and the dummy “Papers in the top 5%” are not significant.

The number of authors and the number of different organisations participating to a paper are highly correlated (more than 0.9). Even if these two variables are of interest for our study, putting them simultaneously in the regressions gives opposite signs for the coefficients. Inserting them separately leads to consistently positive coefficients. To avoid possible problems of collinearity, we only include the number of authors in our regressions.

Even if we are interested in the Spearman rank-order correlations for our indexes of interdisciplinarity, we first comment the Pearson correlations given in Tables 3.13 and 3.14. Three of our four interdisciplinary measures (Berger-Parker, Simpson and Shannon indexes) are strongly and positively correlated with each other (between 0.88 and 0.97) and negatively correlated with the variables of impact (number of citations, impact factor, papers in the top 5% and 10%). The remaining measure (number of disciplines) is positively correlated with the other indicators of interdisciplinarity (0.52 with the Berger-Parker index for instance), but to a lesser extent. Most importantly, the number of disciplines is not correlated with some measures of impact.

However around 40% of the papers in the database have a unique discipline so they are monodisciplinary (index equal to 1). We calculate the Pearson correlations between our indexes of interdisciplinarity then with the measures of impact for papers having more than one discipline. We report the results in Table 3.15. We obtain similar correlations for the Berger-Parker, the Simpson and the Shannon-Wiener indexes. But remark now that the number of disciplines is significantly and positively correlated with most of the measures of impact. This highlights the potential importance of control variables in our econometric estimations.

We calculate the Spearman rank-order correlation for the papers involving more than one discipline. We report the results in Table 3.16. We obtain similar results: the Berger-Parker, the Simpson and the Shannon indexes are highly correlated (around 0.9) while the number of disciplines is positively but sometimes weakly correlated with other measures

⁷We do not report the minimum and the mean distance given that they are necessarily equal to the maximum distance (subsample of papers with 2 disciplines or less).

	1	2	3	4	5	6	7	8
1 Number of citations	1							
2 Papers in the top 10%	0.43	1						
3 Papers in the top 5%	0.47	0.70	1					
4 Impact Factor	0.36	0.29	0.28	1				
5 Berger-Parker	-0.058	-0.049	-0.047	-0.085	1			
6 Simpson	-0.062	-0.051	-0.049	-0.097	0.95	1		
7 Shannon	-0.059	-0.048	-0.046	-0.099	0.87	0.97	1	
8 Number of disciplines	0.0073	0.0060	0.0063	-0.038	0.46	0.60	0.74	1

Table 3.15: Pearson correlations (papers with at least 2 disciplines)

	1	2	3	4
1 Berger-Parker	1			
2 Simpson	0.97	1		
3 Shannon	0.90	0.96	1	
4 Number of disciplines	0.34	0.49	0.66	1

Table 3.16: Spearman's rank-ordered correlations for papers with at most 2 disciplines, OST nomenclature

of interdisciplinarity (0.34 with the Berger-Parker index for instance).

5 Results

In this section, we report the results of our estimations dealing with the effect of interdisciplinarity on impact. We first choose to run negative binomial regressions given that our dependant variable, the number of citations per article, only takes positive integer values and is over dispersed, so that its variance is greater than its mean. We first examine this relationship for papers with less than two disciplines, before considering the whole database. When we consider the probability that a paper belongs to the 5% or 10% most cited papers, we run logit regressions since the dependent variable can only take two values (0 and 1).

5.1 1-2 disciplines

In this subsection, we consider the sample of papers with at most 2 disciplines and at least four references (161,033 observations). Tables 3.17 and 3.18 display the results of these estimations in four blocks, each of one testing a distinct measure of interdisciplinarity (different orders of the Hill index): the Berger-Parker (regression 1), the Simpson (regression 2), the Shannon-Wiener (regression 3) indicators and the number of disciplines

(regression 4). In each block, regression “b” integrates the impact factor, regression “c” includes the maximal distance between the disciplines involved in the paper as a control, and regression “d” considers both the impact factor and the maximum distance between disciplines.

In all the regressions, the coefficient associated with interdisciplinarity (Berger-Parker, Simpson and Shannon-Wiener indexes) is strongly significant and negative. The coefficient associated to the number of disciplines is not significant in regressions (4a) to (4c), and strongly significant and negative when we include both the impact factor and the maximal distance in our regression. These first results partially validate Proposition 3 of our model, therefore that *ceteris paribus*, increasing interdisciplinarity decreases the impact of research projects.

Given that our subsample only contains papers with less than two disciplines, we can interpret the results as following. Here all the interdisciplinary indexes vary from 1 to 2. A value of 1 implies that the paper is monodisciplinary while a value of 2 indicates that the paper involves two disciplines, both contributing equally. If we consider regressions 1 to 3, monodisciplinary papers have on average 7-13% citations more than those with two disciplines contributing equally.

Remember that we have truncated the distribution of interdisciplinarity since we consider papers with at most 2 disciplines. If it has allowed us to partially conclude on the effect of complexity on scientific output, we have introduced a selection bias. Thus we refrain from giving more interpretations with this subsample. In the next subsection, we reintegrate papers involving more than two disciplines.

Table 3.17: Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 4 references, 1-2 disciplines, OST nomenclature)

	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Berger-Parker	-0.116*** (-8.06)	-0.0884*** (-7.76)	-0.135*** (-8.89)	-0.116*** (-9.40)				
Simpson					-0.0917*** (-7.77)	-0.0732*** (-8.14)	-0.117*** (-9.05)	-0.107*** (-10.53)
Number of authors	0.00127*** (11.92)	0.000636*** (11.02)	0.00125*** (11.86)	0.000620*** (10.83)	0.00127*** (11.92)	0.000636*** (11.03)	0.00125*** (11.83)	0.000615*** (10.79)
International Collaboration	0.448*** (54.61)	0.259*** (41.06)	0.448*** (54.62)	0.259*** (41.11)	0.448*** (54.63)	0.259*** (41.08)	0.449*** (54.67)	0.259*** (41.16)
Maximum h-index	0.0329*** (51.43)	0.0199*** (38.96)	0.0329*** (51.40)	0.0199*** (39.01)	0.0329*** (51.40)	0.0199*** (38.95)	0.0329*** (51.40)	0.0199*** (39.03)
Maximal distance			0.0506** (3.03)	0.0728*** (5.43)			0.0698*** (4.03)	0.0935*** (6.66)
Impact Factor		0.108*** (117.03)		0.108*** (117.08)		0.108*** (117.05)		0.108*** (117.13)
Observations	161033	161033	161033	161033	161033	161033	161033	161033
Log-Likelihood	-498320.6	-476056.2	-498312.2	-476033.3	-498319.4	-476051.2	-498304.9	-476017.0
AIC	996707.2	952180.3	996692.4	952136.7	996704.8	952170.5	996677.8	952104.1
BIC	997036.8	952520.0	997032.0	952486.3	997034.5	952510.1	997017.4	952453.7

t statistics in parentheses

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

Table 3.18: Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 4 references, 1-2 disciplines, OST nomenclature)

	(3a)	(3b)	(3c)	(3d)	(4a)	(4b)	(4c)	(4d)
Shannon	-0.0819*** (-7.25)	-0.0674*** (-7.97)	-0.121*** (-9.13)	-0.119*** (-11.49)				
Number of disciplines					-0.00178 (-0.20)	-0.00679 (-1.02)	-0.00205 (-0.14)	-0.0587*** (-5.24)
Number of authors	0.00127*** (11.94)	0.000638*** (11.06)	0.00124*** (11.82)	0.000611*** (10.76)	0.00128*** (11.97)	0.000642*** (11.06)	0.00128*** (11.97)	0.000627*** (10.89)
International Collaboration	0.448*** (54.63)	0.258*** (41.08)	0.449*** (54.70)	0.259*** (41.21)	0.447*** (54.41)	0.258*** (40.97)	0.448*** (54.28)	0.259*** (41.08)
Maximum h-index	0.0329*** (51.41)	0.0199*** (38.97)	0.0329*** (51.45)	0.0199*** (39.11)	0.0330*** (51.52)	0.0200*** (39.10)	0.0330*** (51.51)	0.0201*** (39.39)
Maximal distance			0.0955*** (5.15)	0.123*** (8.13)			0.000628 (0.02)	0.117*** (5.57)
Impact Factor		0.108*** (117.08)		0.108*** (117.22)				0.108*** (117.10)
Observations	161033	161033	161033	161033	161033	161033	161033	161033
Log-Likelihood	-498324.3	-476052.7	-498301.0	-476002.1	-498374.5	-476095.5	-498374.5	-476072.8
AIC	996714.6	952173.3	996670.0	952074.2	996815.0	952258.9	996817.0	952215.7
BIC	997044.2	952512.9	997009.6	952423.8	997144.6	952598.6	997156.6	952565.3

t statistics in parentheses

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

5.2 1-10 disciplines

Here we consider all the papers for which we are able to identify at least 4 references, whatever the number of disciplines. We remark that the coefficients associated to the Berger-Parker, the Simpson and the Shannon-Wiener indexes are strongly significant and negative, which goes in the same direction than our theoretical results. An increase by a standard deviation of these indexes of interdisciplinarity⁸ decreases the number of citations by 1.6% to 6.7% (Berger-Parker), 1.2% to 7.4% (Simpson) and 0.5% to 7.6% (Shannon-Wiener).

However the number of disciplines has a significant and positive effect on the number of citations whatever the list of controls. Indeed we have considered that the number of disciplines is a continuous variable while it can take only ten values (all the integers between 1 and 10). We have already reported the distribution of the number of disciplines per paper in Table 3.5. To have classes with a similar number of observations, we regroup papers with more than 3 disciplines in a category “3 and more disciplines”. We consider the number of disciplines as a categorical variable and choose monodisciplinary papers (1 discipline) as the reference category. We run regressions with the same list of controls and present the results in Table 3.22. All the coefficients associated to the modality “2 disciplines” are non-significant or significant and negative, meaning that papers with two disciplines have in general less citations than the monodisciplinary ones. However the coefficient associated to the category “3 disciplines and more” is significant and positive which is less expected, except in regression (4d).

To appreciate the magnitude of the effect of interdisciplinarity, we calculate the logarithm of the Hill index and integrate it as an explicative variable in our regressions. We present the results in Table 3.23 and 3.24. The logarithmic transformation allows us to interpret the coefficients in terms of elasticities. We obtain that, whatever the choice of the Hill index, an increase of 10% of interdisciplinarity leads to a decrease between 0.2% and 2.7% of the number of citations.

To test the non-linearity of the relationship between interdisciplinarity and the number of citations, we introduce in our regressions a square term for each index, except for the number of disciplines (not continuous variable). We report the results in Table 3.25 and 3.26. Since the coefficient associated to the square term is positive while

⁸The standard deviations for the Berger-Parker, the Simpson and the Shannon-Wiener indexes are respectively 0.41, 0.57 and 0.66.

	(a)	(b)	(c)	(d)	p90	p95	p99
Berger-Parker	3.52	2.51	2.85	2.37	2	2	2.67
Simpson	4.54	2.63	3.64	2.91	2.25	2.65	3.31
Shannon	5.29	2.45	3.93	3.17	2.55	2.86	3.69

Table 3.19: Turning point calculations for regressions (a) to (d) and 90th, 95th and 99th percentiles

this related to the linear term is negative, both being significant, we deduce that we have a U-relationship between interdisciplinarity and the number of citations. This seems to confirm the second part of Proposition 3 concerning the convexity of scientific output with our complexity index. Nevertheless, to ensure that the number of citations is decreasing with our indexes of interdisciplinarity, we need to compute the turning point and check that it is not significant. We report these values and the 90th, 95th and 99th percentiles of the distribution of the Berger-Parker, Simpson and Shannon indexes in Table 3.19.

We obtain that the value of the Berger-Parker index which minimizes the number of citations is between 2.37 (regression (1d)) and 3.52 (regression (1a)). Given that the 95th percentile is 2, more than 95% of the papers in our database have an interdisciplinary index lower than this threshold. For the Simpson index, the minimal threshold obtained in our regressions is 2.63. Around 95% of the papers have a Simpson index below this value. For the Shannon-Wiener, the threshold is between 2.45 and 5.29. Around 90% of the papers have a Shannon index below this threshold in regression (b), and more than 95% in the other regressions. Globally, the turning point is only engendered by extreme values and therefore can't be interpreted. Interdisciplinarity decreases scientific impact, but less and less rapidly, so that the highest penalty on impact appears when the project begins to be complex, *i.e.* when the interdisciplinarity indexes are just above 1. We illustrate the predictions on the number of citations in Figure 3.5 according to regressions (1d), (2d) and (3d) from Tables 3.25 and 3.26.

So far, we have tested the relationship between interdisciplinarity and the number of citations. In the following, we consider the effect of interdisciplinarity on the probability that a paper belongs to the top 10% or top 5% most cited papers of their category. Since the dependant variable is a dummy, we run Logit regressions. We use the same list of controls as previously except the impact factor, otherwise logistic regressions do not converge. We report the results in Table 3.27 for the top 10% and in Table 3.28 for the top 5%. All the coefficients associated with the Berger-Parker, the Simpson and the Shannon-Wiener indexes are significant and negative. To appreciate the strenght of the

effect, let's take an example. A monodisciplinary paper published in 2013 in Physics (all other variables at means) has 11.07% chances to belong to the top 10% most cited papers. A paper published in 2013 in the same journal but with two disciplines contributing equally has 9.43% chances to belong to the top 10% most cited papers. Remark again that the coefficient associated to the number of disciplines is positive and significant. For the same reasons than previously, we run regressions considering that the number of disciplines is a categorical variable. We present the results in Table 3.29. The first two columns deal with the dummy "Papers in the top 10%", the last two ones with the dummy "Papers in the top 5%". We consider "1 discipline" as the reference category. In all the regressions, the coefficients associated to "2 disciplines" are significant and negative while this related to "3 disciplines and more" is non significant and/or negative, except in the first regression.

The results go in the same direction than with the number of citations. The more a paper is interdisciplinary, the less chances he has to belong to the 10% or 5% most cited papers.

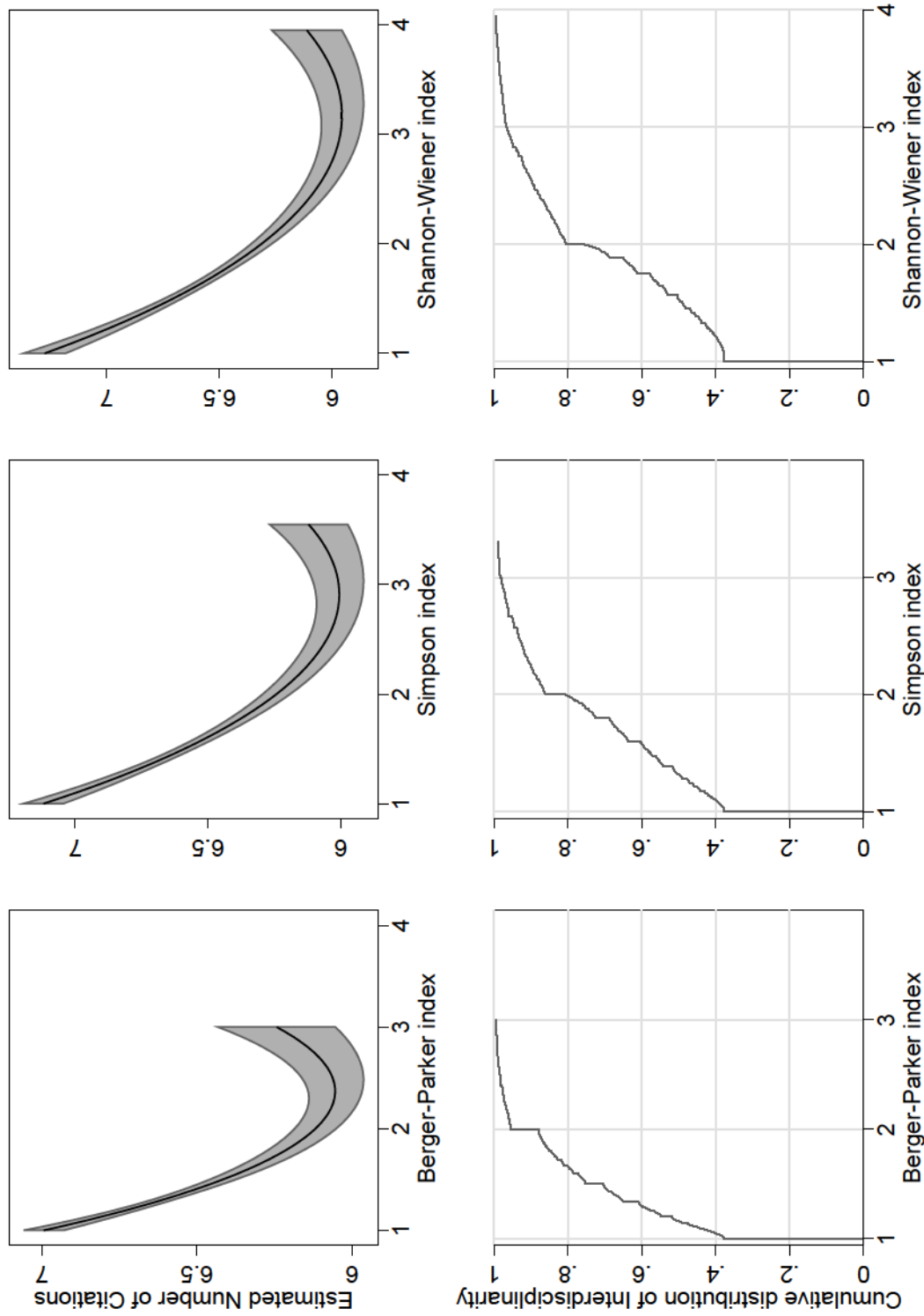


Figure 3.5: Predictions on the number of citations (regressions (1d), (2d) and (3d), Tables 3.25 and 3.26) and cumulative distributions of the interdisciplinary indexes

Table 3.20: Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 4 references, 1-10 disciplines, OST nomenclature)

	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Berger-Parker	-0.123*** (-17.01)	-0.0390*** (-6.55)	-0.170*** (-19.52)	-0.0983*** (-13.40)				
Simpson					-0.0840*** (-16.21)	-0.0217*** (-5.17)	-0.134*** (-19.98)	-0.0776*** (-13.71)
Number of authors	0.00107*** (13.22)	0.000570*** (10.99)	0.00104*** (13.20)	0.000544*** (10.62)	0.00107*** (13.21)	0.000571*** (11.00)	0.00104*** (13.16)	0.000541*** (10.59)
International Collaboration	0.415*** (63.30)	0.255*** (49.52)	0.416*** (63.39)	0.256*** (49.83)	0.415*** (63.22)	0.255*** (49.50)	0.416*** (63.39)	0.256*** (49.86)
Maximum h-index	0.0327*** (60.07)	0.0200*** (44.80)	0.0327*** (60.50)	0.0200*** (45.37)	0.0327*** (60.01)	0.0201*** (44.86)	0.0326*** (60.43)	0.0200*** (45.40)
Maximal distance			0.112*** (9.15)	0.142*** (13.93)			0.146*** (11.03)	0.162*** (14.56)
Impact Factor		0.103*** (125.81)		0.104*** (126.51)		0.103*** (125.70)		0.103*** (126.54)
Observations	250164	250164	250164	250164	250164	250164	250164	250164
Log-Likelihood	-772885.9	-740790.9	-772789.6	-740593.3	-772896.1	-740800.6	-772754.5	-740578.3
AIC	1545837.8	1481649.8	1545647.3	1481256.7	1545858.2	1481669.3	1545577.0	1481226.5
BIC	1546182.0	1482004.4	1546001.9	1481621.7	1546202.4	1482023.9	1545931.6	1481591.6

t statistics in parentheses

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

Table 3.21: Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 4 references, 1-10 disciplines, OST nomenclature)

	(3a)	(3b)	(3c)	(3d)	(4a)	(4b)	(4c)	(4d)
Shannon	-0.0595*** (-12.90)	-0.00750* (-2.04)	-0.120*** (-18.16)	-0.0700*** (-12.56)				
Number of disciplines					0.0216*** (6.22)	0.0359*** (12.50)	0.0588*** (9.55)	0.0420*** (8.10)
Number of authors	0.00108*** (13.23)	0.000572*** (11.02)	0.00104*** (13.15)	0.000540*** (10.57)	0.00108*** (13.32)	0.000567*** (11.03)	0.00110*** (13.46)	0.000569*** (11.09)
International Collaboration	0.415*** (63.12)	0.255*** (49.47)	0.416*** (63.33)	0.256*** (49.84)	0.417*** (63.12)	0.256*** (49.66)	0.417*** (63.02)	0.256*** (49.64)
Maximum h-index	0.0328*** (60.15)	0.0201*** (45.00)	0.0327*** (60.55)	0.0200*** (45.52)	0.0333*** (61.08)	0.0203*** (45.53)	0.0331*** (60.69)	0.0203*** (45.44)
Maximal distance			0.174*** (11.85)	0.179*** (14.36)			-0.136*** (-7.26)	-0.0222 (-1.44)
Impact Factor		0.104*** (125.68)		0.104*** (126.60)		0.104*** (126.15)		0.104*** (125.81)
Observations	250164	250164	250164	250164	250164	250164	250164	250164
Log-Likelihood	-772962.1	-740814.9	-772798.7	-740595.0	-773052.0	-740662.1	-772996.7	-740660.2
AIC	1545990.2	1481697.9	1545665.4	1481260.1	1546170.0	1481392.2	1546061.5	1481390.5
BIC	1546334.4	1482052.5	1546020.0	1481625.1	1546514.2	1481746.8	1546416.1	1481755.5

t statistics in parentheses

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

Table 3.22: Negative binomial estimates for the effect of the number of disciplines (categorical variable) on the number of citations (papers with at least 4 references, 1-10 disciplines, OST nomenclature)

	(4a)	(4b)	(4c)	(4d)
1 discipline	0 (.)	0 (.)	0 (.)	0 (.)
2 disciplines	-0.0118 (-1.38)	-0.0166* (-2.57)	0.000892 (0.07)	-0.0638*** (-6.57)
3 disciplines and more	0.0202* (2.09)	0.0611*** (7.98)	0.0434* (2.19)	-0.0244 (-1.57)
Number of authors	0.00108*** (13.35)	0.000567*** (11.13)	0.00109*** (13.36)	0.000559*** (10.94)
International Collaboration	0.417*** (63.12)	0.255*** (49.67)	0.417*** (63.05)	0.256*** (49.75)
Maximum h-index	0.0332*** (61.16)	0.0202*** (45.68)	0.0332*** (60.99)	0.0203*** (45.92)
Maximal distance			-0.0293 (-1.37)	0.108*** (6.26)
Impact Factor		0.104*** (126.27)		0.104*** (126.36)
Observations	250164	250164	250164	250164
Log-Likelihood	-773078.2	-740675.1	-773076.4	-740644.8
AIC	1546224.4	1481420.3	1546222.9	1481361.7
BIC	1546579.0	1481785.3	1546587.9	1481737.1

t statistics in parentheses

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

Table 3.23: Negative binomial estimates for the effect of the logarithm of interdisciplinarity on the number of citations (papers with at least 4 references, 1-10 disciplines, OST nomenclature)

	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
log Berger-Parker	-0.193*** (-16.64)	-0.0639*** (-6.86)	-0.285*** (-19.93)	-0.176*** (-14.68)				
log Simpson					-0.145*** (-15.47)	-0.0423*** (-5.71)	-0.259*** (-20.53)	-0.167*** (-15.77)
Number of authors	0.00107*** (13.21)	0.000570*** (10.99)	0.00104*** (13.17)	0.000540*** (10.58)	0.00108*** (13.21)	0.000571*** (11.00)	0.00103*** (13.11)	0.000536*** (10.53)
International Collaboration	0.415*** (63.31)	0.255*** (49.52)	0.416*** (63.46)	0.256*** (49.88)	0.415*** (63.22)	0.255*** (49.50)	0.416*** (63.48)	0.256*** (49.94)
Maximum h-index	0.0327*** (60.11)	0.0200*** (44.81)	0.0327*** (60.58)	0.0200*** (45.42)	0.0327*** (60.10)	0.0200*** (44.87)	0.0326*** (60.59)	0.0200*** (45.48)
Maximal distance			0.132*** (10.42)	0.158*** (14.86)			0.176*** (12.70)	0.190*** (16.09)
Impact Factor		0.103*** (125.77)		0.103*** (126.59)		0.103*** (125.64)		0.103*** (126.62)
Observations	250164	250164	250164	250164	250164	250164	250164	250164
Log-Likelihood	-772879.9	-740787.3	-772755.5	-740559.7	-772896.3	-740796.0	-772711.5	-740523.4
AIC	1545825.8	1481642.6	1545579.1	1481189.3	1545858.7	1481659.9	1545491.0	1481116.9
BIC	1546170.0	1481997.2	1545933.7	1481554.4	1546202.9	1482014.5	1545845.6	1481481.9

t statistics in parentheses

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

Table 3.24: Negative binomial estimates for the effect of the logarithm of interdisciplinarity on the number of citations (papers with at least 4 references, 1-10 disciplines, OST nomenclature)

	(3a)	(3b)	(3c)	(3d)
log Shannon	-0.106*** (-11.96)	-0.0183** (-2.64)	-0.259*** (-19.00)	-0.175*** (-15.46)
Number of authors	0.00108*** (13.23)	0.000572*** (11.02)	0.00103*** (13.10)	0.000534*** (10.48)
International Collaboration	0.415*** (63.12)	0.255*** (49.46)	0.416*** (63.43)	0.256*** (49.95)
Maximum h-index	0.0328*** (60.25)	0.0201*** (44.99)	0.0327*** (60.75)	0.0200*** (45.63)
Maximal distance			0.220*** (13.85)	0.225*** (16.57)
Impact Factor		0.104*** (125.62)		0.103*** (126.73)
Observations	250164	250164	250164	250164
Log-Likelihood	-772968.2	-740812.4	-772751.1	-740523.6
AIC	1546002.3	1481692.7	1545570.2	1481117.2
BIC	1546346.5	1482047.3	1545924.8	1481482.2

t statistics in parentheses

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

Table 3.25: Negative binomial estimates for the effect of interdisciplinarity (square term) on the number of citations (papers with at least 4 references, 1-10 disciplines, OST nomenclature)

	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Berger-Parker	-0.235*** (-6.40)	-0.118*** (-4.02)	-0.431*** (-10.76)	-0.365*** (-10.96)				
Berger-Parker ²	0.0334** (3.23)	0.0235** (2.81)	0.0756*** (6.98)	0.0770*** (8.52)				
Simpson					-0.148*** (-6.69)	-0.0852*** (-4.91)	-0.313*** (-12.66)	-0.271*** (-13.41)
Simpson ²					0.0163** (3.11)	0.0162*** (3.86)	0.0430*** (7.87)	0.0466*** (10.39)
Number of authors	0.00107*** (13.21)	0.000570*** (10.99)	0.00104*** (13.16)	0.000539*** (10.57)	0.00108*** (13.21)	0.000571*** (11.01)	0.00103*** (13.11)	0.000535*** (10.53)
International Collaboration	0.415*** (63.32)	0.255*** (49.52)	0.416*** (63.49)	0.256*** (49.91)	0.415*** (63.25)	0.255*** (49.51)	0.416*** (63.52)	0.256*** (49.98)
Maximum h-index	0.0327*** (60.10)	0.0200*** (44.82)	0.0327*** (60.64)	0.0200*** (45.52)	0.0327*** (60.03)	0.0201*** (44.89)	0.0326*** (60.59)	0.0200*** (45.60)
Maximal distance			0.133*** (10.42)	0.164*** (15.14)			0.175*** (12.63)	0.193*** (16.36)
Impact Factor		0.103*** (125.80)		0.104*** (126.63)		0.103*** (125.68)		0.103*** (126.66)
Observations	250164	250164	250164	250164	250164	250164	250164	250164
Log-Likelihood	-772878.4	-740786.3	-772754.9	-740548.9	-772888.4	-740791.3	-772706.2	-740509.0
AIC	1545824.9	1481642.6	1545579.9	1481169.8	1545844.8	1481652.6	1545482.4	1481090.1
BIC	1546179.5	1482007.7	1545944.9	1481545.3	1546199.4	1482017.7	1545847.4	1481465.5

t statistics in parentheses

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

Table 3.26: Negative binomial estimates for the effect of interdisciplinarity (square term) on the number of citations (papers with at least 4 references, 1-10 disciplines, OST nomenclature)

	(3a)	(3b)	(3c)	(3d)
Shannon	-0.0995*** (-5.34)	-0.0568*** (-3.87)	-0.300*** (-13.45)	-0.269*** (-14.76)
Shannon ²	0.00940* (2.32)	0.0116*** (3.55)	0.0382*** (8.81)	0.0424*** (11.91)
Number of authors	0.00108*** (13.23)	0.000573*** (11.03)	0.00103*** (13.09)	0.000534*** (10.48)
International Collaboration	0.415*** (63.15)	0.255*** (49.47)	0.416*** (63.48)	0.256*** (49.99)
Maximum h-index	0.0328*** (60.15)	0.0201*** (45.03)	0.0327*** (60.75)	0.0200*** (45.78)
Maximal distance			0.222*** (14.04)	0.232*** (17.08)
Impact Factor		0.104*** (125.65)		0.104*** (126.76)
Observations	250164	250164	250164	250164
Log-Likelihood	-772957.5	-740806.4	-772735.0	-740499.3
AIC	1545983.0	1481682.8	1545540.1	1481070.6
BIC	1546337.7	1482047.8	1545905.1	1481446.1

t statistics in parentheses

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

Table 3.27: Logit estimates for the effect of interdisciplinarity on the probability that the paper belongs to the top 10% most cited papers (papers with at least 4 references, 1-10 disciplines, OST nomenclature)

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Berger-Parker	-0.139*** (-8.79)	-0.228*** (-12.43)						
Simpson			-0.0898*** (-8.19)	-0.179*** (-13.06)				
Shannon					-0.0579*** (-6.17)	-0.161*** (-12.30)		
Number of disciplines							0.0436*** (7.44)	0.0835*** (7.66)
Number of authors	0.000546*** (13.74)	0.000519*** (13.05)	0.000548*** (13.79)	0.000516*** (12.96)	0.000551*** (13.86)	0.000516*** (12.97)	0.000548*** (13.78)	0.000560*** (14.03)
International Collaboration	0.614*** (52.80)	0.616*** (52.92)	0.614*** (52.78)	0.616*** (52.89)	0.614*** (52.77)	0.615*** (52.87)	0.615*** (52.89)	0.615*** (52.86)
Maximum h-index	0.0446*** (55.90)	0.0447*** (55.98)	0.0446*** (55.89)	0.0446*** (55.88)	0.0447*** (56.05)	0.0447*** (55.96)	0.0452*** (56.74)	0.0451*** (56.56)
Maximal distance		0.200*** (9.75)		0.242*** (11.04)		0.281*** (11.47)		-0.144*** (-4.32)
Observations	250164	250164	250164	250164	250164	250164	250164	250164
Log-Likelihood	-104693.7	-104646.4	-104699.1	-104638.5	-104713.9	-104648.5	-104705.6	-104696.2
AIC	209451.5	209358.8	209462.2	209343.0	209491.7	209363.1	209475.2	209458.5
BIC	209785.3	209703.0	209796.0	209687.2	209825.5	209707.3	209809.0	209802.7

t statistics in parentheses

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

Table 3.28: Logit estimates for the effect of interdisciplinarity on the probability that the paper belongs to the top 5% most cited papers (papers with at least 4 references, 1-10 disciplines, OST nomenclature)

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Berger-Parker	-0.212*** (-9.87)	-0.325*** (-13.01)						
Simpson			-0.140*** (-9.44)	-0.255*** (-13.78)				
Shannon					-0.0965*** (-7.64)	-0.233*** (-13.25)		
Number of disciplines							0.0430*** (5.60)	0.0857*** (5.97)
Number of authors	0.000582*** (14.19)	0.000551*** (13.40)	0.000584*** (14.25)	0.000547*** (13.29)	0.000588*** (14.35)	0.000547*** (13.29)	0.000585*** (14.28)	0.000597*** (14.51)
International Collaboration	0.697*** (44.96)	0.698*** (45.05)	0.696*** (44.94)	0.698*** (45.03)	0.696*** (44.93)	0.697*** (45.01)	0.697*** (45.01)	0.696*** (44.99)
Maximum h-index	0.0495*** (50.57)	0.0495*** (50.61)	0.0494*** (50.54)	0.0494*** (50.49)	0.0496*** (50.69)	0.0495*** (50.56)	0.0502*** (51.40)	0.0501*** (51.27)
Maximal distance		0.245*** (9.15)		0.304*** (10.63)		0.364*** (11.41)		-0.154*** (-3.50)
Observations	250164	250164	250164	250164	250164	250164	250164	250164
Log-Likelihood	-68380.5	-68338.9	-68385.0	-68329.0	-68401.1	-68336.8	-68415.2	-68409.1
AIC	136825.0	136743.7	136834.0	136724.1	136866.2	136739.6	136894.4	136884.1
BIC	137158.7	137087.9	137167.8	137068.3	137199.9	137083.8	137228.2	137228.3

t statistics in parentheses

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

Table 3.29: Negative binomial estimates for the effect of the number of disciplines (categorical variable) on the probability that the paper belongs to the top 10% and top 5% most cited papers (papers with at least 4 references, 1-10 disciplines, OST nomenclature)

	Top90	Top90	Top95	Top95
1 discipline	0 (.)	0 (.)	0 (.)	0 (.)
2 disciplines	-0.0355* (-2.43)	-0.0764*** (-3.32)	-0.0503** (-2.64)	-0.128*** (-4.23)
3 disciplines and more	0.0504** (3.22)	-0.0230 (-0.65)	0.0320 (1.56)	-0.108* (-2.29)
Number of authors	0.000551*** (13.84)	0.000547*** (13.72)	0.000589*** (14.36)	0.000583*** (14.19)
International Collaboration	0.615*** (52.87)	0.615*** (52.90)	0.697*** (45.00)	0.697*** (45.04)
Maximum h-index	0.0451*** (56.64)	0.0452*** (56.68)	0.0501*** (51.31)	0.0503*** (51.41)
Maximal distance		0.0926* (2.31)		0.176*** (3.31)
Observations	250164	250164	250164	250164
Log-Likelihood	-104712.8	-104710.2	-68419.6	-68414.1
AIC	209491.7	209488.4	136905.2	136896.3
BIC	209835.9	209843.0	137249.4	137250.9

t statistics in parentheses

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

6 Robustness checks

To test the robustness of our findings, we conduct a number of alternative analyses. We have considered papers for which we are able to identify at least 4 references. Increasing this threshold could reduce the proportion of papers having all their references in a unique discipline. We select papers for which we are able to identify at least 6 references and run exactly the same regressions. We report the results from Tables 3.35 to 3.50. We obtain very similar results compared to previously. In Table 3.36, we present the results of the regressions run with our subsample of papers involving less than two disciplines. All the coefficients associated to the number of disciplines are negative and significant, while some of them were non significant with four references. Now considering the whole sample of papers with more than 6 references, we can see in Table 3.39 that all the coefficients associated to “2 disciplines” and “3 disciplines or more” are negative, except one in regression (4b). However for the last category, the values of the coefficients are similar or higher than for the category two disciplines, whereas we expected to find lower ones.

Instead of computing the indexes of interdisciplinarity with the OST nomenclature, we use the CWTS classification which is more granular with 35 disciplines. We calculate the Spearman rank-order correlations that we report in Table 3.30. Again the ranking of projects according to the Berger-Parker, the Simpson and the Shannon-Wiener indexes are close given that the correlation coefficients are greater than 0.9, while this will be less the case for the ranking according to the number of disciplines.

	1	2	3	4
1 Berger-Parker	1			
2 Simpson	0.98	1		
3 Shannon	0.91	0.97	1	
4 Number of disciplines	0.53	0.64	0.78	1

Table 3.30: Spearman’s rank-ordered correlations for papers with at least 2 disciplines, CWTS nomenclature

We calculate the interdisciplinary indexes with the CWTS classification and we run exactly the same regressions as before. We report the results from Tables 3.51 to 3.66. Results are unchanged concerning the Berger-Parker, the Simpson and the Shannon-Wiener

indexes. We observe again strongly significant and negative coefficients associated to these variables. For the number of disciplines, we still have mixed results. In Table 3.56, we consider the subsample of papers with at most 2 disciplines. We observe that the coefficient associated to the number of disciplines is not significant whatever the regression (4a to 4d), meaning that papers with two disciplines don't have necessarily less citations than the monodisciplinary ones. Considering the whole database, we can see in Table 3.57 that the same coefficients are negative and significant. However for papers with 3 disciplines, we can't reject the assumption that they have as many citations as the monodisciplinary ones. With 4 disciplines or more, we obtain positive and significant coefficients, meaning that those articles have on average 0.06% more citations than those with 1 discipline. The effect is very small, but probably reflects the existence of a U-shaped relationship between the number of disciplines and the number of citations.

We also use the Web of Science disciplinary classification (subject categories), which is very granular with 252 disciplines. We don't consider the subsample of papers with at most 2 disciplines and at least 4 references, given that we only obtain around 10,000 observations. We run the same regressions as before with all the papers having at least 4 references. In most of the regressions, we obtain non-significant coefficients. Given that we have 252 disciplines, perhaps the threshold of the number of references is insufficient to properly determine the disciplinary profile of the paper. To address this situation, we consider papers with at least 6 references and report the results from Tables 3.68 to 3.76. We obtain very similar results compared to the previous disciplinary classifications. We also try other thresholds (7, 8, 9, 10 references). Results are unchanged.

Wang et al. (2015) has found that the time citation window can change the results. We have used a 3-year citation window. Maybe the delay of recognition is more important for interdisciplinary papers. To deal with this issue, we consider a 5-year citation window. Naturally this reduces the number of observations: our database now covers the period 1999 to 2011. Since the 5-year number of citations is correlated at more than 0.95 with the 3-year number of citations, results are very similar. We only report the results of our regressions with the OST disciplinary classification in Table 3.48.

We also include authors' fixed effects in our regressions. We have more than 770,000 participations of 27,000 French authors between 1999 and 2013. We want to capture authors-specific unobserved and time-invariant factors that could influence the number of citations. Each author can publish more than one paper over the period. If an author is

involved in more than one project, then the observations are not necessarily independent. We decide to cluster both by author and research project. We use the command *reghdfe* in Stata which allows to incorporate a huge number of fixed effects in an OLS model. Our dependent variable is the logarithm of the number of citations⁹. The total number of participations of authors in our database is greater than 5,000,000 so we control approximately for 15% of the individual characteristics of authors. We first consider the OST and CWTS disciplinary nomenclature (Tables 3.92 and 3.94). For the Berger-Parker, the Simpson and the Shannon index, the results still go in the same direction as before. For the number of disciplines, we observe a positive effect on impact. However, we run the same regressions without fixed effects (Tables 3.93 and 3.95). The number of disciplines has still a positive effect. In fact, the type of econometric model seems to be the cause of the different results. Without fixed effects, we obtain a negative coefficient with the negative binomial model while we obtain a positive one with the OLS model. Anyway, the fixed effects don't seem to be the cause of the sign change. We finally consider the WoS disciplinary classification and report the results in Table 3.96. We remark that we obtain positive signs for the Simpson and the Shannon index coefficients, and also for the number of disciplines. We run the same regressions without fixed effects and we present the results in Table 3.97. We obtain positive signs for the Shannon-Wiener and the number of disciplines. Again, the econometric model seems to be the cause since we obtain negative signs for the Shannon coefficient with a negative binomial model for instance. Note that the Simpson index coefficient is positive in the model with fixed effects while negative in the model without fixed effects.

Finally, we consider the impact factor as the dependent variable. We estimate OLS regressions and report the results in Tables 3.49 and 3.50 (OST nomenclature), Tables 3.65 and 3.66 (CWTS nomenclature), and Tables 3.78 and 3.79 (WoS nomenclature). Sometimes we obtain opposite results when we include the maximum distance between the disciplines. In the correlation matrix (Table 3.67), the maximal distance and the indexes of interdisciplinarity are correlated at more than 0.5. Without the maximal distance, the results are similar to these obtained with the number of citations.

⁹ $\ln(1+\text{number of citations})$

7 Is interdisciplinarity increasing over time?

In the general introduction, we have presented two graphs indicating that interdisciplinarity, a form of complexity of research projects, is increasing over time. We deal with two databases to support this assertion. The first one contains around 400,000 research articles in academic journals published by at least a French researcher between 1999 and 2013. The second one is the full Web of Science between 2013 and 2017 with approximately 8,000,000 publications. Each database has its own advantages and disadvantages: the French one considers a sufficiently long period to observe trends but is geographically limited to one country. The second database has a huge number of observations all over the world, but on a shorter period.

This section¹⁰ aims at showing that interdisciplinarity is really increasing over time. We conduct different robustness checks to this purpose.

First, this observation does not depend on the choice of the disciplinary classification. We present in Figures 3.6 to 3.11 the evolution of interdisciplinarity according to four Hill indexes (Berger-Parker, Simpson, Shannon and number of disciplines) computed with the OST, CWTS and Web of Science disciplinary nomenclatures. The y -axis is in logarithmic scale. Whatever the database, the chosen index and the disciplinary classification, interdisciplinarity is increasing over time at a constant rate. We calculate the annual average growth rates with the full Web of Science database and report the results in Table 3.31. We observe that interdisciplinarity increases between 1% and 3.5% per year on average.

	OST(11 disc)	CWTS (35 disc)	WoS (252 disc)
Berger-Parker	0.88%	1.22%	1.71%
Simpson	1.42%	1.87%	2.54%
Shannon	1.81%	2.3%	2.90%
Number of disciplines	2.62%	3.14%	3.34%

Table 3.31: Annual average growth rate of interdisciplinarity, Web of Science database (8,000,000 observations, 2013-2017)

¹⁰This section investigates whether interdisciplinarity really increases over time.

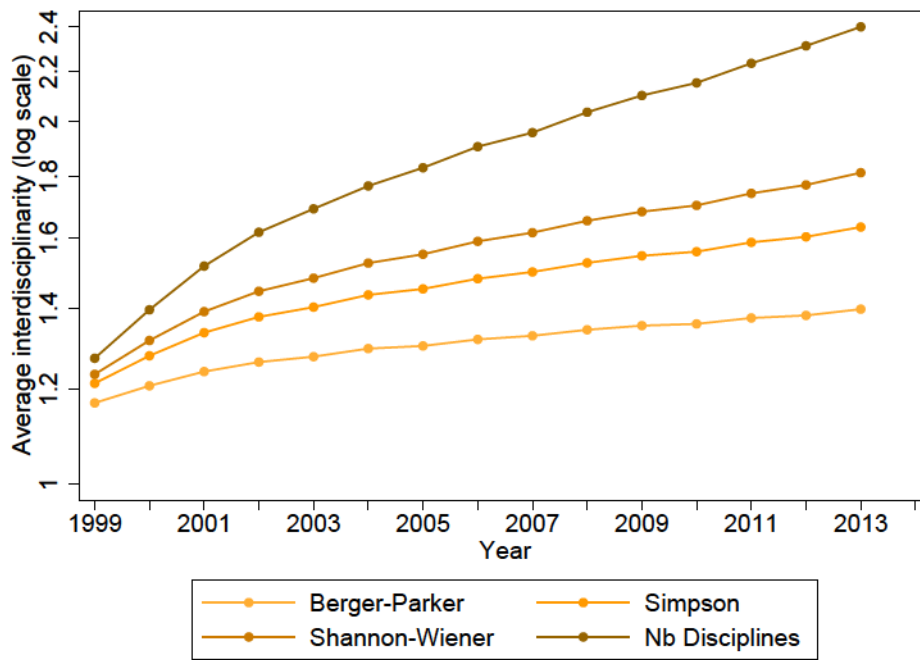


Figure 3.6: Evolution of the interdisciplinarity of research articles published by at least a French researcher between 1999 and 2013 (400,000 observations), OST disciplinary nomenclature.

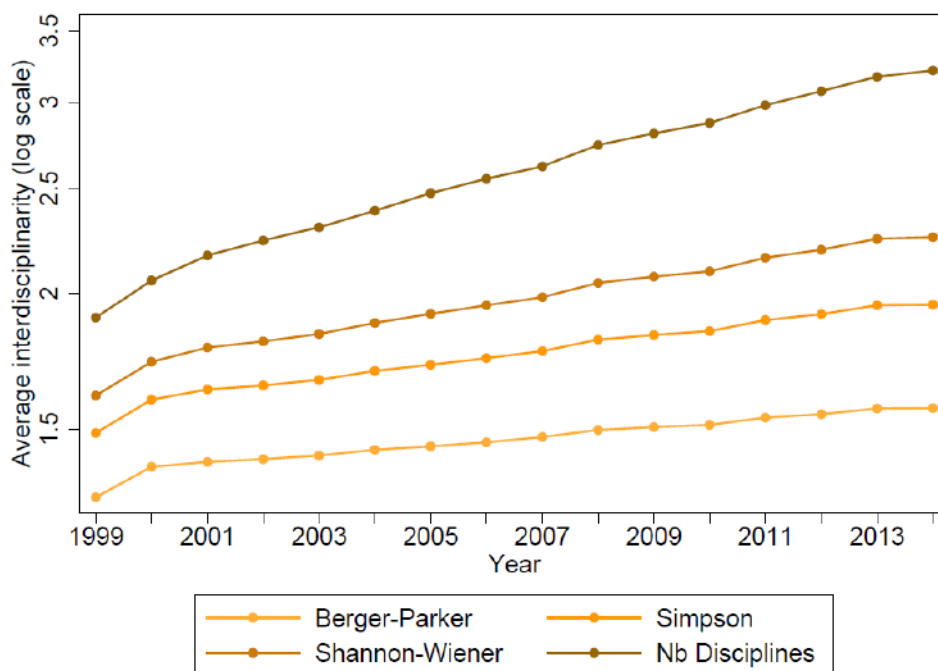


Figure 3.7: Evolution of the interdisciplinarity of research articles published by at least a French researcher between 1999 and 2013 (400,000 observations), CWTS disciplinary nomenclature.

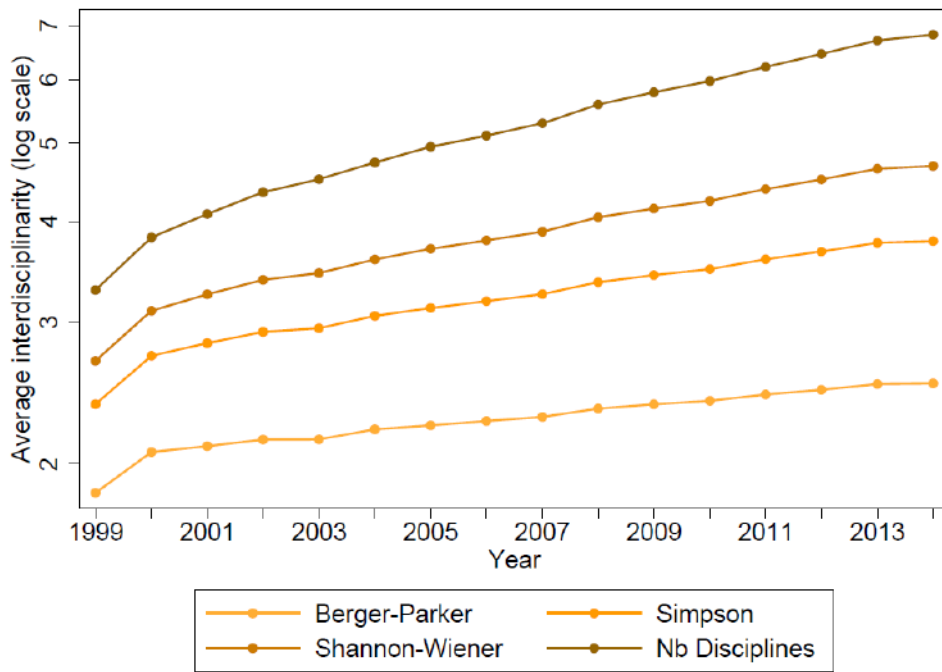


Figure 3.8: Evolution of the interdisciplinarity of research articles published by at least a French researcher between 1999 and 2013 (400,000 observations), WoS disciplinary nomenclature.

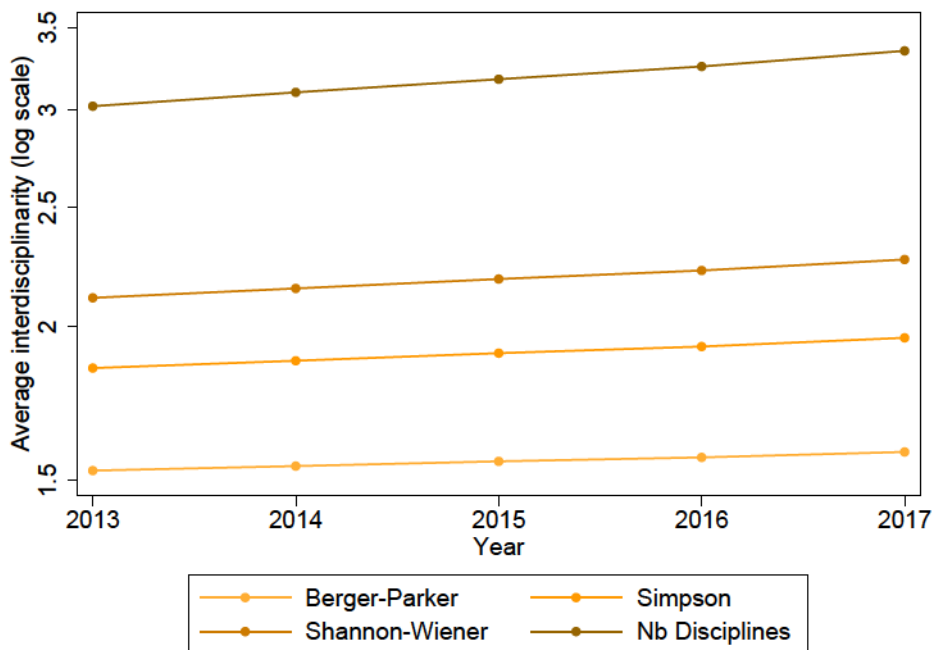


Figure 3.9: Evolution of the interdisciplinarity of research articles between 2013 and 2017, Web of Science database (8,000,000 observations), OST disciplinary nomenclature.

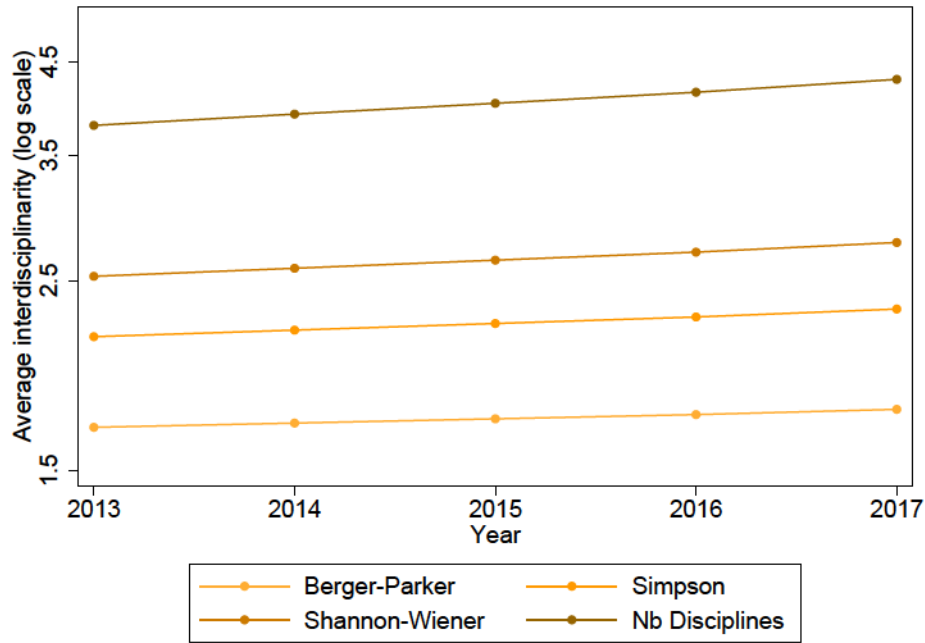


Figure 3.10: Evolution of the interdisciplinarity of research articles between 2013 and 2017, Web of Science database (8,000,000 observations), CWTS disciplinary nomenclature.

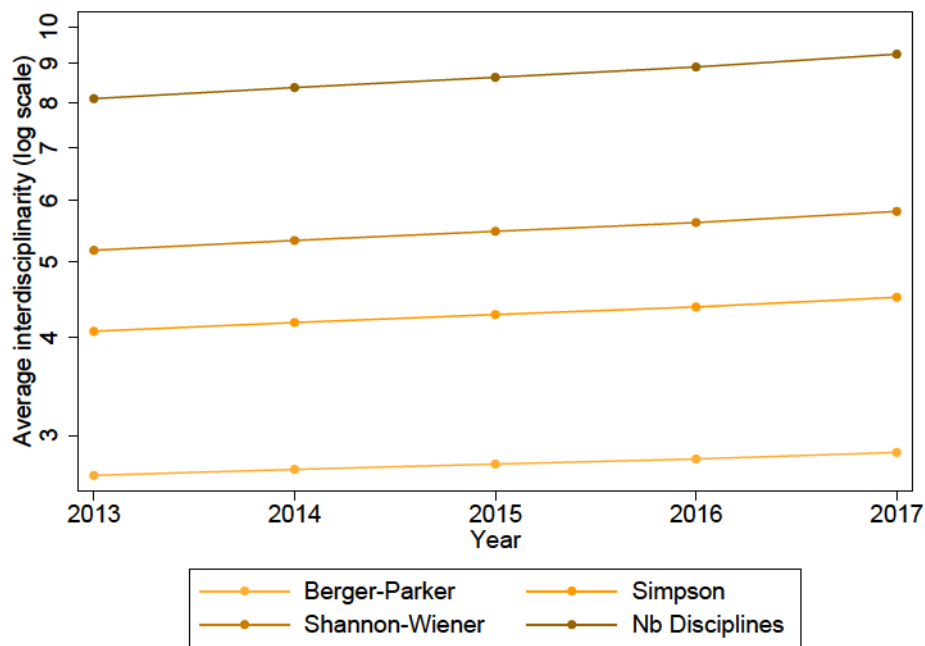


Figure 3.11: Evolution of the interdisciplinarity of research papers between 2013 and 2017, Web of Science database (8,000,000 observations), WoS disciplinary nomenclature.

We need to verify if the growth of interdisciplinarity that we observe is not an artefact. To calculate our complexity indexes, we associate to each reference the disciplines of the journal in which the paper has been published. A first source of bias for the increase of our interdisciplinary indexes could be that journals are attached to more and more disciplines over time. However, this does not hold here since our correspondance table between journals and disciplines is time invariant.

A second source of bias could be that the scientific production is becoming more and more egalitarian between the disciplines over time. Researchers may be more influenced by works coming from other fields and consequently cite more these papers. This could lead to an increase of the interdisciplinary indexes, while the complexity of problems researchers have to face remains unchanged. To deal with this issue, we compute disciplinary diversity indexes of the scientific production (Berger-Parker, Simpson and Shannon-Wiener indexes) and their annual growth rates (g_{BP} , g_{SI} and g_{SW}). We present the results in Table 3.32a (OST nomenclature), Table 3.32b (CWTS nomenclature) and Table 3.32c (WoS nomenclature). Since our database only covers five years, we need to be cautious about the conclusions. Nevertheless the results based on the different disciplinary classifications don't seem to validate this assumption. According to the OST nomenclature (Table 3.32a), we obtain an annual average growth rate between 0.58% and 1.37% for the disciplinary diversity of the scientific production over the period 2013-2017. However this observation is mainly due to the jump of the different indexes between 2016 and 2017, not a regular trend. If we consider the period 2013-2016, this is stable or even declining. With the CWTS nomenclature, the Berger-Parker and the Simpson indexes fluctuate while the Shannon index raises. Again, the increase of the disciplinary diversity of scientific production is mainly driven by the strong jump between 2016 and 2017, except for the Shannon index. According to the latter, even if scientific production is becoming more and more egalitarian between the disciplines over time, it is insufficient to explain the growth of this interdisciplinary index. Indeed, Table 3.32b indicates an annual growth rate of 1.18% while our complexity index grows at 2.3% per year on average. Finally, according to the Web of Science disciplinary classification (Table 3.32c), the Berger-Parker and the Simpson indexes fluctuate a lot but decrease on average while the Shannon index is stable over time. Or we have shown in Table 3.31 that the interdisciplinary indexes at the paper level increase at an annual rate between 1.71% and 3.34%. Thus we really can't conclude that the growth of interdisciplinary indexes is driven by the increase of the

disciplinary diversity of the scientific production.

A third source of bias is that in our database, the average number of references per paper is raising over time (Table 3.33), which could lead to a growth of our interdisciplinary indexes. To address this situation, we estimate tobit models where the dependant variables are the different interdisciplinary indexes expressed in logarithm, so as they belong to $[0; \ln(m)]$. In all the regressions, observations with the logarithm of interdisciplinarity equal to 0 (monodisciplinary papers) are left-censored. For the explanatory variables, we are mainly interested in the year dummies coefficients. We use 1999 as the reference category. If interdisciplinarity is really increasing, then the year dummies coefficients must be significant, positive and increasing in our regressions, even when we control for the number of references. If this is an artefact from the data, they will be non-significant. We successively run tobit regressions with the OST, CWTS and WoS disciplinary classifications and report the results from Tables 3.80 to 3.91. In an overwhelming majority of the regressions, the year dummies coefficients are positive, significant and growing over time. With the WoS disciplinary classification, and only when interdisciplinarity is measured with the number of disciplines (Table 3.91), we observe that the coefficients are negative and decreasing between 2000 and 2003 before increasing and becoming positive in the remaining period. Nevertheless the number of observations considerably raises between 1999 (3,705) and 2003 (24,414) in our database as shown in Table 3.33, potentially indicating a selection bias at the beginning of the period. Given that the results go in the expected direction for the other years, we can conclude that the growth of interdisciplinarity that we observe can't just be explained by the increase of the number of references that we are able to identify.

year	Berger Parker	g_{BP}	Simpson	g_{SI}	Shannon	g_{SW}
2013	3.002		5.972		7.777	
2014	3.019	+0.57%	5.995	+0.39%	7.783	+0.07%
2015	3.031	+0.40%	6.023	+0.47%	7.801	+0.23%
2016	2.992	-1.29%	5.957	-1.1%	7.764	-0.47%
2017	3.171	+5.98%	6.283	+5.47%	7.959	+2.51%
avg		+1.37%		+1.28%		+0.58%

(a) OST nomenclature

year	Berger Parker	g_{BP}	Simpson	g_{SI}	Shannon	g_{SW}
2013	3.757		9.484		17.03	
2014	3.805	+1.28%	9.586	+1.08%	17.08	+0.29%
2015	3.829	+0.63%	9.719	+1.39%	17.27	+1.11%
2016	3.795	-0.89%	9.707	-0.12%	17.35	+0.46%
2017	4.046	+6.61%	10.36	+6.73%	17.85	+2.88%
avg		+1.87%		+2.23%		+1.18%

(b) CWTS nomenclature

year	Berger Parker	g_{BP}	Simpson	g_{SI}	Shannon	g_{SW}
2013	33.54		100.5		145.1	
2014	30.61	-8.74%	97.70	-2.79%	143.5	-1.10%
2015	32.63	+6.57%	99.54	+1.89%	145.0	+1.05%
2016	30.78	-5.67%	99.29	-0.25%	144.4	-0.41%
2017	28.54	-7.28%	100.0	+0.72%	145.2	+0.55%
avg		-3.96%		-0.12%		+0.02%

(c) WoS nomenclature

Table 3.32: Disciplinary diversity of scientific production, Web of Science database (8,000,000 observations, 2013-2017)

8 Discussion

In chapter 2, we have theoretically shown that the productivity of research effort for each project $i = 1, \dots, g$ is

$$\alpha_i^* = \frac{\Omega_i}{\Gamma_i}.$$

	Observations	Min	Median	Max	Mean	S.D.
1999	3705	1	1	21	1.48	1.07
2000	12195	1	2	41	2.18	1.89
2001	18070	1	2	60	3.01	2.73
2002	22011	1	3	77	3.87	3.64
2003	24414	1	3	75	4.58	4.15
2004	26569	1	4	104	5.22	4.86
2005	28468	1	4	129	5.88	5.53
2006	29009	1	5	115	6.40	5.96
2007	30244	1	5	200	6.99	6.66
2008	32538	1	6	700	7.62	8.43
2009	33215	1	6	275	8.37	8.20
2010	34133	1	7	847	8.95	9.59
2011	34493	1	7	259	9.52	9.12
2012	33464	1	8	227	10.3	9.47
2013	32663	1	9	335	11.1	10.6
Total	395191	1	5	847	7.23	7.88

Table 3.33: Number of identified references per paper per year

Now let's suppose that research projects have similar characteristics ($\Omega_i = \Omega$ and $\Gamma_i = \Gamma \forall i$) which evolve over time. Let define α_t the productivity, Ω_t the team efficiency and Γ_t the project complexity at time t . Then

$$\alpha_t = \frac{\Omega_t}{\Gamma_t},$$

so that the research productivity growth evolves over time according to

$$g_\alpha = g_\Omega - g_\Gamma, \tag{3.5}$$

with $g_\alpha = \frac{\dot{\alpha}}{\alpha}$, $g_\Omega = \frac{\dot{\Omega}}{\Omega}$ and $g_\Gamma = \frac{\dot{\Gamma}}{\Gamma}$.

If the team efficiency growth rate g_Ω is lower than the complexity growth rate g_Γ , estimated between 1% and 3.5% per year, then research productivity is decreasing over time. If we reasonably assume $g_\Omega \geq 0$, our model can explain the decline of research productivity until an upper bound of 3.5% per year. Naturally research productivity may be affected by many effects, and interdisciplinarity is not the only existing form of complexity. Nevertheless our results don't seem unrealistic compared to the estimates of Bloom et al. (2017) reported in Table 3.34.

Scope	Average annual growth rate
Agregate economy	-5.3%
Moore's law	-6.8%
Agriculture (seeds)	-5.5%
New molecular entities	-3.5%
Disease mortality	-5.6%
Compustat firms	-11.1%

Table 3.34: Average annual growth rate of research productivity

Our study may overestimate the growth rate of complexity over time. Indeed we measure interdisciplinarity with Hill indexes of order β . From Proposition 16, the complexity indicator is the Hill index of order $1 - \sigma$. Or σ is not necessarily constant over time. Since our measures of interdisciplinarity are increasing and given that β is constant, this implies that the contributions of the different disciplines to the output become more and more egalitarian over time. Given that they depend on the vector ω and the value of σ (equation (3.3)), if only the vector ω changes over time (σ constant), then the order of our indicator of complexity remains the same for the whole period. Since we have represented the Hill index for some values of β from Figures 3.6 to 3.11, and that we systematically obtain a constant growth rate whatever β , then the complexity indicator (Hill index of order $1 - \sigma$) increases at a constant rate. On the contrary, if σ changes over time, then the order of the complexity index also changes. Similarly to the proof given in Appendix 8.1 part(3) of chapter 2, we can show that given a vector λ , the Hill index of order $1 - \sigma$ is increasing with σ . If we assume that σ is constant over time while it is decreasing (more complementarity between inputs), then we overestimate the complexity indicator and potentially its growth rate.

On the contrary, our study may underestimate the growth rate of complexity over time. In our database, we only observe publications in scientific journals. However many papers are rejected and not published. According to our theory, those with a higher Γ/Ω are more likely not to belong to our databases. We probably have a selection bias, and we may underestimate the average complexity and its growth rate.

In this chapter, we have shown that interdisciplinarity is increasing at the world level which could signal an increase of the complexity of the knowledge production function, leading to a decrease of the research productivity. Naturally comes the question of why

researchers deal with more complex problems if they experience lower production efficiency doing so. Here we discuss some hypotheses. A first explanation could be that knowledge is finite, so that researchers have first treated the most obvious ideas. The remaining ones are more complex, leading to more labor consumption and finally to a lower research productivity. Globalization can be a second explanation: the competition between researchers has considerably risen so that interesting problems in existing fields become rapidly scarce. Consequently scientists search for new areas and deal with more and more complex problems. New communication technologies (internet, search engines) can also explain the increase of complexity. They ease the diffusion and the contribution of knowledge from other fields. Research policies in favor of interdisciplinarity help to improve team efficiency and allow researchers to treat more complex problems than before. The increase of complexity is perhaps a temporary phenomenon: academic disciplines have evolved independently for several decades and the new communication technologies have allowed researchers to import concepts, skills and methods from other fields. When they will be integrated by the discipline, this phenomenon could stabilize. For further research, we want to verify the evolution of interdisciplinarity on a longer period, in particular before 1999. Nevertheless data are often incomplete before this date which complicates the study.

9 Conclusion

In this chapter, we have tested the relation between interdisciplinarity and impact. Our estimates show that papers with a higher index of interdisciplinarity have less citations and less chances to belong to the 10% or 5% most cited papers. The Berger-Parker, the Simpson and the Shannon indexes give similar results while the number of disciplines is sometimes non-significant or has a U-relationship with the measures of impact. We have conducted an important number of robustness checks compared to the literature. We have considered three disciplinary classifications, different measures of impact (3-year and 5-year number of citations, dummies if the paper belongs to the 10% or 5% most cited articles, journal impact factor). We have also integrated authors' fixed effects in some of our regressions. Globally, the results still prevail.

We have also shown that interdisciplinarity is increasing over time at a constant rate (between 1% and 3.5% per year), which suggests that the complexity of research projects

is growing. This empirical observation has not been established yet. We have discussed the effect on the research productivity and compared our results with [Bloom et al. \(2017\)](#).

For further research, we want to estimate the value of σ in order to be able to select the “right” order of the Hill index. We would like to test the effect of the competition given that we have theoretically shown in Chapter 2 that it can exert a positive or a negative effect on impact. Empirically we would like to see if the interdisciplinarity of French authors is increasing over time, or if they are still specialized and participate more and more to interdisciplinary projects. We would like to test if the productivity and the output at the author level are decreasing with the complexity of the projects in which a researcher is involved. This opens new questions but also new challenges, since it requires data at the author level.

10 Appendix

10.1 Proof of Proposition 16

The contribution of field of expertise j to the output of research project i at equilibrium is given in (3.3). As a reminder,

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

We deduce that

$$\sum_{s=1}^{m_i} \omega_{is}^{\frac{1}{1-\sigma_i}} = \frac{\omega_{i1}^{\frac{1}{1-\sigma_i}}}{\lambda_{i1}^*} = \dots = \frac{\omega_{im_i}^{\frac{1}{1-\sigma_i}}}{\lambda_{im_i}^*}.$$

For each field of expertise s , we obtain the following relation with any field of expertise j

$$\omega_{is} = \left(\frac{\lambda_{is}^*}{\lambda_{ij}^*} \right)^{1-\sigma_i} \omega_{ij}. \quad (3.6)$$

Summing (3.6) $\forall s = 1, \dots, m_i$, and since we have $\sum_{s=1}^{m_i} \omega_{is} = 1$, we have

$$\begin{aligned} 1 &= \sum_{s=1}^{m_i} \left(\frac{\lambda_{is}^*}{\lambda_{ij}^*} \right)^{1-\sigma_i} \omega_{ij} \\ &= \frac{\sum_{s=1}^{m_i} \lambda_{is}^{*1-\sigma_i}}{\lambda_{ij}^{*1-\sigma_i}} \omega_{ij}. \end{aligned}$$

We deduce that

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

Replacing ω_{ij} in the expression of Γ_i we finally obtain:

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

10.2 Proof of Proposition 17

Consider a vector $\lambda_i = (\lambda_{i1}, \lambda_{i2})$ with $\lambda_{i1} \geq \lambda_{i2}$. The Hill index of order β of vector λ_i is

$$H_i^\beta(\lambda_i) = (\lambda_{i1}^\beta + \lambda_{i2}^\beta)^{\frac{1}{1-\beta}}.$$

Given that $\lambda_{i1} + \lambda_{i2} = 1$, we can write the last expression as follows

$$H_i^\beta(\lambda_i) = (\lambda_{i1}^\beta + (1 - \lambda_{i1})^\beta)^{\frac{1}{1-\beta}}.$$

To compare two vectors λ_1 and λ_2 , it is enough to study the sign of the derivative of H_i^β with respect to λ_{i1} and to show that it is the same whatever the value of $\beta \geq 0$. We obtain

$$\frac{\partial H_i^\beta}{\partial \lambda_{i1}} = \frac{\beta}{1-\beta} (\lambda_{i1}^\beta + (1 - \lambda_{i1})^\beta)^{\frac{1}{1-\beta}-1} (\lambda_{i1}^{\beta-1} - (1 - \lambda_{i1})^{\beta-1}).$$

If $\beta \in [0; 1[$ then $\frac{\beta}{1-\beta} \geq 0$ and $(\lambda_{i1}^{\beta-1} - (1 - \lambda_{i1})^{\beta-1}) \leq 0$ so $\frac{\partial H_i^\beta}{\partial \lambda_{i1}} \leq 0$.

If $\beta \in]1; +\infty]$ then $\frac{\beta}{1-\beta} \leq 0$ and $(\lambda_{i1}^{\beta-1} - (1 - \lambda_{i1})^{\beta-1}) \geq 0$ so $\frac{\partial H_i^\beta}{\partial \lambda_{i1}} \leq 0$.

The sign of $\frac{\partial H_i^\beta}{\partial \lambda_{i1}}$ doesn't depend on β . Thus the ordering of two vectors λ_i ($i = 1, 2$) is the same whatever β .

10.3 More than 6 references

Table 3.35: Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 6 references, 1-2 disciplines, OST nomenclature)

	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Berger-Parker	-0.130*** (-7.56)	-0.110*** (-8.37)	-0.140*** (-7.89)	-0.130*** (-9.38)				
Simpson					-0.103*** (-7.24)	-0.0893*** (-8.46)	-0.117*** (-7.82)	-0.114*** (-9.95)
Number of authors	0.00155*** (9.51)	0.000822*** (8.98)	0.00154*** (9.47)	0.000804*** (8.88)	0.00155*** (9.51)	0.000822*** (8.99)	0.00154*** (9.45)	0.000800*** (8.87)
International Collaboration	0.437*** (45.64)	0.259*** (35.26)	0.438*** (45.64)	0.260*** (35.31)	0.437*** (45.65)	0.259*** (35.27)	0.438*** (45.67)	0.260*** (35.34)
Maximum h-index	0.0293*** (43.18)	0.0187*** (34.01)	0.0294*** (43.16)	0.0187*** (34.04)	0.0293*** (43.14)	0.0186*** (33.98)	0.0293*** (43.15)	0.0187*** (34.06)
Maximal distance			0.0281 (1.44)	0.0586*** (3.69)			0.0438* (2.20)	0.0759*** (4.63)
Impact Factor		0.0967*** (94.73)		0.0968*** (94.79)		0.0967*** (94.73)		0.0968*** (94.81)
Observations	111366	111366	111366	111366	111366	111366	111366	111366
Log-Likelihood	-360639.6	-345418.2	-360637.6	-345407.1	-360638.1	-345414.4	-360633.7	-345397.1
AIC	721345.1	690904.3	721343.2	690884.2	721342.3	690896.7	721335.4	690864.2
BIC	721662.6	691231.4	721670.3	691220.9	721659.8	691223.8	721662.5	691200.9

t statistics in parentheses

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

Table 3.36: Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 6 references, 1-2 disciplines, OST nomenclature)

	(3a)	(3b)	(3c)	(3d)	(4a)	(4b)	(4c)	(4d)
Shannon	-0.0966*** (-7.03)	-0.0849*** (-8.34)	-0.123*** (-8.03)	-0.125*** (-10.68)				
Number of disciplines					-0.0261* (-2.46)	-0.0232** (-2.87)	-0.0528** (-2.96)	-0.0897*** (-6.64)
Number of authors	0.00156*** (9.52)	0.000824*** (9.01)	0.00153*** (9.43)	0.000794*** (8.85)	0.00157*** (9.58)	0.000833*** (9.02)	0.00155*** (9.54)	0.000806*** (8.88)
International Collaboration	0.437*** (45.66)	0.259*** (35.27)	0.438*** (45.70)	0.260*** (35.38)	0.437*** (45.52)	0.259*** (35.16)	0.438*** (45.48)	0.260*** (35.28)
Maximum h-index	0.0293*** (43.14)	0.0186*** (33.98)	0.0293*** (43.19)	0.0187*** (34.12)	0.0294*** (43.26)	0.0187*** (34.08)	0.0295*** (43.41)	0.0189*** (34.45)
Maximal distance			0.0702*** (3.34)	0.105*** (6.02)			0.0606 (1.93)	0.150*** (5.94)
Impact Factor		0.0967*** (94.75)		0.0968*** (94.88)		0.0968*** (94.70)		0.0969*** (94.90)
Observations	111366	111366	111366	111366	111366	111366	111366	111366
Log-Likelihood	-360638.1	-345413.5	-360628.2	-345384.7	-360681.0	-345455.3	-360677.5	-345427.8
AIC	721342.2	690895.1	721324.4	690839.4	721428.0	690978.6	721423.1	690925.6
BIC	721659.7	691222.2	721651.5	691176.1	721745.5	691305.7	721750.2	691262.3

t statistics in parentheses

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

Table 3.37: Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 6 references, 1-10 disciplines, OST nomenclature)

	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Berger-Parker	-0.141*** (-16.79)	-0.0549*** (-7.92)	-0.166*** (-16.96)	-0.101*** (-12.06)				
Simpson					-0.0999*** (-16.88)	-0.0354*** (-7.30)	-0.132*** (-17.75)	-0.0799*** (-12.56)
Number of authors	0.00129*** (10.54)	0.000713*** (9.91)	0.00127*** (10.59)	0.000687*** (9.67)	0.00129*** (10.51)	0.000714*** (9.92)	0.00126*** (10.55)	0.000683*** (9.64)
International Collaboration	0.405*** (54.87)	0.257*** (43.50)	0.405*** (54.84)	0.257*** (43.71)	0.404*** (54.80)	0.257*** (43.48)	0.405*** (54.83)	0.257*** (43.74)
Maximum h-index	0.0294*** (50.81)	0.0188*** (38.78)	0.0295*** (51.12)	0.0188*** (39.32)	0.0294*** (50.69)	0.0188*** (38.79)	0.0294*** (51.04)	0.0188*** (39.32)
Maximal distance			0.0650*** (4.80)	0.115*** (9.93)			0.0979*** (6.75)	0.135*** (10.79)
Impact Factor		0.0923*** (103.78)		0.0924*** (104.11)		0.0923*** (103.64)		0.0924*** (104.10)
Observations	184620	184620	184620	184620	184620	184620	184620	184620
Log-Likelihood	-592861.3	-569880.1	-592834.9	-569775.6	-592850.2	-569885.4	-592797.3	-569758.1
AIC	1185788.6	1139828.3	1185737.8	1139621.2	1185766.4	1139838.8	1185662.5	1139586.1
BIC	1186122.7	1140172.6	1186082.1	1139975.6	1186100.5	1140183.1	1186006.8	1139940.5

t statistics in parentheses

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

Table 3.38: Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 6 references, 1-10 disciplines, OST nomenclature)

	(3a)	(3b)	(3c)	(3d)	(4a)	(4b)	(4c)	(4d)
Shannon	-0.0782*** (-14.97)	-0.0223*** (-5.30)	-0.123*** (-16.94)	-0.0757*** (-12.17)				
Number of disciplines					0.00436 (1.13)	0.0236*** (7.33)	0.0343*** (4.95)	0.0242*** (4.15)
Number of authors	0.00130*** (10.52)	0.000715*** (9.93)	0.00126*** (10.54)	0.000681*** (9.61)	0.00130*** (10.60)	0.000711*** (9.93)	0.00133*** (10.65)	0.000711*** (9.95)
International Collaboration	0.404*** (54.70)	0.257*** (43.44)	0.405*** (54.77)	0.257*** (43.73)	0.407*** (54.59)	0.257*** (43.54)	0.406*** (54.56)	0.257*** (43.52)
Maximum h-index	0.0294*** (50.77)	0.0188*** (38.88)	0.0294*** (51.11)	0.0188*** (39.39)	0.0300*** (51.79)	0.0191*** (39.51)	0.0299*** (51.55)	0.0191*** (39.52)
Maximal distance			0.134*** (8.31)	0.159*** (11.29)			-0.114*** (-5.30)	-0.00235 (-0.13)
Impact Factor		0.0923*** (103.58)		0.0924*** (104.16)		0.0928*** (103.97)		0.0928*** (103.72)
Observations	184620	184620	184620	184620	184620	184620	184620	184620
Log-Likelihood	-592887.9	-569901.4	-592807.4	-569758.6	-593069.8	-569864.6	-593038.7	-569864.6
AIC	1185841.9	1139870.9	1185682.9	1139587.2	1186205.6	1139797.1	1186145.4	1139799.1
BIC	1186176.1	1140215.1	1186027.2	1139941.6	1186539.8	1140141.4	1186489.7	1140153.5

t statistics in parentheses

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

Table 3.39: Negative binomial estimates for the effect of the number of disciplines (categorical variable) on the number of citations (papers with at least 6 references, 1-10 disciplines, OST nomenclature)

	(4a)	(4b)	(4c)	(4d)
1 discipline	0 (.)	0 (.)	0 (.)	0 (.)
2 disciplines	-0.0336** (-3.24)	-0.0342*** (-4.35)	-0.0383* (-2.54)	-0.0945*** (-8.21)
3 disciplines and more	-0.0282* (-2.52)	0.0257** (2.84)	-0.0370 (-1.58)	-0.0857*** (-4.61)
Number of authors	0.00131*** (10.62)	0.000714*** (10.02)	0.00131*** (10.60)	0.000702*** (9.84)
International Collaboration	0.406*** (54.63)	0.257*** (43.58)	0.406*** (54.61)	0.257*** (43.66)
Maximum h-index	0.0300*** (51.77)	0.0191*** (39.62)	0.0300*** (51.73)	0.0192*** (39.89)
Maximal distance			0.0110 (0.44)	0.139*** (6.87)
Impact Factor		0.0928*** (103.88)		0.0929*** (103.99)
Observations	184620	184620	184620	184620
Log-Likelihood	-593058.9	-569852.8	-593058.7	-569813.8
AIC	1186185.8	1139775.6	1186187.4	1139699.7
BIC	1186530.1	1140130.0	1186541.8	1140064.2

t statistics in parentheses

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

Table 3.40: Negative binomial estimates for the effect of the logarithm of interdisciplinarity on the number of citations (papers with at least 6 references, 1-10 disciplines, OST nomenclature)

	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Number of citations								
log Berger-Parker	-0.224*** (-16.78)	-0.0924*** (-8.51)	-0.279*** (-17.44)	-0.180*** (-13.35)				
log Simpson					-0.177*** (-16.46)	-0.0692*** (-8.01)	-0.256*** (-18.25)	-0.171*** (-14.45)
Number of authors	0.00129*** (10.52)	0.000713*** (9.91)	0.00126*** (10.56)	0.000683*** (9.64)	0.00130*** (10.50)	0.000714*** (9.92)	0.00126*** (10.52)	0.000678*** (9.59)
International Collaboration	0.405*** (54.89)	0.257*** (43.51)	0.406*** (54.90)	0.257*** (43.77)	0.404*** (54.81)	0.257*** (43.49)	0.405*** (54.90)	0.257*** (43.82)
Maximum h-index	0.0294*** (50.82)	0.0188*** (38.77)	0.0295*** (51.19)	0.0188*** (39.36)	0.0294*** (50.75)	0.0188*** (38.79)	0.0294*** (51.16)	0.0188*** (39.39)
Maximal distance			0.0832*** (5.97)	0.130*** (10.89)			0.125*** (8.31)	0.161*** (12.28)
Impact Factor		0.0922*** (103.72)		0.0924*** (104.14)		0.0922*** (103.54)		0.0924*** (104.13)
Observations	184620	184620	184620	184620	184620	184620	184620	184620
Log-Likelihood	-592846.4	-569872.3	-592805.7	-569746.0	-592841.3	-569876.2	-592762.4	-569712.0
AIC	1185758.8	1139812.5	1185679.4	1139562.0	1185748.5	1139820.4	1185592.9	1139494.1
BIC	1186092.9	1140156.8	1186023.7	1139916.4	1186082.7	1140164.6	1185937.2	1139848.5

t statistics in parentheses

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

Table 3.41: Negative binomial estimates for the effect of the logarithm of interdisciplinarity on the number of citations (papers with at least 6 references, 1-10 disciplines, OST nomenclature)

	(3a)	(3b)	(3c)	(3d)
log Shannon	-0.145*** (-14.27)	-0.0481*** (-5.94)	-0.267*** (-17.70)	-0.188*** (-14.77)
Number of authors	0.00130*** (10.51)	0.000716*** (9.94)	0.00125*** (10.51)	0.000674*** (9.53)
International Collaboration	0.404*** (54.71)	0.257*** (43.44)	0.405*** (54.86)	0.257*** (43.83)
Maximum h-index	0.0295*** (50.85)	0.0188*** (38.88)	0.0294*** (51.27)	0.0188*** (39.50)
Maximal distance			0.178*** (10.33)	0.204*** (13.48)
Impact Factor		0.0923*** (103.49)		0.0924*** (104.22)
Observations	184620	184620	184620	184620
Log-Likelihood	-592886.7	-569894.4	-592765.7	-569695.4
AIC	1185839.4	1139856.8	1185599.4	1139460.8
BIC	1186173.6	1140201.1	1185943.7	1139815.3

t statistics in parentheses

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

Table 3.42: Negative binomial estimates for the effect of interdisciplinarity (square term) on the number of citations (papers with at least 6 references, 1-10 disciplines, OST nomenclature)

	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Berger-Parker	-0.343*** (-7.92)	-0.214*** (-6.18)	-0.467*** (-10.06)	-0.417*** (-10.88)				
Berger-Parker ²	0.0605*** (4.95)	0.0478*** (4.80)	0.0875*** (6.91)	0.0923*** (8.76)				
Simpson					-0.200*** (-7.92)	-0.128*** (-6.40)	-0.312*** (-11.28)	-0.280*** (-12.36)
Simpson ²					0.0252*** (4.25)	0.0235*** (4.91)	0.0434*** (7.07)	0.0483*** (9.57)
Number of authors	0.00129*** (10.52)	0.000713*** (9.91)	0.00126*** (10.56)	0.000682*** (9.63)	0.00130*** (10.50)	0.000714*** (9.93)	0.00126*** (10.51)	0.000678*** (9.60)
International Collaboration	0.405*** (54.92)	0.257*** (43.53)	0.406*** (54.94)	0.257*** (43.81)	0.404*** (54.85)	0.257*** (43.51)	0.406*** (54.95)	0.257*** (43.86)
Maximum h-index	0.0294*** (50.83)	0.0188*** (38.80)	0.0295*** (51.24)	0.0189*** (39.47)	0.0294*** (50.69)	0.0188*** (38.81)	0.0294*** (51.15)	0.0188*** (39.49)
Maximal distance			0.0862*** (6.16)	0.138*** (11.37)			0.124*** (8.27)	0.165*** (12.55)
Impact Factor		0.0922*** (103.75)		0.0924*** (104.20)		0.0922*** (103.60)		0.0924*** (104.20)
Observations	184620	184620	184620	184620	184620	184620	184620	184620
Log-Likelihood	-592843.4	-569866.5	-592800.6	-569729.2	-592835.6	-569870.0	-592757.9	-569699.1
AIC	1185754.9	1139803.1	1185671.1	1139530.4	1185739.2	1139810.0	1185585.8	1139470.2
BIC	1186099.2	1140157.5	1186025.5	1139894.9	1186083.5	1140164.4	1185940.2	1139834.7

t statistics in parentheses

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

Table 3.43: Negative binomial estimates for the effect of interdisciplinarity (square term) on the number of citations (papers with at least 6 references, 1-10 disciplines, OST nomenclature)

	(3a)	(3b)	(3c)	(3d)
Shannon	-0.145*** (-6.89)	-0.0906*** (-5.35)	-0.304*** (-12.23)	-0.278*** (-13.64)
Shannon ²	0.0156*** (3.44)	0.0158*** (4.29)	0.0384*** (7.94)	0.0431*** (10.83)
Number of authors	0.00130*** (10.52)	0.000717*** (9.95)	0.00125*** (10.50)	0.000674*** (9.54)
International Collaboration	0.404*** (54.74)	0.257*** (43.46)	0.406*** (54.91)	0.257*** (43.88)
Maximum h-index	0.0294*** (50.75)	0.0189*** (38.90)	0.0294*** (51.25)	0.0188*** (39.61)
Maximal distance			0.180*** (10.43)	0.209*** (13.88)
Impact Factor		0.0923*** (103.54)		0.0924*** (104.28)
Observations	184620	184620	184620	184620
Log-Likelihood	-592877.7	-569888.6	-592754.6	-569678.1
AIC	1185823.5	1139847.2	1185579.2	1139428.1
BIC	1186167.7	1140201.6	1185933.7	1139792.7

t statistics in parentheses

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

Table 3.44: Logit estimates for the effect of interdisciplinarity on the probability that the paper belongs to the top 10% most cited papers (papers with at least 6 references, 1-10 disciplines, OST nomenclature)

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Berger-Parker	-0.153*** (-8.51)	-0.215*** (-10.51)						
Simpson			-0.104*** (-8.37)	-0.169*** (-11.22)				
Shannon					-0.0767*** (-7.24)	-0.159*** (-11.17)		
Number of disciplines							0.0226*** (3.52)	0.0470*** (3.97)
Number of authors	0.000580*** (11.36)	0.000558*** (10.92)	0.000582*** (11.39)	0.000554*** (10.84)	0.000585*** (11.45)	0.000552*** (10.80)	0.000585*** (11.44)	0.000593*** (11.57)
International Collaboration	0.613*** (46.87)	0.615*** (46.96)	0.613*** (46.85)	0.614*** (46.94)	0.613*** (46.83)	0.614*** (46.92)	0.614*** (46.93)	0.614*** (46.92)
Maximum h-index	0.0412*** (48.10)	0.0413*** (48.23)	0.0412*** (48.04)	0.0413*** (48.13)	0.0412*** (48.13)	0.0413*** (48.15)	0.0419*** (48.92)	0.0418*** (48.82)
Maximal distance		0.146*** (6.46)		0.186*** (7.73)		0.233*** (8.75)		-0.0906* (-2.45)
Observations	184620	184620	184620	184620	184620	184620	184620	184620
Log-Likelihood	-81893.4	-81872.5	-81894.7	-81864.9	-81903.7	-81865.6	-81924.1	-81921.1
AIC	163850.7	163811.1	163853.4	163795.9	163871.4	163797.2	163912.1	163908.1
BIC	164174.8	164145.2	164177.4	164130.0	164195.4	164131.4	164236.2	164242.3

t statistics in parentheses

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

Table 3.45: Logit estimates for the effect of interdisciplinarity on the probability that the paper belongs to the top 5% most cited papers (papers with at least 6 references, 1-10 disciplines, OST nomenclature)

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Berger-Parker	-0.222*** (-9.20)	-0.318*** (-11.57)						
Simpson			-0.150*** (-9.04)	-0.249*** (-12.38)				
Shannon					-0.109*** (-7.77)	-0.234*** (-12.35)		
Number of disciplines							0.0299*** (3.60)	0.0494** (3.21)
Number of authors	0.000605*** (11.64)	0.000575*** (11.03)	0.000607*** (11.68)	0.000570*** (10.93)	0.000612*** (11.77)	0.000568*** (10.89)	0.000610*** (11.74)	0.000616*** (11.82)
International Collaboration	0.696*** (40.37)	0.698*** (40.45)	0.696*** (40.34)	0.697*** (40.43)	0.695*** (40.33)	0.697*** (40.41)	0.696*** (40.41)	0.696*** (40.40)
Maximum h-index	0.0465*** (44.59)	0.0467*** (44.72)	0.0465*** (44.53)	0.0466*** (44.61)	0.0466*** (44.62)	0.0466*** (44.62)	0.0473*** (45.44)	0.0473*** (45.38)
Maximal distance		0.220*** (7.52)		0.276*** (8.91)		0.345*** (10.04)		-0.0721 (-1.49)
Observations	184620	184620	184620	184620	184620	184620	184620	184620
Log-Likelihood	-54471.9	-54443.7	-54473.6	-54434.2	-54484.8	-54434.9	-54509.1	-54508.0
AIC	109007.8	108953.5	109011.2	108934.4	109033.6	108935.8	109082.2	109081.9
BIC	109331.8	109287.7	109335.2	109268.5	109357.6	109270.0	109406.2	109416.1

t statistics in parentheses

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

Table 3.46: Negative binomial estimates for the effect of the number of disciplines (categorical variable) on the probability that the paper belongs to the top 5% and top 10% most cited papers (papers with at least 6 references, 1-10 disciplines, OST nomenclature)

	Top90	Top90	Top95	Top95
1 discipline	0 (.)	0 (.)	0 (.)	0 (.)
2 disciplines	-0.0617*** (-3.61)	-0.138*** (-5.27)	-0.0592** (-2.69)	-0.179*** (-5.26)
3 disciplines and more	-0.00853 (-0.48)	-0.148*** (-3.67)	-0.00166 (-0.07)	-0.221*** (-4.20)
Number of authors	0.000590*** (11.53)	0.000582*** (11.38)	0.000616*** (11.83)	0.000606*** (11.63)
International Collaboration	0.613*** (46.90)	0.614*** (46.95)	0.696*** (40.38)	0.697*** (40.44)
Maximum h-index	0.0417*** (48.80)	0.0419*** (48.93)	0.0472*** (45.33)	0.0474*** (45.50)
Maximal distance		0.174*** (3.86)		0.274*** (4.64)
Observations	184620	184620	184620	184620
Log-Likelihood	-81920.9	-81913.4	-54509.7	-54498.9
AIC	163907.7	163894.9	109085.3	109065.9
BIC	164241.9	164239.1	109419.5	109410.2

t statistics in parentheses

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

Table 3.47: Negative binomial estimates for the effect of interdisciplinarity on the 5-year number of citations (papers with at least 4 references, 1- 10 disciplines, OST nomenclature)

	(1d)	(2d)	(3d)	(4d)
Berger-Parker	-0.0708*** (-9.56)			
Simpson		-0.0541*** (-9.56)		
Shannon			-0.0454*** (-8.23)	
Number of disciplines				0.0471*** (8.24)
Number of authors	0.000460*** (7.78)	0.000458*** (7.76)	0.000459*** (7.76)	0.000481*** (8.18)
International Collaboration	0.227*** (43.69)	0.227*** (43.71)	0.227*** (43.70)	0.227*** (43.69)
Maximum h-index	0.0179*** (41.60)	0.0179*** (41.67)	0.0180*** (41.84)	0.0181*** (41.58)
Maximal distance	0.152*** (14.94)	0.165*** (14.97)	0.172*** (14.16)	-0.00937 (-0.61)
Impact Factor	0.109*** (138.34)	0.109*** (138.27)	0.109*** (138.28)	0.109*** (138.22)
Observations	222688	222688	222688	222688
Log-Likelihood	-782045.8	-782043.0	-782060.0	-782036.4
AIC	1564157.6	1564151.9	1564186.0	1564138.7
BIC	1564497.9	1564492.3	1564526.3	1564479.1

t statistics in parentheses

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

Table 3.48: Negative binomial estimates for the effect of interdisciplinarity on the 5-year number of citations (papers with at least 4 references, 1- 10 disciplines, OST nomenclature)

	(1d)	(2d)	(3d)	(4d)
Berger-Parker	-0.0708*** (-9.56)			
Simpson		-0.0541*** (-9.56)		
Shannon			-0.0454*** (-8.23)	
Number of disciplines				0.0471*** (8.24)
Number of authors	0.000460*** (7.78)	0.000458*** (7.76)	0.000459*** (7.76)	0.000481*** (8.18)
International Collaboration	0.227*** (43.69)	0.227*** (43.71)	0.227*** (43.70)	0.227*** (43.69)
Maximum h-index	0.0179*** (41.60)	0.0179*** (41.67)	0.0180*** (41.84)	0.0181*** (41.58)
Maximal distance	0.152*** (14.94)	0.165*** (14.97)	0.172*** (14.16)	-0.00937 (-0.61)
Impact Factor	0.109*** (138.34)	0.109*** (138.27)	0.109*** (138.28)	0.109*** (138.22)
Observations	222688	222688	222688	222688
Log-Likelihood	-782045.8	-782043.0	-782060.0	-782036.4
AIC	1564157.6	1564151.9	1564186.0	1564138.7
BIC	1564497.9	1564492.3	1564526.3	1564479.1

t statistics in parentheses

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

Table 3.49: OLS estimates for the effect of interdisciplinarity on the journal impact (papers with at least 4 references, 1-10 disciplines, OST nomenclature)

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Berger-Parker	-0.448*** (-15.88)	-0.424*** (-12.45)						
Simpson			-0.346*** (-17.69)	-0.369*** (-14.28)				
Shannon					-0.283*** (-16.70)	-0.336*** (-13.39)		
Number of disciplines							-0.0147 (-1.16)	0.249*** (10.69)
Number of authors	0.00237*** (24.53)	0.00238*** (24.52)	0.00237*** (24.56)	0.00236*** (24.40)	0.00238*** (24.66)	0.00236*** (24.41)	0.00240*** (24.70)	0.00248*** (25.21)
International Collaboration	1.563*** (64.94)	1.562*** (64.91)	1.561*** (64.90)	1.562*** (64.90)	1.561*** (64.86)	1.561*** (64.87)	1.562*** (64.89)	1.560*** (64.81)
Maximum h-index	0.125*** (54.39)	0.125*** (54.34)	0.125*** (54.20)	0.125*** (54.20)	0.125*** (54.27)	0.125*** (54.27)	0.127*** (55.24)	0.126*** (54.82)
Maximal distance		-0.0571 (-1.34)		0.0657 (1.41)		0.149** (2.86)		-0.951*** (-14.62)
Observations	250163	250163	250163	250163	250163	250163	250163	250163
Log-Likelihood	-789828.4	-789827.5	-789803.7	-789802.6	-789817.8	-789813.4	-789937.0	-789836.4
AIC	1579718.9	1579719.0	1579669.4	1579669.3	1579697.5	1579690.8	1579936.0	1579736.8
BIC	1580042.2	1580052.8	1579992.7	1580003.0	1580020.8	1580024.5	1580259.3	1580070.5

t statistics in parentheses

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

Table 3.50: OLS estimates for the effect of the number of disciplines (categorical variable) on the journal impact factor (papers with at least 4 references, 1-10 disciplines, OST nomenclature)

	(1)	(2)
1 discipline	0 (.)	0 (.)
2 disciplines	0.296*** (9.92)	0.893*** (18.47)
3 and more disciplines	-0.0208 (-0.66)	1.058*** (15.49)
Number of authors	0.00239*** (24.65)	0.00246*** (24.94)
International Collaboration	1.560*** (64.86)	1.553*** (64.64)
Maximum h-index	0.127*** (55.24)	0.125*** (54.43)
Maximal distance		-1.358*** (-17.86)
Observations	250163	250163
Log-Likelihood	-789852.4	-789708.7
AIC	1579768.8	1579483.5
BIC	1580102.5	1579827.6

t statistics in parentheses

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

10.4 CWTS nomenclature

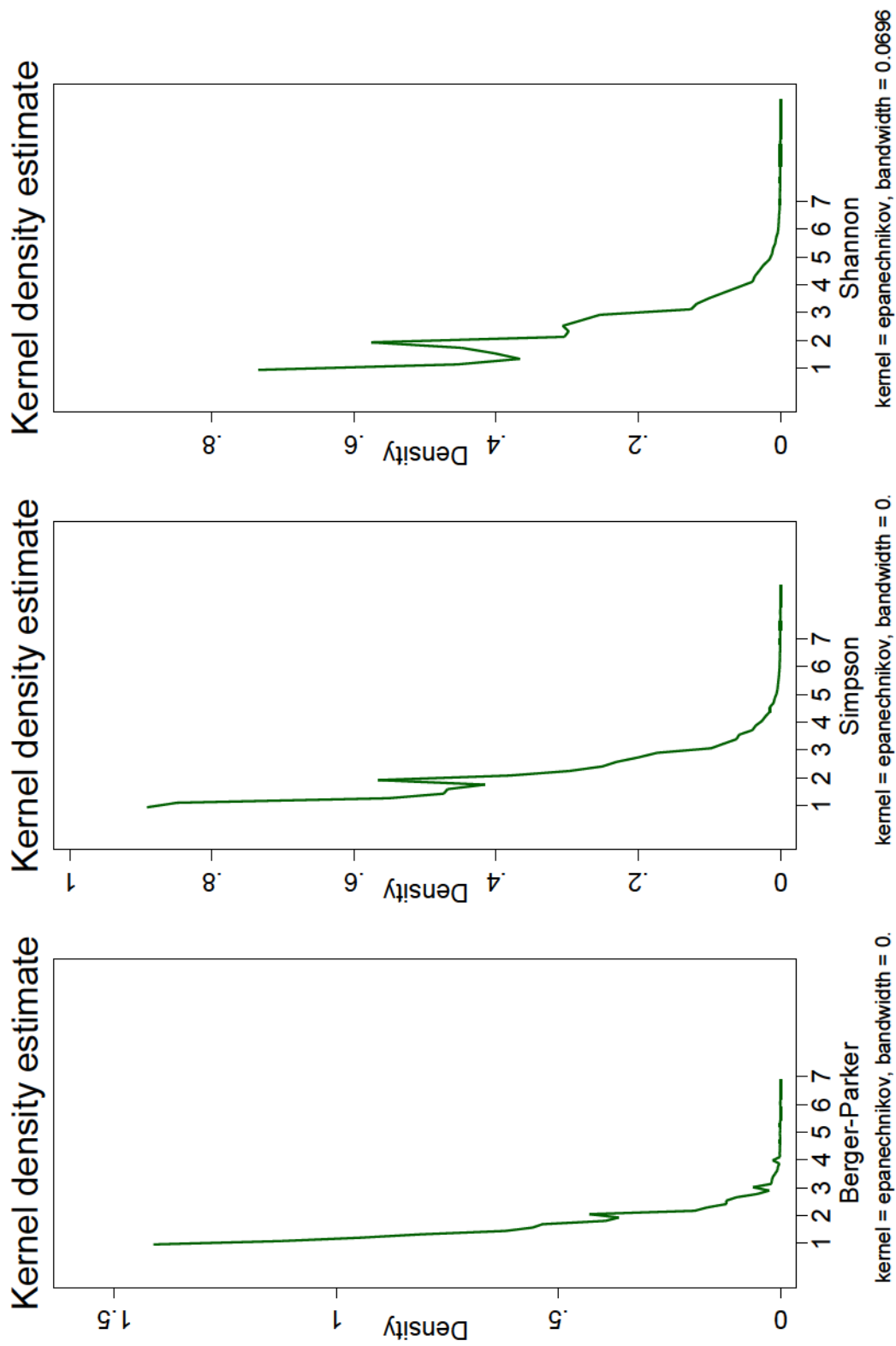


Figure 3.12: Distribution of Interdisciplinarity according to the Berger-Parker, the Simpson and the Shannon indexes (CWTS nomenclature, papers with at least 4 references)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 Number of citations	1													
2 Papers in the top 10%	0.53	1												
3 Papers in the top 5%	0.57	0.71	1											
4 Impact Factor	0.36	0.29	0.27	1										
5 Number of references	0.23	0.17	0.14	0.15	1									
6 Berger-Parker	-0.061	-0.050	-0.039	-0.059	-0.22	1								
7 Simpson	-0.071	-0.058	-0.045	-0.064	-0.25	0.95	1							
8 Shannon	-0.077	-0.064	-0.049	-0.067	-0.27	0.90	0.98	1						
9 Number of disciplines	-0.063	-0.058	-0.041	-0.054	-0.19	0.64	0.77	0.87	1					
10 Number of authors	0.10	0.065	0.070	0.084	0.043	0.045	0.045	0.042	0.033	1				
11 Number of different organisations	0.21	0.14	0.14	0.17	0.20	-0.023	-0.032	-0.040	-0.035	0.86	1			
12 International Collaboration	0.18	0.16	0.13	0.20	0.22	-0.091	-0.11	-0.12	-0.12	0.14	0.30	1		
13 Maximum h-index	0.16	0.097	0.091	0.14	0.40	-0.12	-0.13	-0.14	-0.11	0.083	0.21	0.17	1	
14 Maximal distance	-0.041	-0.052	-0.031	-0.055	-0.097	0.43	0.53	0.61	0.74	0.031	0.0035	-0.10	-0.024	1

Table 3.51: Correlation matrix (papers with at least four references, 1-2 disciplines, CWTS nomenclature)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 Number of citations	1															
2 Papers in the top 10%	0.44	1														
3 Papers in the top 5%	0.47	0.70	1													
4 Impact Factor	0.35	0.30	0.29	1												
5 Number of references	0.29	0.16	0.15	0.17	1											
6 Berger-Parker	-0.031	-0.037	-0.035	-0.029	-0.061	1										
7 Simpson	-0.022	-0.032	-0.030	-0.017	0.13	0.94	1									
8 Shannon	-0.0058	-0.019	-0.019	0.00030	0.23	0.86	0.97	1								
9 Number of disciplines	0.049	0.024	0.021	0.047	0.45	0.63	0.79	0.90	1							
10 Number of authors	0.085	0.053	0.054	0.042	-0.0046	-0.020	-0.021	-0.019	-0.0094	1						
11 Number of different organisations	0.16	0.11	0.11	0.096	0.056	-0.039	-0.038	-0.033	-0.010	0.91	1					
12 International Collaboration	0.11	0.12	0.10	0.13	0.052	-0.042	-0.041	-0.037	-0.019	0.088	0.20	1				
13 Maximum h-index	0.12	0.088	0.085	0.13	0.26	0.0077	0.037	0.068	0.14	0.081	0.16	0.11	1			
14 Minimal distance	-0.024	-0.018	-0.0056	-0.088	-0.16	-0.23	-0.29	-0.33	-0.35	0.041	0.063	-0.042	-0.026	1		
15 Mean distance	-0.021	-0.023	-0.013	-0.089	0.087	0.29	0.34	0.37	0.40	0.040	0.053	-0.054	0.020	0.49	1	
16 Maximal distance	-0.011	-0.020	-0.015	-0.047	0.17	0.45	0.52	0.58	0.62	0.024	0.023	-0.047	0.047	0.15	0.89	1

Table 3.52: Correlation matrix (papers with at least four references, 1-35 disciplines, CWTS nomenclature)

Table 3.53: Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 4 references, 1-2 disciplines, CWTS nomenclature)

	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Berger-Parker	-0.128*** (-7.78)	-0.0957*** (-7.01)	-0.127*** (-7.29)	-0.114*** (-7.68)				
Simpson					-0.107*** (-7.90)	-0.0790*** (-7.40)	-0.114*** (-7.69)	-0.103*** (-8.60)
Number of authors	0.00220*** (11.67)	0.00114*** (8.93)	0.00220*** (11.68)	0.00114*** (8.89)	0.00220*** (11.66)	0.00114*** (8.95)	0.00220*** (11.66)	0.00114*** (8.89)
International Collaboration	0.437*** (45.75)	0.284*** (38.60)	0.437*** (45.72)	0.284*** (38.62)	0.437*** (45.77)	0.284*** (38.62)	0.437*** (45.76)	0.284*** (38.65)
Maximum h-index	0.0310*** (40.37)	0.0204*** (33.50)	0.0310*** (40.37)	0.0204*** (33.53)	0.0310*** (40.39)	0.0204*** (33.51)	0.0310*** (40.39)	0.0204*** (33.54)
Maximal distance			-0.00311 (-0.17)	0.0443** (3.04)			0.0164 (0.87)	0.0624*** (4.11)
Impact Factor		0.0974*** (78.47)		0.0974*** (78.45)		0.0974*** (78.47)		0.0974*** (78.46)
Observations	124227	124227	124227	124227	124227	124227	124227	124227
Log-Likelihood	-382210.2	-367596.6	-382210.2	-367590.1	-382202.9	-367592.7	-382202.3	-367581.0
AIC	764534.5	735309.2	764536.4	735298.2	764519.8	735301.5	764520.5	735279.9
BIC	765089.1	735873.6	765100.8	735872.2	765074.4	735865.8	765084.9	735854.0

t statistics in parentheses

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

Table 3.54: Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 4 references, 1-2 disciplines, CWTS nomenclature)

	(3a)	(3b)	(3c)	(3d)	(4a)	(4b)	(4c)	(4d)
Shannon	-0.101*** (-7.80)	-0.0725*** (-7.23)	-0.119*** (-7.95)	-0.110*** (-9.10)				
Number of disciplines					-0.0167 (-1.68)	-0.00679 (-0.88)	0.0159 (1.01)	-0.0201 (-1.62)
Number of authors	0.00220*** (11.65)	0.00114*** (8.96)	0.00219*** (11.64)	0.00113*** (8.88)	0.00221*** (11.70)	0.00115*** (9.00)	0.00222*** (11.75)	0.00114*** (8.97)
International Collaboration	0.437*** (45.78)	0.284*** (38.62)	0.438*** (45.78)	0.285*** (38.68)	0.437*** (45.63)	0.284*** (38.51)	0.436*** (45.58)	0.284*** (38.55)
Maximum h-index	0.0310*** (40.43)	0.0204*** (33.54)	0.0310*** (40.46)	0.0204*** (33.60)	0.0312*** (40.50)	0.0205*** (33.61)	0.0311*** (40.35)	0.0205*** (33.58)
Maximal distance			0.0406* (2.05)	0.0859*** (5.30)			-0.0742** (-2.77)	0.0302 (1.39)
Impact Factor		0.0973*** (78.47)		0.0974*** (78.48)		0.0974*** (78.48)		0.0975*** (78.43)
Observations	124227	124227	124227	124227	124227	124227	124227	124227
Log-Likelihood	-382201.3	-367593.5	-382197.9	-367574.3	-382254.8	-367629.0	-382248.6	-367627.7
AIC	764516.6	735303.1	764511.8	735266.6	764623.7	735374.0	764613.2	735373.3
BIC	765071.2	735867.4	765076.2	735840.7	765178.3	735938.3	765177.5	735947.4

t statistics in parentheses

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

Table 3.55: Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)

	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Berger-Parker	-0.0798*** (-13.62)	-0.0359*** (-7.17)	-0.109*** (-15.23)	-0.0744*** (-11.95)				
Simpson					-0.0496*** (-12.28)	-0.0182*** (-5.35)	-0.0770*** (-14.51)	-0.0510*** (-11.32)
Number of authors	0.00112*** (13.45)	0.000668*** (12.17)	0.00111*** (13.59)	0.000656*** (12.15)	0.00113*** (13.47)	0.000669*** (12.19)	0.00112*** (13.63)	0.000659*** (12.24)
International Collaboration	0.414*** (63.26)	0.276*** (51.36)	0.415*** (63.40)	0.276*** (51.66)	0.414*** (63.22)	0.276*** (51.33)	0.415*** (63.43)	0.276*** (51.68)
Maximum h-index	0.0311*** (57.74)	0.0202*** (44.32)	0.0311*** (58.17)	0.0202*** (44.88)	0.0311*** (57.67)	0.0202*** (44.33)	0.0310*** (58.09)	0.0202*** (44.86)
Maximal distance			0.0965*** (7.40)	0.125*** (11.25)			0.114*** (8.19)	0.136*** (11.44)
Impact Factor		0.0905*** (104.65)		0.0905*** (105.07)		0.0905*** (104.62)		0.0905*** (105.11)
Observations	254710	254710	254710	254710	254710	254710	254710	254710
Log-Likelihood	-791489.1	-762745.9	-791417.3	-762596.5	-791509.8	-762761.4	-791419.4	-762604.4
AIC	1583092.2	1525607.7	1582950.7	1525311.0	1583133.7	1525638.7	1582954.8	1525326.9
BIC	1583687.8	1526213.7	1583556.7	1525927.4	1583729.2	1526244.7	1583560.7	1525943.3

t statistics in parentheses

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

Table 3.56: Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)

	(3a)	(3b)	(3c)	(3d)	(4a)	(4b)	(4c)	(4d)
Shannon	-0.0308*** (-8.68)	-0.00488 (-1.64)	-0.0589*** (-11.56)	-0.0371*** (-8.60)				
Number of disciplines					0.0278*** (8.09)	0.0330*** (11.72)	0.0583*** (10.95)	0.0472*** (10.93)
Number of authors	0.00113*** (13.48)	0.000668*** (12.15)	0.00113*** (13.65)	0.000664*** (12.29)	0.00107*** (13.60)	0.000625*** (11.66)	0.00105*** (13.46)	0.000616*** (11.49)
International Collaboration	0.415*** (63.17)	0.276*** (51.31)	0.415*** (63.40)	0.276*** (51.65)	0.416*** (63.32)	0.276*** (51.75)	0.415*** (63.49)	0.276*** (51.79)
Maximum h-index	0.0311*** (57.69)	0.0203*** (44.40)	0.0311*** (58.03)	0.0202*** (44.85)	0.0314*** (58.98)	0.0204*** (45.66)	0.0313*** (58.95)	0.0204*** (45.56)
Maximal distance			0.119*** (7.75)	0.135*** (10.40)			-0.166*** (-9.61)	-0.0772*** (-5.46)
Impact Factor		0.0905*** (104.59)		0.0906*** (105.07)		0.0906*** (105.56)		0.0905*** (105.85)
Observations	254710	254710	254710	254710	254710	254710	254710	254710
Log-Likelihood	-791567.3	-762779.7	-791484.2	-762646.8	-791497.9	-762550.0	-791383.8	-762519.5
AIC	1583248.5	1525675.5	1583084.4	1525411.5	1583109.8	1525215.9	1582883.6	1525156.9
BIC	1583844.1	1526281.5	1583690.4	1526027.9	1583705.4	1525821.9	1583489.5	1525773.3

t statistics in parentheses

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

Table 3.57: Negative binomial estimates for the effect of the number of disciplines (categorical variable) on the number of citations (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)

	Number of citations	Number of citations	Number of citations	Number of citations
	0	0	0	0
	(.)	(.)	(.)	(.)
1 discipline				
2 disciplines	-0.0293** (-3.02)	-0.0260*** (-3.40)	0.00258 (0.20)	-0.0459*** (-4.38)
3 disciplines	-0.00122 (-0.12)	0.00353 (0.44)	0.0471** (2.86)	-0.0267* (-2.01)
4 and more disciplines	0.0505*** (4.40)	0.0796*** (8.30)	0.113*** (5.72)	0.0410* (2.51)
Number of authors	0.00108*** (13.38)	0.000627*** (11.58)	0.00108*** (13.22)	0.000631*** (11.65)
International Collaboration	0.416*** (63.13)	0.276*** (51.53)	0.415*** (62.92)	0.276*** (51.58)
Maximum h-index	0.0313*** (58.45)	0.0204*** (45.16)	0.0313*** (58.34)	0.0204*** (45.31)
Maximal distance			-0.0740*** (-3.77)	0.0460** (2.81)
Impact Factor		0.0906*** (105.06)		0.0907*** (104.89)
Observations	254710	254710	254710	254710
Log-Likelihood	-791551.1	-762598.8	-791536.8	-762591.9
AIC	1583220.2	1525317.6	1583193.6	1525305.8
BIC	1583836.6	1525944.5	1583820.5	1525943.1

t statistics in parentheses

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

Table 3.58: Negative binomial estimates for the effect of the logarithm of interdisciplinarity on the number of citations (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)

	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
log Berger-Parker	-0.143*** (-13.98)	-0.0692*** (-8.04)	-0.214*** (-16.33)	-0.156*** (-13.90)				
log Simpson					-0.106*** (-12.73)	-0.0457*** (-6.64)	-0.189*** (-16.21)	-0.140*** (-14.24)
Number of authors	0.00112*** (13.46)	0.000669*** (12.18)	0.00111*** (13.63)	0.000655*** (12.18)	0.00113*** (13.48)	0.000670*** (12.21)	0.00112*** (13.69)	0.000659*** (12.30)
International Collaboration	0.414*** (63.29)	0.276*** (51.36)	0.415*** (63.52)	0.276*** (51.75)	0.414*** (63.25)	0.276*** (51.34)	0.415*** (63.58)	0.276*** (51.80)
Maximum h-index	0.0311*** (57.78)	0.0202*** (44.32)	0.0311*** (58.32)	0.0202*** (44.97)	0.0311*** (57.73)	0.0202*** (44.33)	0.0310*** (58.31)	0.0202*** (45.00)
Maximal distance			0.121*** (8.85)	0.147*** (12.55)			0.154*** (10.27)	0.172*** (13.48)
Impact Factor		0.0904*** (104.65)		0.0905*** (105.18)		0.0905*** (104.60)		0.0905*** (105.25)
Observations	254710	254710	254710	254710	254710	254710	254710	254710
Log-Likelihood	-791471.7	-762735.6	-791368.4	-762547.7	-791489.8	-762749.0	-791347.3	-762529.3
AIC	1583057.4	1525587.1	1582852.8	1525213.3	1583093.7	1525613.9	1582810.6	1525176.6
BIC	1583652.9	1526193.1	1583458.8	1525829.8	1583689.2	1526219.9	1583416.6	1525793.0

t statistics in parentheses

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

Table 3.59: Negative binomial estimates for the effect of the logarithm of interdisciplinarity on the number of citations (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)

	(3a)	(3b)	(3c)	(3d)
ln(Shannon)	-0.0724*** (-9.17)	-0.0203** (-3.13)	-0.173*** (-13.81)	-0.130*** (-12.46)
Number of authors	0.00114*** (13.49)	0.000670*** (12.18)	0.00113*** (13.73)	0.000668*** (12.41)
International Collaboration	0.415*** (63.17)	0.276*** (51.31)	0.415*** (63.56)	0.276*** (51.79)
Maximum h-index	0.0312*** (57.76)	0.0203*** (44.38)	0.0311*** (58.31)	0.0202*** (45.01)
Maximal distance			0.173*** (10.22)	0.189*** (13.07)
Impact Factor		0.0905*** (104.55)		0.0906*** (105.24)
Observations	254710	254710	254710	254710
Log-Likelihood	-791554.0	-762774.1	-791411.1	-762566.0
AIC	1583222.0	1525664.2	1582938.2	1525249.9
BIC	1583817.5	1526270.2	1583544.2	1525866.3

t statistics in parentheses

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

Table 3.60: Negative binomial estimates for the effect of interdisciplinarity (square term) on the number of citations (papers with at least 4 references, 1-35 disciplines, CWTs nomenclature)

	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Berger-Parker	-0.201*** (-8.12)	-0.131*** (-6.28)	-0.327*** (-11.67)	-0.290*** (-12.11)				
Berger-Parker ²	0.0318*** (5.21)	0.0250*** (4.77)	0.0549*** (8.43)	0.0542*** (9.63)				
Simpson					-0.130*** (-9.01)	-0.0847*** (-7.01)	-0.232*** (-13.64)	-0.202*** (-14.03)
Simpson ²					0.0170*** (6.06)	0.0141*** (5.88)	0.0307*** (10.23)	0.0299*** (11.64)
Number of authors	0.00112*** (13.46)	0.000668*** (12.19)	0.00111*** (13.63)	0.000654*** (12.18)	0.00113*** (13.48)	0.000670*** (12.21)	0.00112*** (13.70)	0.000658*** (12.29)
International Collaboration	0.414*** (63.32)	0.276*** (51.39)	0.415*** (63.57)	0.276*** (51.80)	0.415*** (63.31)	0.276*** (51.38)	0.415*** (63.65)	0.276*** (51.86)
Maximum h-index	0.0311*** (57.81)	0.0202*** (44.36)	0.0311*** (58.42)	0.0202*** (45.09)	0.0311*** (57.75)	0.0203*** (44.39)	0.0311*** (58.43)	0.0202*** (45.17)
Maximal distance			0.124*** (9.04)	0.153*** (13.01)			0.155*** (10.43)	0.176*** (13.79)
Impact Factor		0.0905*** (104.68)		0.0905*** (105.23)		0.0905*** (104.62)		0.0905*** (105.29)
Observations	254710	254710	254710	254710	254710	254710	254710	254710
Log-Likelihood	-791468.0	-762730.1	-791360.7	-762530.1	-791479.3	-762736.4	-791332.4	-762505.9
AIC	1583051.9	1525578.2	1582839.3	1525180.2	1583074.7	1525590.9	1582782.9	1525131.9
BIC	1583657.9	1526194.6	1583455.7	1525807.0	1583680.6	1526207.3	1583399.3	1525758.8

t statistics in parentheses

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

Table 3.61: Negative binomial estimates for the effect of interdisciplinarity (square term) on the number of citations (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)

	(3a)	(3b)	(3c)	(3d)
Shannon	-0.0926*** (-7.71)	-0.0596*** (-5.95)	-0.203*** (-13.43)	-0.182*** (-14.33)
Shannon ²	0.0118*** (5.58)	0.0104*** (5.81)	0.0247*** (10.76)	0.0249*** (12.68)
Number of authors	0.00114*** (13.50)	0.000670*** (12.17)	0.00113*** (13.75)	0.000667*** (12.40)
International Collaboration	0.415*** (63.24)	0.276*** (51.35)	0.415*** (63.66)	0.276*** (51.86)
Maximum h-index	0.0312*** (57.73)	0.0203*** (44.45)	0.0312*** (58.44)	0.0203*** (45.22)
Maximal distance			0.179*** (10.67)	0.196*** (13.69)
Impact Factor		0.0905*** (104.57)		0.0906*** (105.27)
Observations	254710	254710	254710	254710
Log-Likelihood	-791539.5	-762753.6	-791384.0	-762525.3
AIC	1583195.0	1525625.2	1582886.0	1525170.5
BIC	1583801.0	1526241.6	1583502.4	1525797.4

t statistics in parentheses

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

Table 3.62: Logit estimates for the effect of interdisciplinarity on the probability that the paper belongs to the top 10% most cited papers (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Berger-Parker	-0.136*** (-11.05)	-0.173*** (-12.33)						
Simpson			-0.0867*** (-10.53)	-0.123*** (-12.39)				
Shannon					-0.0573*** (-8.24)	-0.0947*** (-10.38)		
Number of disciplines							0.0284*** (6.56)	0.0711*** (10.69)
Number of authors	0.000628*** (15.70)	0.000619*** (15.47)	0.000634*** (15.83)	0.000625*** (15.61)	0.000639*** (15.95)	0.000636*** (15.86)	0.000598*** (14.91)	0.000581*** (14.48)
International Collaboration	0.627*** (54.84)	0.627*** (54.84)	0.627*** (54.83)	0.627*** (54.82)	0.627*** (54.84)	0.627*** (54.83)	0.628*** (54.89)	0.627*** (54.86)
Maximum h-index	0.0448*** (56.45)	0.0448*** (56.50)	0.0447*** (56.40)	0.0448*** (56.43)	0.0448*** (56.50)	0.0448*** (56.50)	0.0451*** (56.93)	0.0449*** (56.67)
Maximal distance		0.113*** (5.59)		0.142*** (6.61)		0.148*** (6.39)		-0.232*** (-8.39)
Observations	254710	254710	254710	254710	254710	254710	254710	254710
Log-Likelihood	-107568.3	-107552.7	-107574.2	-107552.3	-107596.2	-107575.8	-107609.2	-107573.9
AIC	215248.6	215219.4	215260.3	215218.7	215304.4	215265.6	215330.4	215261.9
BIC	215833.7	215815.0	215845.4	215814.2	215889.5	215861.2	215915.5	215857.4

t statistics in parentheses

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

Table 3.63: Logit estimates for the effect of interdisciplinarity on the probability that the paper belongs to the top 5% most cited papers (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Berger-Parker	-0.196*** (-11.93)	-0.233*** (-12.42)						
Simpson			-0.124*** (-11.31)	-0.161*** (-12.18)				
Shannon					-0.0856*** (-9.29)	-0.125*** (-10.31)		
Number of disciplines							0.0258*** (4.60)	0.0840*** (9.79)
Number of authors	0.000683*** (16.71)	0.000675*** (16.50)	0.000690*** (16.88)	0.000682*** (16.67)	0.000698*** (17.06)	0.000695*** (16.98)	0.000657*** (15.97)	0.000637*** (15.44)
International Collaboration	0.703*** (46.68)	0.703*** (46.67)	0.703*** (46.67)	0.703*** (46.65)	0.703*** (46.68)	0.703*** (46.66)	0.703*** (46.71)	0.703*** (46.69)
Maximum h-index	0.0496*** (51.26)	0.0497*** (51.30)	0.0496*** (51.22)	0.0496*** (51.24)	0.0496*** (51.30)	0.0497*** (51.30)	0.0500*** (51.71)	0.0498*** (51.45)
Maximal distance		0.109*** (4.15)		0.141*** (5.08)		0.152*** (5.04)		-0.318*** (-8.81)
Observations	254708	254708	254708	254708	254708	254708	254708	254708
Log-Likelihood	-71119.7	-71111.2	-71127.6	-71114.7	-71149.4	-71136.7	-71182.9	-71143.9
AIC	142349.5	142334.3	142365.2	142341.5	142408.7	142385.4	142475.8	142399.7
BIC	142924.1	142919.4	142939.9	142926.6	142983.4	142970.5	143050.4	142984.8

t statistics in parentheses

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

Table 3.64: Logit estimates for the effect of the number of disciplines (categorical variable) on the probability that the paper belongs to the top 10% and top 5% most cited papers (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)

	Top90	Top90	Top95	Top95
1 discipline	0 (.)	0 (.)	0 (.)	0 (.)
2 disciplines	-0.0695*** (-4.29)	-0.0397 (-1.79)	-0.0964*** (-4.62)	-0.0667* (-2.32)
3 disciplines	-0.0708*** (-4.08)	-0.0265 (-0.93)	-0.118*** (-5.26)	-0.0741* (-2.01)
4 and more disciplines	0.0379* (2.11)	0.0942** (2.78)	-0.00503 (-0.22)	0.0508 (1.16)
Number of authors	0.000598*** (14.90)	0.000594*** (14.78)	0.000661*** (16.04)	0.000657*** (15.93)
International Collaboration	0.627*** (54.84)	0.627*** (54.83)	0.703*** (46.68)	0.703*** (46.68)
Maximum h-index	0.0451*** (56.89)	0.0450*** (56.76)	0.0500*** (51.69)	0.0499*** (51.58)
Maximal distance		-0.0674 (-1.96)		-0.0673 (-1.50)
Observations	254710	254710	254708	254708
Log-Likelihood	-107596.0	-107594.1	-71167.7	-71166.6
AIC	215308.0	215306.2	142449.5	142449.3
BIC	215914.0	215922.6	143045.0	143055.2

t statistics in parentheses

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

Table 3.65: OLS estimates for the effect of interdisciplinarity on the journal impact (papers with at least 4 references, 1- 35 disciplines, CWTS nomenclature)

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Berger-Parker	-0.285*** (-10.63)	-0.296*** (-9.62)						
Simpson			-0.209*** (-11.58)	-0.240*** (-10.73)				
Shannon					-0.166*** (-10.67)	-0.209*** (-9.94)		
Number of disciplines							0.0110 (0.97)	0.107*** (6.01)
Number of authors	0.00241*** (24.24)	0.00240*** (24.17)	0.00242*** (24.35)	0.00241*** (24.26)	0.00244*** (24.49)	0.00244*** (24.44)	0.00238*** (23.80)	0.00234*** (23.40)
International Collaboration	1.639*** (61.16)	1.639*** (61.15)	1.638*** (61.13)	1.638*** (61.14)	1.638*** (61.14)	1.638*** (61.14)	1.640*** (61.18)	1.640*** (61.13)
Maximum h-index	0.119*** (43.02)	0.119*** (43.02)	0.119*** (42.94)	0.119*** (42.95)	0.119*** (42.97)	0.119*** (42.98)	0.120*** (43.35)	0.120*** (43.16)
Maximal distance		0.0362 (0.72)		0.126* (2.32)		0.179** (3.02)		-0.518*** (-7.48)
Observations	254710	254710	254710	254710	254710	254710	254710	254710
Log-Likelihood	-841036.7	-841036.4	-841026.1	-841023.1	-841034.8	-841029.6	-841089.3	-841058.0
AIC	1682185.3	1682186.8	1682164.3	1682160.2	1682181.6	1682173.3	1682290.6	1682230.1
BIC	1682770.4	1682782.3	1682749.3	1682755.8	1682766.7	1682768.8	1682875.7	1682825.6

t statistics in parentheses

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

Table 3.66: OLS estimates for the effect of the number of disciplines (categorical variable) on the journal impact factor (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)

	(1)	(2)
1 discipline	0	0
2 disciplines	-0.0765	1.537***
3 disciplines	-0.0104	2.125***
4 disciplines	0.0349	2.490***
5 disciplines	0.270**	2.931***
6 disciplines	0.318***	3.105***
7 disciplines	0.385***	3.264***
8-9 disciplines	0.464***	3.444***
9 and more disciplines	0.736***	3.861***
Number of authors	0.00238***	0.00260***
International Collaboration	1.641***	1.621***
Maximum h-index	0.119***	0.115***
Maximal distance		-3.123***
Observations	255203	255203
Log-Likelihood	-842458.6	-841851.2
AIC	1685043.1	1683830.4
BIC	1685701.5	1684499.1

t statistics in parentheses

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

10.5 WoS subject categories

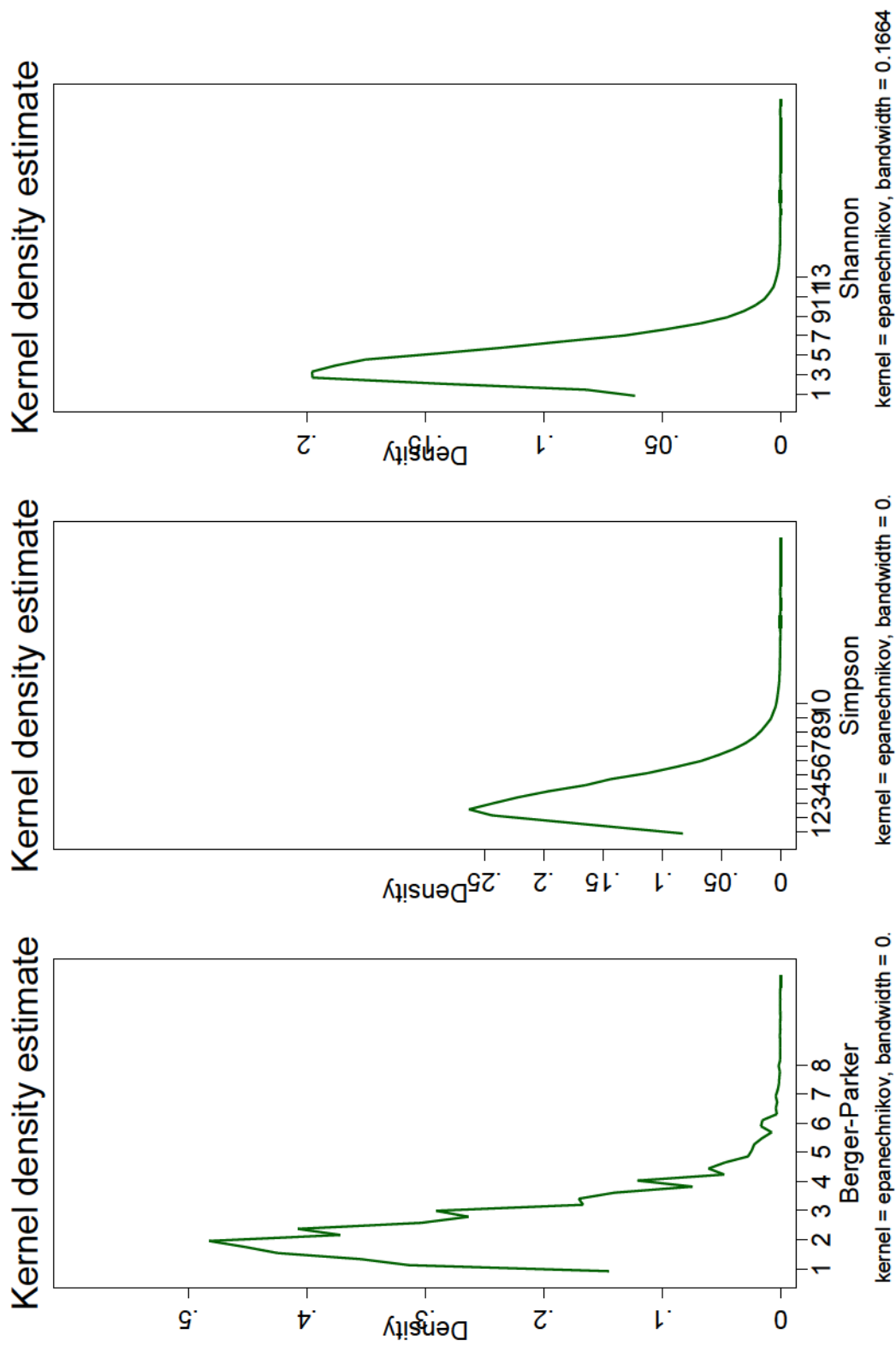


Figure 3.13: Distribution of Interdisciplinarity according to the Berger-Parker, the Simpson and the Shannon indexes (WoS nomenclature, papers with at least 4 references)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1															
2	0.44	1														
3	0.47	0.70	1													
4	0.35	0.30	0.29	1												
5	0.29	0.16	0.15	0.17	1											
6	-0.030	-0.037	-0.034	-0.028	0.063	1										
7	-0.022	-0.031	-0.030	-0.016	0.14	0.94	1									
8	-0.0054	-0.019	-0.019	0.0017	0.23	0.86	0.97	1								
9	0.049	0.024	0.021	0.049	0.46	0.63	0.79	0.90	1							
10	0.085	0.053	0.054	0.042	-0.0046	-0.020	-0.021	-0.019	-0.0094	1						
11	0.16	0.11	0.11	0.096	0.056	-0.039	-0.038	-0.033	-0.010	0.91	1					
12	0.11	0.12	0.10	0.13	0.052	-0.042	-0.041	-0.037	-0.019	0.088	0.20	1				
13	0.12	0.088	0.085	0.13	0.26	0.0077	0.037	0.068	0.14	0.081	0.16	0.11	1			
14	-0.024	-0.019	-0.0056	-0.088	-0.16	-0.23	-0.29	-0.33	-0.35	0.041	0.063	-0.042	-0.026	1		
15	-0.021	-0.023	-0.013	-0.088	0.087	0.29	0.34	0.37	0.40	0.040	0.053	-0.054	0.020	0.49	1	
16	-0.011	-0.020	-0.015	-0.047	0.17	0.45	0.52	0.58	0.62	0.024	0.023	-0.047	0.047	0.15	0.89	1

Table 3.67: Correlation matrix (papers with at least four references, 1-252 disciplines, WoS nomenclature)

Table 3.68: Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)

	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Berger-Parker	-0.0379*** (-11.45)	-0.0205*** (-7.65)	-0.0355*** (-10.10)	-0.0282*** (-9.66)				
Simpson					-0.0190*** (-9.30)	-0.00915*** (-5.61)	-0.0174*** (-7.61)	-0.0146*** (-7.77)
Number of authors	0.00132*** (10.50)	0.000809*** (10.47)	0.00133*** (10.54)	0.000787*** (10.30)	0.00132*** (10.54)	0.000809*** (10.48)	0.00133*** (10.60)	0.000786*** (10.31)
International Collaboration	0.402*** (54.73)	0.273*** (44.63)	0.402*** (54.65)	0.273*** (44.75)	0.402*** (54.78)	0.273*** (44.66)	0.402*** (54.70)	0.273*** (44.79)
Maximum h-index	0.0280*** (49.12)	0.0188*** (38.32)	0.0279*** (49.08)	0.0188*** (38.65)	0.0280*** (49.15)	0.0188*** (38.31)	0.0280*** (49.12)	0.0188*** (38.68)
Maximal distance			-0.0306 (-1.61)	0.0995*** (6.14)			-0.0304 (-1.53)	0.102*** (5.97)
Impact Factor		0.0812*** (89.09)		0.0814*** (89.00)		0.0812*** (89.21)		0.0815*** (89.09)
Observations	188792	188792	188792	188792	188792	188792	188792	188792
Log-Likelihood	-609379.2	-588319.5	-609376.0	-588277.9	-609407.2	-588336.0	-609404.2	-588294.8
AIC	1218870.4	1176753.0	1218866.0	1176673.8	1218926.4	1176785.9	1218922.5	1176705.6
BIC	1219438.7	1177331.4	1219444.4	1177272.5	1219494.7	1177364.4	1219500.9	1177294.2

t statistics in parentheses

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

Table 3.69: Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)

	(3a)	(3b)	(3c)	(3d)	(4a)	(4b)	(4c)	(4d)
Shannon	-0.00717*** (-3.91)	-0.000661 (-0.46)	-0.00369 (-1.72)	-0.00462** (-2.67)				
Number of disciplines					0.0179*** (11.11)	0.0171*** (12.13)	0.0307*** (16.60)	0.0213*** (12.76)
Number of authors	0.00133*** (10.55)	0.000811*** (10.48)	0.00135*** (10.62)	0.000794*** (10.38)	0.00131*** (10.80)	0.000802*** (10.52)	0.00139*** (10.97)	0.000821*** (10.71)
International Collaboration	0.403*** (54.76)	0.274*** (44.67)	0.403*** (54.67)	0.274*** (44.77)	0.405*** (54.90)	0.274*** (45.24)	0.403*** (55.06)	0.274*** (45.27)
Maximum h-index	0.0281*** (49.24)	0.0189*** (38.40)	0.0281*** (49.14)	0.0189*** (38.72)	0.0280*** (50.03)	0.0188*** (39.55)	0.0276*** (49.36)	0.0186*** (39.37)
Maximal distance			-0.0680** (-3.26)	0.0767*** (4.29)			-0.317*** (-15.01)	-0.104*** (-5.74)
Impact Factor		0.0813*** (89.19)		0.0815*** (88.92)		0.0813*** (89.36)		0.0810*** (88.97)
Observations	188792	188792	188792	188792	188792	188792	188792	188792
Log-Likelihood	-609471.2	-588358.6	-609457.2	-588336.8	-609250.0	-588099.5	-608976.6	-588063.6
AIC	1219056.3	1176833.2	1219030.5	1176789.6	1218611.9	1176312.9	1218067.2	1176245.1
BIC	1219634.8	1177421.8	1219619.1	1177378.2	1219180.2	1176891.4	1218645.7	1176843.9

t statistics in parentheses

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

Table 3.70: Negative binomial estimates for the effect of the number of disciplines (categorical variable) on the number of citations (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)

	(4a)	(4b)	(4c)	(4d)
1 discipline	0 (.)	0 (.)	0 (.)	0 (.)
2 disciplines	-0.0371 (-1.15)	-0.0240 (-1.00)	0.138*** (3.86)	-0.0394 (-1.45)
3 disciplines	-0.0338 (-1.05)	-0.0248 (-1.02)	0.190*** (5.09)	-0.0445 (-1.52)
4 disciplines	-0.0371 (-1.16)	-0.0265 (-1.09)	0.218*** (5.59)	-0.0491 (-1.61)
5 disciplines	-0.0481 (-1.52)	-0.0369 (-1.54)	0.229*** (5.75)	-0.0615* (-1.97)
6 disciplines	-0.0480 (-1.51)	-0.0303 (-1.27)	0.244*** (6.01)	-0.0561 (-1.76)
7 disciplines	-0.0438 (-1.36)	-0.0278 (-1.14)	0.257*** (6.27)	-0.0544 (-1.67)
8-9 disciplines	-0.0280 (-0.87)	-0.00600 (-0.25)	0.284*** (6.93)	-0.0336 (-1.04)
9 and more disciplines	0.101** (3.06)	0.113*** (4.44)	0.429*** (9.98)	0.0842* (2.46)
Number of authors	0.00135*** (10.69)	0.000822*** (10.61)	0.00139*** (10.76)	0.000819*** (10.61)
International Collaboration	0.404*** (54.98)	0.274*** (45.07)	0.404*** (54.82)	0.274*** (45.13)
Maximum h-index	0.0280*** (49.60)	0.0187*** (38.87)	0.0278*** (48.97)	0.0188*** (39.06)
Maximal distance			-0.317*** (-12.93)	0.0280 (1.29)
Impact Factor		0.0813*** (89.40)		0.0814*** (88.34)
Observations	188792	188792	188792	188792
Log-Likelihood	-609297.4	-588133.3	-609147.6	-588131.9
AIC	1218720.8	1176396.5	1218423.2	1176395.7
BIC	1219360.1	1177056.2	1219072.7	1177065.5

t statistics in parentheses

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

Table 3.71: Negative binomial estimates for the effect of the logarithm of interdisciplinarity measured with the number of disciplines on the number of citations (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)

	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
log Berger-Parker	-0.101*** (-11.08)	-0.0553*** (-7.60)	-0.0974*** (-9.89)	-0.0831*** (-10.30)				
log Simpson					-0.0765*** (-9.75)	-0.0394*** (-6.42)	-0.0751*** (-8.30)	-0.0713*** (-9.65)
Number of authors	0.00133*** (10.50)	0.000811*** (10.50)	0.00134*** (10.54)	0.000787*** (10.31)	0.00133*** (10.54)	0.000811*** (10.50)	0.00133*** (10.60)	0.000784*** (10.30)
International Collaboration	0.402*** (54.73)	0.273*** (44.61)	0.402*** (54.66)	0.273*** (44.74)	0.402*** (54.81)	0.273*** (44.63)	0.402*** (54.75)	0.273*** (44.80)
Maximum h-index	0.0279*** (49.06)	0.0188*** (38.27)	0.0279*** (49.05)	0.0188*** (38.60)	0.0280*** (49.06)	0.0188*** (38.24)	0.0280*** (49.10)	0.0188*** (38.63)
Maximal distance			-0.0165 (-0.86)	0.115*** (6.97)			-0.00591 (-0.29)	0.132*** (7.40)
Impact Factor		0.0812*** (89.07)		0.0814*** (89.02)		0.0812*** (89.24)		0.0815*** (89.16)
Observations	188792	188792	188792	188792	188792	188792	188792	188792
Log-Likelihood	-609372.4	-588316.4	-609371.5	-588264.1	-609389.8	-588326.9	-609389.7	-588265.5
AIC	1218856.8	1176746.9	1218857.0	1176644.2	1218893.7	1176769.7	1218893.4	1176647.0
BIC	1219425.1	1177325.3	1219435.5	1177232.8	1219472.1	1177358.3	1219471.9	1177235.6

t statistics in parentheses

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

Table 3.72: Negative binomial estimates for the effect of the logarithm of interdisciplinarity measured with the number of disciplines on the number of citations (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)

	(3a)	(3b)	(3c)	(3d)
log Shannon	-0.0444*** (-5.42)	-0.0148* (-2.33)	-0.0323** (-3.15)	-0.0480*** (-5.74)
Number of authors	0.00133*** (10.55)	0.000812*** (10.49)	0.00135*** (10.63)	0.000789*** (10.34)
International Collaboration	0.403*** (54.81)	0.273*** (44.63)	0.403*** (54.74)	0.273*** (44.79)
Maximum h-index	0.0281*** (49.18)	0.0189*** (38.33)	0.0281*** (49.17)	0.0189*** (38.73)
Maximal distance			-0.0433 (-1.94)	0.118*** (6.05)
Impact Factor		0.0813*** (89.25)		0.0815*** (89.01)
Observations	188792	188792	188792	188792
Log-Likelihood	-609453.6	-588354.1	-609448.9	-588311.7
AIC	1219021.2	1176824.1	1219013.9	1176739.4
BIC	1219599.7	1177412.7	1219602.5	1177328.0

t statistics in parentheses

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

Table 3.73: Negative binomial estimates for the effect of interdisciplinarity (square term) on the number of citations (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)

	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Berger-Parker	-0.0754*** (-5.92)	-0.0420*** (-4.08)	-0.0707*** (-5.42)	-0.0696*** (-6.55)				
Berger-Parker ²	0.00601** (3.21)	0.00345* (2.24)	0.00550** (2.92)	0.00649*** (4.20)				
Simpson					-0.0547*** (-8.14)	-0.0334*** (-5.95)	-0.0541*** (-7.58)	-0.0543*** (-8.96)
Simpson ²					0.00381*** (5.82)	0.00261*** (4.60)	0.00377*** (5.65)	0.00409*** (7.05)
Number of authors	0.00133*** (10.50)	0.000810*** (10.49)	0.00133*** (10.53)	0.000787*** (10.31)	0.00133*** (10.54)	0.000812*** (10.51)	0.00133*** (10.60)	0.000784*** (10.31)
International Collaboration	0.402*** (54.73)	0.273*** (44.62)	0.402*** (54.66)	0.273*** (44.75)	0.402*** (54.83)	0.273*** (44.66)	0.402*** (54.76)	0.273*** (44.82)
Maximum h-index	0.0279*** (49.08)	0.0188*** (38.28)	0.0279*** (49.07)	0.0188*** (38.61)	0.0280*** (49.09)	0.0188*** (38.25)	0.0280*** (49.15)	0.0188*** (38.68)
Maximal distance			-0.0200 (-1.05)	0.112*** (6.84)			-0.00431 (-0.21)	0.131*** (7.45)
Impact Factor		0.0812*** (89.07)		0.0814*** (89.01)		0.0812*** (89.23)		0.0815*** (89.17)
Observations	188792	188792	188792	188792	188792	188792	188792	188792
Log-Likelihood	-609369.8	-588315.9	-609368.5	-588265.8	-609376.1	-588319.2	-609376.0	-588257.1
AIC	1218853.6	1176747.7	1218853.0	1176649.5	1218866.2	1176754.5	1218868.0	1176632.3
BIC	1219432.1	1177336.3	1219441.6	1177248.3	1219444.6	1177343.1	1219456.7	1177231.0

t statistics in parentheses

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

Table 3.74: Negative binomial estimates for the effect of interdisciplinarity (square term) on the number of citations (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)

	(3a)	(3b)	(3c)	(3d)
Shannon	-0.0447*** (-7.84)	-0.0277*** (-5.85)	-0.0409*** (-6.49)	-0.0474*** (-8.68)
Shannon ²	0.00332*** (7.00)	0.00241*** (6.07)	0.00309*** (6.31)	0.00358*** (8.45)
Number of authors	0.00134*** (10.57)	0.000816*** (10.54)	0.00135*** (10.65)	0.000791*** (10.36)
International Collaboration	0.403*** (54.85)	0.273*** (44.63)	0.403*** (54.77)	0.273*** (44.78)
Maximum h-index	0.0281*** (49.18)	0.0189*** (38.31)	0.0281*** (49.25)	0.0189*** (38.78)
Maximal distance			-0.0257 (-1.18)	0.126*** (6.66)
Impact Factor		0.0813*** (89.24)		0.0815*** (89.07)
Observations	188792	188792	188792	188792
Log-Likelihood	-609418.2	-588326.5	-609416.4	-588275.6
AIC	1218950.3	1176769.1	1218948.8	1176669.1
BIC	1219528.8	117357.7	1219537.4	1177267.9

t statistics in parentheses

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

Table 3.75: Logit estimates for the effect of interdisciplinarity on the probability that the paper belongs to the top 10% most cited papers (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Berger-Parker	-0.0510*** (-7.73)	-0.0566*** (-7.97)						
Simpson			-0.0222*** (-5.75)	-0.0296*** (-6.93)				
Shannon					-0.0000920 (-0.03)	-0.00874* (-2.41)		
Number of disciplines							0.0404*** (19.47)	0.0398*** (15.68)
Number of authors	0.000638*** (12.45)	0.000663*** (12.80)	0.000639*** (12.47)	0.000662*** (12.79)	0.000642*** (12.54)	0.000673*** (12.98)	0.000640*** (12.50)	0.000714*** (13.72)
International Collaboration	0.592*** (46.58)	0.623*** (48.47)	0.593*** (46.58)	0.623*** (48.46)	0.594*** (46.70)	0.624*** (48.51)	0.600*** (47.17)	0.626*** (48.64)
Maximum h-index	0.0413*** (49.00)	0.0412*** (48.47)	0.0414*** (49.10)	0.0412*** (48.54)	0.0416*** (49.32)	0.0413*** (48.68)	0.0414*** (49.08)	0.0408*** (48.01)
Maximal distance		-0.0637* (-2.17)		-0.0590 (-1.95)		-0.111*** (-3.52)		-0.435*** (-13.09)
Observations	188791	188791	188791	188791	188791	188791	188791	188791
Log-Likelihood	-85497.2	-84374.8	-85510.8	-84382.8	-85527.5	-84404.1	-85341.9	-84285.7
AIC	171100.4	168861.7	171127.7	168877.6	171161.0	168920.3	170789.7	168683.4
BIC	171638.3	169430.0	171665.5	169445.9	171698.9	169488.6	171327.6	169251.7

t statistics in parentheses

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

Table 3.76: Logit estimates for the effect of interdisciplinarity on the probability that the paper belongs to the top 5% most cited papers (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Berger-Parker	-0.0639*** (-7.41)	-0.0794*** (-8.54)						
Simpson			-0.0298*** (-5.90)	-0.0457*** (-8.20)				
Shannon					-0.00345 (-0.84)	-0.0222*** (-4.72)		
Number of disciplines							0.0480*** (18.24)	0.0383*** (12.03)
Number of authors	0.000676*** (13.07)	0.000690*** (13.23)	0.000677*** (13.08)	0.000687*** (13.18)	0.000681*** (13.17)	0.000696*** (13.35)	0.000681*** (13.16)	0.000742*** (14.20)
International Collaboration	0.658*** (39.68)	0.701*** (41.75)	0.658*** (39.68)	0.701*** (41.73)	0.659*** (39.79)	0.701*** (41.78)	0.668*** (40.29)	0.704*** (41.93)
Maximum h-index	0.0464*** (45.46)	0.0464*** (45.08)	0.0465*** (45.56)	0.0465*** (45.15)	0.0466*** (45.76)	0.0466*** (45.30)	0.0465*** (45.57)	0.0461*** (44.77)
Maximal distance		0.0117 (0.31)		0.0303 (0.78)		-0.0115 (-0.29)		-0.380*** (-8.97)
Observations	188791	188791	188791	188791	188791	188791	188791	188791
Log-Likelihood	-57850.3	-56734.6	-57860.7	-56737.6	-57878.0	-56760.7	-57717.7	-56701.1
AIC	115806.7	113581.2	115827.4	113587.3	115862.0	113633.3	115541.5	113514.2
BIC	116344.6	114149.5	116365.3	114155.6	116399.8	114201.6	116079.4	114082.5

t statistics in parentheses

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

Table 3.77: Logit estimates for the effect of the number of disciplines (categorical variable) on the probability that the paper belongs to the top 10% and top 5% most cited papers (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)

	Top90	Top90	Top95	Top95
1 discipline	0 (.)	0 (.)	0 (.)	0 (.)
2 disciplines	-0.179*** (-4.22)	-0.0349 (-0.70)	-0.102 (-1.90)	-0.00667 (-0.11)
3 disciplines	-0.161*** (-3.94)	0.0151 (0.29)	-0.0533 (-1.03)	0.0546 (0.82)
4 disciplines	-0.174*** (-4.32)	0.0235 (0.43)	-0.0804 (-1.58)	0.0361 (0.52)
5 disciplines	-0.193*** (-4.83)	0.0158 (0.28)	-0.101* (-2.00)	0.0206 (0.28)
6 disciplines	-0.168*** (-4.16)	0.0485 (0.83)	-0.0679 (-1.33)	0.0540 (0.72)
7 disciplines	-0.181*** (-4.38)	0.0304 (0.50)	-0.109* (-2.08)	0.000228 (0.00)
8-9 disciplines	-0.124** (-3.07)	0.0717 (1.18)	-0.0778 (-1.52)	0.00372 (0.05)
9 and more disciplines	0.147*** (3.61)	0.255*** (4.06)	0.286*** (5.56)	0.243** (3.03)
Number of authors	0.000678*** (13.13)	0.000707*** (13.57)	0.000711*** (13.68)	0.000731*** (13.98)
International Collaboration	0.598*** (46.99)	0.625*** (48.62)	0.665*** (40.10)	0.703*** (41.89)
Maximum h-index	0.0412*** (48.86)	0.0410*** (48.25)	0.0463*** (45.43)	0.0464*** (45.02)
Maximal distance		-0.275*** (-6.03)		-0.185** (-3.14)
Observations	188791	188791	188791	188791
Log-Likelihood	-85365.8	-84340.0	-57748.1	-56727.9
AIC	170851.7	168806.1	115616.3	113581.7
BIC	171460.6	169445.4	116225.2	114221.1

t statistics in parentheses

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

Table 3.78: OLS estimates for the effect of interdisciplinarity on the journal impact (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Berger-Parker	-0.133*** (-7.85)	-0.0297 (-1.66)						
Simpson			-0.0707*** (-6.54)	0.00526 (0.44)				
Shannon					-0.0344*** (-3.55)	0.0489*** (4.42)		
Number of disciplines							0.0480*** (6.10)	0.153*** (15.86)
Number of authors	0.00244*** (18.01)	0.00263*** (18.88)	0.00244*** (18.01)	0.00264*** (18.94)	0.00245*** (18.06)	0.00268*** (19.11)	0.00245*** (18.04)	0.00278*** (19.62)
International Collaboration	1.661*** (50.99)	1.652*** (50.69)	1.661*** (50.98)	1.652*** (50.71)	1.662*** (51.02)	1.653*** (50.74)	1.668*** (51.18)	1.652*** (50.76)
Maximum h-index	0.112*** (35.45)	0.110*** (35.12)	0.112*** (35.47)	0.111*** (35.15)	0.112*** (35.57)	0.111*** (35.17)	0.112*** (35.57)	0.108*** (34.46)
Maximal distance		-1.309*** (-14.93)		-1.372*** (-15.10)		-1.570*** (-16.52)		-2.475*** (-23.81)
Observations	188792	188792	188792	188792	188792	188792	188792	188792
Log-Likelihood	-632958.9	-632820.1	-632965.0	-632821.4	-632982.5	-632807.5	-632954.8	-632563.4
AIC	1266029.7	1265754.3	1266040.1	1265754.9	1266074.9	1265729.0	1266019.6	1265240.8
BIC	1266598.0	1266332.7	1266598.2	1266332.2	1266633.1	1266307.4	1266577.7	1265819.3

t statistics in parentheses

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

Table 3.79: OLS estimates for the effect of the number of disciplines (categorical variable) on the journal impact factor (papers with at least 4 references, 1-252 disciplines, WoS nomenclature)

	(1)	(2)
1 discipline	0	0
2 disciplines	-0.0765	1.537***
3 disciplines	-0.0104	2.125***
4 disciplines	0.0349	2.490***
5 disciplines	0.270**	2.931***
6 disciplines	0.318***	3.105***
7 disciplines	0.385***	3.264***
8-9 disciplines	0.464***	3.444***
9 and more disciplines	0.736***	3.861***
Number of authors	0.00238***	0.00260***
International Collaboration	1.641***	1.621***
Maximum h-index	0.119***	0.115***
Maximal distance		-3.123***
Observations	255203	255203
Log-Likelihood	-842458.6	-841851.2
AIC	1685043.1	1683830.4
BIC	1685701.5	1684499.1

t statistics in parentheses

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

10.6 Tobit estimates

Table 3.80: Tobit estimates for the effect of the variable “year” on interdisciplinarity measured with the Berger-Parker index (1-10 disciplines, OST nomenclature)

	(1)	(2)	(3)	(4)	(5)	(6)
1999	0	0	0	0	0	0
2000	0.0283***	0.0252***	0.0246**	0.0246**	0.0245**	0.0242**
2001	0.0514***	0.0449***	0.0460***	0.0460***	0.0457***	0.0450***
2002	0.0663***	0.0570***	0.0605***	0.0605***	0.0600***	0.0590***
2003	0.0732***	0.0630***	0.0689***	0.0688***	0.0687***	0.0674***
2004	0.0889***	0.0778***	0.0874***	0.0874***	0.0873***	0.0861***
2005	0.0918***	0.0798***	0.0929***	0.0929***	0.0928***	0.0914***
2006	0.105***	0.0922***	0.107***	0.107***	0.107***	0.106***
2007	0.111***	0.0983***	0.116***	0.116***	0.116***	0.115***
2008	0.124***	0.110***	0.131***	0.131***	0.131***	0.130***
2009	0.130***	0.117***	0.142***	0.142***	0.142***	0.141***
2010	0.132***	0.119***	0.148***	0.148***	0.148***	0.147***
2011	0.146***	0.132***	0.165***	0.166***	0.166***	0.165***
2012	0.146***	0.134***	0.171***	0.172***	0.172***	0.171***
2013	0.158***	0.133***	0.176***	0.177***	0.177***	0.178***
ln(Number of references)	0.0369***	0.0451***	0.0526***	0.0521***	0.0522	0.0528***
Impact Factor		-0.00378***	-0.00295***	-0.00286***	-0.00277***	-0.00253***
Maximum h-index			-0.00451***	-0.00440***	-0.00435***	-0.00408***
Number of authors				-0.0000829***	-0.0000791***	0.000241***
International Collaboration					-0.0109***	-0.00300*
Number of different organisations						-0.00366***
Observations	395191	395191	390008	390008	389140	389140
Log-Likelihood	-205481.1	-204992.4	-201663.7	-201575.7	-201155.6	-200938.2
AIC	410996.2	410020.8	403365.4	403191.4	402353.2	401920.4
BIC	411181.3	410216.8	403572.0	403408.8	402581.5	402159.6

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

Table 3.81: Tobit estimates for the effect of the variable “year” on interdisciplinarity measured with the Simpson index (1-10 disciplines, OST nomenclature)

	(1)	(2)	(3)	(4)	(5)	(6)
1999	0	0	0	0	0	0
2000	0.0383***	0.0333**	0.0348**	0.0348**	0.0346**	0.0342**
2001	0.0705***	0.0601***	0.0648***	0.0647***	0.0643***	0.0632***
2002	0.0913***	0.0764***	0.0847***	0.0847***	0.0840***	0.0825***
2003	0.102***	0.0857***	0.0976***	0.0975***	0.0973***	0.0953***
2004	0.125***	0.107***	0.124***	0.124***	0.124***	0.122***
2005	0.131***	0.112***	0.135***	0.135***	0.135***	0.133***
2006	0.152***	0.131***	0.157***	0.157***	0.157***	0.155***
2007	0.163***	0.143***	0.173***	0.173***	0.173***	0.172***
2008	0.183***	0.161***	0.196***	0.196***	0.196***	0.194***
2009	0.194***	0.173***	0.215***	0.214***	0.214***	0.213***
2010	0.200***	0.180***	0.227***	0.227***	0.226***	0.226***
2011	0.222***	0.201***	0.255***	0.256***	0.256***	0.254***
2012	0.226***	0.207***	0.267***	0.268***	0.268***	0.267***
2013	0.245***	0.205***	0.276***	0.277***	0.278***	0.278***
ln(Number of references)	0.0930***	0.106***	0.118***	0.117***	0.118***	0.118***
Impact Factor		-0.00608***	-0.00477***	-0.00467***	-0.00452***	-0.00416***
Maximum h-index			-0.00708***	-0.00696***	-0.00688***	-0.00647***
Number of authors				-0.0000942***	-0.0000882***	0.000396***
International Collaboration					-0.0170***	-0.00509***
Number of different organisations						-0.00554***
Observations	395191	395191	390008	390008	389140	389140
Log-Likelihood	-331978.9	-331313.4	-326502.9	-326443.0	-325739.8	-325477.4
AIC	663991.8	662662.8	653043.9	652926.0	651521.6	650998.8
BIC	664176.8	662858.8	653250.5	653143.5	651749.9	651238.0

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

Table 3.82: Tobit estimates for the effect of the variable “year” on interdisciplinarity measured with the Shannon index (1-10 disciplines, OST nomenclature)

	(1)	(2)	(3)	(4)	(5)	(6)
1999	0	0	0	0	0	0
2000	0.0376**	0.0317**	0.0337*	0.0337*	0.0334*	0.0329*
2001	0.0708***	0.0584***	0.0642***	0.0642***	0.0636***	0.0623***
2002	0.0930***	0.0751***	0.0854***	0.0854***	0.0846***	0.0828***
2003	0.105***	0.0855***	0.100***	0.1000***	0.0996***	0.0973***
2004	0.131***	0.109***	0.130***	0.130***	0.130***	0.128***
2005	0.140***	0.117***	0.145***	0.145***	0.145***	0.142***
2006	0.166***	0.142***	0.173***	0.173***	0.172***	0.170***
2007	0.180***	0.156***	0.192***	0.192***	0.192***	0.190***
2008	0.206***	0.180***	0.222***	0.222***	0.222***	0.220***
2009	0.220***	0.195***	0.245***	0.244***	0.244***	0.242***
2010	0.231***	0.206***	0.262***	0.262***	0.262***	0.261***
2011	0.260***	0.235***	0.299***	0.300***	0.299***	0.298***
2012	0.269***	0.246***	0.318***	0.318***	0.318***	0.317***
2013	0.294***	0.246***	0.331***	0.332***	0.332***	0.333***
ln(Number of references)	0.152	0.168***	0.182	0.181***	0.182	0.183***
Impact Factor		-0.00729***	-0.00573***	-0.00565***	-0.00547***	-0.00505***
Maximum h-index			-0.00843***	-0.00833***	-0.00823***	-0.00776***
Number of authors				-0.0000703***	-0.0000627***	0.000499***
International Collaboration					-0.0214***	-0.00761***
Number of different organisations						-0.00642***
Observations	395191	395191	390008	390008	389140	389140
Log-Likelihood	-383305.2	-382568.4	-377130.4	-377104.7	-376278.1	-376006.1
AIC	766644.4	765172.8	754298.9	754249.4	752598.2	752056.2
BIC	766829.5	765368.7	754505.5	754466.9	752826.5	752295.4

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

Table 3.83: Tobit estimates for the effect of the variable “year” on interdisciplinarity measured with the number of disciplines (1-10 disciplines, OST nomenclature)

	(1)	(2)	(3)	(4)	(5)	(6)
1999	0	0	0	0	0	0
2000	0.000315	-0.00680	-0.00396	-0.00401	-0.00473	-0.00518
2001	0.00818	-0.00663	0.00114	0.00116	-0.0000348	-0.00135
2002	0.0179	-0.00347	0.0103	0.0103	0.00902	0.00724
2003	0.0247	0.00109	0.0207	0.0208	0.0197	0.0173
2004	0.0510**	0.0255	0.0525**	0.0524**	0.0512**	0.0490**
2005	0.0677***	0.0401*	0.0759***	0.0759***	0.0748***	0.0722***
2006	0.108***	0.0787***	0.118***	0.118***	0.117***	0.115***
2007	0.125***	0.0962***	0.143***	0.143***	0.142***	0.140***
2008	0.172***	0.142***	0.195***	0.196***	0.195***	0.193***
2009	0.196***	0.166***	0.230***	0.230***	0.229***	0.227***
2010	0.221***	0.191***	0.263***	0.263***	0.262***	0.261***
2011	0.275***	0.244***	0.326***	0.326***	0.325***	0.324***
2012	0.303***	0.275***	0.366***	0.365***	0.364***	0.364***
2013	0.346***	0.289***	0.396***	0.396***	0.396***	0.396***
ln(Number of references)	0.420***	0.439***	0.457***	0.458***	0.459***	0.460***
Impact Factor		-0.00870***	-0.00676***	-0.00683***	-0.00650***	-0.00607***
Maximum h-index			-0.0105***	-0.0106***	-0.0104***	-0.00994***
Number of authors				0.0000662***	0.0000793***	0.000655***
International Collaboration					-0.0369***	-0.0228***
Number of different organisations						-0.00659***
Observations	395191	395191	390008	390008	389140	389140
Log-Likelihood	-513296.1	-512751.9	-506275.9	-506264.2	-505142.1	-504994.4
AIC	1026626.1	1025539.8	1012589.9	1012568.4	1010326.1	1010032.8
BIC	1026811.2	1025735.7	1012796.5	1012785.9	1010554.4	1010271.9

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

Table 3.84: Tobit estimates for the effect of the variable "year" on interdisciplinarity measured with the Berger-Parker index (1-35 disciplines, CWTS nomenclature)

	(1)	(2)	(3)	(4)	(5)	(6)
1999	0	0	0	0	0	0
2000	0.0221*	0.0192*	0.0272*	0.0272*	0.0268*	0.0265*
2001	0.0463***	0.0404***	0.0494***	0.0494***	0.0486***	0.0478***
2002	0.0606***	0.0518***	0.0636***	0.0636***	0.0625***	0.0614***
2003	0.0709***	0.0613***	0.0758***	0.0758***	0.0749***	0.0735***
2004	0.0909***	0.0806***	0.0994***	0.0994***	0.0985***	0.0972***
2005	0.0962***	0.0854***	0.109***	0.109***	0.108***	0.106***
2006	0.111***	0.0997***	0.125***	0.125***	0.124***	0.123***
2007	0.123***	0.111***	0.141***	0.141***	0.140***	0.138***
2008	0.136***	0.124***	0.157***	0.157***	0.156***	0.155***
2009	0.144***	0.132***	0.170***	0.170***	0.169***	0.168***
2010	0.149***	0.137***	0.180***	0.180***	0.179***	0.178***
2011	0.169***	0.157***	0.205***	0.205***	0.204***	0.203***
2012	0.177***	0.166***	0.219***	0.219***	0.218***	0.218***
2013	0.191***	0.168***	0.230***	0.230***	0.230***	0.230***
ln(Number of references)	0.0642***	0.0719***	0.0822***	0.0822***	0.0828***	0.0834***
Impact Factor		-0.00346***	-0.00262***	-0.00263***	-0.00238***	-0.00219***
Maximum h-index			-0.00573***	-0.00574***	-0.00560***	-0.00533***
Number of authors				0.00000540	0.0000152	0.000327***
International Collaboration					-0.0278***	-0.0201***
Number of different organisations						-0.00357***
Observations	400983	400983	395750	395750	394880	394880
Log-Likelihood	-309943.3	-309593.9	-304846.6	-304846.4	-304048.5	-303920.3
AIC	619920.6	619223.8	609731.3	609732.8	608139.0	607884.6
BIC	620106.0	619420.0	609938.2	609950.6	608367.6	608124.1

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

Table 3.85: Tobit estimates for the effect of the variable "year" on interdisciplinarity measured with the Simpson index (1-35 disciplines, CWTS nomenclature)

	(1)	(2)	(3)	(4)	(5)	(6)
1999	0	0	0	0	0	0
2000	0.0338*	0.0289*	0.0405**	0.0404**	0.0397**	0.0393**
2001	0.0706***	0.0605***	0.0745***	0.0746***	0.0733***	0.0721***
2002	0.0900***	0.0750***	0.0941***	0.0941***	0.0923***	0.0908***
2003	0.105***	0.0889***	0.113***	0.113***	0.111***	0.109***
2004	0.134***	0.116***	0.147***	0.147***	0.146***	0.144***
2005	0.146***	0.128***	0.166***	0.166***	0.165***	0.163***
2006	0.169***	0.150***	0.192***	0.192***	0.190***	0.188***
2007	0.188***	0.168***	0.216***	0.217***	0.215***	0.213***
2008	0.213***	0.192***	0.247***	0.247***	0.245***	0.244***
2009	0.226***	0.205***	0.269***	0.270***	0.268***	0.267***
2010	0.236***	0.215***	0.287***	0.287***	0.286***	0.285***
2011	0.270***	0.250***	0.331***	0.331***	0.330***	0.328***
2012	0.286***	0.267***	0.357***	0.356***	0.355***	0.355***
2013	0.313***	0.274***	0.378***	0.377***	0.378***	0.378***
ln(Number of references)	0.161***	0.174***	0.191***	0.192***	0.193***	0.194***
Impact Factor		-0.00587***	-0.00446***	-0.00451***	-0.00412***	-0.00384***
Maximum h-index			-0.00978***	-0.00987***	-0.00965***	-0.00925***
Number of authors				0.0000638***	0.0000796***	0.000538***
International Collaboration					-0.0445***	-0.0332***
Number of different organisations						-0.00524***
Observations	400983	400983	395750	395750	394880	394880
Log-Likelihood	-449534.6	-449031.9	-442708.5	-442692.8	-441593.1	-441455.5
AIC	899103.3	898099.9	885455.1	885425.6	883228.1	882955.1
BIC	899288.6	898296.1	885661.9	885643.3	883456.8	883194.5

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

Table 3.86: Tobit estimates for the effect of the variable "year" on interdisciplinarity measured with the Shannon index (1-35 disciplines, CWTS nomenclature)

	(1)	(2)	(3)	(4)	(5)	(6)
1999	0	0	0	0	0	0
2000	0.0298	0.0237	0.0363*	0.0362*	0.0353*	0.0348*
2001	0.0673***	0.0549***	0.0709***	0.0711***	0.0695***	0.0684***
2002	0.0861***	0.0676***	0.0902***	0.0903***	0.0883***	0.0867***
2003	0.103***	0.0825***	0.111***	0.111***	0.110***	0.108***
2004	0.135***	0.113***	0.150***	0.150***	0.149***	0.147***
2005	0.154***	0.131***	0.178***	0.178***	0.176***	0.174***
2006	0.182***	0.158***	0.209***	0.209***	0.208***	0.205***
2007	0.204***	0.179***	0.239***	0.239***	0.237***	0.235***
2008	0.239***	0.213***	0.280***	0.281***	0.279***	0.277***
2009	0.256***	0.230***	0.309***	0.310***	0.308***	0.307***
2010	0.270***	0.245***	0.333***	0.334***	0.332***	0.331***
2011	0.317***	0.291***	0.391***	0.390***	0.389***	0.388***
2012	0.340***	0.317***	0.428***	0.426***	0.425***	0.424***
2013	0.378***	0.329***	0.458***	0.456***	0.457***	0.457***
ln(Number of references)	0.255***	0.271***	0.293***	0.295***	0.296***	0.297***
Impact Factor		-0.00728***	-0.00552***	-0.00568***	-0.00523***	-0.00493***
Maximum h-index			-0.0122***	-0.0125***	-0.0122***	-0.0118***
Number of authors				0.000195***	0.000213***	0.000681***
International Collaboration					-0.0519***	-0.0403***
Number of different organisations						-0.00535***
Observations	400983	400983	395750	395750	394880	394880
Log-Likelihood	-505700.7	-505117.3	-498041.7	-497930.8	-496724.2	-496615.8
AIC	1011435.3	1010270.6	996121.5	995901.6	993490.4	993275.6
BIC	1011620.6	1010466.8	996328.4	996119.3	993719.0	993515.1

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

Table 3.87: Tobit estimates for the effect of the variable "year" on interdisciplinarity measured with the number of disciplines (1-35 disciplines, CWTS nomenclature)

	(1)	(2)	(3)	(4)	(5)	(6)
1999	0	0	0	0	0	0
2000	-0.0288	-0.0368	-0.0221	-0.0224	-0.0235	-0.0236
2001	-0.0274	-0.0438*	-0.0231	-0.0226	-0.0245	-0.0248
2002	-0.0320	-0.0564**	-0.0258	-0.0253	-0.0275	-0.0278
2003	-0.0252	-0.0518*	-0.0122	-0.0112	-0.0133	-0.0138
2004	0.00911	-0.0195	0.0313	0.0316	0.0294	0.0290
2005	0.0411*	0.0111	0.0758***	0.0762***	0.0741***	0.0736**
2006	0.0839***	0.0519*	0.123***	0.123***	0.121***	0.121***
2007	0.113***	0.0809***	0.163***	0.165***	0.162***	0.161***
2008	0.178***	0.144***	0.237***	0.239***	0.237***	0.236***
2009	0.208***	0.174***	0.283***	0.286***	0.284***	0.284***
2010	0.241***	0.207***	0.330***	0.332***	0.329***	0.329***
2011	0.315***	0.281***	0.419***	0.417***	0.416***	0.415***
2012	0.369***	0.338***	0.492***	0.484***	0.483***	0.483***
2013	0.434***	0.368***	0.546***	0.541***	0.541***	0.541***
ln(Number of references)	0.638***	0.659***	0.691***	0.695***	0.697***	0.697***
Impact Factor		-0.00957***	-0.00714***	-0.00771***	-0.00724***	-0.00717***
Maximum h-index			-0.0169***	-0.0179***	-0.0176***	-0.0175***
Number of authors				0.000711***	0.000732***	0.000836***
International Collaboration					-0.0564***	-0.0538***
Number of different organisations						-0.00119*
Observations	400983	400983	395750	395750	394880	394880
Log-Likelihood	-633762.1	-633229.7	-624888.2	-624108.6	-622735.4	-622732.5
AIC	1267558.2	1266495.3	1249814.4	1248257.1	1245512.7	1245509.0
BIC	1267743.6	1266691.5	1250021.3	1248474.9	1245741.3	1245748.5

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

Table 3.88: Tobit estimates for the effect of the variable “year” on interdisciplinarity measured with the Berger-Parker index (1-252 disciplines, WoS nomenclature)

	(1)	(2)	(3)	(4)	(5)	(6)
1999	0	0	0	0	0	0
2000	0.0330	0.0283	0.0510**	0.0510**	0.0504**	0.0495*
2001	0.0713***	0.0616***	0.0855***	0.0854***	0.0839***	0.0816***
2002	0.0971***	0.0827***	0.112***	0.112***	0.110***	0.107***
2003	0.0997***	0.0840***	0.117***	0.117***	0.116***	0.112***
2004	0.143***	0.126***	0.166***	0.166***	0.165***	0.161***
2005	0.158***	0.140***	0.187***	0.187***	0.186***	0.181***
2006	0.179***	0.160***	0.210***	0.210***	0.209***	0.205***
2007	0.196***	0.177***	0.234***	0.234***	0.232***	0.228***
2008	0.224***	0.203***	0.266***	0.266***	0.264***	0.261***
2009	0.242***	0.222***	0.294***	0.293***	0.292***	0.289***
2010	0.255***	0.235***	0.313***	0.313***	0.311***	0.310***
2011	0.291***	0.271***	0.358***	0.358***	0.357***	0.354***
2012	0.308***	0.290***	0.385***	0.386***	0.385***	0.383***
2013	0.336***	0.297***	0.405***	0.406***	0.408***	0.408***
ln(Number of references)	0.264***	0.276***	0.292***	0.292***	0.294***	0.295***
Impact Factor	-0.00564***		-0.00434***	-0.00429***	-0.00373***	-0.00316***
Maximum h-index			-0.00916***	-0.00907***	-0.00875***	-0.00797***
Number of authors				-0.0000626***	-0.0000395**	0.000860***
International Collaboration					-0.0656***	-0.0434***
Number of different organisations						-0.0103***
Observations	401522	401522	396287	396287	395417	395417
Log-Likelihood	-550880.0	-550598.8	-543320.9	-543311.7	-541926.2	-541605.6
AIC	1101794.0	1101233.7	1086679.7	1086663.4	1083894.3	1083255.2
BIC	1101979.4	1101429.9	1086886.6	1086881.2	1084123.0	1083494.7

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

Table 3.89: Tobit estimates for the effect of the variable “year” on interdisciplinarity measured with the Simpson index (1-252 disciplines, WoS nomenclature)

	(1)	(2)	(3)	(4)	(5)	(6)
1999	0	0	0	0	0	0
2000	0.0381	0.0303	0.0548	0.0548	0.0531	0.0515
2001	0.0821**	0.0661*	0.0949***	0.0948**	0.0916**	0.0873**
2002	0.110***	0.0862***	0.125***	0.125***	0.120***	0.115***
2003	0.113***	0.0872***	0.133***	0.133***	0.130***	0.122***
2004	0.168***	0.140***	0.198***	0.198***	0.195***	0.187***
2005	0.197***	0.168***	0.238***	0.238***	0.234***	0.226***
2006	0.235***	0.204***	0.280***	0.280***	0.277***	0.269***
2007	0.264***	0.233***	0.320***	0.320***	0.315***	0.308***
2008	0.320***	0.286***	0.384***	0.384***	0.380***	0.373***
2009	0.352***	0.320***	0.432***	0.432***	0.428***	0.422***
2010	0.381***	0.348***	0.473***	0.473***	0.469***	0.466***
2011	0.450***	0.417***	0.557***	0.557***	0.555***	0.549***
2012	0.495***	0.464***	0.619***	0.620***	0.618***	0.615***
2013	0.556***	0.492***	0.670***	0.670***	0.673***	0.674***
ln(Number of references)	0.700***	0.721***	0.750***	0.749***	0.752***	0.755***
Impact Factor	-0.00933***	-0.00706***	-0.00706***	-0.00698***	-0.00595***	-0.00488***
Maximum h-index			-0.0161***	-0.0159***	-0.0154***	-0.0139***
Number of authors				-0.0000950***	-0.0000529*	0.00167***
International Collaboration					-0.119***	-0.0768***
Number of different organisations						-0.0197***
Observations	401522	401522	396287	396287	395417	395417
Log-Likelihood	-723976.8	-723652.5	-714656.5	-714647.7	-712855.2	-712359.7
AIC	1447987.6	1447341.0	1429351.1	1429335.3	1425752.3	1424763.5
BIC	1448173.0	1447537.3	1429558.0	1429553.1	1425981.0	1425003.0

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

Table 3.90: Tobit estimates for the effect of the variable “year” on interdisciplinarity measured with the Shannon index (1-252 disciplines, WoS nomenclature)

	(1)	(2)	(3)	(4)	(5)	(6)
1999	0	0	0	0	0	0
2000	-0.00562	-0.0144	0.0105	0.0105	0.00811	0.00606
2001	0.00955	-0.00846	0.0229	0.0228	0.0184	0.0128
2002	0.0189	-0.00785	0.0354	0.0353	0.0295	0.0221
2003	0.0153	-0.0139	0.0392	0.0390	0.0344	0.0244
2004	0.0695*	0.0381	0.105**	0.105**	0.0994**	0.0902**
2005	0.106***	0.0726*	0.154***	0.154***	0.149***	0.138***
2006	0.151***	0.116***	0.204***	0.203***	0.199***	0.189***
2007	0.185***	0.150***	0.251***	0.251***	0.245***	0.236***
2008	0.265***	0.227***	0.340***	0.340***	0.335***	0.326***
2009	0.304***	0.267***	0.398***	0.397***	0.392***	0.385***
2010	0.346***	0.309***	0.455***	0.455***	0.450***	0.445***
2011	0.440***	0.403***	0.567***	0.567***	0.563***	0.556***
2012	0.507***	0.473***	0.654***	0.655***	0.652***	0.649***
2013	0.595***	0.524***	0.731***	0.732***	0.735***	0.736***
ln(Number of references)	1.075***	1.098***	1.133***	1.133***	1.137***	1.141***
Impact Factor		-0.0105***	-0.00782***	-0.00775***	-0.00642***	-0.00504***
Maximum h-index			-0.0190***	-0.0189***	-0.0181***	-0.0162***
Number of authors				-0.0000865***	-0.0000318	0.00217***
International Collaboration					-0.155***	-0.100***
Number of different organisations						-0.0252***
Observations	401522	401522	396287	396287	395417	395417
Log-Likelihood	-784806.8	-784503.0	-774875.2	-774869.8	-772881.0	-772281.3
AIC	1569647.6	1569042.0	1549788.4	1549779.5	1545804.0	1544606.5
BIC	1569833.0	1569238.2	154995.3	154997.3	1546032.6	1544846.0

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

Table 3.91: Tobit estimates for the effect of the variable “year” on interdisciplinarity measured with the number of disciplines (1-252 disciplines, WoS nomenclature)

	(1)	(2)	(3)	(4)	(5)	(6)
1999	0	0	0	0	0	0
2000	-0.186***	-0.194***	-0.168***	-0.168***	-0.172***	-0.174***
2001	-0.304***	-0.320***	-0.285***	-0.285***	-0.292***	-0.299***
2002	-0.374***	-0.398***	-0.348***	-0.348***	-0.356***	-0.365***
2003	-0.415***	-0.442***	-0.378***	-0.378***	-0.386***	-0.398***
2004	-0.376***	-0.405***	-0.327***	-0.326***	-0.335***	-0.347***
2005	-0.336***	-0.366***	-0.271***	-0.271***	-0.279***	-0.292***
2006	-0.287***	-0.319***	-0.216***	-0.216***	-0.224***	-0.236***
2007	-0.240***	-0.272***	-0.154***	-0.154***	-0.164***	-0.175***
2008	-0.108**	-0.143***	-0.0107	-0.0107	-0.0193	-0.0294
2009	-0.0520	-0.0858*	0.0668	0.0668	0.0589	0.0500
2010	0.0290	-0.00475	0.164***	0.164***	0.156***	0.151***
2011	0.168***	0.133***	0.323***	0.323***	0.316***	0.308***
2012	0.288***	0.257***	0.466***	0.466***	0.461***	0.457***
2013	0.434***	0.368***	0.607***	0.607***	0.612***	0.613***
ln(Number of references)	2.088***	2.110***	2.154***	2.154***	2.160***	2.165***
Impact Factor		-0.00968***	-0.00658***	-0.00659***	-0.00465***	-0.00296***
Maximum h-index						
Number of authors						
International Collaboration						
Number of different organisations				0.00000408	0.0000827*	0.00277***
Observations	401522	401522	396287	396287	395417	395417
Log-Likelihood	-892848.6	-892697.9	-882066.8	-882066.8	-879787.0	-879266.9
AIC	1785731.2	1785431.7	1764171.6	1764173.6	1759616.1	1758577.8
BIC	1785916.5	1785628.0	1764378.5	1764391.4	1759844.7	1758817.3

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

10.7 Individual Fixed Effects

Table 3.92: OLS estimates for the effect of interdisciplinarity on the logarithm of the number of citations (papers with at least 4 references, Individual effects, 1-10 disciplines, OST nomenclature)

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Berger-Parker	-0.0416*** (-3.77)	-0.0372*** (-6.02)						
Simpson			-0.0387*** (-4.08)	-0.0297*** (-6.13)				
Shannon					-0.0406*** (-4.10)	-0.0227*** (-4.82)		
1 discipline							0 (.)	0 (.)
2 disciplines							0.0436*** (4.05)	0.0349*** (3.80)
3 disciplines and more								0.0968*** (6.99)
Number of authors	0.000487*** (8.72)	0.000448*** (9.11)	0.000486*** (8.71)	0.000447*** (9.11)	0.000486*** (8.72)	0.000448*** (9.12)	0.000489*** (8.73)	0.000505*** (9.47)
International Collaboration	0.239*** (38.83)	0.222*** (45.38)	0.239*** (38.84)	0.222*** (45.37)	0.239*** (38.84)	0.222*** (45.36)	0.238*** (38.76)	0.247*** (44.97)
Maximal distance	0.0860*** (7.85)	0.120*** (15.89)	0.0932*** (8.10)	0.127*** (15.78)	0.103*** (8.12)	0.129*** (14.47)	0.00412 (0.21)	0.0103 (0.66)
Impact Factor	0.0764*** (67.57)	0.0714*** (66.45)	0.0764*** (67.56)	0.0714*** (66.45)	0.0764*** (67.55)	0.0715*** (66.45)	0.0763*** (67.53)	0.0544*** (33.38)
Observations	300179	476747	300179	476747	300179	476747	300179	486393
Log-Likelihood	-347784.8	-555384.4	-347781.8	-555380.9	-347781.2	-555397.5	-347781.1	-579437.4
AIC	695635.6	1110834.8	695629.6	1110827.9	695628.3	1110861.1	695628.2	1158942.7
BIC	695985.8	1111200.2	695979.8	1111193.4	695978.5	1111226.5	695978.4	1159319.9

t statistics in parentheses

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

Table 3.93: OLS estimates for the effect of interdisciplinarity measured with the number of disciplines on the logarithm of the number of citations (papers with at least 4 references, No individual effects, 1-10 disciplines, OST nomenclature)

	(1)	(2)
1 discipline	0 (.)	0 (.)
2 disciplines	0.0221** (2.76)	0.0278*** (3.82)
3 disciplines and more		0.0868*** (7.76)
Number of authors	0.000401*** (12.32)	0.000385*** (16.19)
International Collaboration	0.246*** (54.01)	0.265*** (62.35)
Maximal distance	-0.0112 (-0.77)	-0.0297* (-2.33)
Impact Factor	0.0824*** (94.83)	0.0621*** (40.32)
Observations	303175	488506
Log-Likelihood	-376200.7	-618909.2
AIC	752467.3	1237886.3
BIC	752817.9	1238263.7

t statistics in parentheses

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

Table 3.94: OLS estimates for the effect of interdisciplinarity on the logarithm of the number of citations (papers with at least 4 references, Individual effects, 1-35 disciplines, CWTS nomenclature)

	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
Berger-Parker	-0.0430*** (-3.29)	-0.0315*** (-6.56)						
Simpson			-0.0421*** (-3.76)	-0.0213*** (-6.01)				
Shannon					-0.0451*** (-3.92)	-0.0116*** (-3.48)		
1 discipline							0 (.)	0 (.)
2 disciplines							0.0456*** (3.91)	0.0173* (1.97)
3 disciplines								0.0505*** (4.49)
4 and more disciplines								0.104*** (7.84)
Number of authors	0.000966*** (7.92)	0.000477*** (9.17)	0.000966*** (7.92)	0.000478*** (9.18)	0.000966*** (7.92)	0.000479*** (9.21)	0.000969*** (7.92)	0.000469*** (8.96)
International Collaboration	0.288*** (33.15)	0.249*** (45.36)	0.288*** (33.16)	0.249*** (45.36)	0.288*** (33.16)	0.249*** (45.35)	0.288*** (33.15)	0.249*** (45.43)
Maximal distance	0.0806*** (6.30)	0.133*** (16.98)	0.0888*** (6.62)	0.137*** (16.64)	0.0991*** (6.82)	0.132*** (14.84)	-0.00146 (-0.07)	0.0287* (2.17)
Impact Factor	0.0529*** (20.95)	0.0540*** (33.06)	0.0529*** (20.94)	0.0540*** (33.06)	0.0529*** (20.94)	0.0540*** (33.06)	0.0529*** (20.94)	0.0540*** (33.13)
Observations	223299	484472	223299	484472	223299	484472	223299	485367
Log-Likelihood	-265925.6	-576936.7	-265921.9	-576942.8	-265920.3	-576969.5	-265918.7	-577857.2
AIC	531965.1	1153989.4	531957.8	1154001.5	531954.6	1154055.1	531951.3	1155834.3
BIC	532553.2	1154632.7	532545.8	1154644.8	532542.6	1154698.3	532539.4	1156499.9

t statistics in parentheses

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

Table 3.95: OLS estimates for the effect of interdisciplinarity measured with the number of disciplines on the logarithm of the number of citations (papers with at least 4 references, No individual effects, 1-35 disciplines, CWTS nomenclature)

	(1)	(2)
1 discipline	0 (.)	0 (.)
2 disciplines	0.0208* (2.31)	0.00260 (0.37)
3 disciplines		0.0413*** (4.45)
4 and more disciplines		0.0910*** (8.23)
Number of authors	0.000622*** (9.74)	0.000294*** (12.76)
International Collaboration	0.270*** (42.50)	0.245*** (60.07)
Maximal distance	-0.0103 (-0.63)	0.00136 (0.12)
Impact Factor	0.0588*** (25.10)	0.0595*** (39.32)
Observations	226569	487452
Log-Likelihood	-289140.2	-612440.8
AIC	578394.3	1225001.5
BIC	578983.2	1225667.3

t statistics in parentheses

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

Table 3.96: OLS estimates for the effect of interdisciplinarity on the logarithm of the number of citations (papers with at least 6 references, Individual effects, 1-252 disciplines, WoS nomenclature)

	(1)	(2)	(3)	(4)
Berger-Parker	-0.00154			
Simpson		0.00457**		
Shannon			0.0119***	
1 discipline				0
2 disciplines				0.0520**
3 disciplines				0.0521**
4 disciplines				0.0733***
5 disciplines				0.0780***
6 disciplines				0.0952***
7 disciplines				0.112***
8-9 disciplines				0.141***
9 and more disciplines				0.233***
Number of authors	0.000586***	0.000586***	0.000585***	0.000581***
International Collaboration	0.247***	0.247***	0.248***	0.248***
Impact Factor	0.0513***	0.0513***	0.0513***	0.0512***
Observations	364863	364863	364863	364876
Log-Likelihood	-432898.5	-432888.4	-432796.0	-432307.5
AIC	865907.0	865886.8	865701.9	864739.0
BIC	866501.4	866481.2	866296.3	865409.1

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

Table 3.97: OLS estimates for the effect of interdisciplinarity measured with the number of disciplines on the logarithm of the number of citations (papers with at least 6 references, No individual effect, 1-252 disciplines, WoS nomenclature)

	(1)	(2)	(3)	(4)
Berger-Parker	-0.0122***			
Simpson		-0.00348*		
Shannon			0.00377**	
1 discipline				0
2 disciplines				-0.0246
3 disciplines				-0.0297
4 disciplines				-0.0221
5 disciplines				-0.0270
6 disciplines				-0.01000
7 disciplines				-0.00108
8-9 disciplines				0.0255
9 and more disciplines				0.117***
Number of authors	0.000399***	0.000399***	0.000400***	0.000407***
International Collaboration	0.264***	0.264***	0.264***	0.265***
Maximal distance				
Impact Factor	0.0574***	0.0575***	0.0575***	0.0574***
Observations	367148	367148	367148	367160
Log-Likelihood	-464699.7	-464726.0	-464720.4	-464254.6
AIC	929509.4	929562.0	929550.8	928633.2
BIC	930104.2	930156.8	930145.5	929303.6

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

Catcode	Subject Category	Discipline(s)	
AA	Acoustics	Physical Sciences	Mechanical Engineering and Aerospace
AC	Automation & Control Systems	Engineering Sciences	Electrical Engineering and Telecommunication
AD	Agriculture, Dairy & Animal Science	Applied Biology - Ecology	Agriculture and Food Science
AE	Agricultural Engineering	Applied Biology - Ecology	Agriculture and Food Science
AF	Agricultural Economics & Policy	Applied Biology - Ecology	Economics and Business
AH	Agriculture, Multidisciplinary	Applied Biology - Ecology	Agriculture and Food Science
AI	Engineering, Aerospace	Engineering Sciences	Mechanical Engineering and Aerospace
AM	Agronomy	Applied Biology - Ecology	Agriculture and Food Science
AQ	Allergy	Medical Research	Clinical Medicine
AY	Anatomy & Morphology	Basic Biology	Biomedical Sciences
AZ	Andrology	Medical Research	Clinical Medicine
BA	Anesthesiology	Medical Research	Clinical Medicine
BD	Biodiversity Conservation	Applied Biology - Ecology	Environmental Sciences and Technology
BF	Anthropology	Human Sciences	Sociology and Anthropology
BI	Archaeology	Human Sciences	History, Philosophy and Religion
BK	Architecture	Human Sciences	Creative Arts, Culture and Music
BM	Area Studies	Human Sciences	Management and Planning
BP	Art	Human Sciences	Creative Arts, Culture and Music
BQ	Humanities, Multidisciplinary	Human Sciences	Creative Arts, Culture and Music
BU	Astronomy & Astrophysics	Sciences of the Universe	Astronomy and Astrophysics
BV	Psychology, Biological	Human Sciences	Psychology
CL	Audiology & Speech-Language Pathology	Medical Research	Clinical Medicine
CN	Behavioral Sciences	Basic Biology	Basic Life Sciences
CO	Biochemical Research Methods	Basic Biology	Basic Life Sciences
CQ	Biochemistry & Molecular Biology	Basic Biology	Basic Life Sciences
CT	Cell & Tissue Engineering	Basic Biology	Basic Life Sciences
CU	Biology	Applied Biology - Ecology	Basic Life Sciences
CX	Biology, Miscellaneous	Applied Biology - Ecology	Biological Sciences
DA	Biophysics	Basic Biology	Basic Life Sciences
DB	Biotechnology & Applied Microbiology	Basic Biology	Basic Life Sciences
DE	Plant Sciences	Applied Biology - Ecology	Basic Life Sciences
DI	Business	Social Sciences	Economics and Business
DK	Business, Finance	Social Sciences	Economics and Business
DM	Oncology	Medical Research	Clinical Medicine
DQ	Cardiac & Cardiovascular Systems	Medical Research	Clinical Medicine

Table 3.98 continued from previous page

Catcode	Subject Category	Discipline OST	Discipline NOWT
DR	Cell Biology	Basic Biology	Basic Life Sciences
DS	Critical Care Medicine	Medical Research	Clinical Medicine
DT	Thermodynamics	Engineering Sciences	Mechanical Engineering and Aerospace
DW	Chemistry, Applied	Chemistry	Chemistry and Chemical Engineering
DX	Chemistry, Medicinal	Medical Research	Basic Medical Sciences
DY	Chemistry, Multidisciplinary	Chemistry	Chemistry and Chemical Engineering
EA	Chemistry, Analytical	Chemistry	Chemistry and Chemical Engineering
EC	Chemistry, Inorganic & Nuclear	Chemistry	Chemistry and Chemical Engineering
EE	Chemistry, Organic	Chemistry	Chemistry and Chemical Engineering
EI	Chemistry, Physical	Chemistry	Chemistry and Chemical Engineering
EN	Cultural Studies	Social Sciences	Sociology and Anthropology
EO	Classics	Human Sciences	Creative Arts, Culture and Music
EP	Computer Science, Artificial Intelligence	Engineering Sciences	Computer Sciences
EQ	Psychology, Clinical	Human Sciences	Psychology
ER	Computer Science, Cybernetics	Engineering Sciences	Computer Sciences
ES	Computer Science, Hardware & Architecture	Engineering Sciences	Computer Sciences
ET	Computer Science, Information Systems	Engineering Sciences	Computer Sciences
EU	Communication	Human Sciences	Information and Communication Sciences
EV	Computer Science, Interdisciplinary Applications	Engineering Sciences	Computer Sciences
EW	Computer Science, Software Engineering	Engineering Sciences	Computer Sciences
EX	Computer Science, Theory & Methods	Engineering Sciences	Computer Sciences
FA	Construction & Building Technology	Engineering Sciences	Civil Engineering and Construction
FE	Criminology & Penology	Social Sciences	Law and Criminology
FF	Emergency Medicine	Medical Research	Clinical Medicine
FI	Crystallography	Chemistry	Physics and Materials Science
FS	Dance	Human Sciences	Creative Arts, Culture and Music
FU	Demography	Human Sciences	Social and Behavioral Sciences, Interdisciplinary
FY	Dentistry, Oral Surgery & Medicine	Medical Research	Clinical Medicine

Table 3.98 continued from previous page

Catcode	Subject Category	Discipline OST	Discipline NOWT
GA	Dermatology	Medical Research	Clinical Medicine
GC	Geochemistry & Geophysics	Sciences of the Universe	Earth Sciences and Technology
GM	Substance Abuse	Medical Research Social Sciences	Health Sciences
GU	Ecology	Applied Biology - Ecology	Environmental Sciences and Technology
GY	Economics	Social Sciences	Economics and Business
HA	Education & Educational Research	Social Sciences	Educational Sciences
HB	Education, Scientific Disciplines	Multidisciplinary Category	Educational Sciences
HE	Education, Special	Social Sciences	Educational Sciences
HF	Ethics	Human Sciences	History, Philosophy and Religion
HI	Psychology, Educational	Human Sciences	Educational Sciences
HL	Health Care Sciences & Services	Medical Research	Health Sciences
HQ	Electrochemistry	Chemistry	Chemistry and Chemical Engineering
HT	Evolutionary Biology	Basic Biology	Basic Life Sciences
HY	Developmental Biology	Basic Biology	Basic Life Sciences
IA	Endocrinology & Metabolism	Medical Research	Clinical Medicine
ID	Energy & Fuels	Engineering Sciences	Energy and Science Technology
IF	Engineering, Multidisciplinary	Engineering Sciences	General and Industrial Engineering
IG	Engineering, Biomedical	Basic Biology	Basic Medical Sciences
IH	Engineering, Environmental	Sciences of the Universe	Environmental Sciences and Technology
II	Engineering, Chemical	Engineering Sciences	Chemistry and Chemical Engineering
IJ	Engineering, Industrial	Engineering Sciences	General and Industrial Engineering
IK	Engineering, Manufacturing	Engineering Sciences	General and Industrial Engineering
IL	Engineering, Marine	Engineering Sciences	Earth Sciences and Technology
IM	Engineering, Civil	Engineering Sciences	Civil Engineering and Construction
IO	Engineering, Ocean	Engineering Sciences	Earth Sciences and Technology
IP	Engineering, Petroleum	Engineering Sciences	Energy and Science Technology
IQ	Engineering, Electrical & Electronic	Engineering Sciences	Electrical Engineering and Telecommunication
IU	Engineering, Mechanical	Engineering Sciences	Mechanical Engineering and Aerospace
IX	Engineering, Geological	Sciences of the Universe	Earth Sciences and Technology
IY	Entomology	Applied Biology - Ecology	Basic Life Sciences

Table 3.98 continued from previous page

Catcode	Subject Category	Discipline OST	Discipline NOWT
JA	Environmental Sciences	Sciences of the Universe	Environmental Sciences and Technology
JB	Environmental Studies	Social Sciences	Environmental Sciences and Technology
JI	Ergonomics	Social Sciences	General and Industrial Engineering
JM	Ethnic Studies	Human Sciences	Sociology and Anthropology
JO	Family Studies	Social Sciences	Sociology and Anthropology
JS	Film, Radio, Television	Human Sciences	Creative Arts, Culture and Music
JU	Fisheries	Applied Biology - Ecology	Basic Life Sciences
JW	Folklore	Human Sciences	Creative Arts, Culture and Music
JY	Food Science & Technology	Applied Biology - Ecology	Agriculture and Food Science
KA	Forestry	Applied Biology - Ecology	Environmental Sciences and Technology
KI	Gastroenterology & Hepatology	Medical Research	Clinical Medicine
KM	Genetics & Heredity	Basic Biology	Basic Life Sciences
KU	Geography	Human Sciences	Environmental Sciences and Technology
KV	Geography, Physical	Sciences of the Universe	Earth Sciences and Technology
KY	Geology	Sciences of the Universe	Earth Sciences and Technology
LE	Geosciences, Multidisciplinary	Sciences of the Universe	Earth Sciences and Technology
LI	Geriatrics & Gerontology	Medical Research	Health Sciences
LJ	Gerontology	Social Sciences	Health Sciences
LQ	Health Policy & Services	Social Sciences	Health Sciences
MA	Hematology	Medical Research	Clinical Medicine
MC	Mathematical & Computational Biology	Basic Biology	Basic Life Sciences
ML	Primary Health Care	Medical Research	Clinical Medicine
MM	History	Human Sciences	History, Philosophy and Religion
MQ	History & Philosophy of Science	Human Sciences	History, Philosophy and Religion
MR	History of Social Sciences	Human Sciences	History, Philosophy and Religion
MU	Horticulture	Applied Biology - Ecology	Basic Life Sciences
MW	Hospitality, Leisure, Sport & Tourism	Social Sciences	Sociology and Anthropology
MY	Psychology, Developmental	Human Sciences	Psychology
NE	Public, Environmental & Occupational Health	Medical Research Social Sciences	Clinical medicine
NI	Immunology	Medical Research	Biomedical Sciences
NM	Industrial Relations & Labor	Social Sciences	Economics and Business
NN	Infectious Diseases	Medical Research	Clinical Medicine
NQ	Psychology, Applied	Human Sciences	Psychology

Table 3.98 continued from previous page

Catcode	Subject Category	Discipline OST	Discipline NOWT
NS	Nanoscience & Nanotechnology	Chemistry	Physics and Materials Science
NU	Information Science & Library Science	Social Sciences	Information and Communication Sciences
OA	Instruments & Instrumentation	Physical Sciences	Instruments and Instrumentation
OE	International Relations	Social Sciences	Political Science and Public Administration
OI	Integrative & Complementary Medicine	Medical Research	Biomedical Sciences
OM	Law	Social Sciences	Law and Criminology
OO	Medical Ethics	Medical Research	History, Philosophy and Religion
OP	Medicine, Legal	Medical Research	Law and Criminology
OR	Asian Studies	Human Sciences	Creative Arts, Culture and Music
OT	Linguistics	Human Sciences	Language and Linguistics
OU	Limnology	Sciences of the Universe	Environmental Sciences and Technology
OX	Literary Theory & Criticism	Human Sciences	Literature
OY	Language & Linguistics	Human Sciences	Language and Linguistics
OZ	Literary Reviews	Human Sciences	Literature
PA	Literature	Human Sciences	Literature
PC	Management	Social Sciences	Management and Planning
PD	Literature, African, Australian, Canadian	Human Sciences	Literature
PE	Operations Research & Management Science	Engineering Sciences	Statistical Sciences
PF	Literature, American	Human Sciences	Literature
PG	Literature, British Isles	Human Sciences	Literature
PH	Literature, German, Dutch, Scandinavian	Human Sciences	Literature
PI	Marine & Freshwater Biology	Sciences of the Universe	Basic Life Sciences
PJ	Materials Science, Paper & Wood	Chemistry	Chemistry and Chemical Engineering
PK	Materials Science, Ceramics	Chemistry	Physics and Materials Science
PM	Materials Science, Multidisciplinary	Chemistry	Physics and Materials Science
PN	Mathematics, Applied	Mathematics	Mathematics
PO	Mathematics, Interdisciplinary Applications	Mathematics	Mathematics
PQ	Mathematics	Mathematics	Mathematics
PS	Social Sciences, Mathematical Methods	Social Sciences	Statistical Sciences
PT	Medical Informatics	Basic Biology	Basic Medical Sciences
PU	Mechanics	Engineering Sciences	Mechanical Engineering and Aerospace

Table 3.98 continued from previous page

Catcode	Subject Category	Discipline OST	Discipline NOWT
PW	Medical Laboratory Technology	Basic Biology	Biomedical Sciences
PY	Medicine, General & Internal	Medical Research	Clinical Medicine
PZ	Metallurgy & Metallurgical Engineering	Engineering Sciences	Physics and Materials Science
QA	Medicine, Research & Experimental	Medical Research	Biomedical Sciences
QC	Literature, Romance	Human Sciences	Literature
QD	Literature, Slavic	Human Sciences	Literature
QE	Materials Science, Biomaterials	Basic Biology	Basic Medical Sciences
QF	Materials Science, Characterization & Testing	Chemistry	Physics and Materials Science
QG	Materials Science, Coatings & Films	Chemistry	Physics and Materials Science
QH	Materials Science, Composites	Chemistry	Physics and Materials Science
QJ	Materials Science, Textiles	Applied Biology - Ecology	Chemistry and Chemical Engineering
QK	Medieval & Renaissance Studies	Human Sciences	History, Philosophy and Religion
QL	Logic	Engineering Sciences	Mathematics
QQ	Meteorology & Atmospheric Sciences	Sciences of the Universe	Earth Sciences and Technology
QU	Microbiology	Basic Biology	Basic Life Sciences
RA	Microscopy	Basic Biology	Instruments and Instrumentation
RB	Robotics	Engineering Sciences	Electrical Engineering and Telecommunication
RE	Mineralogy	Sciences of the Universe	Earth Sciences and Technology
RO	Multidisciplinary Sciences	Multidisciplinary Category	Multidisciplinary Journals
RP	Music	Human Sciences	Creative Arts, Culture and Music
RQ	Mycology	Applied Biology - Ecology	Basic Life Sciences
RT	Clinical Neurology	Medical Research	Clinical Medicine
RU	Neurosciences	Basic Biology	Biomedical Sciences
RX	Neuroimaging	Basic Biology	Biomedical Sciences
RY	Nuclear Science & Technology	Engineering Sciences	Energy and Science Technology
RZ	Nursing	Medical Research Social Sciences	Health Sciences
SA	Nutrition & Dietetics	Basic Biology	Agriculture and Food Science
SD	Obstetrics & Gynecology	Medical Research	Clinical Medicine
SI	Oceanography	Sciences of the Universe	Earth Sciences and Technology
SR	Remote Sensing	Engineering Sciences	Earth Sciences and Technology
SU	Ophthalmology	Medical Research	Clinical Medicine
SY	Optics	Physical Sciences	Physics and Materials Science
TA	Ornithology	Applied Biology - Ecology	Basic Life Sciences
TC	Orthopedics	Medical Research	Clinical Medicine

Table 3.98 continued from previous page

Catcode	Subject Category	Discipline OST	Discipline NOWT
TD	Otorhinolaryngology	Medical Research	Clinical Medicine
TE	Paleontology	Sciences of the Universe	Earth Sciences and Technology
TI	Parasitology	Basic Biology	Clinical Medicine
TM	Pathology	Medical Research	Clinical Medicine
TQ	Pediatrics	Medical Research	Clinical Medicine
TU	Pharmacology & Pharmacy	Medical Research	Biomedical Sciences
UA	Philosophy	Human Sciences	History, Philosophy and Religion
UB	Physics, Applied	Physical Sciences	Physics and Materials Science
UE	Imaging Science & Photographic Technology	Engineering Sciences	Earth Sciences and Technology
UF	Physics, Fluids & Plasmas	Physical Sciences	Physics and Materials Science
UH	Physics, Atomic, Molecular & Chemical	Physical Sciences	Physics and Materials Science
UI	Physics, Multidisciplinary	Physical Sciences	Physics and Materials Science
UK	Physics, Condensed Matter	Physical Sciences	Physics and Materials Science
UM	Physiology	Basic Biology	Biomedical Sciences
UN	Physics, Nuclear	Physical Sciences	Physics and Materials Science
UP	Physics, Particles & Fields	Physical Sciences	Physics and Materials Science
UQ	Planning & Development	Social Sciences	Management and Planning
UR	Physics, Mathematical	Physical Sciences	Physics and Materials Science
UT	Poetry	Human Sciences	Literature
UU	Political Science	Social Sciences	Political Science and Public Administration
UY	Polymer Science	Chemistry	Chemistry and Chemical Engineering
VE	Psychiatry	Medical Research Human Sciences	Clinical medicine
VI	Psychology	Basic Biology	Psychology
VJ	Psychology, Multidisciplinary	Human Sciences	Psychology
VM	Public Administration	Social Sciences	Political Science and Public Administration
VP	Psychology, Psychoanalysis	Human Sciences	Psychology
VS	Psychology, Mathematical	Human Sciences	Psychology
VX	Psychology, Experimental	Human Sciences	Psychology
VY	Radiology, Nuclear Medicine & Medical Imaging	Medical Research	Biomedical Sciences
WC	Rehabilitation	Medical Research Social Sciences	Health Sciences
WE	Respiratory System	Medical Research	Clinical Medicine
WF	Reproductive Biology	Basic Biology	Basic Life Sciences
WH	Rheumatology	Medical Research	Clinical Medicine
WM	Social Issues	Social Sciences	Social and Behavioral Sciences, Interdisciplinary
WQ	Psychology, Social	Human Sciences	Psychology
WU	Social Sciences, Interdisciplinary	Social Sciences	Social and Behavioral Sciences, Interdisciplinary

Table 3.98 continued from previous page

Catcode	Subject Category	Discipline OST	Discipline NOWT
WV	Social Sciences, Biomedical	Social Sciences	Social and Behavioral Sciences, Interdisciplinary
WY	Social Work	Social Sciences	Health Sciences
XA	Sociology	Social Sciences	Sociology and Anthropology
XE	Soil Science	Applied Biology - Ecology	Agriculture and Food Science
XQ	Spectroscopy	Physical Sciences	Chemistry and Chemical Engineering
XW	Sport Sciences	Medical Research	Health Sciences
XY	Statistics & Probability	Mathematics	Statistical Sciences
YA	Surgery	Medical Research	Clinical Medicine
YE	Telecommunications	Engineering Sciences	Electrical Engineering and Telecommunication
YG	Theater	Human Sciences	Creative Arts, Culture and Music
YI	Religion	Human Sciences	History, Philosophy and Religion
YO	Toxicology	Medical Research	Biomedical Sciences
YP	Transplantation	Medical Research	Clinical Medicine
YQ	Transportation	Social Sciences	Electrical Engineering and Telecommunication
YR	Transportation Science & Technology	Engineering Sciences	Electrical Engineering and Telecommunication
YU	Tropical Medicine	Medical Research	Clinical Medicine
YY	Urban Studies	Social Sciences	Environmental Sciences and Technology
ZA	Urology & Nephrology	Medical Research	Clinical Medicine
ZC	Veterinary Sciences	Medical Research	Clinical Medicine
ZD	Peripheral Vascular Disease	Medical Research	Clinical Medicine
ZE	Virology	Basic Biology	Biomedical Sciences
ZK	Women's Studies	Social Sciences	Sociology and Anthropology
ZM	Zoology	Applied Biology - Ecology	Basic Life Sciences
ZQ	Mining & Mineral Processing	Engineering Sciences	Energy and Science Technology
ZR	Water Resources	Sciences of the Universe	Environmental Sciences and Technology

Table 3.98: Table of correspondance between disciplines from the WoS (252), OST (11) and NOWT (35) classifications

General Conclusion

This thesis aims at understanding the increasing complexity of research projects as one of the possible explanations for the fall in researchers' productivity observed over decades. We theoretically define an indicator of complexity of research projects and show among results that this index negatively affects research productivity. To empirically test our model, we consider one form of complexity of research projects, called interdisciplinarity, which requires to combine skills and efforts from different academic disciplines. Given that our theoretical index of complexity contains parameters which are unobservable, we rewrite it with the contributions of each field of expertise to the production. We show that this indicator belongs to the family of Hill diversity indexes, a standard measure of interdisciplinarity. We test our model on a dataset of publications of French authors and find that increasing interdisciplinarity decreases the impact of publications. Given our theoretical results, this probably means that interdisciplinarity depresses the research productivity. Then we consider a worldwide database to observe that interdisciplinarity is increasing over time at an annual growth rate comprised between 1% and 3.5%. We link this observation to the estimates of the decline of research productivity from [Bloom et al. \(2017\)](#) and provide some explanations for the increase of interdisciplinarity.

This thesis also explores how competition affects the researchers' incentives to exert effort, and how this translates to scientific production. Increasing the number or the efficiency of competitors encourages the more efficient teams to exert effort and raises their production. On the contrary, the less efficient ones decrease their effort and output or even no longer participate to the competition. Since selective procedures have been considerably developed in Science, and that researchers have to face uncertainty concerning the prize structure, we propose a theoretical model to address this question. We consider a multiple prize contest where the number of prizes to be awarded is uncertain. We find that whatever the mechanism of selection, complete information maximizes the effort at equilibrium.

This thesis opens new perspectives on the relation between complexity and research productivity. In the general introduction, we present some examples where we observe

a strong decline in research productivity. We suggest to carry out case studies to see if for instance inventions to increase the number of transistors in microprocessors require spending more and more efforts in different disciplines. In a recent paper dealing with the Moore's Law, [Theis and Wong \(2017\)](#) indicate that the National Science Foundation (NSF) and the Semiconductor Research Corporation have announced a new multidisciplinary research program, called "*Energy Efficient Computing: from Devices to Architectures*" (page 48). The synopsis of this program is the following:

There is a consensus across the many industries touched by our ubiquitous computing infrastructure that future performance improvements across the board are now severely limited by the amount of energy it takes to manipulate, store, and critically, transport data. While the limits and tradeoffs for this performance-energy crisis vary across the full range of application platforms, they have all reached a point at which evolutionary approaches to addressing this challenge are no longer adequate.

Truly disruptive breakthroughs are now required, and not just from any one segment of the technology stack. Rather, due to the complexity of the challenges, revolutionary new approaches are needed at each level in the hierarchy. Furthermore, simultaneous co-optimization across all levels is essential for the creation of new, sustainable computing platforms.

These simultaneous technical and organizational challenges have never been as complex or as critically important as they are now. The urgency of solving the multi-disciplinary technical challenges will require new methods of collaboration and organization among researchers.

Therefore, a comprehensive and collaborative approach must be undertaken to maximize the potential for successfully identifying and implementing revolutionary solutions to break through the bottleneck of energy-constrained computational performance. Programmers, system architects, circuit designers, chip processing engineers, material scientists, and computational chemists must all explore these new paths together to co-design an optimal solution path.

Clearly, research projects to improve the computers' performances seem to become more and more complex over time. We would like to verify if our index of complexity is

raising and to which extent this could explain the decline of research productivity in this sector.

Another extension of this thesis is to see if the same relation holds for other forms of research output, like patents for instance. In the affirmative, we would like to calculate the different indexes and their growth rate over time, and to see if the results are comparable to the ones with the scientific publications.

In this thesis we have deduced an index of complexity from a particular class of production functions (constant elasticity of substitution). We would like to consider other production functions in order to find potentially other complexity indexes. Moreover we have focused on one particular form of complexity called interdisciplinarity, but there probably exists many other types of complexity.

Finally, if the increase of complexity seems inevitable, we need to study how to counterbalance the negative effect on research productivity. This probably implies to improve the interactions between researchers coming from different academic disciplines, and to improve management methods to avoid free-riding and coordination costs in teams.

Bibliography

Milton Abramowitz and Irene A Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, volume 55. Courier Corporation, 1964.

James D Adams, Grant C Black, J Roger Clemmons, and Paula E Stephan. Scientific teams and institutional collaborations: Evidence from us universities, 1981–1999. *Research policy*, 34(3):259–285, 2005.

Jonathan Adams. Collaborations: The fourth age of research. *Nature*, 497(7451):557, 2013.

Ufuk Akcigit, Douglas Hanley, and Nicolas Serrano-Velarde. Back to basics: Basic research spillovers, innovation policy and growth. Technical report, National Bureau of Economic Research, 2013.

Robert Albanese and David D Van Fleet. Rational behavior in groups: The free-riding tendency. *Academy of Management review*, 10(2):244–255, 1985.

Kyung Hwan Baik. Effort levels in contests: The public-good prize case. *Economics Letters*, 41(4):363–367, 1993.

Loukas Balafoutas, Brent Davis, and Matthias Sutter. How uncertainty and ambiguity in tournaments affect gender differences in competitive behavior. 2017.

Christoph Bartneck and Servaas Kokkelmans. Detecting h-index manipulation through self-citation analysis. *Scientometrics*, 87(1):85–98, 2010.

Christian Beck and Friedrich Schögl. *Thermodynamics of chaotic systems: an introduction*. Number 4. Cambridge University Press, 1995.

S Keith Berry. Rent-seeking with multiple winners. *Public Choice*, 77(2):437–443, 1993.

Nicholas Bloom, Charles I Jones, John Van Reenen, and Michael Webb. Are ideas getting harder to find? Technical report, National Bureau of Economic Research, 2017.

- Lutz Bornmann and Hans-Dieter Daniel. What do we know about the h index? *Journal of the American Society for Information Science and technology*, 58(9):1381–1385, 2007.
- Nicolas Carayol and Marianne Lanoë. The impact of project-based funding in science: Lessons from the anr experience. Technical report, mimeo, 2017.
- Nicolas Carayol and Mireille Matt. Does research organization influence academic production?: Laboratory level evidence from a large european university. *Research Policy*, 33(8):1081–1102, 2004.
- Nicolas Carayol and Thuc Uyen Nguyen Thi. Why do academic scientists engage in interdisciplinary research? *Research evaluation*, 14(1):70–79, 2005.
- Christian Catalini, Nicola Lacetera, and Alexander Oettl. The incidence and role of negative citations in science. *Proceedings of the National Academy of Sciences*, 112(45):13823–13826, 2015.
- Bernard CK Choi and Anita WP Pak. 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, 2006.
- Subhasish M Chowdhury and Sang-Hyun Kim. A note on multi-winner contest mechanisms. *Economics Letters*, 125(3):357–359, 2014.
- Subhasish M Chowdhury, Dongryul Lee, and Roman M Sheremeta. Top guns may not fire: Best-shot group contests with group-specific public good prizes. *Journal of Economic Behavior & Organization*, 92:94–103, 2013.
- Derek J Clark and Christian Riis. A multi-winner nested rent-seeking contest. *Public Choice*, 87(1-2):177–184, 1996.
- Derek J Clark and Christian Riis. Influence and the discretionary allocation of several prizes. *European Journal of Political Economy*, 14(4):605–625, 1998.
- Iain M Cockburn. Is the pharmaceutical industry in a productivity crisis? *Innovation policy and the economy*, 7:1–32, 2006.
- Jean-Michel Courtault, Bertrand Crettez, and Naila Hayek. Characterization of stochastic dominance for discrete random variable. 2006.

- Robert E Evenson. Patents, r&d, and invention potential: International evidence. *The American Economic Review*, 83(2):463–468, 1993.
- Mary Frank Fox and Catherine A Faver. Independence and cooperation in research: The motivations and costs of collaboration. *The Journal of Higher Education*, 55(3):347–359, 1984.
- Qiang Fu, Jingfeng Lu, and Zhewei Wang. “reverse” nested lottery contests. *Journal of Mathematical Economics*, 50:128–140, 2014.
- Joshua S Gans and George B Shepherd. How are the mighty fallen: Rejected classic articles by leading economists. *Journal of Economic Perspectives*, 8(1):165–179, 1994.
- Amihai Glazer and Refael Hassin. Optimal contests. *Economic Inquiry*, 26(1):133–143, 1988.
- Shaun Goldfinch, Tony Dale, and Karl DeRouen. Science from the periphery: Collaboration, networks and ‘periphery effects’ in the citation of new zealand crown research institutes articles, 1995-2000. *Scientometrics*, 57(3):321–337, Jul 2003. ISSN 1588-2861. doi: 10.1023/A:1025048516769. URL <https://doi.org/10.1023/A:1025048516769>.
- H. Herbertz. Does it pay to cooperate? a bibliometric case study in molecular biology. *Scientometrics*, 33(1):117–122, May 1995. ISSN 1588-2861. doi: 10.1007/BF02020777. URL <https://doi.org/10.1007/BF02020777>.
- Mark O Hill. Diversity and evenness: a unifying notation and its consequences. *Ecology*, 54(2):427–432, 1973.
- Jorge E Hirsch. An index to quantify an individual’s scientific research output. *Proceedings of the National academy of Sciences*, 102(46):16569–16572, 2005.
- Adam B Jaffe. Real effects of academic research. *The American economic review*, pages 957–970, 1989.
- Benjamin F Jones. The burden of knowledge and the “death of the renaissance man”: Is innovation getting harder? *The Review of Economic Studies*, 76(1):283–317, 2009.
- Charles I Jones. R & d-based models of economic growth. *Journal of political Economy*, 103(4):759–784, 1995.

- Nava Kahana and Doron Klunover. A note on poisson contests. *Public Choice*, 165(1-2): 97–102, 2015.
- Eliakim Katz, Shmuel Nitzan, and Jacob Rosenberg. Rent-seeking for pure public goods. *Public Choice*, 65(1):49–60, 1990.
- Martin Kolmar and Hendrik Rommeswinkel. Contests with group-specific public goods and complementarities in efforts. *Journal of Economic Behavior & Organization*, 89: 9–22, 2013.
- Michèle Lamont, Grégoire Mallard, and Joshua Guetzkow. Beyond blind faith: overcoming the obstacles to interdisciplinary evaluation. *Research Evaluation*, 15(1):43–55, 2006.
- Jean O Lanjouw and Mark Schankerman. Patent quality and research productivity: Measuring innovation with multiple indicators. *The Economic Journal*, 114(495):441–465, 2004.
- Erin Leahey, Christine M Beckman, and Taryn L Stanko. Prominent but less productive: The impact of interdisciplinarity on scientists' research. *Administrative Science Quarterly*, 62(1):105–139, 2017.
- Dongryul Lee. Weakest-link contests with group-specific public good prizes. *European Journal of Political Economy*, 28(2):238–248, 2012.
- Jonathan M Levitt and Mike Thelwall. Is multidisciplinary research more highly cited? a macrolevel study. *Journal of the American Society for Information Science and Technology*, 59(12):1973–1984, 2008.
- Wooyoung Lim and Alexander Matros. Contests with a stochastic number of players. *Games and Economic Behavior*, 67(2):584–597, 2009.
- Jingfeng Lu and Zhewei Wang. Axiomatization of reverse nested lottery contests. *Social Choice and Welfare*, 47(4):939–957, 2016.
- Alexander Matros. Rent-seeking with asymmetric valuations: Addition or deletion of a player. *Public Choice*, 129(3-4):369–380, 2006.
- Yizhaq Minchuk and Aner Sela. All-pay auctions with certain and uncertain prizes. *Games and Economic Behavior*, 88:130–134, 2014.

- Gordon E. Moore. Cramming more components onto integrated circuits. *Electronics*, 38 (8):114–117, 1965.
- Johannes Münster. Contests with an unknown number of contestants. *Public Choice*, 129 (3-4):353–368, 2006.
- Roger B Myerson and Karl Wärneryd. Population uncertainty in contests. *Economic Theory*, 27(2):469–474, 2006.
- Isaac Newton and Herbert Westren Turnbull. The correspondence of isaac newton. 1961.
- Fabio Pammolli, Laura Magazzini, and Massimo Riccaboni. The productivity crisis in pharmaceutical r&d. *Nature reviews Drug discovery*, 10(6):428, 2011.
- Gerard Pasterkamp, Joris Rotmans, Dominique de Kleijn, and Cornelius Borst. Citation frequency: A biased measure of research impact significantly influenced by the geographical origin of research articles. *Scientometrics*, 70(1):153–165, 2007.
- Marco Pautasso and Cesare Pautasso. Peer reviewing interdisciplinary papers. *European Review*, 18(2):227–237, 2010.
- Alan Porter and Ismael Rafols. Is science becoming more interdisciplinary? measuring and mapping six research fields over time. *Scientometrics*, 81(3):719–745, 2009.
- Ismael Rafols and Martin Meyer. Diversity and network coherence as indicators of interdisciplinarity: case studies in bionanoscience. *Scientometrics*, 82(2):263–287, 2009.
- Ismael Rafols, Alan L Porter, and Loet Leydesdorff. Science overlay maps: A new tool for research policy and library management. *Journal of the American Society for Information Science and Technology*, 61(9):1871–1887, 2010.
- Linda Reijnhoudt, Rodrigo Costas, Ed Noyons, Katy Börner, and Andrea Scharnhorst. ‘seed+ expand’: a general methodology for detecting publication oeuvres of individual researchers. *Scientometrics*, 101(2):1403–1417, 2014.
- Diana Rhoten and Andrew Parker. Risks and rewards of an interdisciplinary research path, 2004.
- EJ Rinia, Th N Van Leeuwen, HG Van Vuren, and AFJ Van Raan. Influence of interdisciplinarity on peer-review and bibliometric evaluations in physics research. *Research Policy*, 30(3):357–361, 2001.

- Paul M Romer. Endogenous technological change. *Journal of political Economy*, 98(5, Part 2):S71–S102, 1990.
- Michael Rothschild and Joseph E Stiglitz. Increasing risk ii: Its economic consequences. *Journal of Economic theory*, 3(1):66–84, 1971.
- Creso M Sá. ‘interdisciplinary strategies’ in us research universities. *Higher Education*, 55(5):537–552, 2008.
- Paul S Segerstrom. Endogenous growth without scale effects. *American Economic Review*, pages 1290–1310, 1998.
- Per O Seglen. Citations and journal impact factors: questionable indicators of research quality. *Allergy*, 52(11):1050–1056, 1997.
- Louise G Shewan and Andrew JS Coats. Ethics in the authorship and publishing of scientific articles, 2010.
- Dana Sisak. Multiple-prize contests—the optimal allocation of prizes. *Journal of Economic Surveys*, 23(1):82–114, 2009.
- Stergios Skaperdas. Contest success functions. *Economic theory*, 7(2):283–290, 1996.
- Rudi Stracke, Wolfgang Höchtel, Rudolf Kerschbamer, and Uwe Sunde. Optimal prizes in dynamic elimination contests: Theory and experimental evidence. *Journal of Economic Behavior & Organization*, 102:43–58, 2014.
- Stefan Szymanski and Tommaso M Valletti. Incentive effects of second prizes. *European Journal of Political Economy*, 21(2):467–481, 2005.
- Thomas N Theis and H-S Philip Wong. The end of moore’s law: A new beginning for information technology. *Computing in Science & Engineering*, 19(2):41–50, 2017.
- John P Walsh and You-Na Lee. The bureaucratization of science. *Research Policy*, 44(8): 1584–1600, 2015.
- Jian Wang, Bart Thijs, and Wolfgang Glänzel. Interdisciplinarity and impact: Distinct effects of variety, balance, and disparity. *PloS one*, 10(5):e0127298, 2015.
- Jian Wang, Reinhilde Veugelers, and Paula Stephan. Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, 46(8):1416–1436, 2017.

Alfredo Yegros-Yegros, Ismael Rafols, and Pablo D'Este. Does interdisciplinary research lead to higher citation impact? the different effect of proximal and distal interdisciplinarity. *PloS one*, 10(8):e0135095, 2015.

List of Tables

Tableau 2.1	Example of a contest with 6 projects in competition	61
Tableau 3.1	Table of correspondance between σ_i and the diversity indexes of ω_i	89
Tableau 3.2	Contributions of each discipline λ_{ij} to project i	92
Tableau 3.3	Number of citations per paper according to the main discipline (all the database)	98
Tableau 3.4	Cumulative frequency distribution (%) of the number of citations per paper according to the main discipline (all the database) . . .	98
Tableau 3.5	Number of disciplines per paper (OST nomenclature)	99
Tableau 3.6	Number of references per paper (all the database)	101
Tableau 3.7	Type of research document (all the database)	101
Tableau 3.8	Number of references per paper (papers with at least 4 references)	101
Tableau 3.9	Type of research document (papers with at least 4 references) . .	102
Tableau 3.10	Number of authors per paper	106
Tableau 3.11	Example of a matrix of citation flows between disciplines	108
Tableau 3.12	Example of a matrix of distance between disciplines	108
Tableau 3.13	Correlation matrix ¹¹ (papers with at least four references, 1-10 disciplines, OST nomenclature)	109
Tableau 3.14	Correlation matrix ¹² (papers with at least four references, 1-10 disciplines)	109
Tableau 3.15	Pearson correlations (papers with at least 2 disciplines)	111
Tableau 3.16	Spearman's rank-ordered correlations for papers with at most 2 disciplines, OST nomenclature	111
Tableau 3.17	Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 4 references, 1-2 disciplines, OST nomenclature)	113

Tableau 3.18	Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 4 references, 1-2 disciplines, OST nomenclature)	114
Tableau 3.19	Turning point calculations for regressions (a) to (d) and 90 th , 95 th and 99 th percentiles	116
Tableau 3.20	Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 4 references, 1-10 disciplines, OST nomenclature)	119
Tableau 3.21	Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 4 references, 1-10 disciplines, OST nomenclature)	120
Tableau 3.22	Negative binomial estimates for the effect of the number of disciplines (categorical variable) on the number of citations (papers with at least 4 references, 1-10 disciplines, OST nomenclature)	121
Tableau 3.23	Negative binomial estimates for the effect of the logarithm of interdisciplinarity on the number of citations (papers with at least 4 references, 1-10 disciplines, OST nomenclature)	122
Tableau 3.24	Negative binomial estimates for the effect of the logarithm of interdisciplinarity on the number of citations (papers with at least 4 references, 1-10 disciplines, OST nomenclature)	123
Tableau 3.25	Negative binomial estimates for the effect of interdisciplinarity (square term) on the number of citations (papers with at least 4 references, 1-10 disciplines, OST nomenclature)	124
Tableau 3.26	Negative binomial estimates for the effect of interdisciplinarity (square term) on the number of citations (papers with at least 4 references, 1-10 disciplines, OST nomenclature)	125
Tableau 3.27	Logit estimates for the effect of interdisciplinarity on the probability that the paper belongs to the top 10% most cited papers (papers with at least 4 references, 1-10 disciplines, OST nomenclature)	126
Tableau 3.28	Logit estimates for the effect of interdisciplinarity on the probability that the paper belongs to the top 5% most cited papers (papers with at least 4 references, 1-10 disciplines, OST nomenclature)	127

Tableau 3.29	Negative binomial estimates for the effect of the number of disciplines (categorical variable) on the probability that the paper belongs to the top 10% and top 5% most cited papers (papers with at least 4 references, 1-10 disciplines, OST nomenclature) .	128
Tableau 3.30	Spearman's rank-ordered correlations for papers with at least 2 disciplines, CWTS nomenclature	129
Tableau 3.31	Annual average growth rate of interdisciplinarity, Web of Science database (8,000,000 observations, 2013-2017)	132
Tableau 3.32	Disciplinary diversity of scientific production, Web of Science database (8,000,000 observations, 2013-2017)	138
Tableau 3.33	Number of identified references per paper per year	139
Tableau 3.34	Average annual growth rate of research productivity	140
Tableau 3.35	Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 6 references, 1-2 disciplines, OST nomenclature)	146
Tableau 3.36	Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 6 references, 1-2 disciplines, OST nomenclature)	147
Tableau 3.37	Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 6 references, 1-10 disciplines, OST nomenclature)	148
Tableau 3.38	Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 6 references, 1-10 disciplines, OST nomenclature)	149
Tableau 3.39	Negative binomial estimates for the effect of the number of disciplines (categorical variable) on the number of citations (papers with at least 6 references, 1-10 disciplines, OST nomenclature) .	150
Tableau 3.40	Negative binomial estimates for the effect of the logarithm of interdisciplinarity on the number of citations (papers with at least 6 references, 1-10 disciplines, OST nomenclature)	151
Tableau 3.41	Negative binomial estimates for the effect of the logarithm of interdisciplinarity on the number of citations (papers with at least 6 references, 1-10 disciplines, OST nomenclature)	152

Tableau 3.42	Negative binomial estimates for the effect of interdisciplinarity (square term) on the number of citations (papers with at least 6 references, 1-10 disciplines, OST nomenclature)	153
Tableau 3.43	Negative binomial estimates for the effect of interdisciplinarity (square term) on the number of citations (papers with at least 6 references, 1-10 disciplines, OST nomenclature)	154
Tableau 3.44	Logit estimates for the effect of interdisciplinarity on the probability that the paper belongs to the top 10% most cited papers (papers with at least 6 references, 1-10 disciplines, OST nomenclature)	155
Tableau 3.45	Logit estimates for the effect of interdisciplinarity on the probability that the paper belongs to the top 5% most cited papers (papers with at least 6 references, 1-10 disciplines, OST nomenclature)	156
Tableau 3.46	Negative binomial estimates for the effect of the number of disciplines (categorical variable) on the probability that the paper belongs to the top 5% and top 10% most cited papers (papers with at least 6 references, 1-10 disciplines, OST nomenclature)	157
Tableau 3.47	Negative binomial estimates for the effect of interdisciplinarity on the 5-year number of citations (papers with at least 4 references, 1- 10 disciplines, OST nomenclature)	158
Tableau 3.48	Negative binomial estimates for the effect of interdisciplinarity on the 5-year number of citations (papers with at least 4 references, 1- 10 disciplines, OST nomenclature)	159
Tableau 3.49	OLS estimates for the effect of interdisciplinarity on the journal impact (papers with at least 4 references, 1-10 disciplines, OST nomenclature)	160
Tableau 3.50	OLS estimates for the effect of the number of disciplines (categorical variable) on the journal impact factor (papers with at least 4 references, 1-10 disciplines, OST nomenclature)	161
Tableau 3.51	Correlation matrix (papers with at least four references, 1-2 disciplines, CWTS nomenclature)	163
Tableau 3.52	Correlation matrix (papers with at least four references, 1-35 disciplines, CWTS nomenclature)	163

Tableau 3.53	Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 4 references, 1-2 disciplines, CWTS nomenclature)	164
Tableau 3.54	Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 4 references, 1-2 disciplines, CWTS nomenclature)	165
Tableau 3.55	Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)	166
Tableau 3.56	Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)	167
Tableau 3.57	Negative binomial estimates for the effect of the number of disciplines (categorical variable) on the number of citations (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)	168
Tableau 3.58	Negative binomial estimates for the effect of the logarithm of interdisciplinarity on the number of citations (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)	169
Tableau 3.59	Negative binomial estimates for the effect of the logarithm of interdisciplinarity on the number of citations (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)	170
Tableau 3.60	Negative binomial estimates for the effect of interdisciplinarity (square term) on the number of citations (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)	171
Tableau 3.61	Negative binomial estimates for the effect of interdisciplinarity (square term) on the number of citations (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)	172
Tableau 3.62	Logit estimates for the effect of interdisciplinarity on the probability that the paper belongs to the top 10% most cited papers (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)	173

Tableau 3.63	Logit estimates for the effect of interdisciplinarity on the probability that the paper belongs to the top 5% most cited papers (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)	174
Tableau 3.64	Logit estimates for the effect of the number of disciplines (categorical variable) on the probability that the paper belongs to the top 10% and top 5% most cited papers (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)	175
Tableau 3.65	OLS estimates for the effect of interdisciplinarity on the journal impact (papers with at least 4 references, 1- 35 disciplines, CWTS nomenclature)	176
Tableau 3.66	OLS estimates for the effect of the number of disciplines (categorical variable) on the journal impact factor (papers with at least 4 references, 1-35 disciplines, CWTS nomenclature)	177
Tableau 3.67	Correlation matrix (papers with at least four references, 1-252 disciplines, WoS nomenclature)	179
Tableau 3.68	Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)	180
Tableau 3.69	Negative binomial estimates for the effect of interdisciplinarity on the number of citations (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)	181
Tableau 3.70	Negative binomial estimates for the effect of the number of disciplines (categorical variable) on the number of citations (papers with at least 6 references, 1-252 disciplines, WoS nomenclature) .	182
Tableau 3.71	Negative binomial estimates for the effect of the logarithm of interdisciplinarity measured with the number of disciplines on the number of citations (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)	183
Tableau 3.72	Negative binomial estimates for the effect of the logarithm of interdisciplinarity measured with the number of disciplines on the number of citations (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)	184

Tableau 3.73	Negative binomial estimates for the effect of interdisciplinarity (square term) on the number of citations (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)	185
Tableau 3.74	Negative binomial estimates for the effect of interdisciplinarity (square term) on the number of citations (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)	186
Tableau 3.75	Logit estimates for the effect of interdisciplinarity on the probability that the paper belongs to the top 10% most cited papers (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)	187
Tableau 3.76	Logit estimates for the effect of interdisciplinarity on the probability that the paper belongs to the top 5% most cited papers (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)	188
Tableau 3.77	Logit estimates for the effect of the number of disciplines (categorical variable) on the probability that the paper belongs to the top 10% and top 5% most cited papers (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)	189
Tableau 3.78	OLS estimates for the effect of interdisciplinarity on the journal impact (papers with at least 6 references, 1-252 disciplines, WoS nomenclature)	190
Tableau 3.79	OLS estimates for the effect of the number of disciplines (categorical variable) on the journal impact factor (papers with at least 4 references, 1-252 disciplines, WoS nomenclature)	191
Tableau 3.80	Tobit estimates for the effect of the variable “year” on interdisciplinarity measured with the Berger-Parker index (1-10 disciplines, OST nomenclature)	192
Tableau 3.81	Tobit estimates for the effect of the variable “year” on interdisciplinarity measured with the Simpson index (1-10 disciplines, OST nomenclature)	193
Tableau 3.82	Tobit estimates for the effect of the variable “year” on interdisciplinarity measured with the Shannon index (1-10 disciplines, OST nomenclature)	194

Tableau 3.83	Tobit estimates for the effect of the variable “year” on interdisciplinarity measured with the number of disciplines (1-10 disciplines, OST nomenclature)	195
Tableau 3.84	Tobit estimates for the effect of the variable "year" on interdisciplinarity measured with the Berger-Parker index (1-35 disciplines, CWTS nomenclature)	196
Tableau 3.85	Tobit estimates for the effect of the variable "year" on interdisciplinarity measured with the Simpson index (1-35 disciplines, CWTS nomenclature)	197
Tableau 3.86	Tobit estimates for the effect of the variable "year" on interdisciplinarity measured with the Shannon index (1-35 disciplines, CWTS nomenclature)	198
Tableau 3.87	Tobit estimates for the effect of the variable "year" on interdisciplinarity measured with the number of disciplines (1-35 disciplines, CWTS nomenclature)	199
Tableau 3.88	Tobit estimates for the effect of the variable “year” on interdisciplinarity measured with the Berger-Parker index (1-252 disciplines, WoS nomenclature)	200
Tableau 3.89	Tobit estimates for the effect of the variable “year” on interdisciplinarity measured with the Simpson index (1-252 disciplines, WoS nomenclature)	201
Tableau 3.90	Tobit estimates for the effect of the variable “year” on interdisciplinarity measured with the Shannon index (1-252 disciplines, WoS nomenclature)	202
Tableau 3.91	Tobit estimates for the effect of the variable “year” on interdisciplinarity measured with the number of disciplines (1-252 disciplines, WoS nomenclature)	203
Tableau 3.92	OLS estimates for the effect of interdisciplinarity on the logarithm of the number of citations (papers with at least 4 references, Individual effects, 1-10 disciplines, OST nomenclature)	205
Tableau 3.93	OLS estimates for the effect of interdisciplinarity measured with the number of disciplines on the logarithm of the number of citations (papers with at least 4 references, No individual effects, 1-10 disciplines, OST nomenclature)	206

Tableau 3.94	OLS estimates for the effect of interdisciplinarity on the logarithm of the number of citations (papers with at least 4 references, Individual effects, 1-35 disciplines, CWTS nomenclature) .	207
Tableau 3.95	OLS estimates for the effect of interdisciplinarity measured with the number of disciplines on the logarithm of the number of citations (papers with at least 4 references, No individual effects, 1-35 disciplines, CWTS nomenclature)	208
Tableau 3.96	OLS estimates for the effect of interdisciplinarity on the logarithm of the number of citations (papers with at least 6 references, Individual effects, 1-252 disciplines, WoS nomenclature) . .	209
Tableau 3.97	OLS estimates for the effect of interdisciplinarity measured with the number of disciplines on the logarithm of the number of citations (papers with at least 6 references, No individual effect, 1-252 disciplines, WoS nomenclature)	210
Tableau 3.98	Table of correspondance between disciplines from the WoS (252), OST (11) and NOWT (35) classifications	218

List of Figures

Figure 1 Aggregate Data on Growth and Research Effort. Source: Bloom et al. (2017). 2

Figure 2 Moore’s Law - The number of transistors on integrated circuit chips (1971-2016). Source: Wikipedia. 2

Figure 3 Evolution of the interdisciplinarity of research articles published by at least a French researcher over the period 1999-2013, Web of Science database (400,000 observations), OST disciplinary nomenclature. 6

Figure 4 Evolution of the interdisciplinarity of research articles over the period 2013-2017, Web of Science database (8,000,000 observations), OST disciplinary nomenclature. 7

Figure 1.1 Influence of the probability distribution on the number of prizes (first-order and second-order stochastic dominance) on effort . . . 19

Figure 1.2 Probability distribution on the number of prizes 23

Figure 1.3 Total effort as a function of \bar{k} 25

Figure 1.4 Total effort as a function of v 25

Figure 1.5 Mechanism of Clark and Riis (1996) 30

Figure 1.6 Mean preserving spread 34

Figure 1.7 Illustration of the Second-Order Stochastic Dominance on the number of prizes 34

Figure 2.1 Example to illustrate Propositions 12 and 13 66

Figure 2.2 Illustration of Lorenz-dominance 72

Figure 3.1 Hill index when $m_i = 2$ 91

Figure 3.2 Negative correlation between the Simpson and our complexity indexes. 92

Figure 3.3	Distribution of Interdisciplinarity according to the Berger-Parker, the Simpson and the Shannon indexes (OST nomenclature, papers with at least 4 references)	103
Figure 3.4	Disciplinary profile of a paper	104
Figure 3.5	Predictions on the number of citations (regressions (1d), (2d) and (3d), Tables 3.25 and 3.26) and cumulative distributions of the interdisciplinary indexes	118
Figure 3.6	Evolution of the interdisciplinarity of research articles published by at least a French researcher between 1999 and 2013 (400,000 observations), OST disciplinary nomenclature.	133
Figure 3.7	Evolution of the interdisciplinarity of research articles published by at least a French researcher between 1999 and 2013 (400,000 observations), CWTS disciplinary nomenclature.	133
Figure 3.8	Evolution of the interdisciplinarity of research articles published by at least a French researcher between 1999 and 2013 (400,000 observations), WoS disciplinary nomenclature.	134
Figure 3.9	Evolution of the interdisciplinarity of research articles between 2013 and 2017, Web of Science database (8,000,000 observations), OST disciplinary nomenclature.	134
Figure 3.10	Evolution of the interdisciplinarity of research articles between 2013 and 2017, Web of Science database (8,000,000 observations), CWTS disciplinary nomenclature.	135
Figure 3.11	Evolution of the interdisciplinarity of research papers between 2013 and 2017, Web of Science database (8,000,000 observations), WoS disciplinary nomenclature.	135
Figure 3.12	Distribution of Interdisciplinarity according to the Berger-Parker, the Simpson and the Shannon indexes (CWTS nomenclature, papers with at least 4 references)	162
Figure 3.13	Distribution of Interdisciplinarity according to the Berger-Parker, the Simpson and the Shannon indexes (WoS nomenclature, papers with at least 4 references)	178

Contents

1	Tullock contests with an uncertain number of prizes	9
1	Introduction.	9
2	The model.	11
3	Prize allocation mechanisms and equilibrium outcomes	13
3.1	Winner Selection Mechanism (WSM)	13
3.2	Looser Elimination Mechanism (LEM)	15
4	Comparative statics	18
4.1	Distribution of the number of prizes	18
4.2	Number of contestants	25
5	Conclusion	28
6	Appendix	29
6.1	Winner Selection Mechanism (Clark and Riis (1996), subsection 3.1)	29
6.2	Looser Selection Mechanism (Fu et al. (2014), subsection 3.2)	31
6.3	First Order and Second Order Stochastic Dominance	33
6.4	Proofs associated to the Winner Selection Mechanism (WSM)	35
6.5	Proofs associated to the Loser Elimination Mechanism (LEM)	46
2	Team Production Function in Science	53
1	Introduction	53
2	The model	56
3	Complexity of the production function	58
4	Equilibrium outcomes	59
5	Comparative statics	60
5.1	Project i characteristics	60
5.2	Competition intensity	62
6	Extensions	66
7	Conclusion	69

8	Appendix	71
8.1	Proof of Proposition 8	71
8.2	Proof of Proposition 9	75
8.3	Proof of Proposition 10	77
8.4	Proof of Proposition 11	78
8.5	Proof of Proposition 12	79
8.6	Proof of Proposition 13	80
8.7	Proof of Proposition 14	82
8.8	Properties of Υ_i	84
3	Do complex projects lead to higher citation impact?	87
1	Introduction	87
2	From the model to the data	88
3	Literature review	93
4	Data and Methods	95
4.1	Data set	95
4.2	Variables	96
5	Results	111
5.1	1-2 disciplines	111
5.2	1-10 disciplines	115
6	Robustness checks	129
7	Is interdisciplinarity increasing over time?	132
8	Discussion	138
9	Conclusion	141
10	Appendix	143
10.1	Proof of Proposition 16	143
10.2	Proof of Proposition 17	144
10.3	More than 6 references	145
10.4	CWTS nomenclature	161
10.5	WoS subject categories	177
10.6	Tobit estimates	191
10.7	Individual Fixed Effects	204

<i>CONTENTS</i>	245
Liste of figures	241
Table of contents	243

Titre : Compétition, Interdisciplinarité et Équipes dans la Science

Résumé: Cette thèse a pour objectif de comprendre la complexité croissante des projets de recherche comme une des explications possibles de la baisse de productivité des chercheurs observée au fil des décennies. Nous conceptualisons un projet de recherche comme une idée et une équipe de chercheurs. Chaque idée est associée à une fonction de production de connaissances donnée que nous supposons de la forme CES. Les facteurs de production sont les efforts des sous-équipes, chacune dans un domaine d'expertise distinct. Nous montrons théoriquement qu'à l'équilibre du jeu, la production d'une équipe dépend négativement d'un indicateur synthétique qui caractérise sa fonction de production de connaissances que nous appelons "complexité" (disciplinaire) du projet de recherche. Bien que cet indicateur et ses composants ne soient pas directement observables dans les données, nous montrons qu'il est lié à l'indicateur de Hill des contributions des facteurs à la production, une mesure standard de l'interdisciplinarité. Cela nous offre l'occasion de tester empiriquement la complexité croissante de la recherche au cours du temps comme explication de la baisse de productivité des chercheurs. Nous confirmons ces prévisions sur un jeu de données original de près de quatre cent mille projets de recherche sur la période 1999-2013.

Mots clés : Compétition, Interdisciplinarité, Équipes, Science, Contest, Information incomplète.

Title: Competition, Interdisciplinarity and Teams in Science

Abstract: This thesis aims at understanding the increasing complexity of research projects as one of the possible explanations for the fall in researchers' productivity observed over decades. We conceptualize a research project as an idea and a team of researchers. Each idea is associated to a given knowledge production function that we suppose of the CES-form. Production factors are sub-team efforts, each one in a distinct field of expertise. We theoretically show that, at equilibrium, team outcome depends negatively on a synthetic index which characterizes its knowledge production function that we call disciplinary complexity of the research project. Though this index and its components are typically not observable in the data, we show that it is tied to the Hill index of factor contributions to the output, a standard interdisciplinary measurement in our application. This offers an opportunity to test empirically the increasing disciplinary complexity over time of research as an explanation of its decreasing productivity. We confirm those predictions on an original dataset of nearly four hundred thousand research projects over the period 1999-2013.

Keywords : Competition, Interdisciplinarity, Teams, Science, Contest, Incomplete information.