Doctoral Program in Economics
of the Tuscan Universities
jointly held by the
Universities of Firenze, Pisa and Siena

Endogenous Diffusion in Social Networks

Two Cases: Infectious Diseases and Sharing of Knowledge

Doctoral Dissertation
October 6, 2017

Ph.D. Candidate
Alessio Muscillo

Advisor
prof. Paolo Pin
(Bocconi University)

Academic Year 2016/2017

Contents

Introduction

Complex phenomena arising from the interaction of "elemental" pieces have been first studied in physics and biology, where such constitutive particles were given deterministic rules for their behavior. In that context it was already clear that even critical[1] outcomes can result on the aggregate level in situations where agents' behaviors are "mechanic" and "simple".

In recent years, inspired by real-world phenomena, economics and other social sciences have also started to play a role in this very wide strand of research. On the one hand, by introducing degrees of rationality in agents' behaviors and, on the other hand, by allowing heterogeneity in their interactions and responses to endogenous and exogenous *stimuli*.

This kind of reasoning has proven itself of particular success when applied in the context of social networks. Research on such intrinsically complex objects blossomed naturally within the realm of sociology[2], however it was only with the advent of the Internet, with the availability of large databases and the application of mathematical techniques from statistical physics[3] that the field has really started its golden period of prosperity.

In this dissertation we contribute to this strand of literature by focusing on diffusive mechanisms that naturally emerge in the context of social networks. The first example is provided by the contagion of diseases channeled through social contacts, with possible straightforward applications to the cases of diffusion of opinions or of bad habits. The second example under study is that of knowledge diffusion (sharing?), which is not only typical of the academic world but also of innovation-seeking environments, such as that of research-and-development firms, where a collaboration network is naturally constituted by the individuals.

A common feature of these cases is the fact that economic agents can endogenously and dynamically adapt by changing their (local) network of contacts or their response. In both examples, though, the impact of a single agent's action can reverberate through the whole system via its contacts (and its contacts' contacts, and so on). In the context of social networks, then, it becomes particularly challenging to understand how local features (behaviors or inclinations) may propagate, amplify or dissolute when embedded in the whole environment.

[1] [Bak et al., 1988]
[2] [Granovetter, 1973]
[3] [Albert and Barabási, 2002]

One crucial difference with other approaches lies exactly in the fact that "local" neighborhoods can indeed be very different from one another and, moreover, very different from the global situation, which is the outcome at an aggregate level.

Overview of this dissertation The rest of this thesis is structured as follows. The first chapter describes a model of diffusion of a disease between two different locations, where the agents are able to respond and adapt to this menace. A peculiarity of our model is the possibility of agents of deciding where (i.e. with whom) to interact, in the attempt of avoiding contagion while still obtaining the benefits coming from the interactions with other healthy agents. The analytical results show such individual-level behaviors have crucially different outcomes depending on the "world" this agents are living in: in particular, the two globally different systems considered (one, "globalized", where connections between the locations are allowed and the other, "autarkic", where they are forbidden) exhibit crucially different resistance to exogenous shocks in the infection rates. Further research in this field is still needed, as this model is one of the few attempts in the economics literature at trying to embed rational and responsive agents in a dynamical model of diffusion on networks. Applications to systemic risk and systemic resistance can benefit from this kind of research as well as analyses of mechanisms where is prevalent the interplay between local versus global forces.

The second chapter deals with a classic dilemma in the economics and business literature, that of exploration versus exploitation, and relates it to the achievement of results, i.e. to the notion of performance. Specifically, we follow individual scientists over all their career and use their co-authorship and citation networks to map their "knowledge space", in order to measure their propensity to explore, both in terms of new topics and of new collaborations. The econometric results shows that the relationship between exploration and performance tends to exhibit an inverted-U shape, hence supporting the theory that a "sweet" spot where performance is maximized might exist, at least at an individual level.

Further research on this topic is still necessary, for example to understand in depth the relationship existing (if any) between forms of "social exploration" (i.e. exploration in terms of collaborations and social contacts) and "scientific exploration" (i.e. in terms of changes of the subjects studied or fields of expertise). Moreover, the results and techniques developed here can not only be directly applied to bibliometrics studies, but can also be fundamental to give the right incentives (and, possibly, fundings) to encourage long-term innovation-seeking behaviors.

The third chapter tackles the same research question as the second one, but from a different viewpoint: what is the outcome of that analysis when the production units are "aggregated" at the level of (departments of) universities? At this aggregate level, it turns out that, in contrast to what seen in the previous chapter, a U-shaped curve characterizes the relationship between performance and exploration. Moreover, this relationship is also complicated by the effects of resources and size of each university. This complication can be seen of evidence of how, at this level, the interplay between economies of scale and economies of scope can generate an overall complex behavior. In this case too, then, the individual-level and the aggregate-level analysis exhibit once again very different outcomes: this underlines even more the complexity that comes out from the interactions in systems composed by different layers and levels.

Chapter 1

Spreading of an Infectious Disease between Different Locations

1.1 Abstract

We study how the spread of an infection evolves in a population adopting self-protecting behavioral responses which, in turn, affect the evolution of the epidemics. We focus on two interplaying mechanisms: firstly, how the contact network conditions the evolution by only allowing contagion via contacts and, secondly, how the network itself is endogenously modified by the behavioral response triggered by the risk perception.

We formulate a simple model of an infection spreading among farms through the trade of livestock. Farmers are aware of the infection prevalence, which negatively affects the benefit obtained from trading, and can decide to pay an extra cost to trade with farmers in another (less infected) country. Cross-country trading first could decrease the infection rate, but then it further fosters the epidemics by feeding back into the country of origin.

Analytical results show that (a stylized) cross-country globalization might in fact reduce the spread of an epidemic triggered by an exogenous small shock in the infection rate. However, a large shock, even in only one country, may end up infecting all countries, making interconnections a source of systemic fragility.

This work done in collaboration with Paolo Pin. Econometric results used as motivations are provided by Mahdi Gholami and Tiziano Razzolini.

1.2 Introduction

Connections between individuals facilitate the exchange of goods, resources and create benefits. However they also are the mean through which diseases and attacks may spread in a society, making it vulnerable to shocks and menaces. Due to the advances in virtual and physical communications, understanding this tradeoff has become increasingly more necessary as well as complicated.

We study how the spread of an infection evolves in a population adopting self-protecting behavioral responses which, in turn, affect the evolution of the epidemics. We focus on two interplaying mechanisms: firstly, how the contact network conditions the evolution by only allowing contagion via contacts and, secondly, how the network itself is endogenously modified by the behavioral response triggered by the risk perception.

Although in the last years a multidisciplinary literature has dealt with this topic, we choose to especially concentrate our attention on analyzing the effect of behavioral responses consisting on establishing long-range connections, as opposed to short/local ones. We are particularly motivated on the one hand by the Ebola outbreak started in West Africa in 2014, on the other hand by some evidence emerged recently from analyzing an extensive database of livestock trading among farms. In the first case, highly infected individuals easily has had access to international air flights, in the second, a surprisingly high correlation has been found between potentially highly infected animals and international import-export connections.

We formulate a simple model of an infection spreading among farms, that are placed in different locations/countries, through the trade of livestock. Farmers are aware of the infection prevalence, which negatively affects the benefit obtained from trading, and can decide to pay an extra cost to trade with farmers in another (less infected) country. Cross-country trading first could decrease the infection rate, but then it further fosters the epidemics by feeding back into the country of origin.

Due to the simplicity of the model, we are able to obtain a system of ordinary differential equations ruling the evolution of the infection and provide some analytical results. These show that this very stylized cross-country globalization might in fact generate complex dynamics in terms of the co-evolution of the coupled mechanism constituted by the trading network and the infection spread.

When a bounded recovery capacity from infection is assumed (e.g. limited hospitalization for quarantine), small exogenous infection shocks are better controlled when the locations are tied together: in fact, the outflow of infected individuals from the most infected locations contributes to dilute and reduce the epidemic. On the contrary, a large shock, even in only one country, may end up infecting all countries instead of just one. Interestingly enough, the latter is not necessarily the case depending on the parameters, interconnections indeed providing a mean to increase and multiply (own and) systemic recovery capacity.

1.2.1 Motivations

From epidemiology: the 2014 Ebola outbreak A complex epidemic of *Zaire ebolavirus* has been affecting West Africa since approximately December 2013. See [Chowell and Nishiura, 2014] for a detailed review; see also [Thomas et al., 2015] and the web site of the World Bank for some estimates of the damages done to the economies of some African countries.[1]

Directly quoted from the web site of the World Health Organization (WHO)[2]:

> "Countries in equatorial Africa have experienced Ebola outbreaks for nearly four decades. [...] In those outbreaks, geography aided containment. [...] In West Africa [which had never experienced an Ebola outbreak], *entire villages have been abandoned after community-wide spread killed or infected many residents and fear caused others to flee.* [...] West Africa is characterized by a *high degree of population movement* across exceptionally porous borders. Recent studies estimate that population mobility in these countries is *seven times higher than elsewhere in the world.* [...] Population mobility created two significant impediments to control. First, [...] cross-border contact tracing is difficult. Populations readily cross porous borders but outbreak responders do not. Second, *as the situation in one country began to improve, it attracted patients from neighbouring countries* seeking unoccupied treatment beds, thus reigniting transmission chains. [...] The traditional custom of *returning, often over long distances*, to a native village to die and be buried near ancestors is another dimension of population movement that carries an especially high transmission risk. [...]
>
> **Spread by international air travel**
> The importation of Ebola into Lagos, Nigeria on 20 July and Dallas, Texas on 30 September [2014] marked *the first times that the virus entered a new country via air travellers.* These events theoretically placed every city with an international airport at risk of an imported case. The imported cases, which provoked intense media coverage and public anxiety, brought home the reality that all countries are at some degree of risk as long as intense virus transmission is occurring anywhere in the world - especially given the radically *increased interdependence and interconnectedness that characterize this century.*"

Financial networks: small vs. large shocks As recently argued in [Acemoglu et al., 2015], financial networks may exhibit sharp phase transitions:

[1] http://www.worldbank.org/en/region/afr/publication/ebola-economic-analysis-ebola-long-term-economic-impact-could-be-devastating and also http://www.worldbank.org/en/region/afr/publication/the-economic-impact-of-the-2014-ebola-epidemic-short-and-medium-term-estimates-for-west-africa

[2] See http://www.who.int/csr/disease/ebola/one-year-report/factors/en/

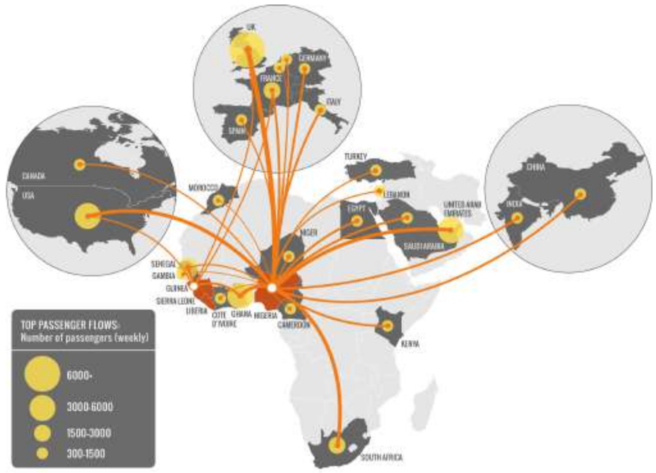


Figure 1.1: Air traffic connections from West African countries to the rest of the world. Source [Gomes et al., 2014]. See also [Halloran et al., 2014].

> “as long as the magnitude of negative shocks affecting financial institutions are sufficiently *small*, a more densely connected financial network (corresponding to a more diversified pattern of interbank liabilities) enhances financial stability. However, beyond a certain point, dense interconnections serve as a mechanism for the propagation of shocks, leading to a more fragile financial system.”

In particular, from the methodological and analytical point of view, it is important to stress that the authors specifically focus on *regular* networks, in which all banks are equal, in order to guarantee that “any variation in the fragility of the system is due to the financial networks structure rather than any heterogeneity in size or leverage among banks.”

It is worth noticing, though, that in this strand of literature the shocks are considered at an individual level: each node (i.e. financial institution, bank and like) has a probability of getting hit by an exogenous and autonomous shock. The authors, then, try to understand how these shocks spread in a financial network and how the network structure itself affects the final outcome. The same kind of reasoning holds for [Bramoullé and Kranton, 2007]. In our cases, however, we will consider shocks that are at a sub-population level.

Livestock trading: infection and long-range connections In an extensive database about the trade network of livestock (cattle) farms in Italy, which has been currently studied by T. Razzolini and M. Gholami, there has been foud evidence suggesting a high correlation between infected farms and long-range trading.

In general, one would expect that the more distant two farms are, the less likely it is that they trade with each other, because of transportation costs etc.

However, what emerges is that in presence of a higher possibility of the cattle to be infected, there is a tendency of trading with far away locations. This could be due to the fact that neighboring farms are subject to the same vets or organizations for detecting diseases and, therefore, close farms would have more information about each others' health status.

Table 1.1 briefly summarizes these results: there is a negative and significant effect of the "distance (in km)" on trading, which is the usual and expected result. However, when the possibility of being infected becomes higher (as measured by the variable "interaction", which mixes "distance" with "sickness test variables"), the sign of the effect becomes positive, meaning that farms tend to send their cattle far away.

Table 1.1: **Distance effect on livestock trading**

Regression
Dependent variable: heads over stock
Observations: 2,147,208
Robust standard errors in parentheses:
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Here we show some of the results obtained in the forthcoming work done by M. Gholami, under the supervision of T. Razzolini. Distance between farms *per se* negatively affects trading; however, when distance interacts with disease-related variables, the sign becomes positive, hence showing a higher propensity to trade with far away locations in case of more likely infected cattle.

Variables	
lag_totalpositive	-0.00634***
	(0.00107)
distance_km	**-1.71e-05*** **
	(2.81e-06)
interaction	**2.79e-05*** **
	(9.34e-06)
2009.year	-3.52e-05
	(0.000146)
2010.year	0.000566
	(0.000471)
2011.year	0.00190***
	(0.000593)
2012.year	0.00351***
	(0.000791)
2013.year	0.00975***
	(0.000875)
2.quarter	0.000658***
	(0.000227)
3.quarter	0.000101
	(0.000165)
4.quarter	0.00342***
	(0.000209)
Constant	0.00762***
	(0.000201)

1.3 The Single-location Model: the Building Block

In this section we develop the building block of the model. We define a system constituted by a single location and describe how the infection evolves in it, as time passes. The dynamics is kept explicitly abstract and simple on purpose, and this has one main reason: the 1-location system so defined is able to recover from small shocks, but unable to do so in case of large shocks (to be defined shortly). This, in turn, will allow us to consider two such systems[3] interacting with each other and, then, evaluate what the effects on this whole 2-location system will be, in terms of resistance to shocks.

Consider a continuum of agents living in one location and susceptible to the infection from a transmittable disease, which can spread through personal contacts with other agents. The intuitive idea is that these agents are trading with each other and that this trade has to happen in meetings among pairs of agents, which is also the mean through which the disease spreads.

Let $x(t)$ denote the fraction of infected individuals at time t. The evolution of this fraction is ruled by the following differential equation, used for example in ecological economics as a development from the classic Bass model (see [Bass, 1969] and [D'Alessandro, 2007]):

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = \nu x(t)(1 - x(t))(x(t) - q), \tag{1.1}$$

where $\nu \in (0, 1)$ is a parameter representing the *contagiousness* of the disease and $q \in (0, 1)$ is a parameter which measures the capacity of the system to control the disease, which we may call *quarantine*. More specifically, we may think of q as the quantity of resources allocated to hospitalize infected individuals as well as to other disease-control measures, which we consider as fixed and exogenous, since they could change but with a longer time-scale with respect to the evolution of the disease.

Remark. Equation (1.1) can be seen as modified susceptible-infected model: we take the probability that an infected individual meets a susceptible one, i.e. $x(1-x)$, and that this meeting results in a new infection with probability ν and then multiply this a factor $(x-q)$ which modifies the sign of the flow of infected according to the fact that the fraction of infectives exceeds or not the threshold q.

The above 1-dimensional dynamical system have only three equilibria in $[0, 1]$ and, in particular, the sign of the right-hand side, as a function of x, gives the direction of the dynamics and, hence, the asymptotic behavior of these equilibria.

PROPOSITION 1.3.1. *The dynamical system* (1.1) *has (only) the following critical points:*

- $x = 0$*, which is an asymptotically stable steady state, called* disease-free equilibrium*;*
- $x = q$*, an unstable equilibrium;*
- $x = 1$*, also an asymptotically stable equilibrium, called* endemic equilibrium*.*

[3] See next section.

Proof. The derivative $\frac{\mathrm{d}x}{\mathrm{d}t}$, which is a cubic function of x, has only three roots $x = 0$, $x = q$ and $x = 1$, where it becomes equal to 0. Moreover, it is strictly negative when $x \in (0, q)$ and strictly positive when $x \in (q, 1)$. $\square$

The same reasoning on the sign of the derivative allows us to easily describe the basins of attraction of the stable points, thus giving the following result.

COROLLARY 1.3.2. *The interval* $[0, q) \subset \mathbb{R}$ *is the basin of attraction of* $x = 0$*, while* $(q, 1]$ *is the basin of* $x = 1$*.*

Remark. In ecology[4], this dynamics sometimes describes the evolution of a species over time. In this context, the analogy is that the species we are considering is that of a bacterium causing the infective disease. The threshold q represents the *critical mass* of infections that the species has to reach and exceed in order to survive: when there are not enough infected individuals, the species cannot proliferate and propagate any more and, eventually, the epidemic dies out. Lastly, it is worth noticing that the susceptible-infected model is based on the so-called *homogeneous mixing* assumption, which states that the meetings between individuals are random, according to their relative proportion in the whole population. This is an assumption that we maintain, in this framework.

Resistance to shocks & policy In this framework, we define a *shock* according to the following intuition: we start from the system being in the disease-free equilibrium $x = 0$ and, then, suppose that at time $t = 0$ there is a sudden and exogenous variation in the infection rate such that $x(0) = x_0 \in [0, 1]$. This initial fraction x_0 of infected is what we will call shock.

The results above clearly imply that if the shock is $x_0 < q$, i.e. below the threshold point, then the system will (asymptotically) return to the disease-free equilibrium, whereas if the shock is larger than q, then the dynamic will converge toward 0, where the whole population is infected. Assuming that the shocks are uniformly distributed over the interval $[0, 1]$, then q exactly measures the ability of the system to recover from a shock, because it measures the length of the basin of attraction $[0, q)$ of the disease-free equilibrium.

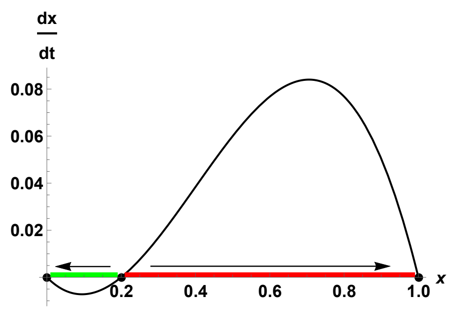
The policy implication here is then straightforward: the more resources can be allocated to control the disease, i.e. the larger q is, the more the system will able to recover from larger and larger shocks in the infection rate.

1.4 The 2-location Model

Starting from the conclusion of the previous section, we want now to extend our analysis: the trading and meeting process will take place both within and across two different locations and, consequently, the same will happen to the spread of the disease. The main question will, then, be how this stylized globalization affects the 2-location systemic resistance to shocks in the infection rates. Is the

[4] See [D'Alessandro, 2007] on this.

Figure 1.2: **Single-location dynamics**



Dynamics defined by equation (1.1), where the parameters are set at $q = 0.2$ and $\nu = 0.8$, the plotted curve represents the right-hand side of (1.1) and the dots the critical points 0, q and 1 respectively. The point $x = q$ is an unstable equilibrium, acting as a threshold and separating the basins of attraction of the two asymptotically stable equilibria $x = 0$ (in green) and $x = 1$ (in red). Small exogenous shocks, i.e. below q, are absorbed, whereas shocks larger than q lead to a fully infected system.

coupled system going to be more resistant to shocks than two separated and autarkic single-locations? What about simultaneous shocks in both locations?

We consider two locations, e.g. two countries, throughout denoted by A and B, both populated by interacting agents, e.g. farmers trading cattle. Agents benefit from interacting with each other but, since there may be a disease spreading, this potential benefit decreases as the infection rate increases among the agents, which accounts for the risk of becoming infected they face. In the attempt to avoid contagion, the agents of one location may be willing to interact with others in the other location, even if to do so they have to pay a higher cost for this long-range interaction, e.g. export costs or trade barriers.[5]

To start with, we restrict our attention to two identical and symmetric locations, where agents are homogeneous and identical in all aspects but in the export cost. In particular, different agents of the same location are assigned different costs to export to the other location, and this may be thought of as reflecting geographical proximity, facility in the contacts with a foreign country, etc.

The key aspect is that using identical locations and identical agents is a normalization that can help guarantee that any variation in the fragility of the coupled system is due to the cross-country connection structure rather than to differences in other characteristics.

1.4.1 The General Model

Consider two intervals $A, B \subset \mathbb{R}$ representing populations of agents living for an infinite time horizon. As a piece of notation, for each agent h we will denote by $H \in \{A, B\}$ her home country and by $E \neq H$ the corresponding other country.

[5] In the appendix, we consider an alternative general model in which trade can happen at the same time within and across country. The results, however, remain essentially the same.

Benefits from interaction and costs Agents benefit from trading/interacting with other agents and, in particular, any agent $h \in H$ receives a gross utility of p_H, when trading in her home country, and a possibly higher gross utility p_E, when instead exporting to the other country. These benefits are assumed to be equal across agents[6] and to be decreasing functions of the current infection rates $x_H(t) \in [0, 1]$ and $x_E(t) \in [0, 1]$. This last assumption reflects the fact that trading becomes riskier as contagion spreads. Formally:

$$p_A = p_A(x_A(t), x_B(t)), \qquad p_B = p_B(x_A(t), x_B(t)),$$

for any time $t \in \mathbb{R}$.

Any generic agent $h \in H$ chooses between two (mutually exclusive) actions, respectively labeled as H and E, which are either "only trading in her home country" or "only exporting to the other country".[7] However, to export to the other country each agent $h \in H$ has to pay an exporting cost $c_h > 0$, which is assumed to be randomly distributed across agents according to a cumulative distribution function F_H. The cost of trading in the home country, instead, is normalized to 0.[8] Depending on the chosen action H or E, agent h's utility at time t is then given by:

$$u_h(t) = \begin{cases} p_H(t), & \text{if trading in } H, \\ p_E(t) - c_h, & \text{if exporting to } E, \end{cases}$$

so that agent $h \in H$ decides to export to E at time t if and only if

$$p_H < p_E - c_h.$$

Remark. Since in our framework export is going to be another channel of transmission for the disease, agent h myopically maximizes her utility by exporting even in presence of a positive exporting cost. So, she does not account for how the disease might possibly spread in the future, thus ultimately modifying her payoff.

Remark. Notice that in our formulation agents make decisions only based on the prices p_A and p_B that they are able to observe in the two markets. In particular, they are not able to observe neither their status (as susceptible or infected) nor the status of the others. In the case of the cattle trade mentioned in the motivation section, this assumption is economically justified as follows: on the one hand, farmers are encouraged to report any situation that may be related to a disease, otherwise incurring in very high fines, while on the other hand, these reports are kept anonymous.

[6] They indeed represent the price at which the trade happens.

[7] According to our notation, then, agents $a \in A$ can export to B and, conversely, agents $b \in B$ to A.

[8] The assumption that the cost c_h's are independent of time t would dramatically decrease the tractability of the model and the same holds for assumptions concerning discounted utility for future times or agents maximizing lifetime expected utility. All this is left for future work.

Since $c_h \sim F_H$, the above expression implies that the fraction of H's agents willing to export to E at time t is given by

$$\mathbb{P}\left\{h \in H : p_H(t) < p_E(t) - c_h\right\} = \mathbb{P}\left\{h \in H : c_h < p_E(t) - p_H(t)\right\} = \\ = F_H\left\{p_E(t) - p_H(t)\right\},$$

or, equivalently, that the fraction of H's agents trading in H and *not* exporting is $1 - F_H\left\{p_E(t) - p_H(t)\right\}$.

Cross-country meetings and flows of infected individuals Let us proceed with the analysis: of the fraction of agents that are exporting from A to B, a subfraction of them given by $x_A \cdot F_A\left\{p_B - p_A\right\}$ is of currently infected agents. Consequently, when these exporting and infected agents meet the fraction of those susceptible in B that remain in B for trade, which is $(1-x_B)(1-F_B\{p_A - p_B\})$, this will give rise to an additional source of infected individuals for country B:[9]

$$\underbrace{x_A \cdot F_A}_{A\text{'s infected exporting to } B} \cdot \underbrace{(1 - x_B) \cdot (1 - F_B)}_{B\text{'s susceptible remaining in } B}.$$

Still another source of infection for B comes from the meetings between B's infected individuals remaining in B with A's susceptible exporting to B:

$$\underbrace{x_B \cdot (1 - F_B)}_{B\text{'s infected remaining}} \cdot \underbrace{(1 - x_A) \cdot F_A}_{A\text{'s susceptible exporting to } B}.$$

However, this additional infective activity due to cross-country interactions is somehow compensated with a reduction in the home country. In particular, now B's within-country spreading cannot follow the single-location equation (1.1) given in Section 1.3: not only because the meetings only happen between B's susceptible and infected agents that are not exporting, but also because we have to subtract the fraction of B's infected agents that are exporting, as an outflow.

$$\underbrace{\nu_B x_B(1 - F_B)(1 - x_B)(1 - F_B)(x_B - q)}_{\text{meetings among } B\text{'s remaining agents resulting in infections}} - \underbrace{x_B F_B}_{\text{outflow of infected}},$$

where $\nu_B \in (0, 1)$ is the contagiousness parameter for B. Analogous reasonings hold symmetrically for A.

By putting all these elements together, we can build a system of coupled differential equations ruling the evolution over time of the infection rates in the two countries. The first line of each equation accounts for the possibly reduced within-country epidemic spreading, whereas the second line accounts for the

[9]For ease of notation, throughout we will write $F_A\{p_B - p_A\} = F_A$ and $F_B\{p_A - p_B\} = F_B$.

additional inflow of infection due to cross-country interactions just described above:[10]

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}x_A = \nu_A\Big[x_A(1-F_A)(1-x_A)(1-F_A)(x_A-q_A)+ \\ \qquad\qquad + x_A(1-F_A)(1-x_B)F_B + (1-x_A)(1-F_A)x_BF_B\Big] - x_AF_A \\ \dfrac{\mathrm{d}}{\mathrm{d}t}x_B = \nu_B\Big[x_B(1-F_B)(1-x_B)(1-F_B)(x_B-q_B)+ \\ \qquad\qquad + x_B(1-F_B)(1-x_A)F_A + (1-x_B)(1-F_B)x_AF_A\Big] - x_BF_B, \end{cases} \tag{1.2}$$

where $\nu_A, \nu_B \in (0,1)$ and $q_A, q_B \in (0,1)$ are the contagiousness and quarantine parameters respectively of location A and B. The system can be algebraically rearranged as follows:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}x_A = \nu_A(1-F_A)\Big[x_A(1-x_A)(x_A-q_A)(1-F_A) + (x_A+x_B-2x_Ax_B)F_B\Big] \\ \qquad\qquad - x_AF_A \\ \dfrac{\mathrm{d}}{\mathrm{d}t}x_B = \nu_B(1-F_B)\Big[x_B(1-x_B)(x_B-q_B)(1-F_B) + (x_A+x_B-2x_Ax_B)F_A\Big] \\ \qquad\qquad - x_BF_B. \end{cases} \tag{1.3}$$

Remark. If cross-country export is not allowed, i.e. when $F_A = F_B = 0$, then system (1.2) reduces to two uncoupled equations, corresponding to two single-location models of the form of equation (1.1), both evolving separately.

Remark. In this model, export only occurs in one direction at time, either from A to B or viceversa. Indeed, suppose that $p_A(t) < p_B(t)$ at a certain time $t \in \mathbb{R}$. Since F_A and F_B are cumulative distributions which are positive only for positive costs, then in such a case $F_B(p_A(t) - p_B(t)) = 0$ while $F_A(p_A(t) - p_B(t)) > 0$. So, there is an outflow of infection in the first equation for x_A and an inflow in the second for x_B.

PROPOSITION 1.4.1. *System (1.2) is well defined in the unit square $(x_A, x_B) \in [0,1]^2$.*

Proof. We want to show that the unit square $[0,1]^2$ is an invariant set under the dynamics defined by system (1.2). In order to do that, we need to the vector field defining the system of equation, i.e. the right-hand side of (1.2) as 2-dimensional function of (x_A, x_B) is "pointing toward the interior" of the square, while restricted on the borders of it. More formally:

- suppose that $x_A = 0$. Then $\dot{x}_A = \nu_A(1-F_A)x_BF_B \geq 0$, for any $x_B \in [0,1]$, as wanted.

[10]For ease of notation, we omit the time t. However, it is worth remembering that F_A and F_B depend on p_A and p_B which, in turn, depend on $x_A(t)$ and $x_B(t)$.

- Suppose, instead, that $x_A = 1$. If, by assumption[11], we have that $F_A = 1$ when $x_A = 1$, then

$$\dot{x}_A = \nu_A(1 - F_A)(1 - x_B)F_B - x_A F_A = -1 < 0,$$

as we wanted.

An analogous and symmetric reasoning shows that $\dot{x}_B \geq 0$, when $x_B = 0$, and that $\dot{x}_B \leq 0$, when $x_B = 1$. □

1.4.2 A Simple Specification: Identical Locations, Linear Utility & Uniform Cost

To keep the analysis as simple as possible and to try to get more analytical results, we restrict our model to a specification of system (1.2) that is possibly the simplest one:[12]

- the two locations A and B are assumed to be identical, from the point of view of the epidemic parameters, so that $\nu_A = \nu_B = \nu \in (0, 1)$ and $q_A = q_B = q \in (0, 1)$;

- the agents' exporting costs $c_h > 0$, for $h \in H$, are uniformly distributed over the interval $[0, 1]$, so that the cumulative distributions are identical and of the form $F_A = F_B = \mathcal{U}(0, 1)$, i.e.

$$F_A(c) = F_B(c) = \begin{cases} 0, & \text{for } c \leq 0 \\ c, & \text{for } c \in [0, 1] \\ 1, & \text{for } c \geq 1. \end{cases}$$

In particular, the maximum and minimum cost are respectively 1 and 0.

- The gross utilities from trading, p_A and p_B, are assumed to depend linearly on the infection rate of the own location:

$$p_A(x_A(t), x_B(t)) := -x_A(t), \qquad p_A(x_A(t), x_B(t)) := -x_B(t).$$

The gross maximum and minimum utility attainable are then normalized to 0 and -1, respectively.

[11] This assumption is intuitive: whenever in A the rate of infection is the maximum, i.e. $x_A = 1$, then all A's agents would be facing the minimum home benefit p_A and thus be willing to export, so $F_A = 1$.

[12] The limitation to the unit interval $[0, 1]$, instead of a generic $[\underline{C}, \overline{C}] \subset \mathbb{R}$, does not change the analysis nor the results. The same holds true for the utility, that could be assumed of the form $p_H = \overline{P}(1 - x_H(t))$, with maximum attainable utility at $\overline{P} > 0$ and minimum at 0. In any case, it should be assumed that $\overline{C} \leq \overline{P}$, i.e. maximum cost to be paid less than maximum benefit obtainable.

With these assumptions in place, we have that[13]

$$F_A = F_A\{x_A - x_B\} = \begin{cases} 0, & \text{if } x_A - x_B < 0 \\ x_A - x_B, & \text{if } 0 \leq x_A - x_B \leq 1 \\ 1, & \text{if } 1 < x_A - x_B \end{cases} = \max\{0, x_A - x_B\},$$

and analogously that $F_B = \max\{0, x_B - x_A\}$. We can then rewrite system (1.2) as follows:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}x_A = \nu(1-\max\{0, x_A - x_B\})\Big[x_A(1-x_A)(x_A-q)(1-\max\{0, x_A - x_B\}) \\ \qquad\qquad + (x_A + x_B - 2x_A x_B)\max\{0, x_B - x_A\}\Big] - x_A \max\{0, x_A - x_B\} \\ \dfrac{\mathrm{d}}{\mathrm{d}t}x_B = \nu(1-\max\{0, x_B - x_A\})\Big[x_B(1-x_B)(x_B-q)(1-\max\{0, x_B - x_A\}) \\ \qquad\qquad + (x_A + x_B - 2x_A x_B)\max\{0, x_A - x_B\}\Big] - x_B \max\{0, x_B - x_A\}. \end{cases} \tag{1.4}$$

Mathematical analysis The system is well defined in $\mathbb{R}^2$, but we will restrict our analysis to the unit square $(x_A, x_B) \in [0,1]^2$, in which the fractions of infected agents make sense. It is continuously differentiable everywhere but the diagonal in $\mathbb{R}^2$, i.e. over $\mathbb{R}^2 \setminus \{(x_A, x_B) \in \mathbb{R}^2 : x_A = x_B\}$, but, thanks to the symmetry of the system due to the assumptions made in this section, we can separate the analysis focusing on three different part: the diagonal, the super-diagonal set and the sub-diagonal.[14]

This allows us to use an *ad hoc* strategy in the analysis and heavily contributes to obtain some explicit results, which are briefly summarized here and formalized in the following propositions: there are two asymptotically stable equilibria, $(x_A, x_B) = (1,1)$ and $(0,0)$, the first corresponding to both countries being fully infected, while the second to both being disease free. There is a third equilibrium, (q,q) which is an unstable saddle point. Its separatrix curves separate the basins of attraction of the asymptotically stable states, as depicted in Figure 1.3. It is worth noting, though, that they are not explicitly characterizable.[15]

PROPOSITION 1.4.2. *System (1.4) is well defined on the plane $\mathbb{R}^2$. Moreover, the unit square $[0,1]^2 \subset \mathbb{R}^2$ is invariant under its dynamics. In addition, it is also symmetric with respect to the diagonal in $\mathbb{R}^2$, which is thus an invariant set together with the super-diagonal and sub-diagonal sets, respectively $\{(x_A, x_B) \in \mathbb{R}^2 : x_A < x_B\}$ and $\{(x_A, x_B) \in \mathbb{R}^2 : x_A > x_B\}$.*

PROPOSITION 1.4.3. *System (1.4) has (only) three equilibria:*

[13] The third case $1 < x_A - x_B$ can never occur, since $x_A, x_B \in [0,1]$.

[14] See the Appendix for the details.

[15] Unfortunately, there is no known way to analytically determine these curves, even in simple dynamical systems, although great progresses have been done with numerical approximations. See, for example, [Cavoretto et al., 2011].

- *$(x_A, x_B) = (0,0)$ and $(1,1)$, which are asymptotically stable states;*
- *$(x_A, x_B) = (q,q)$, which is an unstable saddle point.*

Moreover, the two separatrix curves of the saddle (q,q) are such that the unstable one coincide with the diagonal of the square $\{(x_A, x_B) \in [0,1]^2 : x_A = x_B\}$, while the stable separatrix are part of the boundary of the basins of attraction of the stable equilibria.

As already anticipated, in the analysis that will follow a crucial role will be played by the stable separatrix curve of the saddle (q,q), hereafter denoted by $\mathcal{C}$. From the theory of dynamical systems, it can be shown that $\mathcal{C}$ is partitioned as the image of three distinct trajectories/solutions of system (1.4):

$$\mathcal{C} = \mathcal{C}^- \cup \{q,q\} \cup \mathcal{C}^+.$$

The following result formalizes what is shown in Figure 1.3: depending on the parameters ν and q, as time t increases, the solution $\mathcal{C}^-$ enters the unit square either crossing its border along the segment $[q,1] \times \{0\}$ or along $\{1\} \times [0,q]$ and, eventually, converges toward $\{q,q\}$ as $t \to \infty$. Symmetrically, the same occurs for $\mathcal{C}^+$.

PROPOSITION 1.4.4. *Let $\mathcal{C}$ denote the (unique) stable separatrix of the saddle point (q,q) of system (1.4). The curve $\mathcal{C}$ can be naturally partitioned according to the following three distinct trajectories that compose it:*

$$\mathcal{C} = \{q,q\} \cup \underbrace{(\mathcal{C} \cap \{x_A < x_B\})}_{=: \, \mathcal{C}^-} \cup \underbrace{(\mathcal{C} \cap \{x_A > x_B\})}_{=: \, \mathcal{C}^+}.$$

Then $\mathcal{C}^- \cap [0,1]^2$ is included in $[q,1] \times [0,q]$ and, depending on the parameters q, ν, it either crosses the segment $[q,1] \times \{0\}$ in a point $(\eta, 0)$ or the segment $\{1\} \times [0,q]$ in a point $(1,\zeta)$. A symmetric result holds for $\mathcal{C}^+$.

The previous result is based on the following lemmas.

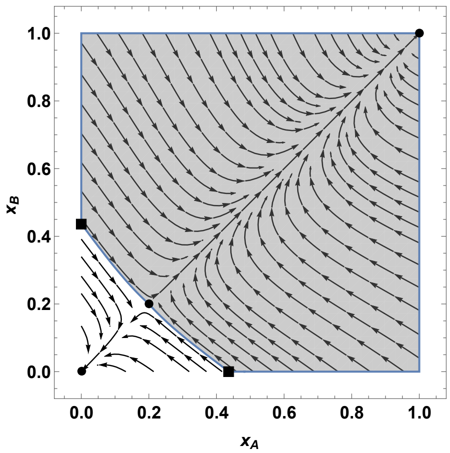
LEMMA 1.4.5. *The components of the eigenvector of the Jacobian corresponding to the stable separatrix $\mathcal{C}$ of the saddle (q,q) are of opposite sign. This implies that $\mathcal{C}^+ \subset [0,q] \times [q,1]$ and $\mathcal{C}^- \subset [q,1] \times [0,q]$ for times t large enough.*

LEMMA 1.4.6. *The signs of the components of the vector field defining system (1.4), when computed in $(x_A, x_B) \in [q,1] \times [0,q]$, are such that $V_A(x_A, x_B) \leq 0$ and $V_B(x_A, x_B) \geq 0$.*

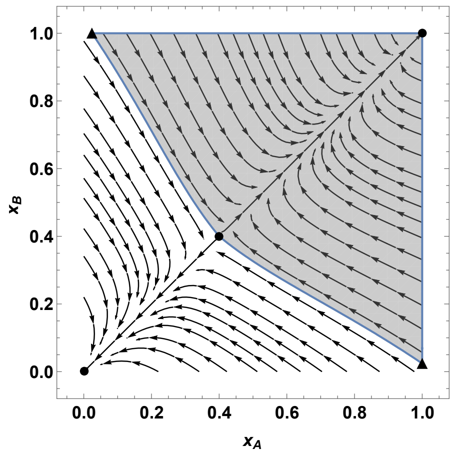
1.4.3 Comparative Statics

Since that we have already observed that the separatrix $\mathcal{C}$ cannot be described analytically, we first compute a linear approximation of it and, then, confirm the results by numerical analysis. Moreover, this will allow us to make a comparative statics analysis that will be then used in the following section.

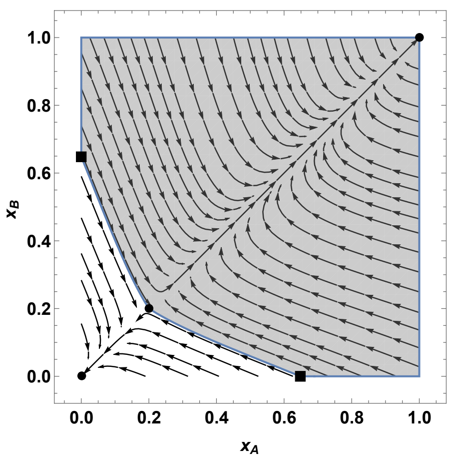
Figure 1.3: **Basins of attraction of the disease-free and fully-endemic equilibria, $(0, 0)$ and $(1, 1)$, for changing epidemic parameters ν, q.**



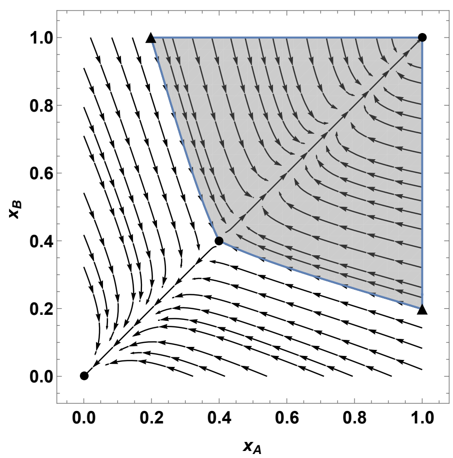
Parameters $\nu = 0.7$, $q = 0.2$.



Parameters $\nu = 0.7$, $q = 0.4$.



Parameters $\nu = 0.4$, $q = 0.2$.



Parameters $\nu = 0.4$, $q = 0.4$.

Monte Carlo simulations in MATHEMATICA to plot the vector field defining system (1.4) and the basins of attraction of the two asymptotically stable states $(x_A, x_B) = (0, 0)$ and $(1, 1)$ (respectively in white and shaded gray). The arrows depict the vector field defining system (1.4) in each point $(x_A, x_B) \in [0, 1]^2$ and confirm that the unit square is invariant and that the same is true for the diagonal and for the super-diagonal and sub-diagonal "triangles". Moreover, from the saddle point (q, q) one can identify the separatrix curves, the unstable separatrix coinciding with the diagonal, while the stable one constituting part of the border of the basins of attraction, thus separating them. Lastly, notice that as quarantine q increases, the system exhibits a larger and larger basin of attraction of the disease-free equilibrium $(0, 0)$, which is intuitively due to the fact that it is easier to recover from infection. The squared dots are $(\eta, 0)$ and $(0, \eta)$, i.e. the intersection points of the separatrix $\mathcal{C}$ with the horizontal and vertical axis mentioned in Proposition 1.4.4. Analogously, the triangular dots are $(1, \zeta)$ and $(\zeta, 1)$.

First-order separatrix We linearize[16] system (1.4) in a neighborhood of the saddle (q,q) and then analyze the eigenvector corresponding to the separatrix $\mathcal{C}$ of its Jacobian matrix, which gives a first-order (linear) approximation of the curve $\mathcal{C}$.

LEMMA 1.4.7. *Define the vector field*[17]

$$\mathbf{V}(x_A,x_B)=\nu\begin{pmatrix}\left[x_A(1-x_A)(x_A-q)(1-(x_A-x_B))^2\right]-x_A(x_A-x_B)\\ x_B(1-x_B)(x_B-q)+(x_A+x_B+2x_Ax_B)(x_A-x_B)\end{pmatrix}.$$

Its Jacobian matrix $J(x_A,x_B)=\left(\frac{\partial\mathbf{V}}{\partial x_A}\middle|\frac{\partial\mathbf{V}}{\partial x_B}\right)$*, when evaluated at* $(x_A,x_B)=(q,q)$ *becomes*

$$J(q,q)=\begin{pmatrix}(1-q)q\nu-q & q\\ 2(1-q)q\nu & -(1-q)q\nu\end{pmatrix},$$

whose eigenvectors are $(1,1)$*, corresponding to the diagonal separatrix, and*

$$\boldsymbol{\xi}=(\xi_A,\xi_B)=\left(-\frac{1}{2(1-q)\nu},1\right),$$

corresponding instead to the separatrix $\mathcal{C}$.[18]

By combining the previous lemma with a symmetric result holding for the analogous super-diagonal vector field, we obtain the following result, describing the linear approximation of the separatrix $\mathcal{C}$. Figure 1.8 shows similarities and differences between $\mathcal{C}$ and $\widetilde{\mathcal{C}}$.

LEMMA 1.4.8. *The separatrix* $\mathcal{C}$ *is linearly approximated*[19] *in* (q,q) *by the two-piece linear curve* $\widetilde{\mathcal{C}}$*, which we respectively call* $\widetilde{\mathcal{C}}^+$ *and* $\widetilde{\mathcal{C}}^-$*, defined by*[20]

$$\widetilde{\mathcal{C}}=\begin{cases}\widetilde{\mathcal{C}}^+: & x_B=\dfrac{1}{-2(1-q)\nu}(x_A-q)+q, \quad \textit{defined for } x_A\in[0,q]\\ \widetilde{\mathcal{C}}^-: & x_B=-2(1-q)\nu(x_A-q)+q, \quad \textit{defined for } x_A\in[q,1].\end{cases}$$

The intersection between $\widetilde{\mathcal{C}}^-$ *with the sub-diagonal boundaries of the unit square* $[0,1]^2$ *is the point*[21]

$$P^-=(P_A^-,P_B^-)=\begin{cases}\left(1,-2\nu(1-q)^2+q\right), & \textit{if } \nu<\dfrac{q}{2(1-q)^2},\\ \left(\dfrac{q}{2\nu(1-q)}+q,0\right), & \textit{if } \nu\geq\dfrac{q}{2(1-q)^2},\end{cases}$$

[16] Notice that we do not consider directly system (1.4), but its sub-diagonal part when it is extended to $\mathbb{R}^2$. By symmetry, the result holds also for the super-diagonal part.

[17] It coincides with the vector field defining system (1.4) in the sub-diagonal when $x_A>x_B$.

[18] Their respective eigenvalues are $q\nu(1-q)>0$ and $-q(1+\nu-q\nu)<0$, for $q,\nu\in(0,1)$.

[19] First-order approximation in a neighborhood of (q,q).

[20] Throughout, the symbols with $\cdot^+$ and $\cdot^-$ will respectively denote objects belonging to the super-diagonal or sub-diagonal part of the unit square $[0,1]^2$.

[21] Of course, by symmetry an analogous point P^+ can be found in the segments $\{0\}\times[q,1]$ and $[0,q]\times\{1\}$, when $\widetilde{\mathcal{C}}^+$ meets the boundary of $[0,1]^2$ in the super-diagonal part.

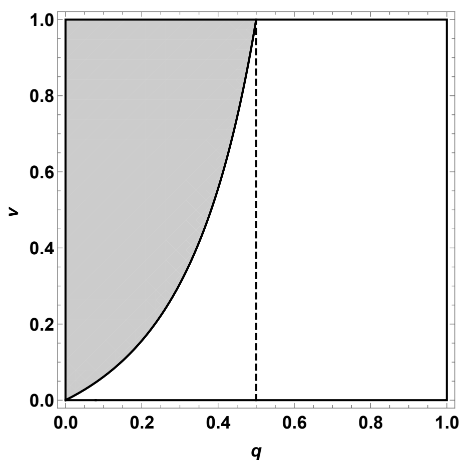
which respectively belong to the segments $\{1\} \times [0, q]$ *and* $[0, q] \times \{0\}$.

Remark. Notice that, provided $q \in (0,1)$ and $\nu \in (0,1)$, the two following conditions are equivalent:

$$\nu < \frac{q}{2(1-q)^2} \qquad \Longleftrightarrow \qquad \frac{1 + 4\nu - \sqrt{8\nu + 1}}{4\nu} < q.$$

Moreover, when $q > 1/2$ then $\nu < \frac{q}{2(1-q)^2}$ for all $\nu \in (0,1)$. Figure 1.4 shows the subregion of the square $(q, \nu) \in (0,1)^2$ where this condition is satisfied.

Figure 1.4: **Condition on the parameters** q **and** ν



Subregions of the square $(q, \nu) \in (0,1)^2$ separated by the curve $\nu = \frac{q}{2(1-q)^2}$. The white area is where $\nu < \frac{q}{2(1-q)^2}$, whereas the gray area is where the opposite inequality holds. In particular, in the white area (respectively, gray) P^- belongs to the vertical segment $\{1\} \times [0, q]$ (resp. horizontal segment $[q, 1] \times \{0\}$) and the area under the curve $\widetilde{\mathcal{C}}^-$ is the trapezoid Q^- (resp. triangle T^-). Lastly, the dashed line is at $q = 1/2$.

Depending on the parameters ν and q, the area under the curve $\widetilde{\mathcal{C}}$ is either a trapezoid or a triangle and is easily computed in the following result. By considering this area as an approximation of the area under the curve $\mathcal{C}$, this will also allow us to make a comparative statics analysis. The results are also shown in Figure 1.9.

LEMMA 1.4.9. *If* $\nu \geq \frac{q}{2(1-q)^2}$, *consider the triangle* $T^- \subset \{(x_A, x_B) \in [0,1]^2 : x_A \geq x_B\}$ *defined as the convex hull in* $\mathbb{R}^2$ *of the following set of vertexes*

$$T^- = \mathrm{Conv}\left(\{(q,q), (q,0), P^-\}\right).$$

If, instead, $\nu < \frac{q}{2(1-q)^2}$, *consider the trapezoid* $Q^- \subset \{(x_A, x_B) \in [0,1]^2 : x_A \geq x_B\}$ *defined by*

$$Q^- = \mathrm{Conv}\left(\{(q,q), (q,0), (1,0), P^-\}\right).$$

The measure of their area is:

$$\mathrm{A}(T^-) = \frac{q \times (P_A^- - q)}{2} = \frac{q^2}{4\nu(1-q)}, \quad \textit{defined whenever } \nu \geq \frac{q}{2(1-q)^2},$$

$$\mathrm{A}(Q^-) = \frac{(1-q) \times (q + P_B^-)}{2} = (1-q)\left(q - \nu(1-q)^2\right), \quad \textit{when } \nu < \frac{q}{2(1-q)^2}.$$

Whenever defined, $q \mapsto [\mathrm{A}(T^-)](q,\nu)$ is always increasing for all ν. Moreover, its derivative with respect to q is:

$$\frac{\partial \mathrm{A}(T^-)}{\partial q} = \frac{q(2-q)}{4\nu(1-q)^2} > 0, \quad \forall\, q,\nu \in (0,1): \ \nu \geq \frac{q}{2(1-q)^2}.$$

The derivative of $\mathrm{A}(Q^-)$ is

$$\frac{\partial \mathrm{A}(Q^-)}{\partial q} = 1 - 2q + 3\nu(1-q)^2, \quad \textit{defined whenever } \nu < \frac{q}{2(1-q)^2},$$

and it is positive if and only if the following condition holds[22]

$$\left\{\frac{2}{7} < q \quad \wedge \quad \frac{1-2q}{3(1-q)^2} < \nu < \frac{q}{2(1-q)^2}\right\} \vee \left\{q < \frac{2}{7} \quad \wedge \quad \left[\nu < \frac{q}{2(1-q)^2} \quad \vee \quad \nu > \frac{1-2q}{3(1-q)^2}\right]\right\}.$$

Now let us compute the ratio between the area under the curve $\widetilde{\mathcal{C}}^-$ and the entire rectangle $[q, 1-q] \times [0, q]$, as in Figure 1.9.

LEMMA 1.4.10. *Let $\widetilde{R}(q,\nu)$ be the ratio between the area under the curve $\widetilde{\mathcal{C}}^-$ and the rectangle $[q,1-q] \times [0,q] \subset [0,1]^2$. Then $\widetilde{R}(q,\nu)$ is the continuous function*[23] *defined by*

$$\widetilde{R}(q,\nu) := \begin{cases} \dfrac{[\mathrm{A}(T^-)](q,\nu)}{q(1-q)} = \dfrac{q}{4\nu(1-q)^2}, & \textit{if } \nu \geq \dfrac{q}{2(1-q)^2} \\ \dfrac{[\mathrm{A}(T^-)](q,\nu)}{q(1-q)} \equiv \dfrac{[\mathrm{A}(Q^-)](q,\nu)}{q(1-q)} = \dfrac{1}{2}, & \textit{if } \nu = \dfrac{q}{2(1-q)^2} \\ \dfrac{[\mathrm{A}(Q^-)](q,\nu)}{q(1-q)} = \dfrac{q-\nu(1-q)^2}{q}, & \textit{if } \nu \leq \dfrac{q}{2(1-q)^2}. \end{cases}$$

The behavior of $\widetilde{R}(q,\nu)$, as a function of the parameters q and ν, is described in the following result and in Figure 1.5.

LEMMA 1.4.11 (Comparative statics on the approximated ratio $\widetilde{R}(q,\nu)$).

The function $\widetilde{R}(q,\nu)$ defined above is bounded in $[0,1]$, moreover its sections $q \mapsto \widetilde{R}(q,\nu)$ are increasing for all $\nu \in (0,1)$, whereas $\nu \mapsto \widetilde{R}(q,\nu)$ are decreasing

[22] The complicated condition derives from the fact that $\frac{1-2q}{3(1-q)^2}$ is decreasing while $\frac{q}{2(1-q)^2}$ is increasing, when $q \in (0,1)$, and they cross in $q = 2/7$.

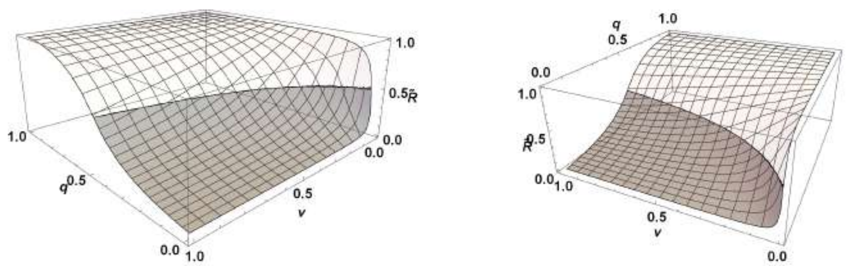
[23] There may be problems for q, ν close to 0, but it should be enough considering $q,\nu > \varepsilon > 0$.

for all $q \in (0,1)$. *Furthermore, whenever defined, its derivatives* $\widetilde{R}(q,\nu)$ *are:*

$$\frac{\partial}{\partial q}\widetilde{R}(q,\nu) = \begin{cases} \dfrac{1+q}{4\nu(1-q)^3} > 0, & \text{if } \nu > \dfrac{q}{2(1-q)^2} \\ \left(\dfrac{1}{q^2} - 1\right)\nu > 0, & \text{if } \nu < \dfrac{q}{2(1-q)^2} \end{cases}$$

$$\frac{\partial}{\partial \nu}\widetilde{R}(q,\nu) = \begin{cases} -\dfrac{q}{4(1-q)^2\nu^2} < 0, & \text{if } \nu > \dfrac{q}{2(1-q)^2} \\ -(1-q)^2/q < 0, & \text{if } \nu < \dfrac{q}{2(1-q)^2}. \end{cases}$$

Figure 1.5: **Plot of the "approximated" ratio** $\widetilde{R}(q,\nu)$



Same plot in 3D of the function $\widetilde{R}(q,\nu)$, for all $(q,\nu) \in (0,1)^2$, from two different viewpoints. The shaded-gray area corresponds to $\widetilde{R}(q,\nu)$ with values q,ν such that $\nu > \frac{q}{2(1-q)^2}$, whereas the white area to those satisfying the opposite inequality (see also Figure 1.4). The plot shows also that $\widetilde{R}(q,\nu)$ is increasing in q and decreasing in ν.

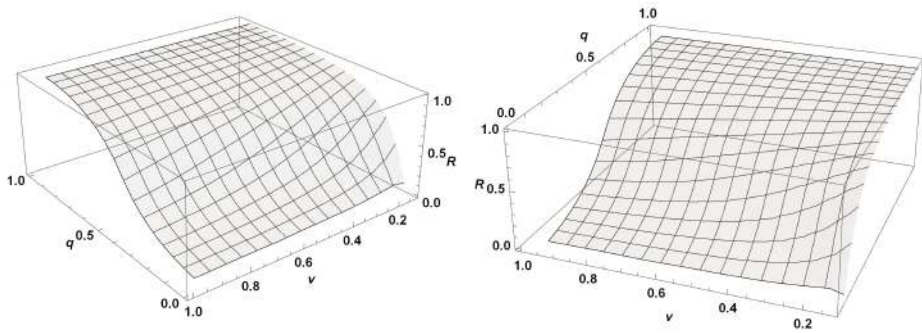
1.5 Discussion

In this section, we focus our attention on the results that can be drawn from the analysis done in Section 1.4.2. We will compare the two following situations:

- the first where no cross-country trade between the two locations is allowed, meaning that they are considered separated and autarkic;
- the second where cross-country trading is instead allowed, as described in the previous section.

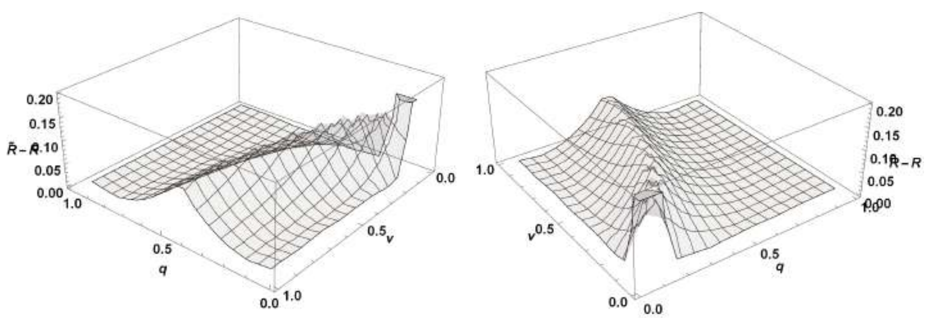
By comparing how a disease differently spreads in these two systems, we aim to analyze how a very "stylized globalization", the second, affects the systemic resistance to potential shocks in the infection rates. We will argue that depending on the intensity and "dimensionality" of the shock, being being "autarkic" or "globalized" may or may not be advantageous.

Figure 1.6: **Plot of the "actual" ratio** $R(q, \nu)$



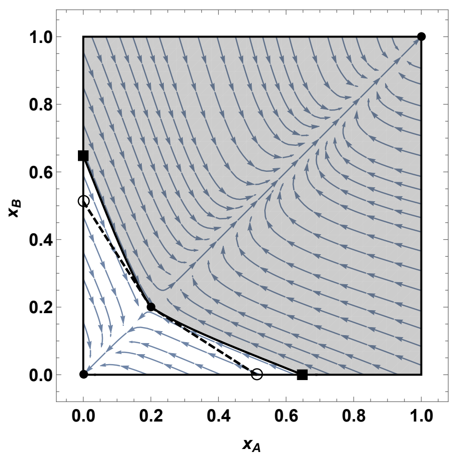
Plot in 3D of the function $R(q, \nu)$, for all $(q, \nu) \in (0, 1)^2$, obtained through Monte Carlo simulations. The plot of $\widetilde{R}(q, \nu)$ of Figure 1.5 resembles this one of $R(q, \nu)$, giving the idea of the goodness of the approximation.

Figure 1.7: **Plot of the difference** $|\widetilde{R}(q, \nu) - R(q, \nu)|$

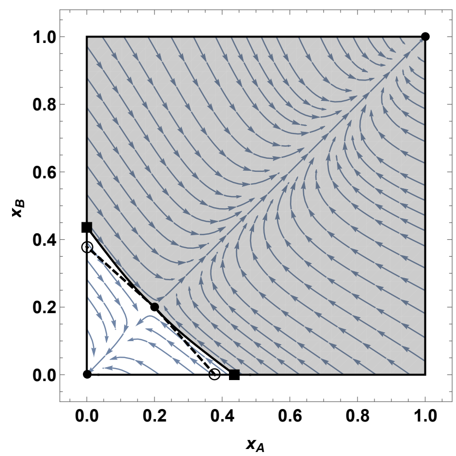


Same plot in 3D of the difference between the "approximated" and "actual" ratio $|\widetilde{R}(q, \nu) - R(q, \nu)|$. With a grid step of 0.05 in both components q and ν, the resulting average difference is 0.03.

Figure 1.8: **Linearized approximated separatrix $\widetilde{\mathcal{C}}$ and comparison with the actual separatrix $\mathcal{C}$**



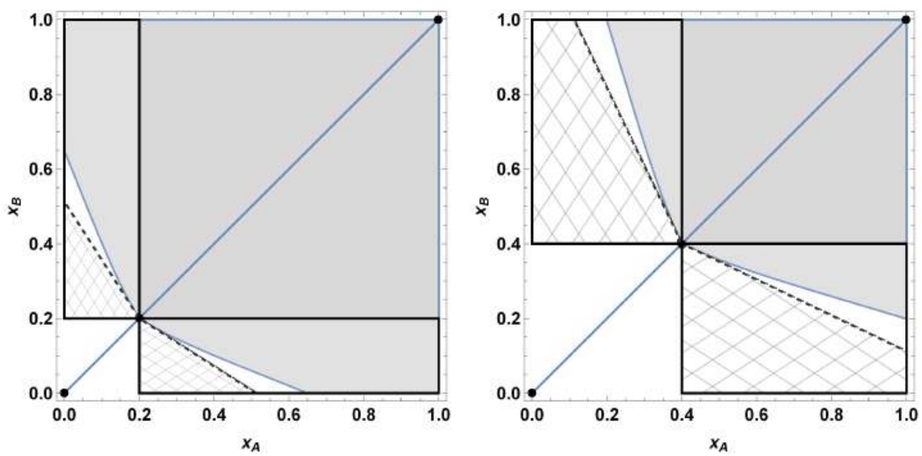
Parameters $\nu = 0.4$, $q = 0.2$.



Parameters $\nu = 0.7$, $q = 0.2$.

Plot of the linearized separatrix $\widetilde{\mathcal{C}}$ (dashed straight line) and comparison with the actual separatrix $\mathcal{C}$ (continuous black curve). $\widetilde{\mathcal{C}}$ is a first-order approximation of $\mathcal{C}$ in a neighborhood of the saddle (q, q) so the approximation is only locally good. The intersection points of $\widetilde{\mathcal{C}}$ and $\mathcal{C}$ with the boundaries of the unit square (respectively, circles and squares) may be quite different depending on the epidemic parameters ν, q.

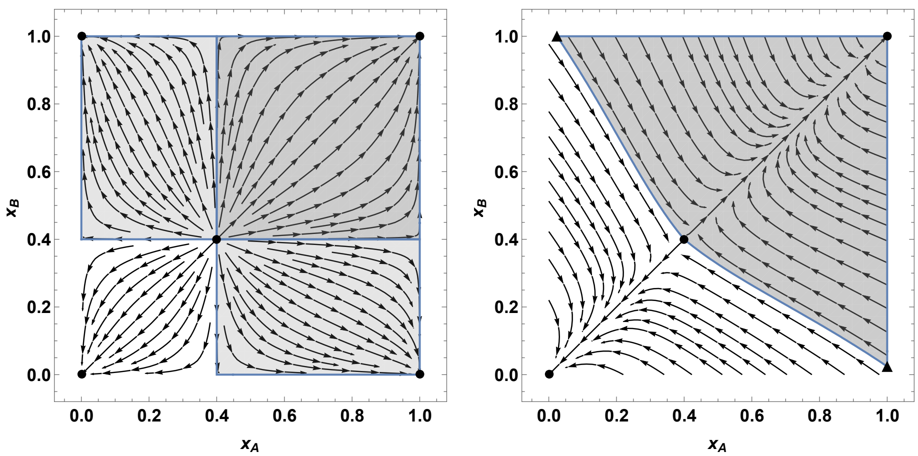
Figure 1.9: **Approximated area of the triangle/trapezoid**



Parameters $\nu = 0.4$, $q = 0.2$. Parameters $\nu = 0.4$, $q = 0.4$.

Approximated areas of the triangles/trapezoids (grid-shaded areas), defined by the linearly approximated $\widetilde{\mathcal{C}}$ (dashed lines), in the rectangles of interest. If $\nu > \frac{q}{2(1-q)^2}$, then the point P^- of Lemma 1.4.8 belongs to the vertical segment $\{1\} \times [0, q]$, generating then the trapezoid Q^- of Lemma 1.4.9 (right). By symmetry with respect to the diagonal, their super-diagonal counterparts P^+ and Q^+. In the opposite case (left), P^- belongs to the horizontal segment $[0, q] \times \{0\}$ and the resulting area is the triangle T^- (by symmetry, T^+).

Figure 1.10: **Comparison between autarkic and interconnected locations**



Autarky, parameters $\nu = 0.7$, $q = 0.4$. Globalized, parameters $\nu = 0.7$, $q = 0.4$.

Dynamics of disease spreading, in case of two autarkic locations (left) and of two globalized locations (right), with same epidemic parameters. In both, in dark gray the basin of attraction of $(1, 1)$, while in white that one of $(0, 0)$. Only the autarkic case exhibits the presence of two partial-endemic asymptotically stable states, $(0, 1)$ and $(1, 0)$, whose basins of attraction are depicted in light gray, where only one of the two locations is fully infected while the other is disease free. A comparison between the white areas can give the idea of the intensity of (correlated or 2-dimensional) shocks that the system can completely absorb.

In particular, we will show that *small shocks* are better absorbed by an interconnected system, independently of their dimensionality: intuitively, the shock is more diluted in a larger system. On the contrary, somehow surprisingly, *large shocks* may or may not have worse consequences when the locations are interconnected, depending on the amount of resources dedicated to having a large recovery parameter q. More specifically, we show that:

- (large and mainly) 1-dimensional shocks, i.e. shocks mainly starting only in 1 location, have worse consequences when the system is globalized, while such shocks would remain “confined” when there is autarky;

- (large) 2-dimensional shocks, i.e. shocks starting simultaneously in both locations, on the contrary, have worse consequences when the system is globalized.

1.5.1 The Case of Autarky

Let us consider two autarkic locations, where no trade is possible between them, and where each is subject to a disease-spread dynamic that behaves according to the single-location model studied in Section 1.3. The evolution over time of the two infection rates $x_A(t)$ and $x_B(t)$ of these two locations A and B can be written as a system of two (uncoupled) differential equations:[24]

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} x_A = \nu x_A(1-x_A)(x_A-q) \\ \dfrac{\mathrm{d}}{\mathrm{d}t} x_B = \nu x_B(1-x_B)(x_B-q). \end{cases} \tag{1.5}$$

The dynamics and results[25] are pictured in Figure 1.10 (left) and summarized in the following proposition.

PROPOSITION 1.5.1. *Given two autarkic locations A and B, system (1.5) has the following properties:*

- *it is symmetric with respect to the diagonal, which is then an invariant set. The super-diagonal and sub-diagonal sets in $\mathbb{R}^2$, $\{(x_A, x_B) \in \mathbb{R}^2 : x_A < x_B\}$ and $\{(x_A, x_B) \in \mathbb{R}^2 : x_A > x_B\}$, are also invariant;*
- *the unit square $[0,1]^2$ is invariant;*
- *the critical points where $\frac{\mathrm{d}x_A}{\mathrm{d}t} = \frac{\mathrm{d}x_B}{\mathrm{d}t} = 0$ are:*
  - *$(0,q)$, (q,q), $(1,q)$, $(q,0)$, $(q,1)$, which all are (unstable) saddle points;*
  - *$(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$, which all are asymptotically stable equilibria.*

Moreover, the separatrix curves of the saddle points are the lines $x_A = 0$, $x_B = 0$, $x_A = q$, $x_B = q$, $x_A = 1$ and $x_B = 1$, and this also makes possible characterizing the basins of attraction of the stable points in $[0,1]^2$:[26]

- *$[0,q)^2$ is the basin of attraction of $(0,0)$;*
- *$[0,q) \times (q,1]$ is the basin of attraction of $(0,1)$;*
- *$(q,1] \times [0,q)$ is the basin of attraction of $(1,0)$;*
- *$(q,1]^2$ is the basin of attraction of $(1,1)$.*

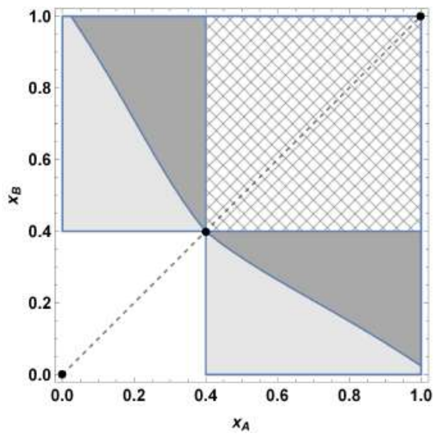
From the previous result, it can also be said that the points $(1,0)$ and $(0,1)$ play a peculiar role, as shown also in Figure 1.10 (left): they represent a situation in which only one of the two locations is fully infected, while the other is disease

[24] For a reasonable comparison with an analogous globalized 2-location model, here we assume the same symmetric epidemic parameters $\nu_A = \nu_B = \nu$ and $q_A = q_B = q$.

[25] See Appendix **??** for details.

[26] Of course, we only consider their intersection with the unit square, which is the sets in which the fraction of infected x_A and x_B make sense.

Figure 1.11: **Systemic resistance to small vs. large shocks**



Comparison by juxtaposition of the areas obtained in Figure 1.10, with same parameters $\nu = 0.7$, $q = 0.4$. The white area, $[0, q)^2$, and the grid-shaded area, $(q, 1]^2$, depict where a hitting shock would produce the same outcome, independently of locations being autarkic or globalized. Light-gray areas measure where shocks result in a partial endemic state, in case of autarkic locations, or where they are instead totally recovered, in case of globalized locations. On the contrary, dark-gray areas are those where shocks result fully infected system, if globalized, whereas only partial infection, if autarkic.

free. In case of autarky, this may happen when the initial point of infection at time $t = 0$ belongs to $[0, q) \times (q, 1]$ or $(q, 1] \times [0, q)$, which will cause the dynamics to convergence toward $(0, 1)$ or $(1, 0)$, respectively. Figure 1.10 (right) also shows that this cannot be the case when the two locations are connected and globalized. This feature will be a key ingredient in the next section about shock analysis.

1.5.2 Shock Analysis

In our framework, we assume that the system is in a disease-free condition when an exogenous "shock" in the infection rates occurs. Analytically, *shocks* are assumed to be uniform at random over the unit square $[0, 1]^2$ and represented by a given initial condition $(x_A(0), x_B(0)) \in [0, 1]^2$ at time $t = 0$, i.e. by a vector

$$\mathbf{s} = (s_A, s_B) := (x_A(0), x_B(0)) .$$

Figures 1.10 and 1.11 show the comparison of the basins of attraction of the point $(1, 1)$, which corresponds to both location being fully infected, when they are considered autarkic or connected. Due to the shape of these basins of attraction, we obtain the following results.

Given a shock $\mathbf{s} = (s_A, s_B)$, if it is large enough or small enough in both components, then the resulting outcome for an autarkic system and for a globalized system will be the same. In particular:[27]

- if $s_A < q$ <u>and</u> $s_B < q$, then both the autarkic system and the globalized system will be able to fully recover;

[27] Notice that assuming that the shock distribution is uniform or continuous, thus atom-less, guarantees that the probability that one component hits 0, q or 1 is zero.

- if $s_A > q$ and $s_B > q$, then both systems will converge to a fully infected endemic state.

In Figure 1.11, these two cases are respectively represented by a shock belonging to the white area and to the grid-shaded area.

On the contrary, the outcome resulting from a shock hitting mainly one location is completely different when the two locations are autarkic or connected. Indeed, consider an almost 1-dimensional shock $\mathbf{s}$ targeting mainly location A[28], that is

$$\mathbf{s} = (s_A, \varepsilon), \quad \text{with} \quad \varepsilon < q < s_A.$$

Let us analyze first the autarkic case. The dynamics will converge to a partial epidemic equilibrium: Proposition 1.5.1 and Figure 1.10 (left) show that such a large shock s_A on A will cause the dynamic to converge to a fully infected location A, while at the same time a small shock ε would be (independently) recovered by location B.

Now, what happens instead when the A and B are connected while facing the same shock $\mathbf{s} = (s_A, \varepsilon)$ as before? Two different situations may arise:

- if s_A is large enough and such that $\mathbf{s} = (s_A, \varepsilon)$ belongs to the basin of attraction of $(1, 1)$, i.e. dark-gray area in Figure 1.11, then the globalized system will end up being fully infected;

- if, instead, s_A is still greater than q but not large enough, then (s_A, ε) will belong to $(0, 0)$'s basin (light gray in Figure 1.11) and so the globalized system will manage to recover from this shock.

This analysis shows that 2-location autarkic system and a 2-location globalized system react very differently in response to large 1-dimensional shocks.

The dark-gray areas in Figure 1.11 are constituted by all those possible shocks that cause the infection to spread to both locations, when they are connected, or to just one location, when autarkic. Assuming a uniform shock distribution, this area then measures the weakness of the system with respect to this kind of shocks and, in addition, it also somehow measures the advantage of an autarkic system over a globalized one.

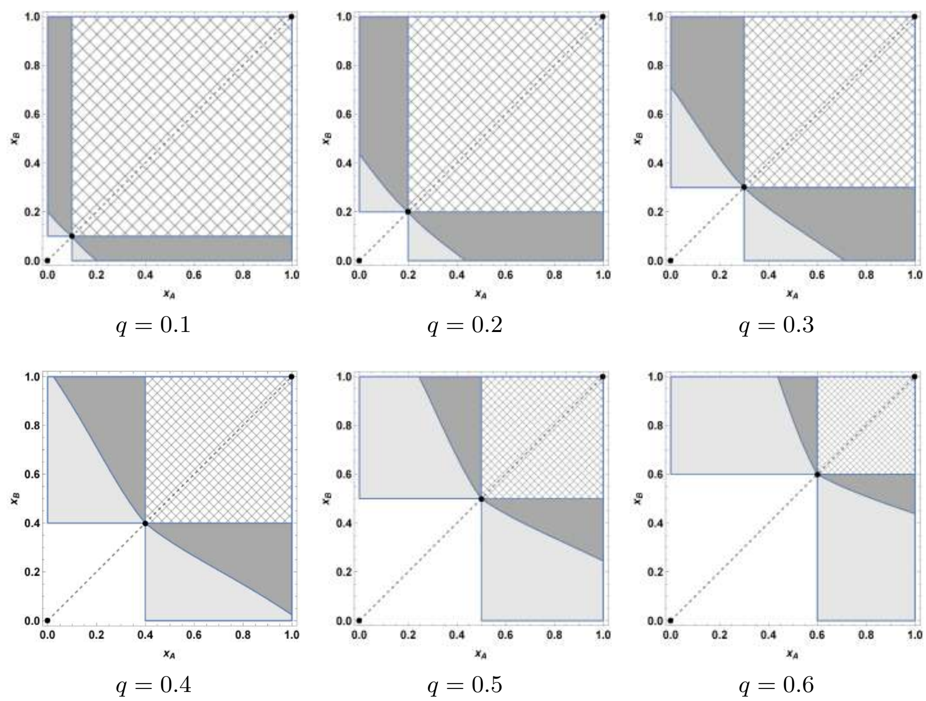
Analogously, but in an opposite way, the light-gray areas in Figure 1.11 capture the advantage of a globalized system over an autarkic one: shocks belonging to these regions are recovered by a connected system, whereas they result in a partial epidemic equilibrium in the autarkic case.

1.5.3 Systemic Resistance & Policy Implications

Understanding the relationship between the two gray areas in Figure 1.11 becomes necessary, because it gives an indication of the relative advantage or disadvantage of an autarkic system over a globalized system. Furthermore, Figure 1.12 also shows that this advantage changes as the exogenous recovery parameter q varies: this turns out to be crucial for policy making.

[28]By symmetry, the argument applies in the same fashion to a shock concentrated in B.

Figure 1.12: **Comparative statics**



Regions of interest when contagiousness $\nu = 0.7$ is kept fixed, while quarantine q, the policy parameter, increases in $(0, 1)$. Notice that as q increases, the ratio of light-gray area to dark-gray area increases as well. This can be interpreted by saying that the more resources are available for recovery, the more advantageous it becomes to be globalized, in terms of systemic resistance to shocks.

One way to address this issue is by analyzing the separatrix curve $\mathcal{C}$ of the saddle (q, q), in case of globalized system, because it forms the boundaries and thus separates the basins of attraction of our the regions of interest. Unfortunately, apart from Proposition 1.4.4, which relies on the "local" information provided by the eigenvector of the linearized system in the neighborhood of the saddle point (q, q) and on the monotonicity of the components of the vector field defining system (1.4) in specific areas, we have to rely on approximated results, due to the impossibility of explicitly describing $\mathcal{C}$ analytically.

Specifically, we first numerically approximate the intersection points between the separatrix and the boundaries of the unit square and, then, numerically measure the gray areas and determine their relative ratio, which, as already observed, is key to understanding whether a globalized system is shock-resistance superior to an autarkic system, given the same parameters q and ν.

(Numerical) comparative statics Concerning the approximation of the intersection points, it is worth remembering that they depend on the epidemic parameters q and ν. With the same notation already used in Proposition 1.4.4 and Figure 1.11, we call them $(x_A, x_B) = (\eta(q,\nu), 0)$ and $(1, \zeta(q,\nu))$. This numerical analysis is shown in Figure 1.13:

- holding fixed $\nu \in (0,1)$, whenever $\mathcal{C}$ crosses the segment $[q,1] \times \{0\}$ in the point $(\eta(q,\nu), 0)$, then $q \mapsto \eta(q,\nu)$ is increasing in q and spans from 0 to 1. Moreover, $\eta(q,\nu) > q$;
- analogously, when q exceeds a certain threshold[29], then $\mathcal{C}$ crosses the segment $\{1\} \times [0,q]$ in the point $(1, \zeta(q,\nu))$; moreover, $q \mapsto \zeta(q,\nu)$ is increasing, spanning from 0 to 1 and always satisfies $\zeta(q,\nu) < q$.

Let us now turn to the relative advantage/disadvantage of an autarkic system versus a globalized system, especially when subject to 1-dimensional[30] shocks. We have already observed that, assuming a uniform distribution of shocks in $[0,1]^2$, the areas in light gray and dark gray of Figures 1.11 and 1.12 measure the extent to which an autarkic system or a globalized system is relatively more or less subject to (1-dimensional[31]) shocks.

Holding fixed the contagiousness[32] ν, as the recovery parameter q increases, the light-gray areas expand while the dark-gray areas shrink. According to our previous interpretation, an expansion of the light-gray areas means that it becomes more likely that a 1-dimensional shock lead the autarkic system to a partial endemic equilibrium, while a corresponding reduction of the dark-gray areas[33] means that a globalized system becomes more able to recover from shocks. This, in turn, means that the larger it is the available level of quarantine q, the more convenient it becomes to be in a globalized system relative to an autarkic one. In this respect, Figure 1.12 shows how the light-gray and dark-gray areas change, as the quarantine q changes[34]. This analysis is also shown in Figure 1.14, where we plot the percentage of the rectangle $[q,1] \times [0,q]$ which is occupied by the dark-gray area. By using the shock analysis done above, as q increases, we observe that having a connected 2-location system becomes more and more advantageous and resistant overall than an autarkic 2-location system.

This conclusion directly translates in terms of policy: if the available quarantine level q can be taken large enough[35], then allowing cross-country import-export is beneficial and preferable, from the point of view of the systemic resistance to infection shocks. On the contrary, two autarkic countries constitute

[29] Threshold that corresponds to $q = 0.39$ in Figure 1.13.

[30] Meaning, shocks that mainly start from a single location, of the form $\mathbf{s} = (\varepsilon, s_B)$ or (s_A, ε), with $\varepsilon \approx 0$.

[31] Meaning, single-located.

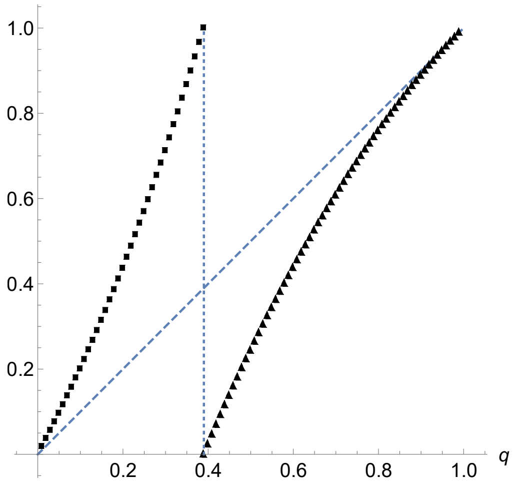
[32] Which we consider as a disease-related parameter, thus not subject to policy making.

[33] Since it corresponds to an expansion of the white recovery area, for a globalized system.

[34] While contagiousness ν is kept fixed, because we think of it as a disease-related parameter, not subject to policy making.

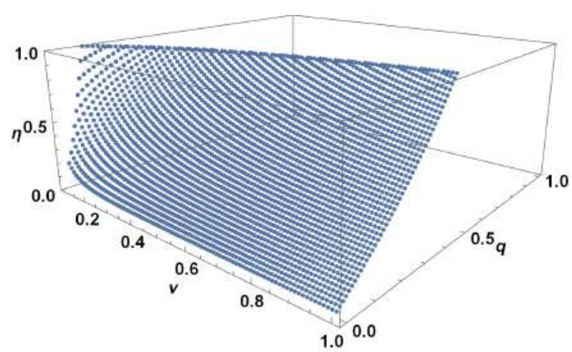
[35] Notice that q is exogenously given in our framework.

Figure 1.13: **Intersection between separatrix $\mathcal{C}$ and boundaries of the unit square** $[0,1]^2$



Plot of the intersection points $q \mapsto \eta(q, 0.7)$ (squares) and $q \mapsto \zeta(q, 0.7)$ (triangles), with fixed $\nu = 0.7$. As q increases, the separatrix $\mathcal{C}$ first crosses the horizontal segment $[q, 1] \times \{0\}$ in $(\eta(q, \nu), 0)$, then as q exceeds a certain threshold ($q = 0.39$ in this case, signaled by the dotted vertical line), $\mathcal{C}$ starts crossing the boundary in the vertical segment $\{1\} \times [0, q]$ in the point $(1, \zeta(q, \nu))$. Lastly, the dashed line is the diagonal, showing that indeed $\eta > q$ while $\zeta < q$.

Plot in 3D of the intersection point $\eta(q, \nu)$ (on the vertical axis) as a function of both epidemic parameters $(q, \nu) \in (0, 1)^2$. All sections $\eta(\cdot, \nu)$ and $\eta(q, \cdot)$ are increasing. In particular, the section $\eta(\cdot, 0.7)$ is plotted above (squares).

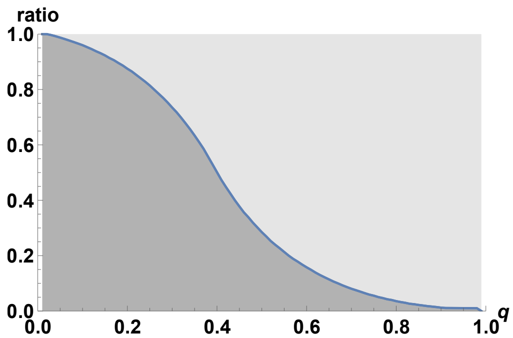


a more resistant system against infection shocks when only a small level of quarantine q is available.

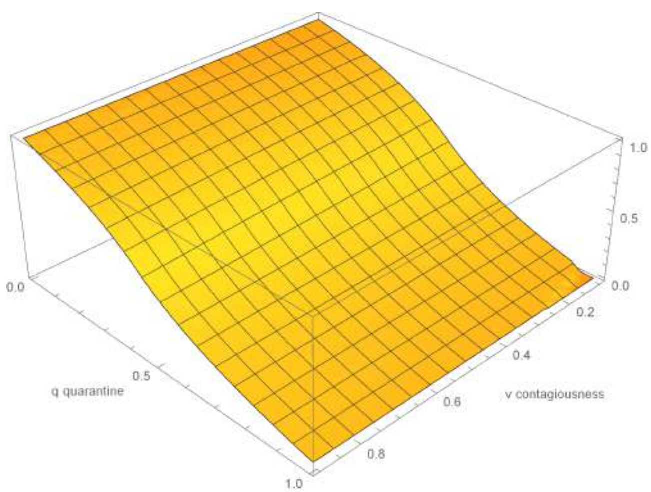
1.6 Conclusions

Starting from a very simple model of epidemic diffusion among homogeneous agents, we consider the case in which two identical countries are inhabited by such agents. The agents at random interact with each other in pairs to obtain benefits and, by doing so, they also spread a contagious disease among them, which lowers the attainable gain from trade. As a response to the infection risk, agents can choose to pay an extra cost to interact with the agents present in the other country, establishing then a stylized form of cross-country import-export trade. By assuming that both countries have a (limited and fixed) quarantine to intervene against the infection, we are also able to introduce the possibility

Figure 1.14: **Ratio between the gray areas**



Plot of the ratio between the dark-gray area and the sum of the dark-gray plus light-gray areas (i.e. $[q,1]\times[0,q]\cup[0,q]\times[q,1]$) numerically obtained as a function of $q\in(0,1)$, holding fix $\nu=0.7$. As expected from Figure 1.12, as the quarantine q increases, the light-gray area increases at the expenses of the dark-gray area, meaning that a globalized system becomes more and more convenient relative to an autarkic one.



Plot in 3D of the ratio of the dark-gray area to the sum of the dark- plus light-gray areas, numerically obtained as a function of both parameters $(q,\nu)\in(0,1)^2$. The plot above is, then, a section of this one corresponding to $\nu=0.7$.

of recovery, i.e. of reducing the infection rate.

Holding fixed the epidemic parameters, we compare the resistance to exogenous shocks in infection rates of the "autarkic" system, in which the two countries are assumed to be independent, with the one of the "globalized" system, where instead cross-country interaction is allowed. Overall, globalized systems result more "extreme" in their reaction to shock than autarkic systems, as a result of the two countries being connected: on the one hand, the globalized system has a larger "recovery capacity" for facing small shocks, on the other hand, it has a larger area where both countries are completely infected. In particular, the main possibility which is precluded to globalized systems with respect to autarkic ones is a situation in which only one country is infected while the other is not.

On the contrary, "autarkic" systems offer a wider spectrum of possible outcomes, resulting from infection shocks and, in particular, they also exhibits partial endemic equilibria in which one location is fully infected while the other is disease free.

By comparing how an autarkic system and a globalized system behave in response to shocks, we are able to understand their similarities and, most of all, their differences. The main result of this shock-resistance analysis is that the behavior of the two systems is substantially different exactly when they are subject to "1-dimensional large shocks": when infection shocks hit mainly one location and only slightly the other, a globalized system either fully recovers or becomes fully infected, while an autarkic system can exhibit partial endemic equilibria. Depending on the amount of resources allocated to recovery, as measured by the quarantine level q in our framework, a globalized system may be preferable when large quarantine levels are available, whereas an autarkic system is preferable in case of low quarantine.

1.7 Appendix: Related Literature

Models studying epidemics have been considered for many decades now and the literature is wide and multidisciplinary. Here we briefly review it by accounting for the fact that scholars from different scientific fields have approached the problem from possibly different perspectives: the first models come from (mathematical) biology, whereas computer scientists and social scientists have only come after. In particular, some of the more recent studies have already tried to include human behavioral responses in the evolution of the epidemic spreading, as it is also explained in [Manfredi and D'Onofrio, 2013].

From mathematical biology, some standard models are those concerning epidemics spreading among different locations (also called *patches*). The authors usually consider a system of ordinary differential equations describing the evolution of the infection rate in the patches. Just to name a few papers:

- [Hsieh et al., 2007], [Wang and Mulone, 2003] and [Wang and Zhao, 2004] analyze the effect of population dispersal or migration, for example. Another standard model of a multi-city spreading epidemic is given in [Arino and Van den Driessche, 2003].

- [Brauer and van den Driessche, 2001] uses instead a constant flow of immigrants. [Castillo-Chavez and Yakubu, 2001] are among the firsts to use a discrete approach (instead of a continuous one).

- [Jin and Wang, 2005] goes on the same line of research opened by [Wang and Zhao, 2004] but it compares more carefully the basic reproduction numbers R_0 obtained when the patches when isolated or connected.

- [Cui et al., 2008] and [Sun et al., 2011] include *media coverage* affecting the probability of getting infected, because it may contribute to inform people about prevention measures to adopt. Other cases of transport-related mechanisms, see [Cui et al., 2006], consider infections occurring in crowded environments, such as trains, airplanes etc.

More recently, the attention has focused also on models including behavioral aspects that may affect the evolution of the spread of an epidemic. Moreover, since one of the main behavioral responses consists in the adaptation of the contact network of the individuals (social distancing or spontaneous social exclusion, etc), this means that there has been an increase need of models capable of distinguish among heterogeneous contact patterns.

More specifically, this has automatically translated in an increase attention in models of epidemics in social networks, where the dynamics of both phenomena can be studied at the same time: on the one hand how the contact network adapts to . Some useful reviews of these studies may be [Wang et al., 2015], [Wang and Li, 2014] and [Funk et al., 2010].

In this direction, some other interesting reads may be the following.

- [Reluga, 2009] and [Reluga, 2010] propose models for epidemics using differential game theory also for modeling social distancing as a response to infection risk.

- Another game-theoretical approach is followed in [Poletti et al., 2012]: the authors explicitly model risk perception and micro-found the agents' uncoordinated decision of increasing social distance by means of evolutionary game theory.

- Other scholars have tried to use a more data-driven approach that also makes use of simulations for taking care of heterogeneity in the local-global contact structure and of behavioral responses etc. Some examples may be [Hufnagel et al., 2004], [Meloni et al., 2011] and [Colizza and Vespignani, 2008], [Funk et al., 2009], [Fenichel et al., 2011], [Sahneh et al., 2012] and [Cherif et al., 2015].

- Some authors focus on the complex dynamics that may arise when the co-evolution of the infection and the network are considered together. To name just a few:

  - in [Zanette and Risau-Gusmán, 2008] the authors also use bifurcation theory to study the stability and behavior of equilibria;

  - in [Gross et al., 2006] the authors find a peculiar dynamics called *bistability*, i.e. the simultaneous coexistence of different stable equilibria;

  - some fluctuating dynamics are described in [Shaw and Schwartz, 2008].

Lastly, we focus on that strand of literature more belonging to economics and social science. To mention just a few, in this part we can refer to [Goyal and Vigier, 2015] or [Galeotti and Rogers, 2015], concerning more strictly game-theoretical foundations and epidemics, whereas to [Ehrhardt et al., 2006] if it concerns the dynamics and an approach based on mean-field analysis. In the economic field, of course many efforts have been done in order to study the optimal allocation of resources to effectively fight and prevent epidemics. Among the last ones, see [Dimitri, 2015] and [Horan et al., 2015].

Chapter 2

Exploration and Exploitation in Science

2.1 Abstract

Since its introduction, by [March, 1991], the *exploration-exploitation dilemma* has been extensively studied in many fields and different contexts. Here, we focus on scientists and researchers: in the pursuit of science, how do they behave to expand their knowledge? If one were to follow them on the map of their personal "knowledge space", would one find that they wander erratically, exploring this space or that, once they find a gold vein, they keep digging there? How do they allocate their limited time and resources between these two behaviors? Is there even an optimal allocation? Does this allocation change over their career, perhaps allowing to distinguish the behavior of junior and senior researchers? Have there been changes over time, perhaps pushed by changes in academia and scientific production in the past decades or by a more specialized and competitive world?

In this paper, we follow individual scientists over all their careers and use their co-authorship (and citation) network to map "what they know", i.e. as a proxy of their knowledge space and of its evolution over time. On this network we then introduce an embeddedness-based measure of the exploration-exploitation ratio for each scientist and compare it with her productivity (i.e. number of papers written, citations received). Preliminary results show that there is an optimal exploration-exploitation ratio that is chosen by scientists who excel in terms of performance and productivity.

This work is done in collaboration with Alessandro Vespignani and Matteo Chinazzi.

2.2 Introduction

Research question and context The aim of this paper is to investigate the following research questions: do scientists choose how to behave, in terms of exploration and exploitation (a.k.a. specialization and diversification)? And, in case they do, is this choice linked to and/or affecting (and how) their productivity and performance?

To clarify why to study science and scientists. We see several peculiar features belonging to such an analysis and useful to justify it: given the abundance of detailed data, it is possible to follow many scientists for decades and for their whole career; moreover, their behavior is under study when the stakes are high (i.e. their jobs, careers), which is often not possible in experiments, for example. Lastly, the possibility of performing cross- and within-field comparisons can be particularly useful to shed light on some behaviors and peculiarities that are typical of some sectors.

Where in the existing literature This paper can be located at the intersection of several strands of research:[1]

- exploration-exploitation dilemma for individuals, groups, firms and organizations;[2]
- advantages and drawbacks of specialization (versus diversification), in the field of science and, additionally, as well-established issues in economics literature;[3]
- literature on innovations, patenting, etc;[4]
- growing literature on the so-called science of science[5], and particularly in its branch within economics[6].

Main ideas and methods A summary of the main methods and ideas used is the following:

- Each author is assumed to have a knowledge space. This knowledge space is, in turn, described by two dimensions: extension (area?) and depth.
- When two scientists co-author a paper, their knowledge spaces interact. How? Roughly speaking, we assume that:

[1] More on literature is in the Appendix.

[2] [Hills et al., 2015]

[3] [Levitt and March, 1988], [Levinthal and March, 1993]

[4] [Gilsing et al., 2008]

[5] [DeSolla Price, 1965], [Newman, 2004], [Uzzi and Spiro, 2005], [Wuchty et al., 2007], [Radicchi et al., 2008], [Petersen et al., 2010], [Uzzi et al., 2013], [Foster et al., 2015], [Boudreau et al., 2016], [Clauset et al., 2017].

[6] [Hudson, 1996], [Fafchamps et al., 2010], [Hamermesh, 2013], [Ductor et al., 2014], [Abramitzky, 2015], [Ductor, 2015], [Seltzer and Hamermesh, 2017].

- where the two spaces overlap, this combination of strengths results in an deepened knowledge (for both);
- where, instead, the two spaces do not overlap, the non-overlapping part of one scientist contributes to expanding and widening the other's knowledge.

In other words, when two knowledge-wise "similar" scientists decide to collaborate, their knowledge mainly becomes deeper in that common topic/area. On the contrary, when the two authors are "different", then their resulting knowledge is wider.

- The overlapping part accounts for an "exploitative" behavior (or specialized), while the non-overlapping one for a more "exploratory" choice.

Practically, a *co-authorship network* is created, where a link exists whenever two authors have co-authored a paper. So, the set of co-authors that a scientist has (had so far) is used as a proxy for her (current) knowledge space. Then, two such sets, respectively corresponding to two authors, will overlap when there are common co-authors. In network terminology, when two co-authors have their respective networks of co-authors very embedded in each other, then they are considered to have a "similar" knowledge space.

2.3 Data

The data come from the *American Physical Society* (APS) datasets and consist of information about over 575,000 different articles published on APS journals[7], written by almost 300,000 different authors over the course of 120+ years, dating from 1893 to 2015. More precisely, all the pieces of information available for every article are the following:

- a unique identification code, i.e. *Digital Object Identifier* (DOI) of the article;
- date and year of publication;
- names of the authors;
- list of cited papers in the references that were published in APS journals;[8]
- authors' affiliations*;[9]
- topic/subject classification codes, i.e. *Physics and Astronomy Classification Scheme* (PACS)*.

Some describing statistics are shown in Section 2.9.

[7] The American Physical Society publishes various international research journals (currently 13), considered among the top journals in the field of physics.

[8] If a reference does not belong to an APS journal, it is not contained in this dataset.

[9] The fields denoted by an asterisk, *, are sometimes incomplete or missing.

2.3.1 Data Filtering

We want to keep track of publications and co-authors of every author, over the course of her entire career. In order to do so, we have to limit the set of authors to those who have published consistently over time and have had the possibility of having sufficiently long career. We implement these two filters by limiting the set of studied authors to those who:

- are in a **cohort** (i.e. the year of first publication) from 1957 to 1975. This is because before the 1950s there may be problems due to numerosity and after 1975 every author could not possibly have a career exceeding 40 years (since the last recorded year in the dataset is 2015);

- have an **activity rate** throughout their career of at least 20%, where the activity rate is defined as the ratio

$$\frac{\text{number of years with a publication}}{\text{number of years of career}},$$

  meaning that we require that every selected author has written at least one paper every 5 years.[10]

Since we are comparing authors that may have entered the "academic" system within a range of almost 20 years, we perform some checks to verify that they are enough homogeneous. This concern is especially true because science is a (exponentially) growing field (see figures in Section 2.9.1) which exhibits inflation[11] as well as institutional and conventional changes, so comparing authors from very different "eras" could be misleading.

In particular, Figures 2.10, 2.12, 2.16 and 2.25 show that across the cohorts 1957 to 1975 the considered distributions of career duration, activity rate and total paper produced are similar enough.[12]

2.4 Methods

Here we briefly describe how we compute the fundamental measures of this paper, detailed better in the following sections.

Authors' Productivity and Performance The concept of *productivity* is linked to the number of papers produced by each author. To control for the number of authors collaborating to each paper we compute the *per-capita number of papers* written by an author, i.e. where every paper is worth $\frac{1}{\text{no. of its authors}}$ instead of just 1.

[10]It may be important to recall that here we can only keep track of papers published in APS journals, so unfortunately we cannot exclude that some author has published somewhere else.

[11]Also, see [Pan et al., 2016].

[12]A proper survival analysis is left as future work.

The concept of *performance*, instead, although very related (and correlated[13]) to that of productivity, is based on the citations that (the papers written by) an author receives. Since the bare citation count has been proven to be an inappropriate measure[14], in order to control for different citation patterns that might exist across different *fields* and the different age the papers have, we renormalize a paper's citations by comparing them to across all other papers in similar fields and written in the same year.

Moreover, due to "citation inflation", we further have to inflation-adjust the citations (or *discount* them) by accounting for the year in which the citation is received: roughly speaking, a citation received in 1970 is worth more than one received in 1990 because, due to the increase in number of publications (and, hence, of citations), the probability of being cited in 1990 is higher than it was in 1970.

Authors' Exploration-Exploitation The notions of exploration and exploitation (also respectively interpretable as *diversification* and *specialization*) rely on the concept of *knowledge space*. Since here we use an author's co-authors as a proxy of her knowledge space[15], then it is assumed to evolve according to the collaborations she undertakes over time.

If these collaborations are with others that are "close" to one's network, then they are assumed to account for further exploitation/specialization. Conversely, if an author co-authors a paper with others that are relatively "distant" from her network, then this is seen as a sign of exploration/diversification. The notion of distance in this context will be measured by the embeddedness of an author in the network of her collaborators/co-authors.

2.5 Authors' Productivity

Let i be an *author* and let P_i^a (and P_i^t and P_i^y) be the *papers* written by i at academic *age* of a (and, respectively, at time t or in year y)[16], where $a \in \{0, \dots, \text{career}_i\}$. Defining an author's **cohort** as the year of her first appearance/publication, then the **academic age** of i in year y is accordingly defined as $a := y - \text{cohort}_i$.

The **productivity of i at age a** is defined as the per-capita number of papers written by i at age a, as follows:

$$\text{Prod}_i^a := \sum_{p \in P_i^a} \frac{1}{\text{no. authors of } p}.$$

Accordingly, by summing across the whole career of i, one obtains the **total**

[13]See Figure 2.28.

[14]See [Radicchi et al., 2008].

[15]Crucially, mapping the knowledge space also using the papers that an author cites over time is left as future work.

[16]Throughout we will try to keep this notation consistent.

productivity of i:

$$\mathrm{Prod}_i := \sum_{a=0}^{\mathrm{career}_i} \mathrm{Prod}_i^a.$$

2.6 Authors' Performance

2.6.1 Performance as Inflation-adjusted Citations

Let c_p^t be the **per-capita citations** that paper p receives at time t, i.e. the number of citations received by p from papers published exactly in year t divided by the number of p's authors[17]:

$$c_p^t := \frac{\text{no. citations to } p \text{ received at time } t}{\text{no. authors of } p}.$$

Let, also, C^t be the **total number of per-capita citations** produced in year t, i.e. the citations generated by all papers published in t:

$$C^t := \sum_p c_p^t,$$

where the summation goes across all papers p.

Since we want to compute an inflation-adjusted counting of citations, we take a *base year*, denoted by $t = \star$, so that $C^\star$ is the number of citations produced by all papers published in the base year. In practice, we will take 2015 as base year. The **current-value count of per-capita citations** of paper p in year y (with respect to the base year) is:

$$\widetilde{c}_p^y := \sum_{t \geq y} c_p^t \frac{C^\star}{C^t},$$

which, in words, means that we normalize all future citations that p receives from year y on.

Remark. Notice that when y is exactly p's publication year, then $\widetilde{c}_p^y$ is the *inflation-adjusted* ***total*** *number of per-capita citations* of p. This will be the basic unit of our analysis: by considering i's papers over time, one can reconstruct i's performance over time.

As done in the previous section, for every author i let P_i^a be the set of i's papers published when i's age is exactly a. Notice that P_i^a only contains papers published in year $y = \mathrm{cohort}_i + a$. The **inflation-adjusted per-capita citations** received by author i at age a are:

$$k_i^a := \sum_{p \in P_i^a} \widetilde{c}_p^{\,\mathrm{cohort}_i + a}.$$

In words, to compute k_i^a:

[17] With this definition, it is obvious that a paper cannot receive citations in years preceding its publication.

- first, consider all papers produced in year $y = \text{cohort}_i + a$;
- then, for every one of these papers, take the discounted per-capita citations that they will ever accumulate over future times.

Lastly, i**'s relative performance at age** a is defined by:[18]

$$K_i^a := \frac{k_i^a}{\langle k_j^a \rangle_{j:\, \text{cohort}_j = \text{cohort}_i}} \qquad \text{or} \qquad K_i^a := \frac{k_i^a - \min_j(k_j^a)}{\max_j(k_j^a) - \min_j(k_j^a)}, \\ \text{or} \qquad K_i^a := \frac{k_i^a - \langle k_j^a \rangle_{j:\, \text{cohort}_j = \text{cohort}_i}}{\sigma_{j:\, \text{cohort}_j = \text{cohort}_i}} \tag{2.1}$$

where the average $\langle \cdot \rangle$, $\min_j$, $\max_j$ and standard deviation σ are computed across all authors j who are in i's cohort.

2.7 Authors' Exploration and Exploitation

In this section we describe how to measure authors' exploration and exploitation. A fundamental assumption here is that the "amounts" of exploration and exploitation of every author (at every time) sum up to 1, meaning that allocating resources to one thing implies subtracting them to the other. With this assumption in place, once computed the exploitation of an author, then 1 minus that quantity will automatically give her exploration, and vice versa.

Remark. It is worth noticing that viewing exploration and exploitation as "substitutes" may not necessarily be the only paradigm present in the literature and way of thinking about this issue. It is not hard to persuade oneself that the two "dimensions" may be rather independent and, then, could also be viewed as "complements". In this respect, see [Uzzi et al., 2013], where "novel" and "conventional" papers/ideas are treated as complementary and independent, both contributing together to yielding a higher performance.

This section is organized as follows:

- first, we introduce a "naïve method" to compute exploration and exploitation. Intuitively, fix a given author, then we simply keep track of her past co-authors: every time a paper with some co-authors is produced, a co-author will contribute to increasing the given author's exploitation if they both had already collaborated previously (in other papers). If, instead, it is the first time they are co-authoring a paper together, then the author's exploration will increase.[19]

[18] The various alternative definitions should be evaluated carefully especially to try to assess what impact they have on the following analysis.

[19] Both of them will experience a symmetric increase in exploration or exploitation. In the following "weighted", non-naïve versions of this measurement, this increase is not necessarily symmetric, depending on how many co-authors each of them have.

- Then, we extend the previous idea by accounting not only for authors who already know each other, but also by considering how establishing a link between two authors is relatively "hard" and "unlikely": if two authors are "distant", in terms of how many common co-authors they both know, then a collaboration between them is more unlikely, than if they both knew almost only the same people. In this context, the extent to which two authors are *distant* is measured by the embeddedness of an author's neighbors into the other's.

An alternative extension, which is common to both the above-mentioned approaches, involves considering co-authors as "already known at time t" only if there has been a collaboration in the previous τ years, with τ being typically 3, 5 or 10 years. In other words, at every time, one has to keep track only of co-authors in the previous *rolling time-window*.

2.7.1 Keeping Track of One's Past Co-authors

Here we describe the naïve method we use to compute an exploration-exploitation ratio: we keep track of one's co-authors over time and define the exploitation ratio as the fraction of times one has co-authored with *already-known* authors across her total number of collaborations. Accordingly, the exploration ratio will then be its complement to 1. But first we need to clarify the notation.

Co-authorship Networks: Notation

Let $V = \{1, \ldots, n\}$ be the set of all authors (vertices). Let A_i^t be the set of all i's co-authors (neighbors) at time t, including i herself. A link/edge $ij \in E^t$ between two authors i and j is present if and only if i and j have co-authored a paper in year t, i.e. $i \in A_j^t$ and $j \in A_i^t$. The pairs $\{(V, E^t)\}_t$ then form a sequence of graphs/networks, where the links are evolving over time.

Weighing Collaboration Intensity: weighted graphs Let i and j be two co-authors of paper p, published at time t. Then we define the **weight** of link ij at time t as:

$$w_{ij}^t = \sum_{p \in P_i^t \cap P_j^t} \frac{1}{\text{no. authors of } p},$$

where, with the same notation of the previous section, P_i^t denotes the set of papers written by i in year t. Intuitively, the weight w_{ij}^t corresponds to the **per-capita collaborations** that i and j have in common at time t: the more co-authors a paper have, the weaker the relationships among its co-authors are. Accordingly, let $\widetilde{E}^t = \{w_{ij} : i, j \in V\}$ be the set of weighted links. Then the pairs $\{(V, \widetilde{E}^t)\}_t$ form a sequence of weighted graphs.

Naïve Measure of Exploration-Exploitation

With the previous notation, let us further define:

$$\Delta_i^t := \bigcup_{s \leq t} A_i^s \qquad \text{and} \qquad \Delta_i^{\tau,t} := \bigcup_{s=\max(0,t-\tau)}^{t} A_i^s,$$

for a fixed τ, that can be usually though of as being 3, 5 or 10 years. These respectively represent all co-authors that i has ever had up to time t and those who have been collaborators in the previous τ years.

Then, we respectively define the **exploitation** of author i at time t and the exploration of i at time t in the **τ-rolling window** as:

$$e_i^t := \frac{|\Delta_i^{t-1} \cap A_i^t|}{|A_i^t|} \qquad \text{and} \qquad e_i^{\tau,t} := \frac{|\Delta_i^{\tau,t-1} \cap A_i^t|}{|A_i^t|}.$$

Notice that[20] $e_i^t \in (0,1]$. Correspondingly, the **exploration** of author i at time t is $(1 - e_i^t)$.

Naïve Measure on Weighted Networks Here we extend the previous notion to the case of a weighted co-authorship network. In particular, in the case of weighted graph we define:

$$\widetilde{e}_i^t := \frac{\displaystyle\sum_{s \leq t-1} \left(\sum_{j \in A_i^s \cap A_i^t} w_{ij}^s \right)}{\displaystyle\sum_{j \in A_i^t} w_{ij}^t} \qquad \text{and} \qquad \widetilde{e}_i^{\tau,t} := \frac{\displaystyle\sum_{s=\max(0,t-\tau)}^{t-1} \left(\sum_{j \in A_i^s \cap A_i^t} w_{ij}^s \right)}{\displaystyle\sum_{j \in A_i^t} w_{ij}^t}.$$

Notice that in case of a unweighted network, i.e. where $w_{ij}^s \in \{0,1\}$, then this definitions corresponds to the previous ones of e_i^t and $e_i^{\tau,t}$, with the main difference that in the unweighted case, if i and j have had many collaborations in the previous (τ) years, then they will always only count as a single collaboration.

2.7.2 Embeddedness-based Exploitation and Exploration

As done so far, let us consider a weighted network, where scientists are the nodes and where links weight w_{ij}, between two authors i and j, represents the strength of their co-authorship relation, that is, how many times they have co-authored a paper over the total number of co-authors they had in those same collaborations. On these kind of weighted graphs we want to define a measure of embeddedness which will ultimately represent our measure of exploitation.

[20] Recall that by definition A_i^t always contains i itself. In particular, if i has never had co-authors, then $e_i^t = 1$.

Embeddedness on Weighted Graphs

Let G be a symmetric weighted graph, $w_{ij} \equiv w_{ji} \in (0,1]$ denote the weight of the link ij between the nodes i and j and, lastly, let N_i denote the set of i's neighbors (excluding i itself). With a notation consistent with the one introduced in the previous section, the authors will be labeled with numbers so that $i \in V = \{1, \ldots, n\}$, where n is the total number of authors.

We, then, define two measures of **embeddedness** of j in i's network, as follows:[21]

$$\begin{aligned} \mu_i(j) &:= \frac{\sum_{k \in N_i \cap N_j} w_{ik}}{\sum_{h \in N_i} w_{ih}} =: \frac{NUM_i(j)}{DEN_i}, \\ \nu_i(j) &:= \frac{\sum_{k \in N_i \cap N_j} w_{jk}}{\sum_{h \in N_j} w_{jh}} =: \frac{NUM_j(i)}{DEN_j}. \end{aligned} \tag{2.2}$$

Then we consider their weighted averages across i's neighbors:

$$\begin{aligned} \mathcal{M}_i &:= \sum_{j \in N_i} \left[\left(\frac{w_{ij}}{\sum_{k \in N_i} w_{ik}} \right) \mu_i(j) \right] \equiv \sum_{j \in N_i} \frac{w_{ij}}{DEN_i} \cdot \mu_i(j), \\ \mathcal{N}_i &:= \sum_{j \in N_i} \left[\left(\frac{w_{ij}}{\sum_{k \in N_i} w_{ik}} \right) \nu_i(j) \right] \equiv \sum_{j \in N_i} \frac{w_{ij}}{DEN_i} \cdot \nu_i(j) \end{aligned} \tag{2.3}$$

Remark. The above-defined notions of embeddedness are slightly different but turn out to be highly correlated, as shown in Figure 2.29.[22]

Embeddedness in Matricial Form The previous formulas can be written in matricial form, which is (computationally) very convenient. Let $W = (w_{ij})_{i,j \in G}$ be the (symmetric) weighted adjacency matrix and let $A = (a_{ij})_{i,j}$ be its binary version, i.e. where $a_{ij} = 1$ if and only if $w_{ij} > 0$ and $a_{ij} = 0$ whenever $w_{ij} = 0$. As a convention, let us consider every vector[23] as a column, so that a matrix X can be written as sequence of its columns:

$$X = (\mathbf{X}_1 | \ldots | \mathbf{X}_n) .$$

As a further piece of notation, let $\widetilde{W}$ denote the column-stochastic version of W, that is,

$$\widetilde{W} = \left(\widetilde{\mathbf{W}}_1 | \ldots | \widetilde{\mathbf{W}}_n \right) := \left(\frac{\mathbf{W}_1}{DEN_1} \middle| \ldots \middle| \frac{\mathbf{W}_n}{DEN_n} \right),$$

which satisfies the property that the sum over each column is always equal to 1. Now, recall that

$$\mathcal{M}_i = \sum_{j \in N_i} \frac{w_{ij}}{DEN_i} \frac{NUM_i(j)}{DEN_i}.$$

[21] Notice that $\nu_j(i) = \mu_j(i)$.

[22] Their use and interpretation are still to be determined, which is why here we use both.

[23] For clarity, vectors are in bold, while matrices in normal font.

Then, it can be written as scalar product[24] between the two following vectors:

- of the vector[25] $\left(\frac{w_{ij}}{DEN_i}\right)_{j=1,\ldots,n} = \left(\frac{w_{ij}}{DEN_i}\right)_{j\in N_i} = \widetilde{\mathbf{W}}_i$, which is the i-th column of $\widetilde{W}$;
- and of the vector[26]

$$\left(\frac{NUM_i(j)}{DEN_i}\right)_{j\in N_i} = \left(\frac{\sum_{k\in N_i\cap N_j} w_{ik}}{DEN_i}\right)_j = \left(\frac{\mathbf{W}_i\cdot\mathbf{A}_j}{DEN_i}\right)_j = \\ = \left(\widetilde{\mathbf{W}}_i\cdot A_j\right)_j = A^T\widetilde{\mathbf{W}}_i = A\widetilde{\mathbf{W}}_i$$

Summing up:[27]

$$\mathcal{M}_i = \widetilde{\mathbf{W}}_i\cdot A\,\widetilde{\mathbf{W}}_i = \widetilde{\mathbf{W}}_i^T A\,\widetilde{\mathbf{W}}_i.$$

With an analogous reasoning,

$$\mathcal{N}_i = \sum_{j\in N_i}\frac{w_{ij}}{DEN_i}\frac{NUM_j(i)}{DEN_j}$$

can be written as scalar product of:

- the column $\widetilde{\mathbf{W}}_i$, exactly as above;
- and the vector

$$\left(\frac{NUM_j(i)}{DEN_j}\right)_j = \left(\frac{\sum_{k\in N_i\cap N_j} w_{jk}}{DEN_j}\right)_j = \left(\frac{\mathbf{W}_j\cdot\mathbf{A}_i}{DEN_j}\right)_j = \left(\widetilde{\mathbf{W}}_j\cdot\mathbf{A}_i\right)_j = \widetilde{W}^T\mathbf{A}_i,$$

so that

$$\mathcal{N}_i = \widetilde{\mathbf{W}}_i\cdot\widetilde{W}^T\mathbf{A}_i = \widetilde{\mathbf{W}}_i^T\,\widetilde{W}^T\mathbf{A}_i.$$

From the previous formulas, the vectors of all nodes' embeddedness, $\mathcal{M} = (\mathcal{M}_i)_i$ and $\mathcal{N} = (\mathcal{N}_i)_i$, can be respectively obtained as the diagonal of the matrices

$$\mathcal{M} = \operatorname{diag}\left(\widetilde{W}^T A\,\widetilde{W}\right) \qquad \text{and} \qquad \mathcal{N} = \operatorname{diag}\left(\widetilde{W}^T\,\widetilde{W}^T A\right).$$

[24] Given two vectors $\mathbf{a} = (a_1,\ldots,a_n)$ and $\mathbf{b} = (b_1,\ldots,b_n)$, their scalar product is $\mathbf{a}\cdot\mathbf{b} = \sum_j a_i b_i$.

[25] By construction, $w_{ij} = 0$ iff $j \notin N_i$.

[26] The last pieces are a matrix-vector multiplication, and A^T indicates the transpose of A, which are equal by symmetry.

[27] Here, the last product is the matrix row-column multiplication.

Embeddedness in Matricial Form: Advanced Optimization Crucially, in our case the matrices A, W and $\widetilde{W}$ are large but **sparse**: indeed, since $n \approx 300,000$, then their dimension is almost $300,000 \times 300,000$, but every author has only a small number of collaborators compared to the total number of authors.

Remark. Even if we are interested in a selected and small sample of authors, it may not be true that their co-authors' are also in that sample. This is important since, to compute the embeddedness, one needs to know every co-author's neighboring network. In our case, even if the sample of selected authors ("born" in cohorts from 1957 to 1975, who have at least 20 years of career and 20% of activity rate) consists of around 30,000 authors, however their co-authors are a *giant component* and basically cover the entire network.

Writing $\mathcal{M}$ as the diagonal of $\widetilde{W}^T A \widetilde{W}$ is a convenient representation, but it is not as convenient in terms of computation. Indeed, the sparsity of the matrix $\widetilde{W}^T A \widetilde{W}$ is similar to that of A^3. Although A is sparse, A^3 is unlikely to be sparse because it contains information about the nodes connected by 3-step paths in the network. In a random network of $300,000$ nodes and same density of the one we have, the average minimum path length is around 4-5 steps[28], meaning that it is very likely that 2 randomly chosen nodes have a path between them of length 3, which implies that A^3 is not likely to be sparse. An analogous reasoning holds for $\mathcal{N}$.

On the contrary, computing $\mathcal{M}_i = \widetilde{\mathbf{W}}_i^T A \widetilde{\mathbf{W}}_i$ for every author i, implies doing this computation for every author of our interest[29] and does not necessarily "waste" the sparsity of A. The same holds for $\mathcal{N}_i = \widetilde{\mathbf{W}}_i^T \widetilde{W}^T \mathbf{A}_i$.

It turns out that $\mathcal{M}_i$ and $\mathcal{N}_i$ can also be computed by doing element-wise multiplications of matrices[30], which is more efficient because it is a sparsity-preserving operation. In particular:

$$\mathcal{M}_i = \sum_{j=1,\dots,n} \left((\widetilde{W}^T A) \circ \widetilde{W}^T \right)_{ij},$$
$$\mathcal{N}_i = \sum_{j=1,\dots,n} \left(\widetilde{W}^T \circ (A \widetilde{W}) \right)_{ij}.$$

This representation is computationally more advantageous, because it contains only "squares" of sparse matrices, i.e. $\widetilde{W}^T A$ and $A \widetilde{W}$, while the element-wise multiplication preserves the sparsity.
To conclude, in order to compute $\mathcal{M}$, one just has to:

- multiply two sparse matrices $\widetilde{W}^T A$ and element-wise multiply $(\widetilde{W}^T A) \circ \widetilde{W}^T$;

[28] That is, of the order of $\log(n)$.

[29] In our case, the authors of interest are those in cohorts 1957-1975 with a long career and consistent pattern of publications, which reduces their number to around $30,000$ authors, i.e. a tenth of all authors.

[30] Given two $(n \times n)$-matrices $A = (a_{ij})_{i,j}$ and $B = (b_{ij})_{i,j}$, the **element-wise multiplication** (a.k.a. Hadamard product or Schur product) is defined as the matrix $A \circ B = (a_{ij} b_{ij})_{i,j}$.

- then, for every row/author i, take the sum of the elements of $(\widetilde{W}^T A) \circ \widetilde{W}^T$ on row i, which yields $\mathcal{M}_i$.

Analogously, every $\mathcal{N}_i$ is computed as the sum of all the elements of $\widetilde{W}^T \circ (A\,\widetilde{W})$ on the i-th row.

Exploitation as Embeddedness

According to our interpretation, the embeddedness $\mu_i(j)$ (or $\nu_i(j)$) of j in i's network gives a measure of how "exploitative" is their relationship: the higher the embeddedness $\mu_i(j)$, the closer j's contacts are to i's.

So far we have defined the embeddedness on a static weighted graph. Here we want to incorporate also the dynamic structure that is given by the evolution over time of the co-authorship network. With the notation introduced in the previous section, let $G^t = (V, \widetilde{E}^t)$ be the weighted network of (per-capita) collaborations at time t, where $V = \{1, \dots, n\}$ are the nodes and $\widetilde{E}^t = \{w^t_{ij}\}_{i,j \in V}$ are the weighted links.[31]

By juxtaposing the various time layers G^t we can keep track of the evolution of the system over time. In particular, let us define the "summation" of links as

$$\omega^t_{ij} := \sum_{s \le t} w^s_{ij} \qquad \text{and} \qquad \omega^{\tau,t}_{ij} := \sum_{s=\max(0,t-\tau)}^{t} w^s_{ij}.$$

Then, let us define the graphs

$$\Gamma^t := (V, \Omega^t) \qquad \text{and} \qquad \Gamma^{\tau,t} := (V, \Omega^{\tau,t}),$$

where $\Omega^t = \{\omega^t_{ij}\}_{i,j \in V}$ and $\Omega^{\tau,t} = \{\omega^{\tau,t}_{ij}\}_{i,j\in V}$, respectively.

With this notation in place, then we define the **exploitation of author** i **at time** t as her average embeddedness $\mathcal{M}^t_i$ (or $\mathcal{N}^t_i$) in the graph Γ^t, as defined in Equation 2.3. Correspondingly, her exploitation in rolling time-windows is the average embeddedness of i in the graph $\Gamma^{\tau,t}$.

Remark. As done with performance in Equation 2.1, since our goal is to compare authors from different cohorts and at different (academic) ages, it is reasonable to define and analyze the **relative exploitation** of author i at age a, defined[32] by[33]

$$m^a_i := \frac{\mathcal{M}^a_i}{\langle \mathcal{M}^a_j \rangle_{j:\,\mathrm{cohort}_j = \mathrm{cohort}_i}} \quad \text{or} \quad m^a_i := \frac{m^a_i - \min_j(\mathcal{M}^a_j)}{\max_j(\mathcal{M}^a_j) - \min_j(\mathcal{M}^a_j)}$$
$$\text{or} \qquad m^a_i := \frac{\mathcal{M}^a_i - \langle \mathcal{M}^a_j \rangle_{j:\,\mathrm{cohort}_j = \mathrm{cohort}_i}}{\sigma_{j:\,\mathrm{cohort}_j = \mathrm{cohort}_i}} \tag{2.4}$$

[31] According to this definition, isolated nodes may correspond either to authors who have no publication in year t or to authors who are "born" (i.e. whose cohort is) after year t.

[32] Similarly to the previous cases, also here the various definitions should be analyzed carefully and their impact on the successive analysis is still to be evaluated.

[33] With a notation used already previously, for any quantity x and every author i, $x^t = x^a$ when $t = \mathrm{cohort}_i + a$.

In this way, i's exploitation is comparable across all ages and all cohorts, because it has been renormalized and compared to that one of all other authors in the same cohort and at the same age.

2.8 Econometric Analysis

2.8.1 Justification & Discussion

Figures 2.1, 2.2 and 2.3 show on the horizontal axis the exploration-exploitation ratio and on the vertical axis the performance (in logarithmic scale). Each point is a pair $(\mathrm{Perf}_{i,a}, \mathrm{Expl}_{i,a})$, i.e. performance and exploitation ratio of author i at academic age a. Younger ages in red-orange colors, while older ones in yellow-green.

Both performance and exploitation ratio are computed in *relative* terms, i.e. comparing every author with others of her cohort and at the same age level. The only difference with Equations 2.1 and 2.4 is that here both quantities are further averaged over 5-year windows of age, meaning that the quantity plotted for age 5-10 for author i, say, is the average of that quantity obtained in those years.

Holding fixed a given academic age, the points seem to be placed in a bell-shaped cloud where, in particular, the peaks of performance correspond to a certain (optimal?) range in the exploitation ratio. The following econometric analysis is aimed at capturing a linear positive trend and, possibly, an inverted-U shaped corresponding to those peaks. An analogous effect is also present when plotting authors' productivity versus exploitation (see Figure 2.30.)

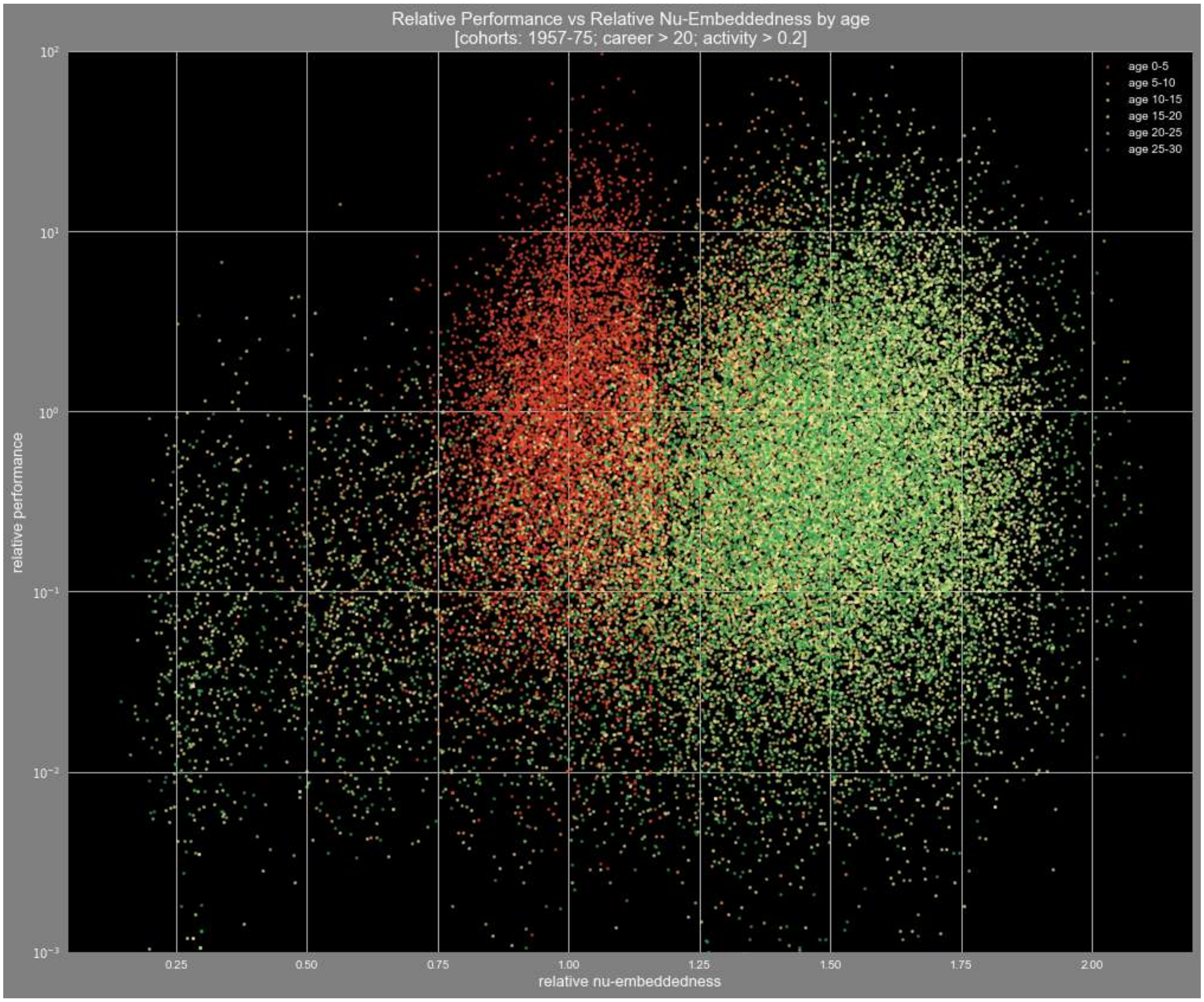


Figure 2.1: **Performance vs. Exploitation**

2.8.2 Panel with Fixed Effects

The econometric model is as follows:

$$\log \mathrm{Perf}_{i,a} = \beta_1 \cdot \mathrm{Expl}_{i,a} + \beta_2 \cdot \left(\mathrm{Expl}_{i,a}\right)^2 + \alpha_i + \varepsilon_{i,a}, \tag{2.5}$$

where $\log \mathrm{Perf}_{i,a}$ is the logarithm of (relative) performance of author i at (academic) age a, $\mathrm{Expl}_{i,a}$ is her (relative) exploitation, α_i is the fixed effect of individual i and $\varepsilon_{i,a}$ are the error terms.

The idea behind the model is that:

- the author's fixed effect reflects her innate ability;
- the goal of the quadratic form (in exploration-exploitation) is to capture a peak in the performance corresponding to an appropriate choice of the exploration-exploitation ratio.

A slight refinement of the model can be considered by controlling for age-specific dummies D_a:

$$\log \mathrm{Perf}_{i,a} = \beta_1 \cdot \mathrm{Expl}_{i,a} + \beta_2 \cdot \left(\mathrm{Expl}_{i,a}\right)^2 + \sum_{a=0}^{\text{max age}} \gamma_a \cdot D_a + \alpha_i + \varepsilon_{i,a}. \quad (2.6)$$

The results for the above regressions in Table 2.1 show that exploitation affects positively performance and, most of all, exploitation squared is also significant and negative.[34]

As a first check, the *Breusch-Pagan Lagrange-multiplier test* for random effects gives a p-value below 0.001, meaning that an OLS model would not be appropriate and, hence, supporting the use of a proper panel-data model. Additionally, the *Hausman test* for fixed-versus-random effects also yields a p-value below 0.001, which assures that the correct model to use is the fixed effects model, since the coefficients differ significantly from those obtained with random effects.

2.8.3 Dynamic Panel

The econometric model with l lags on the dependent variable is as follows:

$$\begin{aligned} \log \mathrm{Perf}_{i,a} = \gamma_1 \cdot \log \mathrm{Perf}_{i,a-1} + \ldots + \gamma_l \cdot \log \mathrm{Perf}_{i,a-l} + \\ + \beta_1 \cdot \mathrm{Expl}_{i,a} + \beta_2 \cdot \left(\mathrm{Expl}_{i,a}\right)^2 + \alpha_i + \varepsilon_{i,a}, \end{aligned}$$

where $\log \mathrm{Perf}_{i,a}$ is the logarithm of (relative) performance of author i at (academic) age a, $\mathrm{Expl}_{i,a}$ is her (relative) exploration-exploitation ratio, α_i is the fixed effect of individual i and $\varepsilon_{i,a}$ are the error terms.

With respect to the previous section, the addition of past performances as regressors is meant to capture their influence on the current performance with the idea that a scientist's reputation is somehow sticky and a good past performance can enable access to more resources.

To a practical level, it is perhaps worth recalling that both performance and exploration-exploitation ratio are computed in "relative terms" with respect to other authors in the same cohort and at the same age. Moreover, the time/age a is in 5-year time windows. In particular, a lag of 1-2 windows will cover the

[34]Recall that Age is in 5-year time windows, so Age $= a$ means being from $5a$ to $5(a+1)$ academic years of age.

Table 2.1: **Fixed-effects (within) Regression**
Robust Std. Err. for clustering on Author, dependent variable: **LogPerf**

	Fixed Effects	Fixed Effects & Ages
Embed	3.822***	2.516***
	(0.000)	(0.000)
Embed^2	-1.511***	-0.680***
	(0.000)	(0.000)
Age=0		0
		(.)
Age=1		-0.798***
		(0.000)
Age=2		-1.019***
		(0.000)
Age=3		-1.157***
		(0.000)
Age=4		-1.228***
		(0.000)
Age=5		-1.261***
		(0.000)
Age=6		-1.237***
		(0.000)
Age=7		-1.267***
		(0.000)
Age=8		-1.263***
		(0.000)
Age=9		-1.276***
		(0.000)
Age=10		-1.198***
		(0.000)
Age=11		-1.105***
		(0.000)
Constant	-3.254***	-2.033***
	(0.000)	(0.000)
Observations	58512	58512
R^2	0.027	0.113
Adjusted R^2	0.027	0.113
R^2 between	0.051	0.057
R^2 overall	0.031	0.083
ρ	0.416	0.436

p-values in parentheses
* $p < 0.1$, ** $p < 0.01$, *** $p < 0.001$
ρ is the fraction of variance in the error due to individual-specific effects

previous 5-10 years, so we deem that having 1 or 2 lags be an appropriate choice, so $l = 1, 2$.

Since it is a dynamic panel with fixed effects and lags on the dependent variable, we adopt the *Arellano-Bond estimator*, which is based on the crucial assumption that the errors $\varepsilon_{i,a}$ are serially uncorrelated and, broadly speaking, uses variables observed in previous periods as instruments. One further thing to notice is that the panel is *unbalanced*, because an author's performance in a certain year is present only if she has written at least one paper in that year and if that paper has received at least one citation (at any time).

Table 2.2 and 2.3 show, respectively, the results for the 2SLS (or 1-step) estimator and the optimal (or 2-step GMM) estimator. The latter is useful to obtain a more efficient estimation, since the model is overidentified.

Table 2.4 shows the *Arellano-Bond test*, to check whether the errors are serially uncorrelated, is always passed, independently of the number of lags used as regressors (L) and the maximum number of lags used as instruments (ML). Indeed, as expected, the first order differences are correlated, while for higher orders there is no correlation.

Lastly, it is also possible to implement a test for the validity of overidentifying instruments (so-called *Hansen-Sargan test*). Notice, though, that it cannot be performed with robust standard errors, which makes it a much less relevant test in the context of a dynamic panel. The test is, in particular, useful to determine the validity of the additional lags used as instruments: under the null hypothesis that all instruments are valid, rejection is interpreted as indicating that at least one of them is not a valid instrument. To the extent to which the test is not very reliable, due to the absence of robust standard errors, Table 2.5 suggests that the estimations with $l = 2$ lags should be preferred.

Table 2.2: **Arellano-Bond dynamic panel-data estimation**
1-Step Results, Robust Std. Err. adjusted for clustering on Author, dependent variable: **LogPerf**

	L(1)ML(1)	L(1)ML(2)	L(1)ML(3)	L(2)ML(1)	L(2)ML(2)	L(2)ML(3)
L.LogPerf	0.137***	0.142***	0.145***	0.208**	0.180***	0.180***
	(0.000)	(0.000)	(0.000)	(0.004)	(0.000)	(0.000)
L2.LogPerf				0.0551	0.0322***	0.0319***
				(0.358)	(0.001)	(0.001)
Embed	2.333***	2.378***	2.407***	2.121***	2.199***	2.215***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Embed^2	-0.578***	-0.594***	-0.604***	-0.320*	-0.350*	-0.355*
	(0.000)	(0.000)	(0.000)	(0.050)	(0.017)	(0.015)
Constant	-2.926***	-2.952***	-2.972***	-3.025***	-3.107***	-3.120***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	32686	32686	32686	24314	24314	24314

p-values in parentheses
* $p < 0.1$, ** $p < 0.01$, *** $p < 0.001$
L = lags of dependent variable used as regressor
ML = max number of lags of dependent variable allowed as instruments

Table 2.3: **Arellano-Bond dynamic panel-data estimation**
2-Step Results, WC-Robust Std. Err. adjusted for clustering on Author, dependent variable: **LogPerf**

	L(1)ML(1)	L(1)ML(2)	L(1)ML(3)	L(2)ML(1)	L(2)ML(2)	L(2)ML(3)
L.LogPerf	0.139***	0.141***	0.143***	0.217**	0.183***	0.183***
	(0.000)	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)
L2.LogPerf				0.0603	0.0336***	0.0333***
				(0.316)	(0.000)	(0.000)
Embed	2.326***	2.344***	2.370***	2.104***	2.202***	2.197***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Embed^2	-0.575***	-0.580***	-0.588***	-0.312*	-0.347*	-0.346*
	(0.000)	(0.000)	(0.000)	(0.056)	(0.017)	(0.018)
Constant	-2.920***	-2.933***	-2.950***	-3.006***	-3.111***	-3.105***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	32686	32686	32686	24314	24314	24314

p-values in parentheses
* $p < 0.1$, ** $p < 0.01$, *** $p < 0.001$
L = lags of dependent variable used as regressor
ML = max number of lags of dependent variable allowed as instruments

Table 2.4: **Arellano-Bond test for zero autocorrelation in first-differenced errors**

	1-Step					
	L(1)ML(1)	L(1)ML(2)	L(1)ML(3)	L(2)ML(1)	L(2)ML(2)	L(2)ML(3)
order 1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
order 2	0.2311	0.1897	0.1589	0.7423	0.9338	0.9242
order 3	0.2866	0.2964	0.3035	0.6131	0.1738	0.1843
	2-Step					
	L(1)ML(1)	L(1)ML(2)	L(1)ML(3)	L(2)ML(1)	L(2)ML(2)	L(2)ML(3)
order 1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
order 2	0.2124	0.1867	0.1709	0.6926	0.9780	0.9560
order 3	0.2895	0.2947	0.2979	0.6592	0.1871	0.1909

H_0 null hypothesis: no autocorrelation
L = lags of dependent variable used as regressor
ML = max number of lags of dependent variable allowed as instruments

Table 2.5: **Sargan test of overidentifying restrictions**

	1-Step					
	L(1)ML(1)	L(1)ML(2)	L(1)ML(3)	L(2)ML(1)	L(2)ML(2)	L(2)ML(3)
p-value	0.0551	0.0069	0.0011	0.2249	0.0824	0.0777
	2-Step					
	L(1)ML(1)	L(1)ML(2)	L(1)ML(3)	L(2)ML(1)	L(2)ML(2)	L(2)ML(3)
p-value	0.0813	0.0132	0.0049	0.2778	0.1263	0.1426

H_0 null hypothesis: overidentifying restrictions are valid

L = lags of dependent variable used as regressor

ML = max number of lags of dependent variable allowed as instruments

2.9 Figures

2.9.1 General Descriptive Statistics

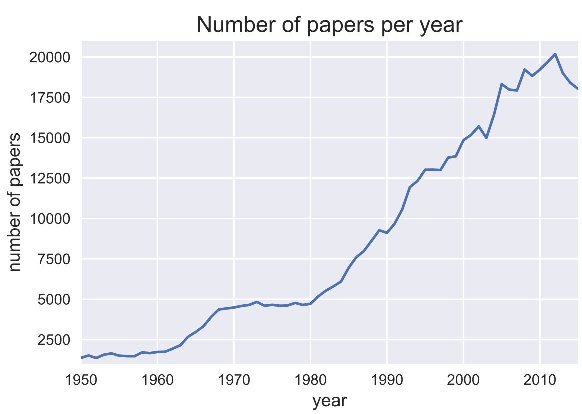


Figure 2.4: **Number of papers in each year**

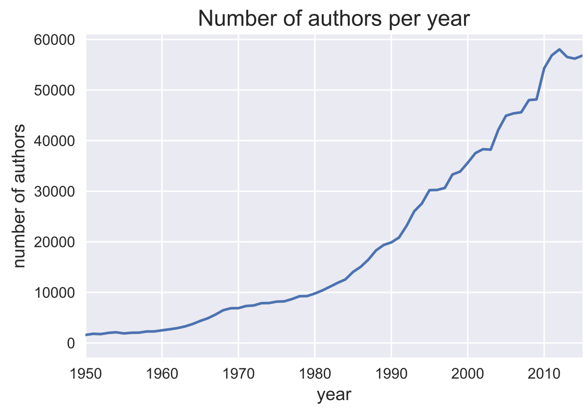


Figure 2.5: **Number of authors present in each year**

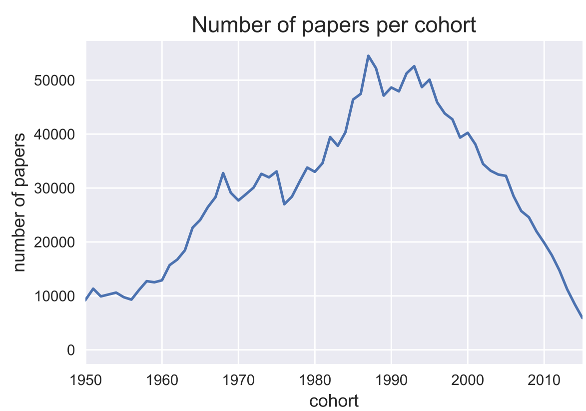


Figure 2.6: **Number of papers written by each cohort**

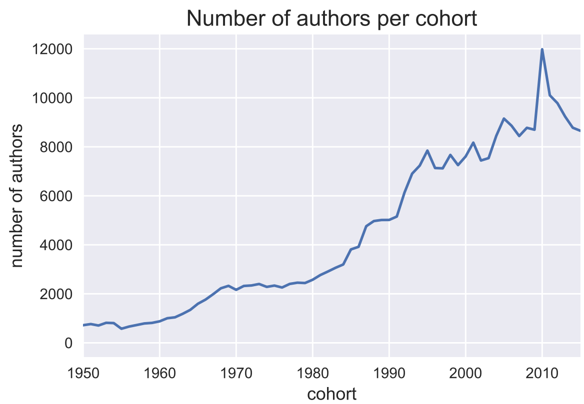


Figure 2.7: **Number of authors present in each cohort**

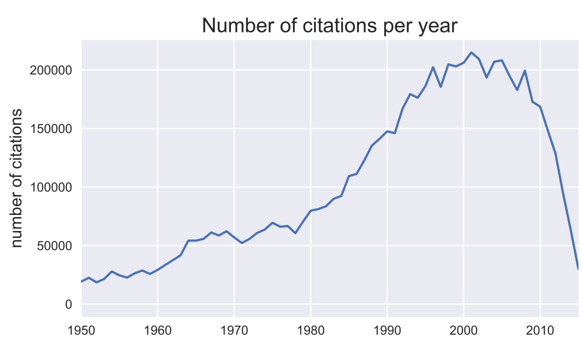


Figure 2.8: **Number of citations produced each year**

,

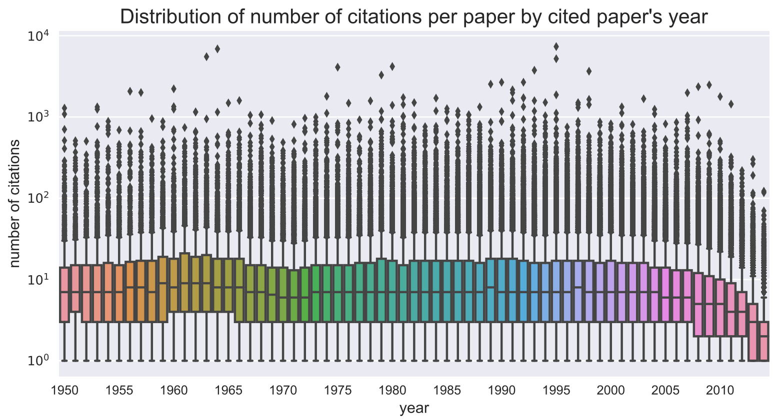


Figure 2.9: **Number of per-paper citations by cited paper's year**

2.9.2 Descriptive Statistics on Career Duration

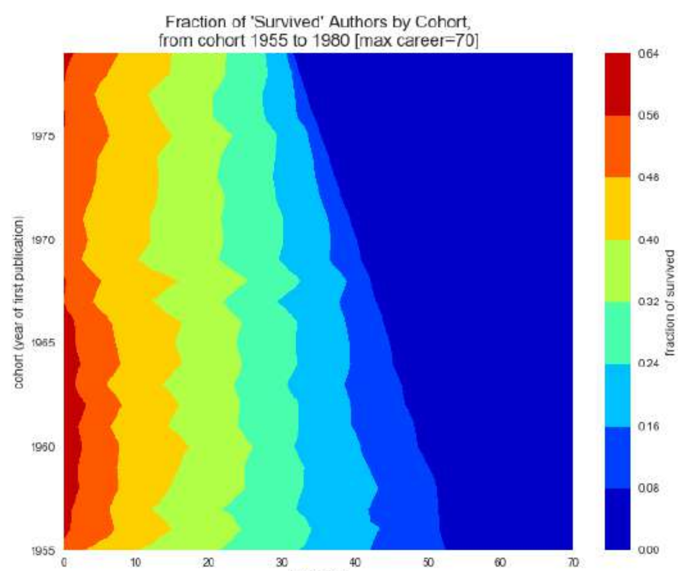


Figure 2.10: **Fraction of "survived" authors by cohort** The fraction of authors that "survive" after a certain career age, by cohort. The plot shows that the duration of the careers remains roughly regular across different cohorts. Notice that authors of cohort 1975 cannot have a career length exceeding 40 years (data until 2015), which explains why the dark blue area is wider for younger cohorts.

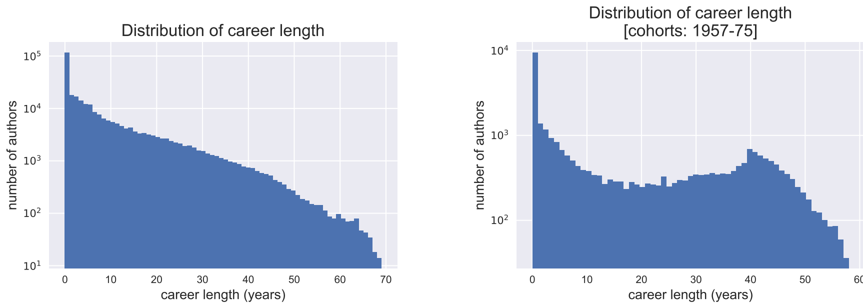
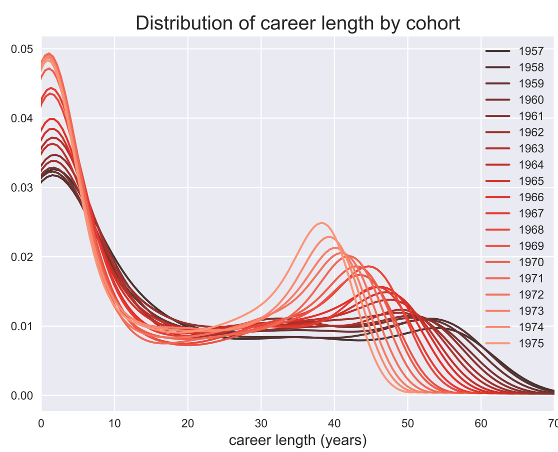


Figure 2.11: **Distribution of Career Length** On the left, the career length distribution of all authors (in log-scale). Many authors have short careers (many have just 1 publication, i.e. 1 year of career). On the right, the distribution for authors in cohorts 1957-1975. These authors can be roughly separated in two different classes: those with short careers (less than 15-20 years) and those with longer careers. The peak at 40 years is due to censoring as explained in Figure 2.12.

Figure 2.12: **Comparison of Career Length Distributions for Cohorts 1957-1975** Due to censoring in year 2015 (last registered year) authors of younger cohorts (brighter reds) cannot possibly have careers exceeding 40 years, which is why the distribution peaks earlier. (NB: the poor form of the curves is due to the numerical approximation done by the computer.)



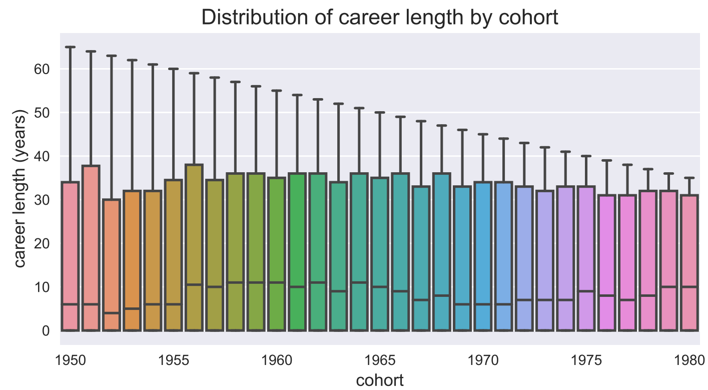


Figure 2.13: **Distribution of career length by cohort**

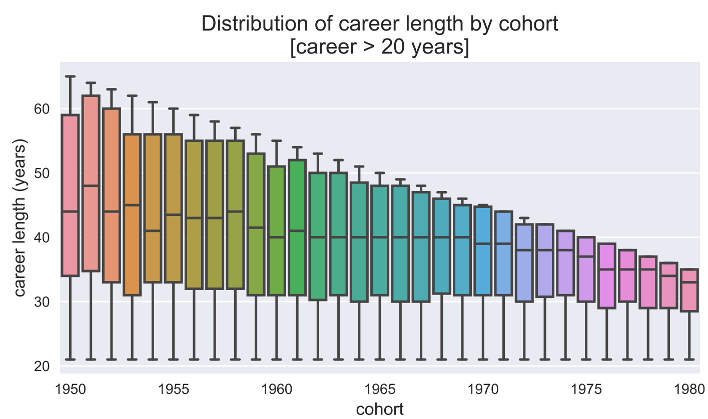


Figure 2.14: **Distribution of career length by cohort** conditional on having a career of at least 20 years. The decreasing trend is, of course, due to the censoring of the data in year 2015.

2.9.3 Descriptive Statistics on Activity Rate

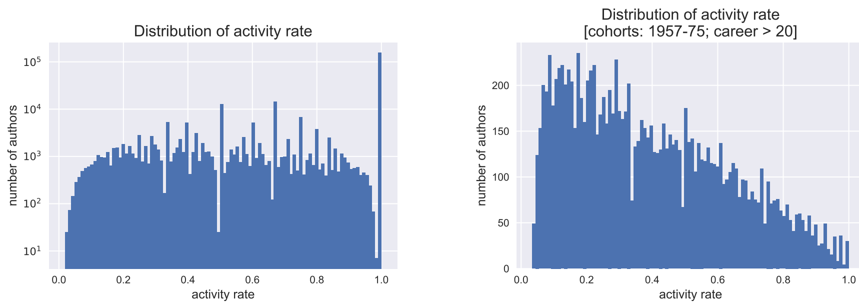


Figure 2.15: **Distribution of activity rate** On the left, the career length distribution of all authors (in log-scale). All authors with only 1 publication have, by construction, an activity rate of 100%. On the right, the distribution for authors in cohorts 1957-1975 with at least 20 years of career. Recall that activity rate $= \frac{\text{no. years with a publication}}{\text{career length}}$.

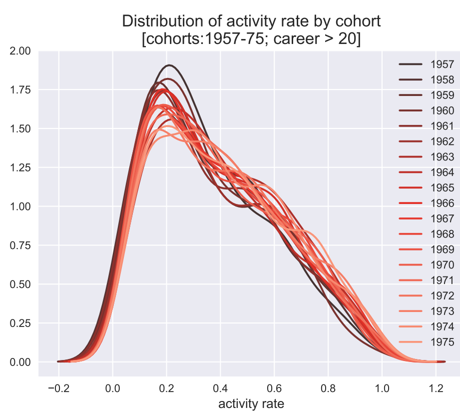


Figure 2.16: **Comparison of Activity Rate Distributions** for authors of cohorts 1957-1975 and at least 20 years of career.

2.9.4 Descriptive Statistics on Number of Authors per Paper

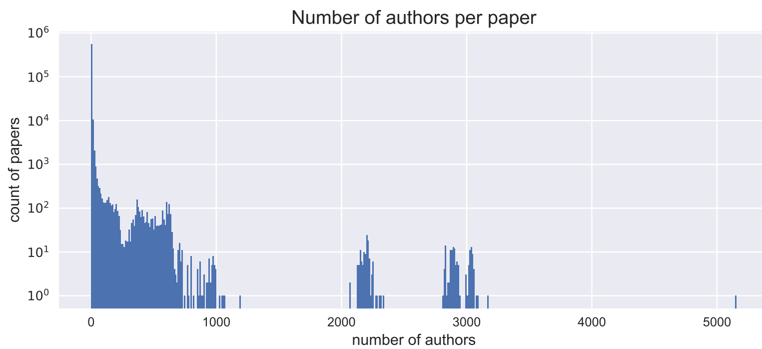


Figure 2.17: **Distribution of number of authors per paper** Hundreds of papers (mainly from experimental fields or from the field of high energy physics) have hundreds or thousands of authors. This is to be taken into account when evaluating an author's performance: the bare citation count cannot be used.

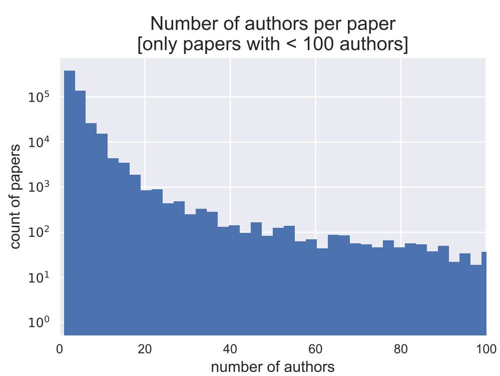


Figure 2.18: **Distribution of number of authors per paper** (zoom-in of previous figure for readability).

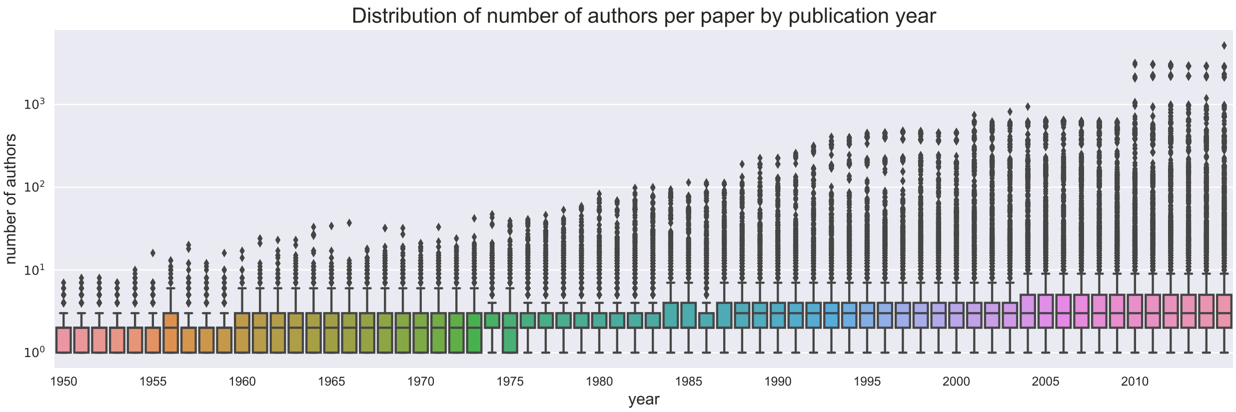


Figure 2.19: **Distribution of number of authors per paper by publication year** There is a clear, increasing trend showing that over time the number of collaborations has increased (see also figure below).

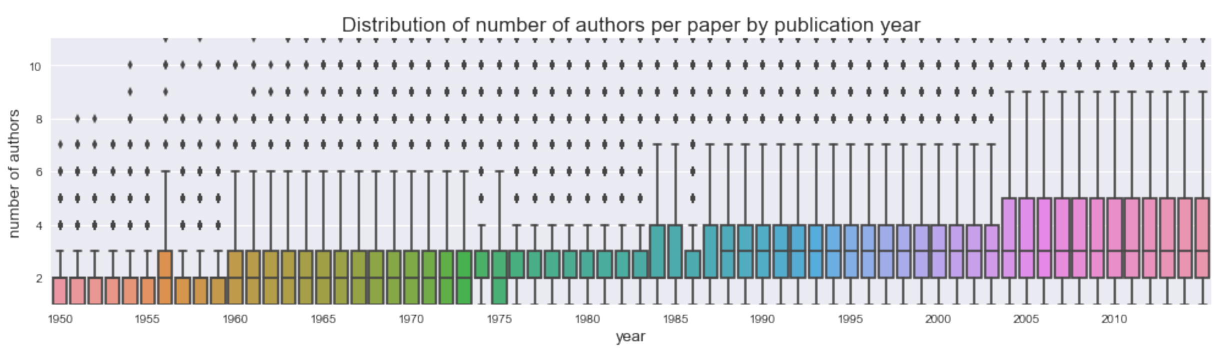


Figure 2.20: **Distribution of number of authors per paper by publication year** (zoom-in of previous figure for readability).

2.9.5 Descriptive Statistics on Team- vs. Single-authored Papers

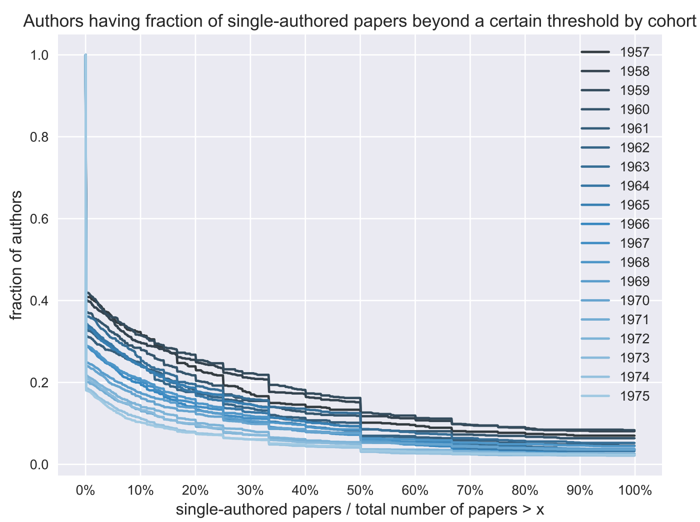


Figure 2.21: **Complementary cumulative distribution of single-to-team fraction** Cumulative distribution of the fraction $\frac{\text{single-authored papers}}{\text{all papers}}$ for every author, by cohort. The plot shows that the majority of authors has just few single-authored papers. Moreover, there is a clear trend by which authors of younger cohorts tend to have fewer single-authored papers per capita. See also [Wuchty et al., 2007].

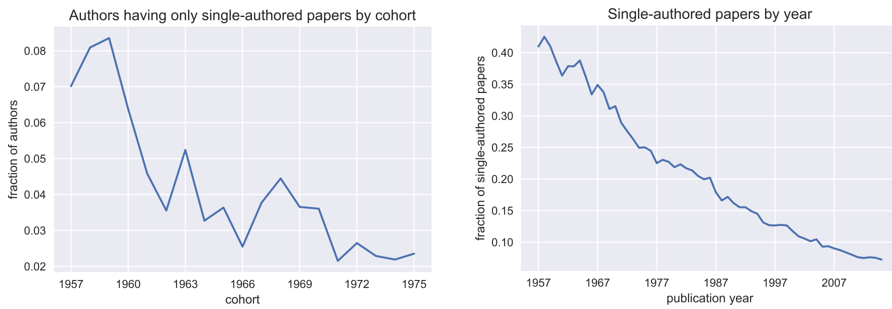


Figure 2.22: **Authors with no co-author & Single-authored papers** On the left, the fraction of authors by cohort who have only published single-authored papers, hence who have no co-author. On the right, the fraction of single-authored papers by year. Although authors with no coauthor account for a small fraction of all the authors (around 2%-8%, depending on cohort), the overall fraction of single-authored papers, when considering all years, amounts to 16% (not shown the in picture).

2.9.6 Descriptive Statistics on Total Papers per Author

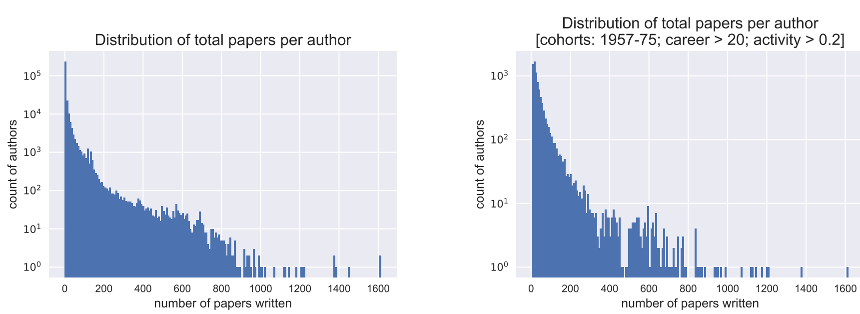


Figure 2.23: **Distribution of total papers written by each author** On the right, the distribution conditional on authors belonging to cohorts 1957-75, with at least 20 years of career and with an activity rate of 20%.

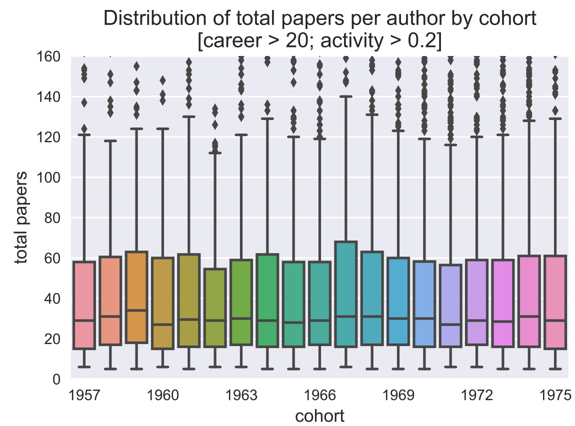


Figure 2.24: **Distribution of total papers per author by cohort**, conditional on having a career of at least 20 years and an activity rate of 20%. The median is consistently around 30 papers per author written throughout one's career. The figure is cut at 160 papers, but there are many outliers not shown for readability.

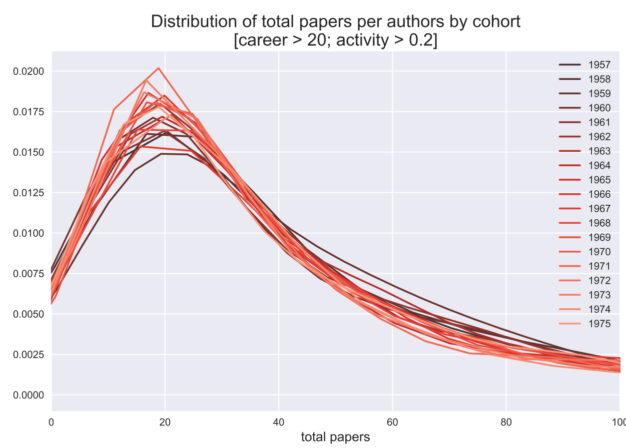


Figure 2.25: **Distribution of total papers per author by cohort**, conditional on having a career of at least 20 years and an activity rate of 20%. The number of total paper is cut at 100 for readability.

2.9.7 Descriptive Statistics on Fields and Topics

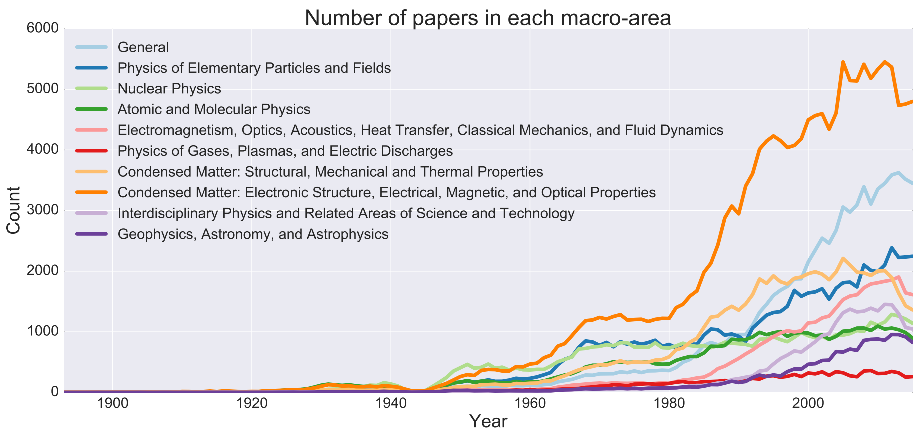


Figure 2.26: **Number of paper in each macro-area**, as determined by the *PACS* codes.

2.9.8 Descriptive Statistics on Co-authors

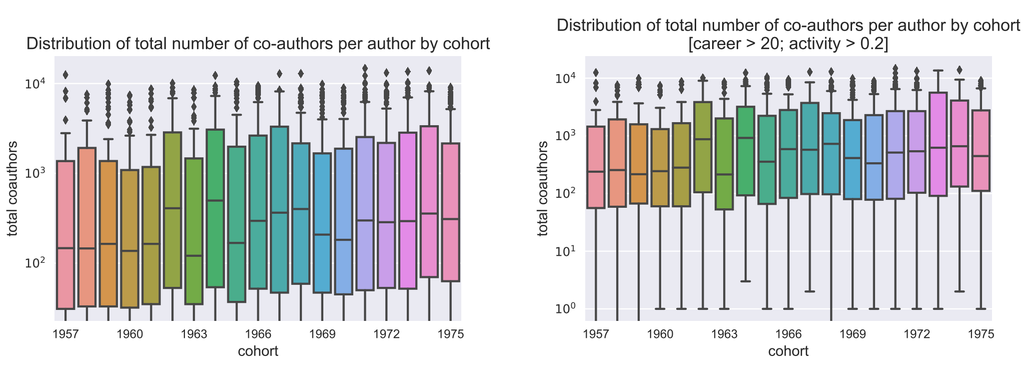


Figure 2.27: **Distribution of total number of co-authors per author by cohort**

Table 2.6: **Statistics on total number of co-authors per author**

	Total number of co-authors	
	across all authors in cohorts 1957-75	only those with career > 20 and activity > 0.2
count	31530	8728
mean	158.42	488.57
std	805.67	1388.40
min	0	0
25%	2	26
50%	6	62
75%	29	208
max	14725	14725

2.9.9 Statistics on Productivity, Performance & Embeddedness

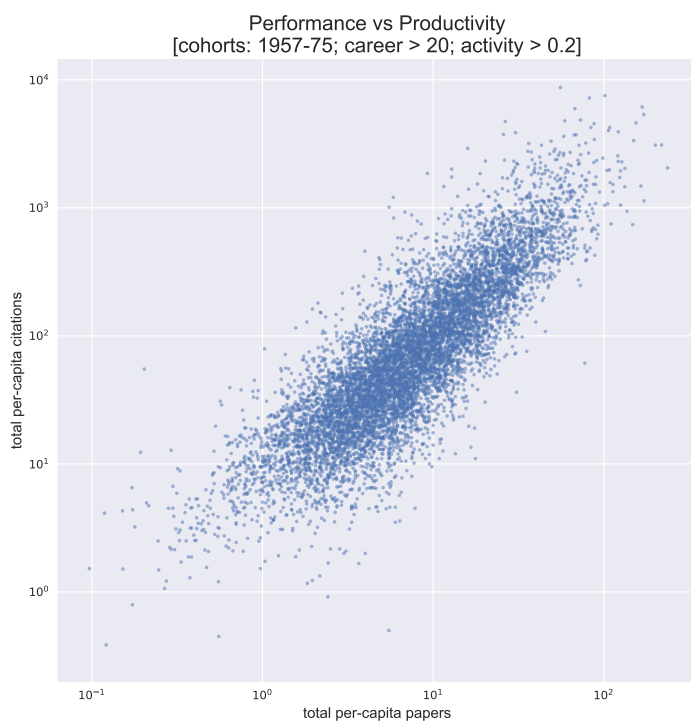


Figure 2.28: **Performance vs. Productivity**

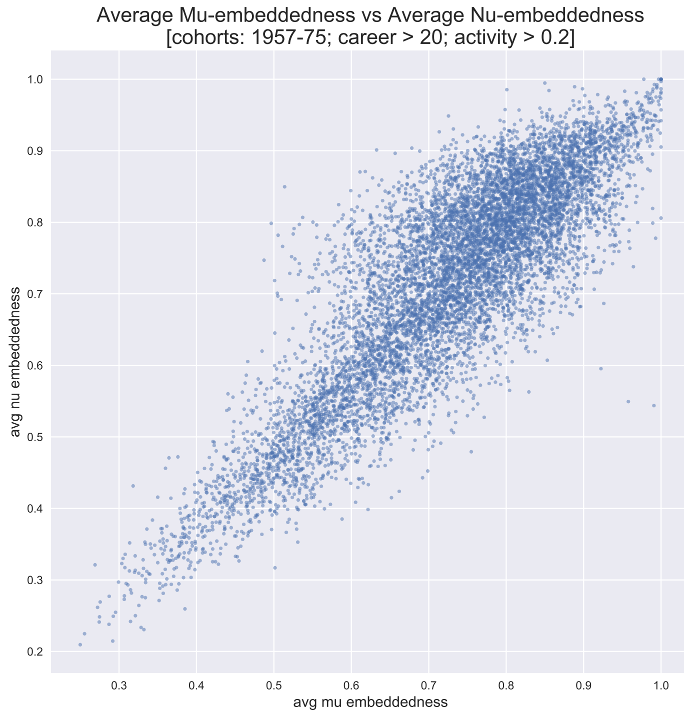


Figure 2.29: **Embeddedness computed as $\mathcal{M}_i$ vs. $\mathcal{N}_i$** averaged over an author's entire career. See Equation 2.3 for details.

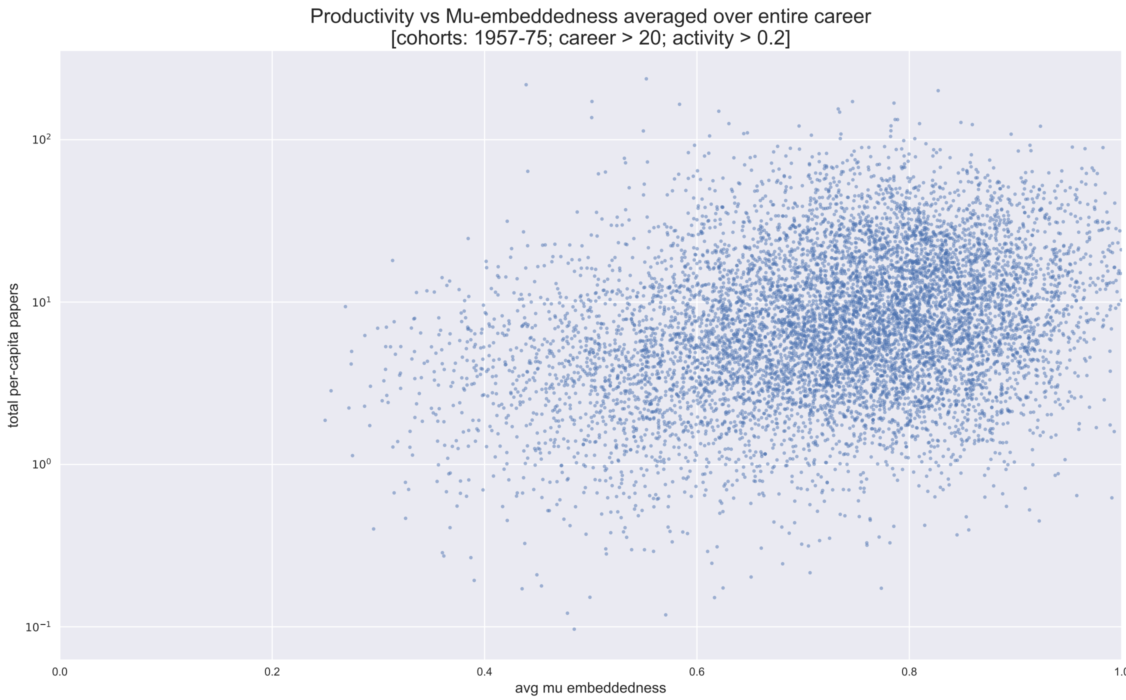


Figure 2.30: **Productivity vs. Exploitation** Total productivity, computed as total per-capita papers published by an author, versus exploitation ratio, computed as $\mathcal{M}_i$ and averaged over an author's entire career. The plot shows an effect similar to the one captured in Figure 2.1.

2.10 Appendix: Existing Literature

The literature may roughly be divided in two different strands, corresponding to the fields of management/business and behavioral/psychological studies:

- **organizational learning**: how firms tend to allocate their resources between the two (exploration or exploitation) and how this depends on various factors, such as embeddedness in the network of firms, complementarities with other connected firms etc.
  - It may be worth noting that in this framework, it is a crucial assumption that resources are scarce and, as a result, exploitation and exploration are "substitutes". In our case, does this assumption still hold?
  - sometimes the interest is also in understanding what is the relation between the duality exploration-exploitation and the duality "specialization vs. diversification". In our case there may be a relation with "distinctive competences"[35].
- **collective learning**: how groups of people, possibly connected in a network, separately and "in parallel" face and solve problems and how their performance (individual and collective) are affected by the network and depend on the shape and properties of the "problem/solution space".[36]
  - many interesting issues here are about individuals' rationality and how different assumptions lead to different predictions. In our case, what could be assume on scientists' rationality and/or level of knowledge and/or of strategic thinking?

It may be worth noticing that on the one hand the exploration-exploitation trade off is also well present in other fields[37], ranging from neurobiology, to ecology, to computer science, to political science, as recently pointed out in [Hills et al., 2015], but, on the other hand, the standard business literature also makes full use of other behavioral methods and approaches, as review in [Dosi et al., forthcoming].

2.10.1 Classics in Business Literature: Organizational Learning & Ambidexterity

The debate over organizational ambidexterity concerns how organizations (e.g. firms) can optimally and efficiently allocate their resources between "exploring"

[35] See March's quotation below.

[36] Notice that often in this framework the solution space is exogenous and fixed, while in "research" (as well as in patenting and innovating) it can be endogenous. In other words, there may be no "true" or "best" solution. To overcome this problem, is it enough to assume that the problem space is multidimensional, so that there is no "natural" order in the goodness of solutions?

[37] Sometime under a different name.

and "exploiting", in order to be performing well in today's business as well as to be adaptable to tomorrow's changing demand and environment.

The notion was introduced first by [Duncan, 1976][38] but then further and essentially developed by [March, 1991]. In March's words:

> "Exploration includes things captured by terms such as search, variation, risk taking, experimentation, play, flexibility, discovery, innovation. Exploitation includes such things as refinement, choice, production, efficiency, selection, implementation, execution. Adaptive systems that engage in exploration to the exclusion of exploitation are likely to find that they suffer the costs of experimentation without gaining many of its benefits. They exhibit too many undeveloped new ideas and too little distinctive competence. Conversely, systems that engage in exploitation to the exclusion of exploration are likely to find themselves trapped in suboptimal stable equilibria."

Moreover, since "compared to returns from exploitation, returns from exploration are systematically less certain", then March's claim/thesis is that "adaptive processes characteristically improve exploitation more rapidly than exploration". Is this true also in academic publishing?

Exploration-Exploitation and Governance The business literature, reviewed in [O'Reilly and Tushman, 2013] and [Dosi et al., forthcoming], has also focused in understanding the role of different forms of governance (more or less hierarchical or vertically/horizontally integrated or modular) in conditioning the exploration-exploitation trade off. In our framework this may give us a hint: considering clusters of the citation network as analogous to modules in an organization.

Learning Myopia and Lock-in Phenomenon As summarized in [Lee et al., 2003], the literature also exhibits two views about the exploration-exploitation topic:

- when a firm accumulates enough experience in a technology, it tends to be trapped by this technology. This phenomenon is labeled "competency trap" in [Levitt and March, 1988] and "learning myopia" in [Levinthal and March, 1993];
- the "lock-in" argument is based on the fact that in presence of "network externalities"[39], an already well-established technology or paradigm can inhibit the adoption of an innovation due to the excessive inertia of (potential) adopters.

[38] I could not find this paper.

[39] NB: here we use the terminology "network externalities" as it is done in the industrial organization literature, see for example [Katz and Shapiro, 1985] and [Shapiro and Varian, 2013].

Exploration-Exploitation and Environmental Volatility One of the standard results stemming from [March, 1991] is that in a dynamical environment, allocating a good amount of resources to exploration is a better response than sticking to exploitation patterns. However other studies, such as [Posen and Levinthal, 2012], suggest that

> "environmental change is not a self-evident call for strategies of greater exploration. Indeed, under some conditions the appropriate response to environmental change is a renewed focus on exploiting existing knowledge and opportunities."

The authors focus on "the relationship between turbulence and the level of exploration" and find that

> "under conditions of very infrequent environmental change, a strategy stressing adaptation via exploration and new knowledge accretion is optimal; but as change becomes more frequent, the optimal response is, in fact, one of greater inertia."

2.10.2 Measuring Exploration, Exploitation and Citation Bias

[Gilsing et al., 2008]

The authors try to explain how firms explore (in terms of number of patents) starting from *technological distance*, *global network density* and (betweenness) *centrality*, while controlling for things such as age, firm size, R&D investments and geographical location.

- Network links represent alliances among firms in the previous 5 years (with respect to the patenting date).
- Technological distance between two firms is measured by considering the (Pearson) correlation between their technology "profiles".

Notice, however, that here the network is assumed somehow exogenous with respect to the exploring activities, while it could be considered endogenous to the extent that agents choose alliances in order to explore more effectively.

[Foster et al., 2015]

The authors explicitly address the problem of the balance or tension between tradition and innovation in scientific strategies. By using a database about papers in biomedical sector, where papers are assigned to chemical components, they are able to create a time-ordered sequence of weighted networks. For every year t, they examine the abstracts and define a network as follows:

- each chemical is a node;

- if two chemicals appear in the same abstract in year t, they add a link at time $t+1$;
- repeated links result in more weight on the same link.

The network formed with this procedure has clusters of related chemicals, so that each link can be subdivided in different kinds, corresponding to different "strategies":

- links within the same cluster are considered consolidation, i.e. *tradition*;
- links bridging between two clusters are considered *innovation*;
- links such that one end is a new chemical are considered *extreme innovation.*

In this framework, then, a strategy corresponds to a link between two chemicals (so far).

In order to estimate the parameters that represent the relative proportion of the various strategies, the authors consider a generative model where the researchers are modeled as a representative agent who allocates different probabilities to the different strategies. Since this probabilities depend on some bias parameters, by maximizing the likelihood that this representative agent will generate the observed history, it is then possible to estimate the bias parameters.

Assigning a strategy to every paper So far, strategies have been identified with chemicals. The authors then proceed by defining (arbitrarily) an "extremely innovative" paper one containing at least one extreme innovation and so on. The relationship between strategy, risk (captured by the surprisal) and reward (in terms of citations) is then assessed as follows: citations are assumed to follow a negative binomial with mean μ. Then they estimate a model of the form

$$\log(\mu) = \beta_0 + \beta_1 \cdot \text{surprisal} + \beta_2 \cdot \text{year} + \varepsilon,$$

or

$$\log(\mu) = \beta_0 + \beta_1 \cdot \text{extr. innovation} + \beta_2 \cdot \text{innovation} + \beta_3 \cdot \text{tradition} + \varepsilon.$$

The coefficients are highly significant and a paper deploying a extremely innovative strategy has around 50% more citations, on average, than a more conservative paper.

Optimal strategy Assuming that citations are distributed according to a negative binomial distribution and fit the model

$$\log(\mu) = \beta_0 + \beta_1 \cdot f(\text{extr. innov.}) + \beta_2 \cdot f(\text{extr. innov.})^2 + \varepsilon,$$

where f(extr. innov.) refers to the proportion of chemical relationships in article that represent an extreme innovation, the authors obtain that the optimal strategy to maximize citations is a innovation-to-tradition ratio around 40%-to-60%, which is in disagreement with [Uzzi et al., 2013] and proves that innovation is highly rewarded.

2.10.3 Sociology Literature

Perhaps unsurprisingly, in the sociology literature some ideas are tackled from different viewpoints and sometimes lead to different conclusions. It is partially the case of the diffusion of (creative) ideas, as analyzed in [Uzzi and Spiro, 2005][40], where the authors point out that "[r]epetead ties can lower innovation costs by spreading the risk of experimentation over the long term". It is worth noticing that in the contexts analyzed there, the good being spread is non-rival, which may be an essential difference with respect to the business literature.

2.10.4 Measuring Creativity

[Uzzi et al., 2013]

By comparing the observed frequency with the frequency distribution created with a randomized citation network, the authors generate a z score for each co-citation pair. The idea is that this normalized measure describes "conventional" pairs, i.e. observed more often than they appear in the randomized network, if the coefficient is above 0 and, conversely, "novel" pairs if they have z scores below 0.

Moreover, for each paper they consider its references' z scores and then examine: (i) the paper's median z score and (ii) the paper's 10th-percentile z score, which captures the more unusual combinations. The first result they draw is that science typically relies on highly conventional combinations, also when the analysis is replicated on a field-by-field basis and across time.

The second result concerns the probability of a "hit" paper (operationalized as those papers in the upper 5th percentile of citations in the whole dataset). The authors consider two distinct "dimensions", novelty and conventionality, to explain the probability of a "hit" paper. High (median) novelty or high (median) conventionality alone can guarantee an average probability of writing a "hit" paper, but, perhaps surprisingly, it turns out that the two dimensions are not opposing factors since that a high level on both of them guarantees a significantly higher probability of a "hit" paper. In this respect, yet another finding is that the peak impact is for papers in the 85th-95th percentile of median conventionality in their references.

The third result is about the higher likelihood of team-authored papers to exhibit novel elements with respect to solo-authored papers. More specifically, by using the 2-dimensional decomposition outlined above, the authors show that

[40] While citing also classic ideas about embeddedness by Granovetter.

team works are not only more novel but rather are more inclined to incorporate this high novelty without giving up much of their conventionality.

[**Zhang et al., 2016**]

The authors keep track of all papers accessed for reading from Indiana University IP addresses and then use this "accumulated knowledge" to estimate how this affects the output papers (i.e. papers written subsequently by Indiana University affiliates).

Defining creativity The authors consider a *general creativity definition* as combination of two factors:

- $d_{i,j}$, measuring the **disconnectedness** of two papers i and j;
- $r_{i,j}$, the **rarity**, capturing the probability that both i and j were previously already cited together in other papers.

They consider, then, a **creativity score** of the pair (i,j) as $\varphi_{i,j} := d_{i,j} \cdot r_{i,j}$. Then a paper k's overall **creativity** is the aggregation[41] of these scores across all papers $\mathcal{C}_k$ cited by k:

$$\phi_k := Aggr\left(\{\varphi_{i,j}\}_{i,j\in\mathcal{C}_k}\right).$$

More specifically, to compute the the rarity of the co-citation (i,j) till year t, they define it as

$$r^t_{i,j} := \frac{1}{1+\log_2\left(c^t_{i,j}+1\right)},$$

where $c^t_{i,j}$ is the number of co-citations of (i,j) till year t.

Concerning, instead, the disconnectedness $d_{i,j}$ between two referenced papers i and j, they use the average dissimilarity of their topics and use the fact that they have multi-layer tags for each paper, corresponding to different topics. In particular, topics/tags are encoded in a network which consists in a 4-level hierarchy, which allows them to first compute a level-wise similarity and then an average dissimilarity measure, where the passages between different layers are discounted by a certain factor.

How much of creativity is explained by the papers read by its authors? Not having the papers actually read by the individual scientists at their disposal[42], the authors compare the reference set in year t with the reading set in previous years t' by computing the Jaccard's coefficient of the two. They are able to prove that: (i) the impact of reading papers over future publications decays over time; (ii) recently read papers have a stronger influence.

[41]E.g. average, median, etc...

[42]The online requests for the reading papers are anonymous, so the reading papers always are considered overall the Indiana University population.

Disentangling creativity and bridges between papers Consider two papers i and j. Then their creativity score $\varphi_{i,j}$ can represent the difficulty of connecting the pair (i, j). A third paper x, if adequately located in the middle between i and j, could reduce this difficulty, so that the path i–x–j is "easier" than i–j. In terms of creativity scores, the authors introduce a metric to describe the impact of x on the creativity required to connect (i, j):

$$\Delta_x^{i,j} = \varphi_{i,j} - \min\left(\varphi_{i,j}, \max(\varphi_{i,x}, \varphi_{x,j})\right).$$

This definition can be generalized to the case of n papers $x_1, ..., x_n$ bridging the gap from i to j. Intuitively, $\Delta_x^{i,j}$ is called "preparation", while $\varphi_{i,j} - \Delta_x^{i,j}$ "inspiration".

By taking the maximum over x, the authors are able to decompose the creativity of a paper k as sum of preparation plus inspiration as follows:

$$\text{Creativity} \equiv Aggr\left(\{\varphi_{i,j}\}_{i,j\in\mathcal{C}_k}\right),$$

$$\text{Inspiration} := Aggr\left(\left\{\varphi_{i,j} - max_x\Delta_x^{i,j}\right\}_{i,j\in\mathcal{C}_k}\right)$$

so that, as a result:

$$\text{Preparation} = \text{Creativity} - \text{Inspiration}.$$

2.11 Appendix: Performance as Field-adjusted Citations

In order to evaluate the impact of different exploration-exploitation strategies on productivity and performance, we first need some measures of productivity and performance that will allow us to compare careers across different times and fields.

We adopt the same approach developed by [Radicchi et al., 2008] and [Petersen et al., 2010], whose aim is to have a citation-based measure of performance that is homogeneous and, then, comparable across fields and time. To do so, every paper's citation is weighted to control for field biases, paper's age and number of authors.

In order to control for field biases, every author has to be assigned to a field[43] on the basis of her published papers. In turn, every paper needs to have a field, defined by the PACS codes. Since, however, not all papers have PACS codes, we devise an "artificial" assignment of PACS codes to extend the classification also to those papers ...

Label Propagation

Let $\tilde{A}$ be the (symmetric) adjacency matrix of the citation network, where the element $\tilde{A}_{ij}$ is 1 if and only the paper i cites paper j *or* is cited by paper j, being 0 otherwise.

[43]Or, rather, to a combination of fields. Every author is, in fact, described by a combination of the 10 main topics defined by the PACS codes in which she has published.

Labeled Papers The possible labels correspond to the 10 macro-areas defined by the PACS codes, so that for every labeled paper i, we can define a *field vector* $\tilde{\mathbf{F}}_i = (F_{i,1}, \ldots, F_{i,10}) \in \{0,1\}^{10}$ such that $\tilde{F}_{i,f}$ is equal to 1 if and only if the paper i is assigned to the field f. Let us $\mathbf{F}_i$ be the row normalized version of $\tilde{\mathbf{F}}_i$, i.e.

$$\mathbf{F}_i = \left(\frac{F_{i,1}}{\sum_f F_{i,f}}, \ldots, \frac{F_{i,10}}{\sum_f F_{i,f}} \right) \in [0,1]^{10},$$

such that $\sum_f F_{i,f} = 1$, for every paper i.

Unlabeled Papers Intuitively, we design a procedure by which every unlabeled paper can "inherit" an average of the labels of its neighboring papers. More formally, for every paper i

$$F_{i,f} = \frac{1}{|N_i|} \sum_{j \in N_i} F_{j,f},$$

where N_i indicates the set of i's neighboring papers (hence, $|N_i|$ is i's degree). This intuition is justified by the strong correlation exhibited for labeled papers, which confirms that a paper tends to have a field vector highly correlated with those of its neighbors.

Label Propagation Algorithm Let A be the row normalized version of $\tilde{A}$. Without loss of generality, we can assume that the papers are ordered in a way such that the first l papers are labeled while the rest u papers are unlabeled, where $l + u$ amounts to the total number of papers present in the dataset. According to this distinction, A and F can be subdivided in blocks

$$A = \begin{pmatrix} A_{LL} & A_{LU} \\ A_{UL} & A_{UU} \end{pmatrix}, \qquad F = \begin{pmatrix} \mathbf{F}_1 \\ \vdots \\ \mathbf{F}_l \\ \mathbf{F}_{l+1} \\ \vdots \\ \mathbf{F}_{l+u} \end{pmatrix} = \begin{pmatrix} F_L \\ F_U \end{pmatrix},$$

where A and F have dimensions $(l+u) \times (l+u)$ and $(l+u) \times 10$, respectively.

The propagation algorithm stems from [Zhu and Ghahramani, 2002] and works as follows:

1. for every labeled paper i, let $\mathbf{F}_i^{(0)}$ be its exact field vector. For every other unlabeled paper, instead, initialize $\mathbf{F}_i^{(0)}$ at random;
2. for every step $n \in \mathbb{N} \setminus \{0\}$, define $F^{(n)} = AF^{(n-1)}$;
3. clamp the labeled data, i.e. re-assign the submatrix $F_L^{(n)} = F_L^{(0)}$;

4. repeat from step 2 until the process converges.

LEMMA 2.11.1. *If there are no isolated nodes and every unlabeled node has at least one link to a labeled node, then the algorithm converges*[44] *and it has a unique limit solution. Consequently, the initialization of the unlabeled papers is inconsequential.*

Proof. The unknown of interest is the submatrix F_U. Writing the iterative step in blocks, $F^{(n)} = AF^{(n-1)}$ for $F_U^{(n)}$ becomes:

$$\begin{pmatrix} F_L^{(n)} \\ F_U^{(n)} \end{pmatrix} = \begin{pmatrix} A_{LL} & A_{LU} \\ A_{UL} & A_{UU} \end{pmatrix} \begin{pmatrix} F_L^{(n-1)} \\ F_U^{(n-1)} \end{pmatrix} \quad \Rightarrow \quad F_U^{(n)} = A_{UL} \underbrace{F_L^{(n-1)}}_{\equiv F_L, \forall n} + A_{UU} F_U^{(n-1)}.$$

By taking the limit on both sides for $n \to \infty$, we obtain that the unknown of interest, $F_U = F_U^{(\infty)}$, is the fixed point solution of

$$F_U = A_{UL}F_L + A_{UU}F_U, \quad \text{i.e.} \quad (I - A_{UU})F_U = A_{UL}F_L.$$

Given the hypothesis on A, the submatrix $(I - A_{UU})$ is invertible because the sum of the off-diagonal elements of every row is strictly less than 1 while the diagonal elements are all 1, so the Gershgorin circle theorem guarantees that 0 cannot be an eigenvalue. The unique solution is then $F_U = (I - A_{UU})^{-1} A_{UL} F_L$. □

Remark. Since there is no possible way to infer labels for isolated and unlabeled papers, uniform labels are assigned to them.

Performance across Sub-disciplines

As usual, we here evaluate the productivity of an author by measuring her citations. However, since it is well known that "comparing bare citation number is inappropriate"[45], we adopt the approach proposed by [Radicchi et al., 2008] and [Petersen et al., 2010], which allows to account for field biases, age of the paper and number of authors.

Method by [Radicchi et al., 2008] (RFC): Field Normalization The method proposed is to define a *weighted citation number of paper* i, belonging to field f, as follows:[46]

$$RFC_i = \frac{c_i}{\langle c_f \rangle},$$

where c_i are the citation received by i from paper of a given year and $\langle c_f \rangle$ is the average number of citations received by papers belonging to the same field f in the same year.

[44] The convergence should be guaranteed because the spectral radius of A_{UU} is less than 1.
[45] [Radicchi et al., 2008]
[46] RFC stands for Radicchi-Fortunato-Castellano

Our method: Revealed Comparative Advantage, version I Our method, instead, uses the notion of comparative advantage, developed in economics, as a starting point. Given a paper i, we consider all the citing papers in a given year and keep track of their fields[47]. This allows us to build the probability distribution of i's citing fields $\mathbf{F}^i = (F^i_1, ..., F^i_{10})$, where F^i_f accounts for the fraction of i's citations coming from field f.

For every paper i and every field j, we define the *reveal comparative advantage of paper i in field f in year y*[48]

$$RCA^i_f = \frac{c_{i,f} \big/ \sum_f c_{i,f}}{\sum_j c_{j,f} \big/ \sum_j \sum_f c_{j,f}},$$

where $c_{j,f} := c_j \cdot F^j_f$ is the number of j's citations in year y from papers belonging to field f.

Our method: version II With the same notation of the previous paragraph, for every paper i, we define the *weighted number of citations for i in the year y*

$$WC_i = \sum_f \frac{c_{i,f}}{\langle c_f \rangle},$$

where $c_{i,f} = c_i \cdot F^i_f$ and $\langle c_f \rangle$ is the average number of citations for papers in field f, received in year y.

[47]Here by “field” we mean the macro-areas defined by the 10 macro-subjects given by the PACS code (see also Figure 2.26).

[48]In the summations, f varies over the fields $f \in \{1, ..., 10\}$, while j across all the papers.

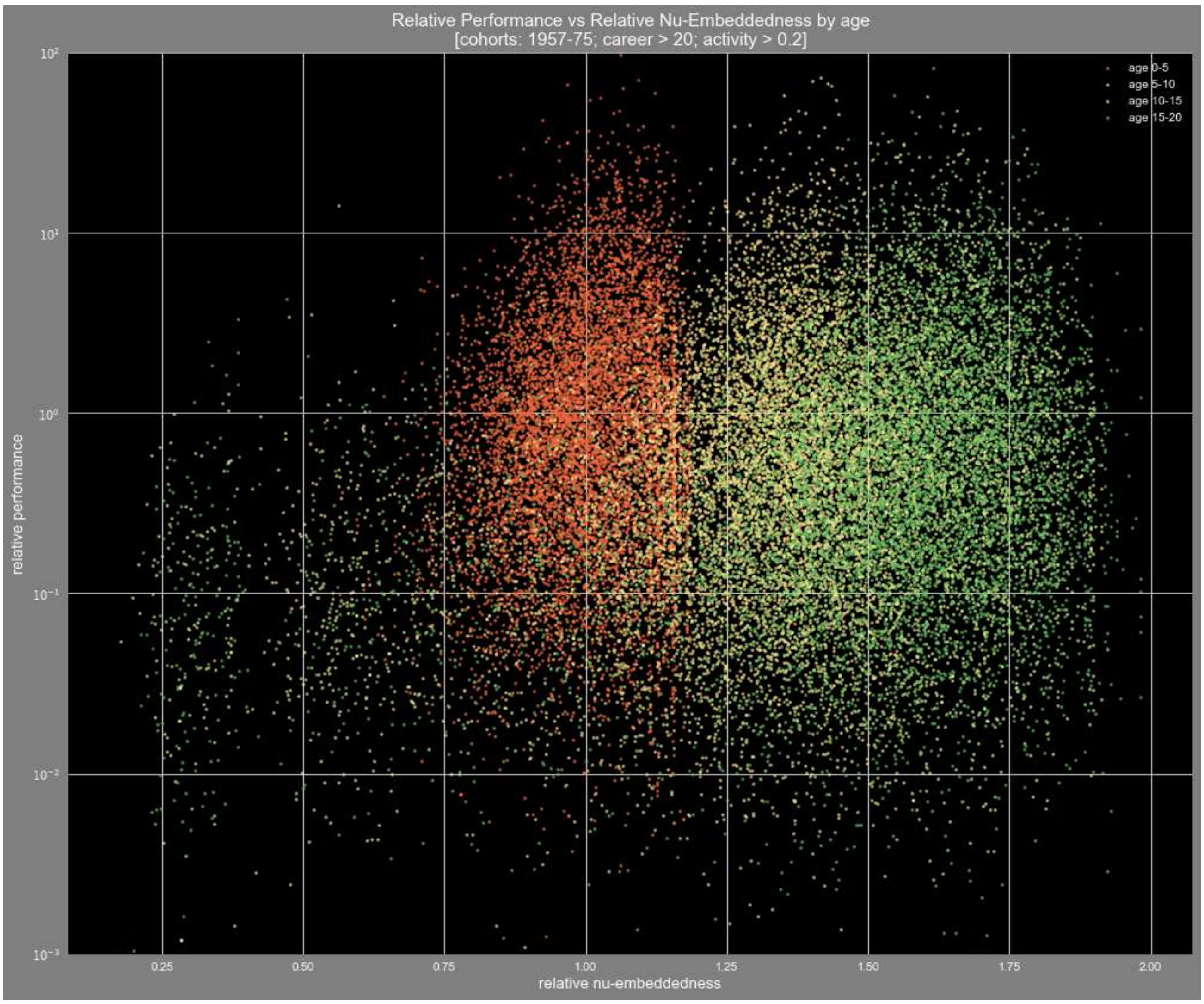


Figure 2.2: **Performance vs. Exploitation**

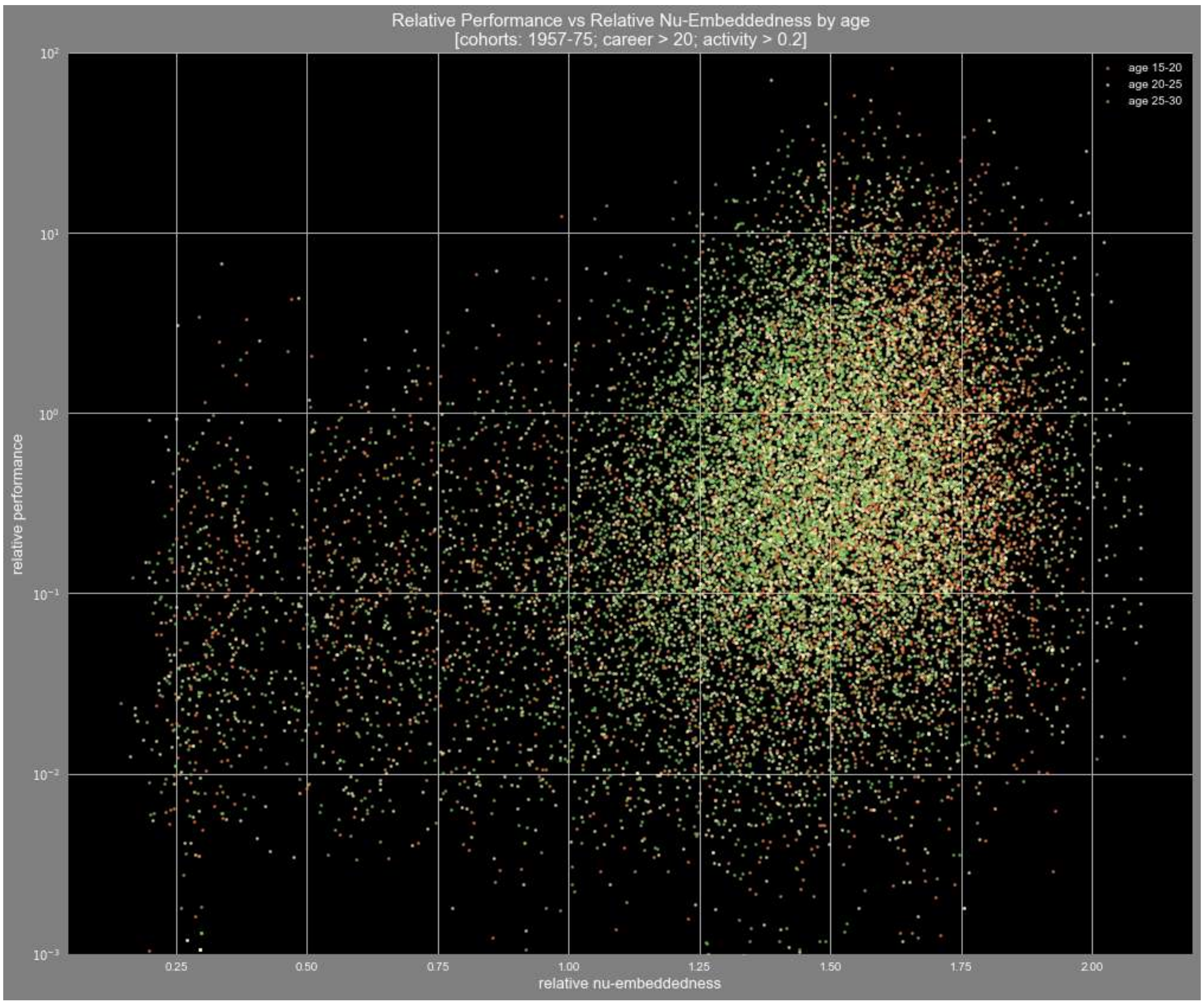


Figure 2.3: **Performance vs. Exploitation**

Chapter 3

Exploration-vs-Exploitation Dilemma & Performance for Top Academic Institutions

3.1 Abstract

By keeping track over time of the publications made by scientists belonging to different universities, we show a significant U-shaped relationship between institutions' performance and their exploration among the various topics/fields, over the years. This relationship consistently holds also when controlling for institution's size, academic reputation, resources, etc (proxied by rankings and other measures).

The results highlight a possible "tension" between economies of scale (favored by an increase in size) and economies of scope (possibly favored by an increase in differentiation/exploration) in academic production at an aggregate level. Moreover, they suggest and delineate two optimal behaviors for achieving high performance: high differentiation versus high specialization. Noticeably, though, the effect of exploration on performance changes in sign and magnitude together with the size of the institution.

This analysis sheds light on the (non-monotonic) relationship existing between the exploration-exploitation dilemma (a.k.a. diversification-specialization) and its impact on performance, at an aggregate level. At the same time, it keeps the door open to models and explanations for how to join individual-level behaviors and aggregate outcomes.

This work is done in collaboration with Alessandro Vespignani and Matteo Chinazzi.

3.2 Introduction

Motivation and context The present paper falls within that strand of literature that has emerged following [March, 1991], among the first to point out the importance of the so-called *exploration-versus-exploitation dilemma*. Intuitively, the "tension" formalized by the dilemma is easily explained: on the one hand, exploring is risky and potentially highly rewarding, while on the other hand, exploiting is seen as less risky and capable of yielding lower but safer and consistent benefits.

Here we first focus on how to measure and quantify the "amount of exploration" and then we try to asses its impact on performance. We do this at an "aggregate" level, meaning that here the production units are entire (departments of) universities. We keep track over time of their publications so that we can build a measure of exploration in terms of (sub)fields and topics treated in each paper and we can also measure performance through the number of citations received.

This approach entails a couple of important features: the first is that with such an aggregation the results should be less prone to biases and fluctuations, which may well be present in the data, with respect to studying the single individuals. The second is that this analysis can be seen as part of the literature that has focused on organizational and collective learning.

Related literature In terms of literature, this paper is located at the intersection of several (sub)fields:

- research on patenting and innovation;[1]
- advantages and drawbacks of specialization (versus diversification), in the field of science and, additionally, as a well-established issue in economics;[2]
- literature on the so-called science of science[3], and particularly in its branch within the realm of economics[4];
- exploration-exploitation dilemma studied not only for individuals but also in groups and organizations.[5]

This last point may be of particular relevance in this context. It is established that a wide range of institutions have their effect on innovation and, although here we take an agnostic approach as to whether exploring is more or less desirable than exploiting (at least concerning academic research as quantified and

[1] As in [Gilsing et al., 2008] and [Gilsing and Nooteboom, 2006]

[2] [Levitt and March, 1988], [Levinthal and March, 1993]

[3] [DeSolla Price, 1965], [Newman, 2004], [Uzzi and Spiro, 2005], [Wuchty et al., 2007], [Radicchi et al., 2008], [Petersen et al., 2010], [Uzzi et al., 2013], [Foster et al., 2015], [Boudreau et al., 2016], [Clauset et al., 2017].

[4] [Hudson, 1996], [Fafchamps et al., 2010], [Hamermesh, 2013], [Ductor et al., 2014], [Abramitzky, 2015], [Ductor, 2015], [Seltzer and Hamermesh, 2017].

[5] [Hills et al., 2015]

measured here), still we think that it is important to understand the mechanisms at play and how resources, say, affect the choices of departments in universities and possibly their performance and impact on society.

Results (and possible explanations) The econometric analysis shows a significant U-shaped relationship existing between performance and exploration, which implies that the highest performances are attained when more "extreme" behaviors are adopted: high specialization and high differentiation pay off more than "intermediate" choices.

However when controls for institutions' resources and size are in place, this relationship becomes less straightforward. On the one hand, a significant and positive effect of size on performance supports that economies of scale seem to be a mechanism at play. On the other hand, when the institution is large enough, then economies of scope seem also to kick in and have a positive impact on performance.

Further research is still needed, but possible explanations for these non-monotonic results could be found in the interplay between individual behaviors and aggregate outcomes. In the academic context, "exploring" and (possibly) sharing different findings and interests can be seen as a (local) public good, shared among and enjoyed by researchers. Some models, such as the so-called *law of the few* by [Galeotti and Goyal, 2010], stress that in some settings, where individuals are ex-ante identical, slight cost advantages in information acquisition can have dramatic effects in terms of aggregate outcomes and also in terms of different roles played by the agents.

The rest of the paper is organized as follows: the first two sections are respectively a description of the data and of the methods and measures used. Then, the results of the econometric analysis are shown while the last section draws the conclusions and the prospectives for the future work.

3.3 Data

The data come from the *American Physical Society* (APS) datasets and consist of information about 302,654 papers published in APS journals from 1985 to 2009. Instead of being associated with authors, every paper is associated with a number of geographical "affiliations", corresponding to the urban areas of the universities of the authors.[6] These different 2216 cities/locations belong to 177 different countries or US states.[7] Throughout this paper we will use indifferently the terms *urban area*, *institution*, *city* or *affiliation*. We also use the data already used in the previous chapter[8], particularly for what concerns the measurement of citations.

[6]For more details, see section "Methods" in [Zhang et al., 2013].

[7]Ranging from Alabama to Albania, to Estonia, to Germany (and German Democratic Republic), to Yugoslavia and to Zimbabwe.

[8]See section "Data" of chapter 2.

It is worth noticing that the particular constraints on our data are mainly due to two different reasons: the first is the somewhat intrinsic difficulty in applying methods for disambiguation and geolocalization of authors and universities. The second is that although the papers recorded in the APS dataset virtually span from 1893 to 2016, those actually attached to PACS codes are only concentrated in a few decades, and only on these one can perform a field/topic classification.[9]

3.3.1 Data Filtering

It is reasonable to assume that the papers recorded in the American Physical Society dataset may exhibit a bias toward American authors as well as institutions, perhaps due to geographical proximity and other similar factors. This bias may also be more significant for older years, especially before the Soviet Union dissolution. For this reason, in our analysis we focus only on American institutions and cities. Around 25% of the cities in the sample are in the U.S. (558 out of 2216).

Moreover, since this database only contains papers published in the APS journals, it can only partially be considered as an exhaustive record of all papers published by a single institution. Therefore, one can only keep track of institutions that have published in APS consistently (and abundantly) throughout the years.

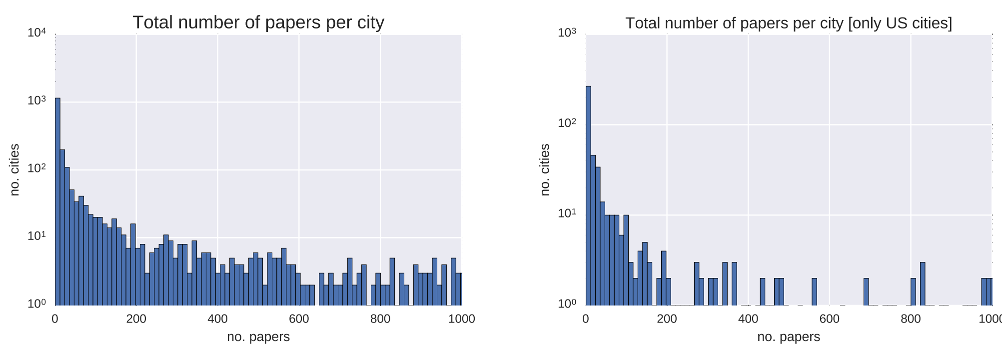


Figure 3.1: **Distribution of total number of papers per city**

In order not to get a time discretization that is too fine grained, we group the years in 3-year time windows, so that the 24 years from 1985 to 2009 are subdivided into 8 periods of 3 years each.

The subsample of institutions under consideration, then, consists first of all of only U.S. cities/institutions. Then we further restrict our attention to those are active in every one of these 3-year time periods (i.e. that have published at least one paper). Lastly, we focus on those that have a minimum number of papers published in any 3-year time window that is above the median (which is 19 papers in 3 years). The result of this selection is: 92 U.S. cities/institutions.

[9] Unless one relies on label propagation techniques, as mentioned in the previous chapter.

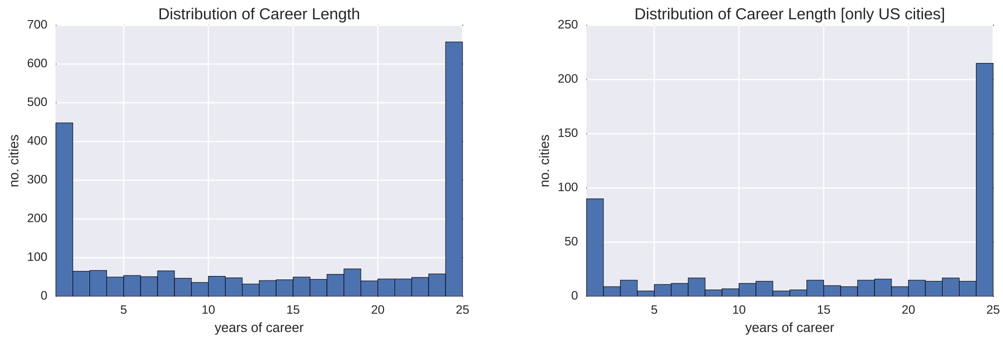


Figure 3.2: **Distribution of career length of every city** The *career length* of a city is defined as the number of years spanning from the year of its first appearance in the database to the last year.

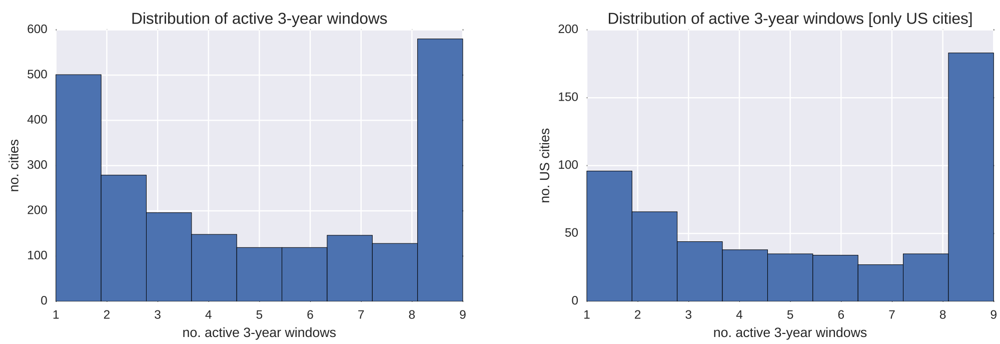


Figure 3.3: **Distribution of active 3-year periods of every city** The number of *active years* of a city is defined as the number of years in which a city has published. Analogously, when the time unit is 3 years, instead of a single year.

3.4 Methods

In this section we describe the fundamental measures that we will use later in the econometric analysis.

3.4.1 Measuring Productivity and Performance

The notion of *productivity* in this context is merely defined as the number of papers written by an agent (i.e. a urban area, constituted by a university/institution/city). The *performance* of an agent is given by the number of citations received by that agent which are relative to papers published in that particular year.

Formally, let p be a paper and let C_p be the set of all papers that cite p (independently of their year of publication). Let P_c^t denote the set of papers published by city c at time t. Then the performance of city c at time t is given by:[10]

$$\mathrm{Perf}_c^t = \sum_{p \in P_c^t} |C_p|.$$

It is worth noting that the database in our possession only allows us to keep track of citations coming from (and going to) papers in the APS journals, therefore every other citation and reference to papers in other journals is omitted and, hence, not accountable.

As a last remark, notice that it is reasonable to assume that the two measures, productivity and performance, are very correlated: the higher the number of papers you write, the more likely it is that you will receive a citation. Indeed, in our data the correlation between the two is around 95%, which is why production is not used in the econometric analysis.

3.4.2 Measuring Exploration (and Exploitation)

Following [Jia et al., 2017], we consider the PACS codes attached to a paper as a classification in terms of topics and subjects, so that every paper can be associated with a vector of 79 components[11] which describes it in those terms.

In particular, we take the 79 PACS codes corresponding to the categories obtained when considering only the classification up to 4 digits.[12]

More formally, we assign to each PACS code a number from 0 to 78. Then, given a paper p, let $K_p = \{k_1, ..., k_n\} \subseteq \{0, ..., 78\}$ be its codes (possibly with repetition). Then the k-component of the **vector of topics** of p is defined by:

$$(\mathbf{F}_p)_k := \frac{|\{k' \in K_p : \, k' = k\}|}{|K|}, \qquad \text{for } k = 0, ..., 78$$

so that $\mathbf{F}_p \in \mathbb{R}^{79}$.

Intuitively, $\mathbf{F}_p$ represents the probability distribution of topics or (sub)fields studied in p. This approach has two advantages: on the one hand, every field touched by p is taken into account (by receiving a positive weight instead of 0), on the other hand, if p has more codes corresponding to a specific area, then that area will receive a more substantial weight.

The main idea, then, is that by keeping track of one's papers and – crucially – of their topical distributions over time, we are able to understand whether an agent is entering into fields that it had not touched before and/or with a different focus or intensity.

[10] Throughout this paper, the notation $|A|$ denotes the cardinality of a set A.

[11] The PACS codes of 2010 edition are divided in 10 macro-areas, each one containing different subcategories. For details, see https://publishing.aip.org/publishing/pacs/pacs-2010-regular-edition.

[12] Since the complete codes are 6-digit strings but we only consider the first 4 of them (i.e. the first 2 levels of classification), then in our analysis the codes $AA.BB.cc$ and $AA.BB.dd$ are both classified as the same code, i.e. $AA.BB$. In this way a paper can have repeated codes.

Formally, let c be a urban area/city/institution and let P_c^t be the set of papers ascribable[13] to c and published at time t. Let $\mathbf{F}_c^t$ be the vector of topics of c at time t, defined by:

$$\mathbf{F}_c^t := \frac{1}{\sum_{p \in P_c^t} |K_p|} \cdot \sum_{p \in P_c^t} \mathbf{F}_p.$$

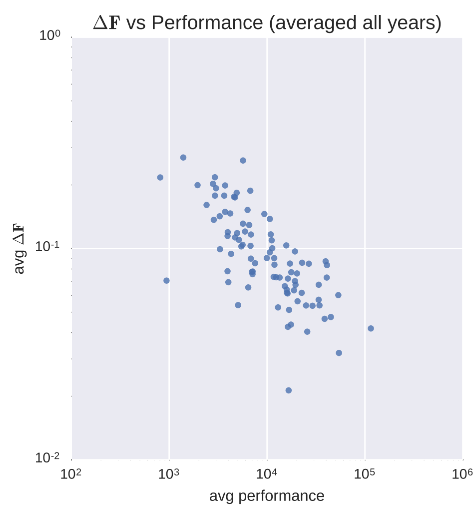
Intuitively, $\mathbf{F}_c^t$ is the distribution of topics treated by c at time t and the trajectory $t \mapsto \mathbf{F}_c^t$ in $\mathbb{R}^{79}$ describes the evolution of c's research interests over time.

To quantify the "change" in interest, we compute the *cosine distance* between every pair of two successive vectors:[14]

$$\Delta \mathbf{F}_c^t = \operatorname{CosDist}\left(\mathbf{F}_c^t, \mathbf{F}_c^{t-1}\right), \qquad \forall\, t \geq 1,$$

where $\operatorname{CosDist}(\mathbf{u}, \mathbf{v}) = 1 - \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|_2 \|\mathbf{v}\|_2}$, for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{79}$.

Figure 3.4: **Relationship between average exploration and performance** The figure shows average exploration ($\Delta\mathbf{F}$) on the vertical axis and average performance on the horizontal one, when both are averaged across all time periods (here the time units are 3-year time windows). The figure seems to suggest a negative relationship between the two average quantities. See also Figure 3.5 for more details.



3.4.3 Additional measures: Ranking and (a proxy for) Size

The first controls that we will use in the econometric analysis are the so-called college and university rankings which give an order of (some) higher education institutions by a combination of factors (such as endowment, research output, internationalization, number of Nobel-awarded scholars in their faculty, academic reputation, teacher-to-student ratio, etc.)

The rankings that we consider are:

[13] A paper can be ascribed to multiple institutions at the same time, because it can have multiple authors

[14] Notice that the cosine distance only depends on the angle between vectors, so in the definitions of the vectors $\mathbf{F}$ the re-normalization to be a distribution and sum up to 1 across components could be omitted.

- *Shanghai Rankings*, a.k.a. ARWU, in the field of sciences[15];
- *QS World University Rankings* in physics[16];
- *Times Higher Education* in physics[17].

Of the 92 cities under consideration, only 48 are classified in the rankings considered above. Moreover, the only data in our possession about university rankings are for years not exactly corresponding to those in the APS database. However, a certain "stickiness" of these rankings may be assumed, especially when taking into account the width of the rank brackets provided.

The second control that we will use is a proxy for the size of an institution. Unfortunately, from our data it is not possible to directly associate the names of the authors to their affiliation, which, in turn, makes it impossible to estimate exactly the "numerosity" or the "size" of the institutions.[18]

Starting from the information present in our dataset, as a (admittedly rough) proxy for an institution's size/numerosity, we devise a measure that is defined as follows. Let p be a paper and let auth_p and aff_p respectively denote the number of authors of p and the number of affiliations of p's authors. The **numerosity** of p is defined by:

$$N_p := \frac{\mathrm{auth}_p}{\mathrm{aff}_p}.$$

The numerosity of a urban area c at time t is then obtained by summing the numerosity of all its papers:

$$N_c^t := \sum_{p \in P_c^t} N_p,$$

where P_c^t denotes the set of all papers produced by c at time t, as already done above.

Remark. The somehow obscure construction of this measure N depends on the fact that, at the present moment, the data in our possession consist of two different datasets: the first one linking every paper to its authors, and the second linking every paper to its PACS codes and to the (geolocalization of the) corresponding urban areas, as described above. Unfortunately, then, the only direct connection between authors and institutions/areas are the papers, which very often have many authors with different affiliation(s) each. Hence, with the

[15]From http://www.shanghairanking.com/FieldSCI2016.html

[16]From https://www.topuniversities.com/university-rankings/university-subject-rankings/2014/physics

[17]From https://www.timeshighereducation.com/world-university-rankings/2017/world-ranking#!/page/0/length/-1/locations/US/subjects/3060/sort_by/rank/sort_order/asc/cols/stats

[18]This problem cannot be circumvent with the current data but may be solved with future versions or different datasets. Still, there are various factors to take into account: even assuming perfect disambiguation of authors' names, many authors still have different affiliations at the same time and they are not always all reported in every paper. In addition, many change often affiliation throughout their career.

current data, it is impossible to exactly discriminate which author is affiliated to which institution.

It may be worth noting that the numerosity N_c^t is strongly related to the number of papers produced by c (i.e. the measure called *production*), which is, in turn, reasonable to assume to be related to c's size. However, it is clear that *numerosity* has many limitations: for example, say that a paper has 10 authors, 8 of whom are affiliated to A while only 2 affiliated to B. Then both institutions A and B will receive an equal contribution of 5 to their numerosity, due to that paper, even if A has provided more authors and it may be a lot larger than B.

3.5 Econometric Analysis

From our data, we consider a panel which is organized as follows:

- although each paper has its year of publication (from 1985 to 2009), we decide to use 3-year windows as basic time unit, in order for an institution to accumulate enough papers and to limit the fluctuations in the topics treated;
- the *exploration* of city c at time t is measured by $\Delta\mathbf{F}_c^t$, as defined above;
- the *performance* of city c at time t is described by $\log\left(\mathrm{Perf}_c^t\right)$. The use of the logarithm is justified by the fact that citations are often shown to follow super-linear (or even power law) patterns, due to rich-get-richer phenomena;
- the *numerosity*, interpreted as a proxy for an institution's size at time t, is also taken into account;
- lastly, the time trend will be considered, in order to control for the accumulation of citations over time and, also, for the so-called "inflation" in authors' performance.[19]

When the *Breusch-Pagan test* and the *Hausmann test* are performed, the first confirms the inappropriateness of the Ordinary Least Square model with respect to a proper panel (random effect) model, while the second one supports the use of the Fixed Effects model over the Random Effects model.

3.5.1 Fixed Effects Model

The econometric model we estimate is the following:

$$\log \mathrm{Perf}_c^t = \alpha_c + \mathbf{x}_c^t \cdot \boldsymbol{\beta} + \varepsilon_c^t, \qquad \text{for all cities } c \text{ and times } t,$$

where the α_c are the city-specific effects, which are permitted to be correlated with the regressors $\mathbf{x}_c^t$, and ε_c are the idiosyncratic errors. Depending on the

[19] See [Pan et al., 2016].

Table 3.1: **Descriptive statistics**

Variable		Mean	Std. Dev.	Min	Max
log Perf	overall	8.993	1.158	5.209	11.985
	between		.993	6.429	11.585
	within		.603	6.309	10.520
Expl	overall	.102	.082	.005	.524
	between		.053	.019	.238
	within		.063	-.101	.431
Numer	overall	942.639	976.732	29.916	7847.104
	between		727.215	69.582	4285.984
	within		655.956	-828.130	4503.759

model, the regressors $\mathbf{x}$ are those listed in Table 3.2, specifically: *exploration* and its square, *numerosity*[20], their *interaction* Expl $\times$ Numer, and the time trend.

Table 3.2 shows that:

- as expected, *time* has always a significant and positive effect on performance: as time passes, citations accumulate (and, possibly, inflate);
- reasonably, the *size* of an institution, proxied by Numer, has also a positive[21] and significant effect on performance: this may also be due to the presence of *economies of scale*;
- *exploration* has a U-shaped effect: a negative effect for small values that becomes positive for larger values. This effect remains consistent also when all controls are in place;
- the *interaction* term Expl $\times$ Numer is also significantly positive, when present.
- Lastly, the number of observation is 644 because of the multiplication of the 92 cities under observation by the 7 3-year periods in which the period from 1985 to 2009 is divided. Notice that exploration $\mathbf{\Delta F}^t$ is defined for $t \geq 1$, by construction.

Overall, the results can be interpreted as follows:

- in terms of the direct performance-exploration relationship, two different behaviors seem to pay more: very high specialization (i.e. low exploration) or relatively very high diversification;

[20] Models with Numer^2 were also considered, but the effect were very close to 0 (even considering the relatively large scale of the variable Numer, see Table 3.1), although significant.

[21] In spite of the relatively small numbers in its coefficient, notice that Numer takes values of thousands.

Table 3.2: **Fixed-effects** (within) Regression with Robust Standard Errors for clustering on City, dependent variable: log Performance.

	log Perf	log Perf	log Perf	log Perf	log Perf	log Perf	log Perf	log Perf	log Perf
Expl	-2.974***	-1.143**	-7.971***	-3.617***	-1.649***	-1.047**	-2.541***	-1.802***	-5.529***
	(0.000)	(0.004)	(0.000)	(0.000)	(0.000)	(0.005)	(0.000)	(0.000)	(0.000)
Time		0.141***		0.130***		0.0960***		0.0900***	0.0783**
		(0.000)		(0.000)		(0.000)		(0.000)	(0.001)
Expl^2			14.27***	6.647**					8.995***
			(0.000)	(0.002)					(0.000)
Numer					0.000384***	0.000198***	0.000307***	0.000148**	0.0000955*
					(0.000)	(0.001)	(0.000)	(0.005)	(0.031)
Expl×Numer							0.00177*	0.00143*	0.00212**
							(0.019)	(0.021)	(0.003)
Constant	9.483***	8.730***	9.747***	8.914***	8.940***	8.692***	8.993***	8.750***	9.031***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	644	644	644	644	644	644	644	644	644
R^2	0.147	0.415	0.226	0.430	0.381	0.449	0.406	0.466	0.490
Adjusted R^2	0.145	0.413	0.224	0.428	0.379	0.447	0.403	0.462	0.486

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

- because of the signs and significance of the main terms and the interaction term, one can say that:
  - when exploration is increasing, the marginal effect of size/numerosity is also increasing. Since the effect of the size is always positive anyway, this means that it becomes larger and larger, as exploration increases. Intuitively, in presence of high diversification (i.e. high exploration), a larger institution seems to have a more beneficial effect. This could be due to the presence of *economies of scope*.
  - analogously, when the size is increasing, the marginal effect of exploration is also increasing. However, it is important to notice that it starts from being negative and then becomes positive. Indeed, assuming for example, the absence of Expl^2[22]:

$$\frac{\mathrm{d}\,(\log \mathrm{Perf})}{\mathrm{d}\,(\mathrm{Expl})} = -1.802 + 0.00143 \times \mathrm{Numer},$$

so that it is positive if and only if Numer ≥ 1260, which is a value obtained in the data.[23] This means that there is a change in the sign of the effect of exploration on performance, depending on an

[22] I.e. second-to-last column in Table 3.2.
[23] Remember that the overall mean of Numer is 942 with a standard deviation of 976.

institution's size. Analogously, when also Expl^2 is considered, we have that

$$\frac{\mathrm{d}\,(\log \mathrm{Perf})}{\mathrm{d}\,(\mathrm{Expl})} = -5.529 + 8.995 \times \mathrm{Expl} + 0.00212 \times \mathrm{Numer} \geq 0$$

if and only if $\mathrm{Numer} \geq \frac{1}{0.00212}\,(5.529 - 2 \times 8.995 \times \mathrm{Expl})$ and taking the overall average $\langle \mathrm{Expl} \rangle = 0.102$, this effect is positive if and only if $\mathrm{Numer} \geq 1742$, which is also a number attainable in the data.

The results highlighted here suggest that there may be a "tension" between economies of scale (favored by an increase in size) and economies of scope (possibly favored by an increase in differentiation/exploration). A possible explanation for this can be that when the dimension of the institution is small, then it is hard to exploit the positive effect of the economies of scope, making then become more specialized a better option. On the contrary, for larger institutions, the combined effect of economies of scale and scope make an exploratory/differentiating behavior more beneficial.

3.5.2 Auxiliary analysis: Random Effects Model

As an additional/auxiliary analysis, we also consider random effects models, where the individual-specific effects are not permitted to be correlated with the other regressors. This allows us to use also time-invariant regressors, such as the ranking[24] of the institution, although it reduces the dimension of the sample because it is available only for 48 cities out of 92.

The results are reported in Tables 3.3, 3.4 and 3.5 and seem to confirm what already said in the previous section, with the fixed-effects models. Notice that Rank can also be considered as a proxy for an institution's size, its resources and, possibly, other institution-specific effects. As a last remark, it may be interesting noting that not only the signs but also the magnitudes of the coefficients given by the random effects model seem consistent with the estimates provided by the fixed effects model.

3.6 Conclusions and Future Work

In this paper we have assessed the impact of exploratory behaviors on performance. Dealing with academic authors and keeping track of careers and publications spanning several decades permits to define and quantify a notion of exploration which is related to topics and subjects under their study. The analysis supports the existing of a statistically significant relationship between the two quantities which, however, turns out to be non-monotonic (a U-shaped effect) and dependent on other factors and control variables.

The mechanisms operating at the same time are, on the one hand, that of the economies of scale, favored by an institution's increase in size, and on the

[24]The regressions are made when considering the average among the 3 ranking considered.

Table 3.3: **Random-effects** (within) Regression with Robust Std. Err. for clustering on City, dependent variable: log Performance.

	log Perf	log Perf	log Perf	log Perf	log Perf	log Perf	log Perf
Expl	-3.454***	-1.614***	-3.467***	-1.231*	-4.425***	-9.400***	-4.681***
	(0.000)	(0.000)	(0.000)	(0.018)	(0.000)	(0.000)	(0.000)
Time		0.135***		0.136***	0.122***		0.120***
		(0.000)		(0.000)	(0.000)		(0.000)
Rank			-0.233***	-0.259***		-0.240***	-0.259***
			(0.000)	(0.000)		(0.000)	(0.000)
Expl^2					7.632***	20.69***	11.12***
					(0.000)	(0.000)	(0.000)
Constant	9.532***	8.804***	10.69***	10.03***	9.009***	10.96***	10.26***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	644	644	336	336	644	336	336

p-values in parentheses

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

Table 3.4: **Random-effects** (within) Regression with Robust Std. Err. for clustering on City, dependent variable: log Performance.
Here a control for institution's size (proxied by Numer) is added.

	log Perf	log Perf	log Perf	log Perf	log Perf	log Perf	log Perf
Expl	-1.920***	-1.623***	-1.590**	-1.277*	-3.510***	-4.764***	-4.151***
	(0.000)	(0.000)	(0.005)	(0.011)	(0.000)	(0.000)	(0.000)
Numer	0.000458***	0.000372***	0.000343***	0.000245***	0.000364***	0.000307***	0.000224***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Time		0.0474*		0.0557**	0.0403*		0.0489*
		(0.011)		(0.009)	(0.027)		(0.011)
Rank			-0.165***	-0.194***		-0.176***	-0.200***
			(0.000)	(0.000)		(0.000)	(0.000)
Expl^2					4.993**	10.32***	9.226***
					(0.008)	(0.000)	(0.000)
Constant	8.889***	8.761***	9.834***	9.811***	8.904***	10.06***	10.02***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	644	644	336	336	644	336	336

p-values in parentheses

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

Table 3.5: **Random-effects** (within) Regression with Robust Std. Err. for clustering on City, dependent variable: log Performance.
Here controls for size and its interaction with exploration are added.

	log Perf	log Perf	log Perf	log Perf	log Perf	log Perf	log Perf
Expl	-3.142***	-2.863***	-2.970***	-2.607**	-6.990***	-7.874***	-7.248***
	(0.000)	(0.000)	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)
Numer	0.000365***	0.000310***	0.000274***	0.000205***	0.000244***	0.000196***	0.000149***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Expl×Numer	0.00240**	0.00226**	0.00162	0.00147	0.00299***	0.00220*	0.00204*
	(0.005)	(0.005)	(0.068)	(0.072)	(0.001)	(0.012)	(0.016)
Time		0.0332		0.0436*	0.0231		0.0326
		(0.067)		(0.047)	(0.205)		(0.108)
Rank			-0.153***	-0.176***		-0.163***	-0.181***
			(0.000)	(0.000)		(0.000)	(0.000)
Expl^2					10.16***	14.34***	13.32***
					(0.000)	(0.000)	(0.000)
Constant	8.951***	8.857***	9.875***	9.852***	9.166***	10.21***	10.17***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	644	644	336	336	644	336	336

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

other hand, that of the economies of scope, which are favored by differentiation in a more exploratory department. Interestingly, the interplay between the two could be used to explain the non-trivial relationship that emerges from our analysis.

Future work The present analysis on the one hand confirms the presence of some significant signals concerning the relationship between exploration and performance, but, on the other hand, also suggests that there are some aspects to be investigated more carefully. Some could be tackled once the access to more comprehensive datasets is achieved: indeed, our plan is to replicate the current analysis with the *Web of Science* database and/or *Microsoft Academic Graph*. The extensiveness of these databases will hopefully allow a deeper comprehension of these phenomena and also, for example, allow cross-subject comparisons.

Another important aspect that could be improved concerns the controls on institutions' size, resources and other somehow exogenous factors (to name but a few: with data about funding and grants from the U.S. National Science Foundation, local GDP, education levels, etc). Such analyses could shed light in the impact of academic research on society, by better framing academic institutions within their specific geo-social environments.

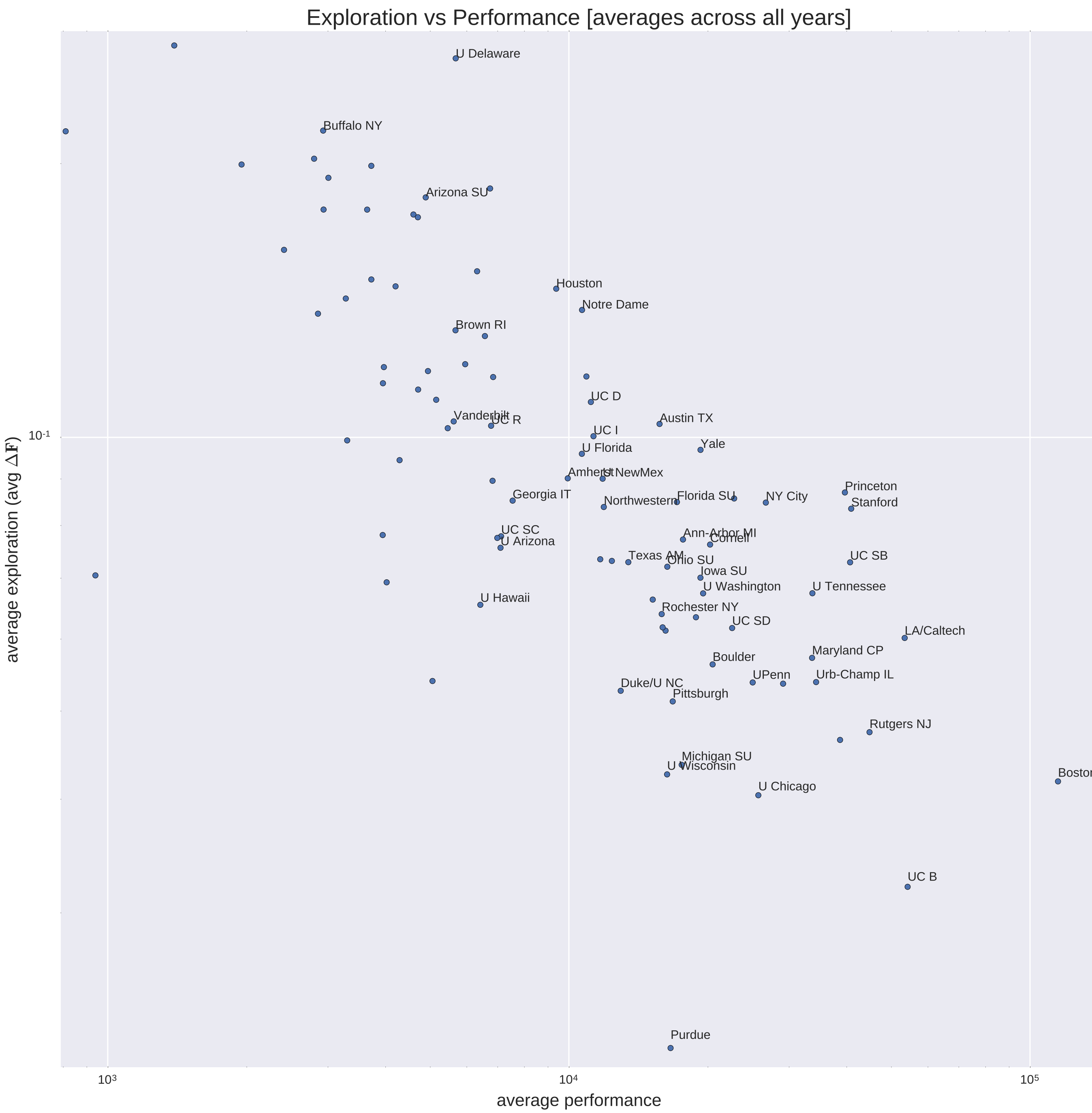
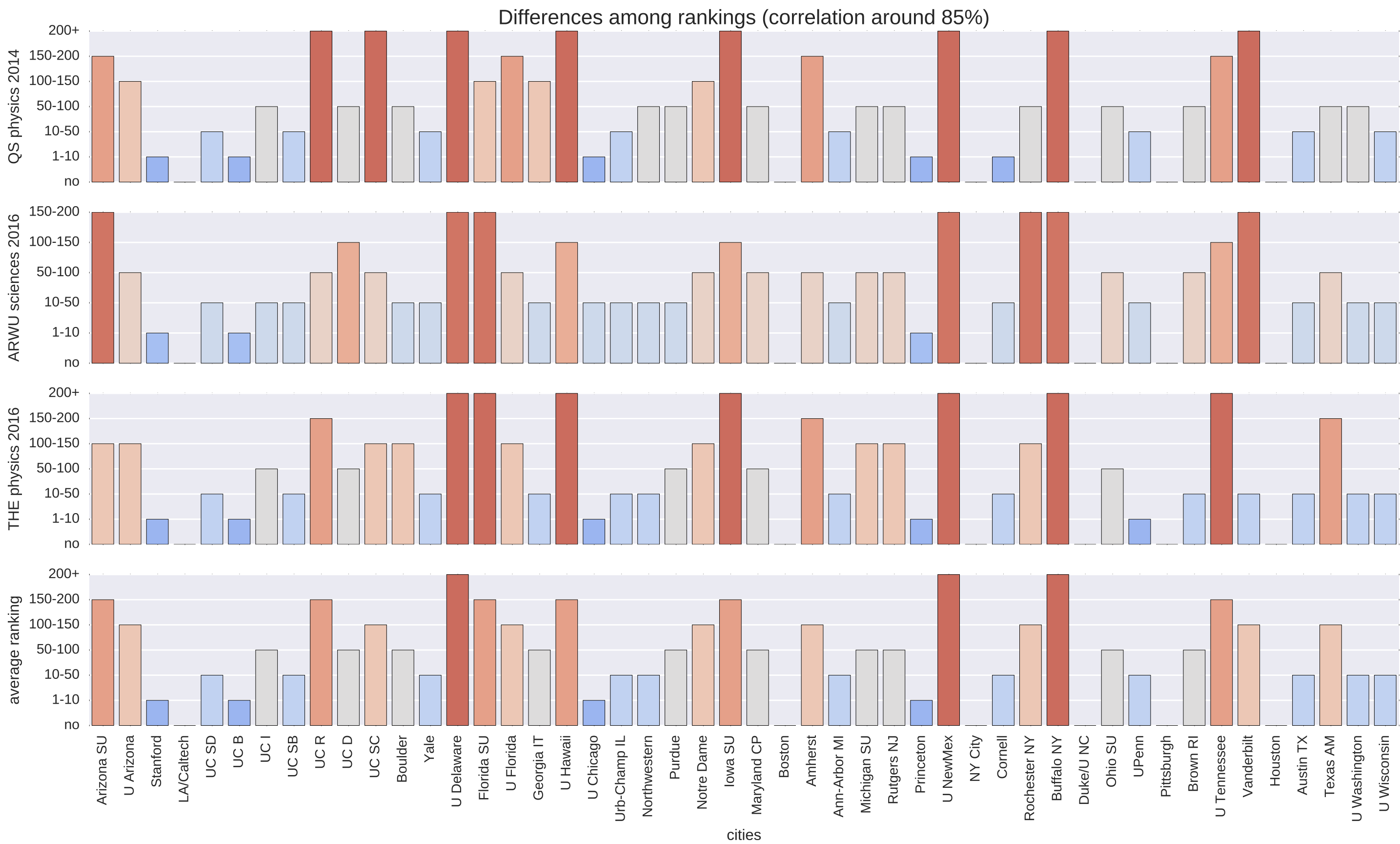


Figure 3.5: **Relationship between average exploration and performance** The figure shows average exploration ($\Delta\mathbf{F}$) on the vertical axis and average performance on the horizontal one, when both are averaged across all time periods (here the time units are 3-year time windows). Of the 92 cities under consideration and plotted here, those with a name are the 48 cities/institutions for which a "rank" is available. The availability of rankings (i.e. presence of names) only in the right-hand side of the figure clearly reflects the existing correlation between highly performing institution and high prestige (Notice that our data on rankings are basically cut at the 200th position.)



Differences among rankings Some areas (associated with "no"), such as Los Angeles/Pasadena, Cambridge/Boston or New York City, cannot be ranked because they contain institutions that are in very different tiers. The numbers on the vertical axes refer to the rank bracket of the corresponding institution/city. The correlation between any two of these rankings is around 85%. The average among these 3 rankings is shown (computed by rounding the average of the three to the closest integer) and is used in the econometric analysis.

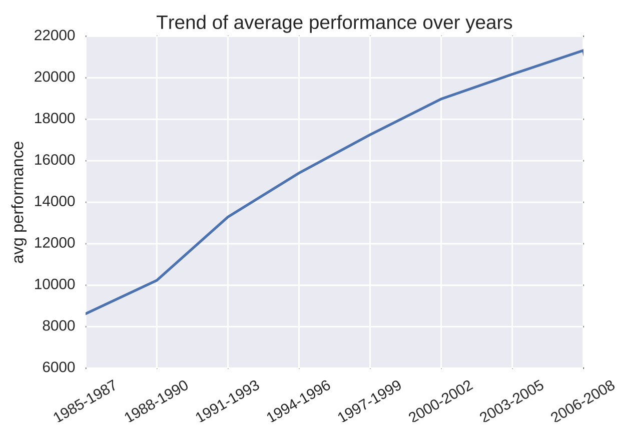


Figure 3.6: **Average performance over the years** The average is computed across all cities. Notice that in our approach a paper published in year t contributes to the performance computed in that specific year. Instead, other approaches may consider the accumulation of the citations over the course of the years.

Bibliography

[Abramitzky, 2015] Abramitzky, R. (2015). Economics and the modern economic historian. Technical report, National Bureau of Economic Research.

[Acemoglu et al., 2015] Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A. (2015). Systemic risk and stability in financial networks. *American Economic Review*, 105(2):564–608.

[Albert and Barabási, 2002] Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47.

[Arino and Van den Driessche, 2003] Arino, J. and Van den Driessche, P. (2003). A multi-city epidemic model. *Mathematical Population Studies*, 10(3):175–193.

[Bak et al., 1988] Bak, P., Tang, C., and Wiesenfeld, K. (1988). Self-organized criticality. *Physical review A*, 38(1):364.

[Bass, 1969] Bass, F. M. (1969). A new product growth for model consumer durables. *Management Science*, 15(5):215–227.

[Boudreau et al., 2016] Boudreau, K. J., Guinan, E. C., Lakhani, K. R., and Riedl, C. (2016). Looking across and looking beyond the knowledge frontier: Intellectual distance, novelty, and resource allocation in science. *Management Science*, 62(10):2765–2783.

[Bramoullé and Kranton, 2007] Bramoullé, Y. and Kranton, R. (2007). Risk-sharing networks. *Journal of Economic Behavior & Organization*, 64(3):275–294.

[Brauer and van den Driessche, 2001] Brauer, F. and van den Driessche, P. (2001). Models for transmission of disease with immigration of infectives. *Mathematical Biosciences*, 171(2):143–154.

[Castillo-Chavez and Yakubu, 2001] Castillo-Chavez, C. and Yakubu, A.-A. (2001). Dispersal, disease and life-history evolution. *Mathematical Biosciences*, 173(1):35–53.

[Cavoretto et al., 2011] Cavoretto, R., Chaudhuri, S., De Rossi, A., Menduni, E., Moretti, F., Rodi, M. C., Venturino, E., Simos, T. E., Psihoyios, G., Tsitouras, C., et al. (2011). Approximation of dynamical system's separatrix curves. In *AIP Conference Proceedings-American Institute of Physics*, volume 1389, page 1220.

[Cherif et al., 2015] Cherif, A., Barley, K., and Hurtado, M. (2015). Homo-psychologicus: Reactionary behavioural aspects of epidemics. *Epidemics*.

[Chowell and Nishiura, 2014] Chowell, G. and Nishiura, H. (2014). Transmission dynamics and control of ebola virus disease (evd): a review. *BMC medicine*, 12(1):196.

[Clauset et al., 2017] Clauset, A., Larremore, D. B., and Sinatra, R. (2017). Data-driven predictions in the science of science. *Science*, 355(6324):477–480.

[Colizza and Vespignani, 2008] Colizza, V. and Vespignani, A. (2008). Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations. *Journal of theoretical biology*, 251(3):450–467.

[Cui et al., 2006] Cui, J., Takeuchi, Y., and Saito, Y. (2006). Spreading disease with transport-related infection. *Journal of Theoretical Biology*, 239(3):376–390.

[Cui et al., 2008] Cui, J.-a., Tao, X., Zhu, H., et al. (2008). An SIS infection model incorporating media coverage. *Rocky Mountain J. Math*, 38(5):1323–1334.

[D'Alessandro, 2007] D'Alessandro, S. (2007). Non-linear dynamics of population and natural resources: The emergence of different patterns of development. *Ecological Economics*, 62(3):473–481.

[DeSolla Price, 1965] DeSolla Price, D. (1965). Networks of scientific papers. *Science*, pages 510–515.

[Dimitri, 2015] Dimitri, N. (2015). The economics of epidemic diseases. *PloS one*, 10(9):e0137964.

[Ductor, 2015] Ductor, L. (2015). Does co-authorship lead to higher academic productivity? *Oxford Bulletin of Economics and Statistics*, 77(3):385–407.

[Ductor et al., 2014] Ductor, L., Fafchamps, M., Goyal, S., and van der Leij, M. J. (2014). Social networks and research output. *Review of Economics and Statistics*, 96(5):936–948.

[Duncan, 1976] Duncan, R. B. (1976). The ambidextrous organization: Designing dual structures for innovation. *The management of organization*, 1:167–188.

[Ehrhardt et al., 2006] Ehrhardt, G. C., Marsili, M., and Vega-Redondo, F. (2006). Phenomenological models of socioeconomic network dynamics. *Physical Review E*, 74(3):036106.

[Fafchamps et al., 2010] Fafchamps, M., Leij, M. J., and Goyal, S. (2010). Matching and network effects. *Journal of the European Economic Association*, 8(1):203–231.

[Fenichel et al., 2011] Fenichel, E. P., Castillo-Chavez, C., Ceddia, M., Chowell, G., Parra, P. A. G., Hickling, G. J., Holloway, G., Horan, R., Morin, B., Perrings, C., et al. (2011). Adaptive human behavior in epidemiological models. *Proceedings of the National Academy of Sciences*, 108(15):6306–6311.

[Foster et al., 2015] Foster, J. G., Rzhetsky, A., and Evans, J. A. (2015). Tradition and innovation in scientists research strategies. *American Sociological Review*, 80(5):875–908.

[Funk et al., 2009] Funk, S., Gilad, E., Watkins, C., and Jansen, V. A. (2009). The spread of awareness and its impact on epidemic outbreaks. *Proceedings of the National Academy of Sciences*, 106(16):6872–6877.

[Funk et al., 2010] Funk, S., Salathé, M., and Jansen, V. A. (2010). Modelling the influence of human behaviour on the spread of infectious diseases: a review. *Journal of the Royal Society Interface*, 7(50):1247–1256.

[Galeotti and Goyal, 2010] Galeotti, A. and Goyal, S. (2010). The law of the few. *The American Economic Review*, 100(4):1468–1492.

[Galeotti and Rogers, 2015] Galeotti, A. and Rogers, B. W. (2015). Diffusion and protection across a random graph. *Network Science*, 3(03):361–376.

[Gilsing and Nooteboom, 2006] Gilsing, V. and Nooteboom, B. (2006). Exploration and exploitation in innovation systems: The case of pharmaceutical biotechnology. *Research Policy*, 35(1):1–23.

[Gilsing et al., 2008] Gilsing, V., Nooteboom, B., Vanhaverbeke, W., Duysters, G., and van den Oord, A. (2008). Network embeddedness and the exploration of novel technologies: Technological distance, betweenness centrality and density. *Research policy*, 37(10):1717–1731.

[Gomes et al., 2014] Gomes, M. F., Piontti, A., Rossi, L., Chao, D., Longini, I., Halloran, M. E., and Vespignani, A. (2014). Assessing the international spreading risk associated with the 2014 west african ebola outbreak. *PLOS Currents Outbreaks*, 1.

[Goyal and Vigier, 2015] Goyal, S. and Vigier, A. (2015). Interaction, protection and epidemics. *Journal of Public Economics*, 125:64–69.

[Granovetter, 1973] Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, 78(6):1360–1380.

[Gross et al., 2006] Gross, T., DLima, C. J. D., and Blasius, B. (2006). Epidemic dynamics on an adaptive network. *Physical review letters*, 96(20):208701.

[Halloran et al., 2014] Halloran, M. E., Vespignani, A., Bharti, N., Feldstein, L. R., Alexander, K., Ferrari, M., Shaman, J., Drake, J. M., Porco, T., Eisenberg, J., et al. (2014). Ebola: mobility data. *Science (New York, NY)*, 346(6208):433–433.

[Hamermesh, 2013] Hamermesh, D. S. (2013). Six decades of top economics publishing: Who and how? *Journal of Economic Literature*, 51(1):162–172.

[Hills et al., 2015] Hills, T. T., Todd, P. M., Lazer, D., Redish, A. D., Couzin, I. D., Group, C. S. R., et al. (2015). Exploration versus exploitation in space, mind, and society. *Trends in cognitive sciences*, 19(1):46–54.

[Horan et al., 2015] Horan, R. D., Fenichel, E. P., Finnoff, D., and Wolf, C. A. (2015). Managing dynamic epidemiological risks through trade. *Journal of Economic Dynamics and Control*, 53:192–207.

[Hsieh et al., 2007] Hsieh, Y.-H., van den Driessche, P., and Wang, L. (2007). Impact of travel between patches for spatial spread of disease. *Bulletin of mathematical biology*, 69(4):1355–1375.

[Hudson, 1996] Hudson, J. (1996). Trends in multi-authored papers in economics. *The Journal of Economic Perspectives*, 10(3):153–158.

[Hufnagel et al., 2004] Hufnagel, L., Brockmann, D., and Geisel, T. (2004). Forecast and control of epidemics in a globalized world. *Proceedings of the National Academy of Sciences of the United States of America*, 101(42):15124–15129.

[Jia et al., 2017] Jia, T., Wang, D., and Szymanski, B. K. (2017). Quantifying patterns of research-interest evolution. *Nature Human Behaviour*, 1:0078.

[Jin and Wang, 2005] Jin, Y. and Wang, W. (2005). The effect of population dispersal on the spread of a disease. *Journal of Mathematical Analysis and Applications*, 308(1):343–364.

[Katz and Shapiro, 1985] Katz, M. L. and Shapiro, C. (1985). Network externalities, competition, and compatibility. *The American economic review*, 75(3):424–440.

[Lee et al., 2003] Lee, J., Lee, J., and Lee, H. (2003). Exploration and exploitation in the presence of network externalities. *Management Science*, 49(4):553–570.

[Levinthal and March, 1993] Levinthal, D. A. and March, J. G. (1993). The myopia of learning. *Strategic management journal*, 14(S2):95–112.

[Levitt and March, 1988] Levitt, B. and March, J. G. (1988). Organizational learning. *Annual review of sociology*, 14(1):319–338.

[Manfredi and D'Onofrio, 2013] Manfredi, P. and D'Onofrio, A. (2013). *Modeling the interplay between human behavior and the spread of infectious diseases*. Springer Science & Business Media.

[March, 1991] March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization science*, 2(1):71–87.

[Meloni et al., 2011] Meloni, S., Perra, N., Arenas, A., Gómez, S., Moreno, Y., and Vespignani, A. (2011). Modeling human mobility responses to the large-scale spreading of infectious diseases. *Scientific reports*, 1.

[Newman, 2004] Newman, M. E. (2004). Coauthorship networks and patterns of scientific collaboration. *Proceedings of the national academy of sciences*, 101(suppl 1):5200–5205.

[O'Reilly and Tushman, 2013] O'Reilly, C. A. and Tushman, M. L. (2013). Organizational ambidexterity: Past, present, and future. *The Academy of Management Perspectives*, 27(4):324–338.

[Pan et al., 2016] Pan, R. K., Petersen, A. M., Pammolli, F., and Fortunato, S. (2016). The memory of science: Inflation, myopia, and the knowledge network. *arXiv preprint arXiv:1607.05606*.

[Petersen et al., 2010] Petersen, A. M., Wang, F., and Stanley, H. E. (2010). Methods for measuring the citations and productivity of scientists across time and discipline. *Physical Review E*, 81(3):036114.

[Poletti et al., 2012] Poletti, P., Ajelli, M., and Merler, S. (2012). Risk perception and effectiveness of uncoordinated behavioral responses in an emerging epidemic. *Mathematical Biosciences*, 238(2):80–89.

[Posen and Levinthal, 2012] Posen, H. E. and Levinthal, D. A. (2012). Chasing a moving target: Exploitation and exploration in dynamic environments. *Management Science*, 58(3):587–601.

[Radicchi et al., 2008] Radicchi, F., Fortunato, S., and Castellano, C. (2008). Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences*, 105(45):17268–17272.

[Reluga, 2009] Reluga, T. C. (2009). An sis epidemiology game with two subpopulations. *Journal of Biological Dynamics*, 3(5):515–531.

[Reluga, 2010] Reluga, T. C. (2010). Game theory of social distancing in response to an epidemic. *PLoS Comput Biol*, 6(5):e1000793.

[Sahneh et al., 2012] Sahneh, F. D., Chowdhury, F. N., and Scoglio, C. M. (2012). On the existence of a threshold for preventive behavioral responses to suppress epidemic spreading. *Scientific reports*, 2.

[Seltzer and Hamermesh, 2017] Seltzer, A. and Hamermesh, D. S. (2017). Co-authorship in economic history and economics: Are we any different? Technical report, National Bureau of Economic Research.

[Shapiro and Varian, 2013] Shapiro, C. and Varian, H. R. (2013). *Information rules: a strategic guide to the network economy*. Harvard Business Press.

[Shaw and Schwartz, 2008] Shaw, L. B. and Schwartz, I. B. (2008). Fluctuating epidemics on adaptive networks. *Physical Review E*, 77(6):066101.

[Sun et al., 2011] Sun, C., Yang, W., Arino, J., and Khan, K. (2011). Effect of media-induced social distancing on disease transmission in a two patch setting. *Mathematical biosciences*, 230(2):87–95.

[Thomas et al., 2015] Thomas, M. R., Smith, G., Ferreira, F. H., Evans, D., Maliszewska, M., Cruz, M., Himelein, K., and Over, M. (2015). The economic impact of ebola on sub-saharan africa: updated estimates for 2015.

[Uzzi et al., 2013] Uzzi, B., Mukherjee, S., Stringer, M., and Jones, B. (2013). Atypical combinations and scientific impact. *Science*, 342(6157):468–472.

[Uzzi and Spiro, 2005] Uzzi, B. and Spiro, J. (2005). Collaboration and creativity: The small world problem 1. *American journal of sociology*, 111(2):447–504.

[Wang and Li, 2014] Wang, L. and Li, X. (2014). Spatial epidemiology of networked metapopulation: An overview. *Chinese Science Bulletin*, 59(28):3511–3522.

[Wang and Mulone, 2003] Wang, W. and Mulone, G. (2003). Threshold of disease transmission in a patch environment. *Journal of Mathematical Analysis and Applications*, 285(1):321–335.

[Wang and Zhao, 2004] Wang, W. and Zhao, X.-Q. (2004). An epidemic model in a patchy environment. *Mathematical Biosciences*, 190(1):97–112.

[Wang et al., 2015] Wang, Z., Andrews, M. A., Wu, Z.-X., Wang, L., and Bauch, C. T. (2015). Coupled disease–behavior dynamics on complex networks: A review. *Physics of life reviews*, 15:1–29.

[Wuchty et al., 2007] Wuchty, S., Jones, B. F., and Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, 316(5827):1036–1039.

[Zanette and Risau-Gusmán, 2008] Zanette, D. H. and Risau-Gusmán, S. (2008). Infection spreading in a population with evolving contacts. *Journal of biological physics*, 34(1-2):135–148.

[Zhang et al., 2013] Zhang, Q., Perra, N., Gonçalves, B., Ciulla, F., and Vespignani, A. (2013). Characterizing scientific production and consumption in physics. *Scientific reports*, 3.

[Zhang et al., 2016] Zhang, X., Wang, D., and Wang, T. (2016). Inspiration or preparation?: Explaining creativity in scientific enterprise. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pages 741–750. ACM.

[Zhu and Ghahramani, 2002] Zhu, X. and Ghahramani, Z. (2002). Learning from labeled and unlabeled data with label propagation.