A supply-side story for a threshold model:

Endogenous growth of open collaborative innovation communities

Francesco Rullani
Department of Economics and Business
LUISS Guido Carli
E-mail: frullani@luiss.it

Lorenzo Zirulia

Department of Economics, University of Bologna
CRIOS, Bocconi University
RCEA
E-mail: lorenzo.zirulia@unibo.it
20 May 2013

Abstract

This paper takes an institutional look (as Dasgupta and David do for Science and Technology) at open collaborative innovation communities (Community). Drawing from the community of practice literature to describe communitarian social processes, we develop a model in which Community is confronted with Technology with respect to its ability to attract researchers. We find that the number of individuals that initially chooses each institution is crucial, as it determines a threshold size that divides the realm of communities doomed to remain small from the set of communities that are able to grow endogenously fast and large. We examine how communities can reach that threshold and discuss this result in light of the strategies firms that invest in communities can apply to exploit this effect. We also discuss how changing the level of openness protection and the importance of the social environment in Community affect innovativeness and find that what really solves any ambiguity in this sense is the way Technology, not Community, is structured. We finally discuss the policy implications of this effect.

***JEL classification**: O31, L86, L88*

***Key Words**: open collaborative communities, science, technology, free and open source*

Introduction

One of the main challenges imposed by the development of the so-called "knowledge society" on economic theory is the assessment of the changes in institutions that enable knowledge production and diffusion. Moving from the production of physical goods to the production of knowledge, in fact, implies a reshaping of the structures upon which the economy has been based.

The economic discourse on institutions connected to knowledge production has its modern origins in the work by Dasgupta and David (1987, 1994), who recognise two main institutional models, namely "Science" and "Technology", whose real manifestations are the academic world for the former and the markets for technology (Arora et al., 2001) for the latter. Science is based on disclosure, rewards from priority, and peer recognition as well as, today, on the public funding of knowledge production. Technology is based on secrecy and/or intellectual property rights (IPR) and is profit-motivated. However, in the "shadows" of this dual system, we have observed the emergence of a series of examples of an open model of knowledge production, what David and Foray (2003) call knowledge-intensive communities. These communities are characterised by a significant number of members who produce and reproduce knowledge in a "public" (often virtual) space in which new information and communication technologies are intensively used to codify and transmit knowledge. In this work, we focus on Baldwin and von Hippel's (2011) version of knowledge-intensive communities, referred to as *open collaborative innovation communities*. In such communities, agents collaboratively develop and openly distribute knowledge without direct external funding or rents assured by usual IPR, which is particularly interesting in our case because of the importance assigned to the concept of openness, i.e., to the fact that anyone can acquire, study, modify, and redistribute the product created collectively. Openness is typical of these communities, but it is not granted. Indeed, many communities have created tools – from legal tools such as licenses to independent organisations whose missions are to manage the community's intellectual property – that can preserve openness against possible appropriation (O'Mahony, 2003).

One of the most prominent examples of open collaborative innovation communities, in terms of economic and social impact, is the free and open source software community. In this community, a large number of individuals (David and Rullani, 2008) spread all over the world (Gonzalez-Barahona et al., 2008) cooperate online to create software and release it openly through the Internet. Anyone can enter the production process and report bugs, propose patches, cooperate with other developers on existing software, or launch new projects; however, thanks to the license scheme adopted by the community (mostly the General Public License, GPL), no one can appropriate the jointly developed software. Openness in this case is preserved via “copyleft”, i.e., via licenses that prevent the appropriation and forcing of developers subsequently acting on the original code to redistribute the improved software under the original open terms.

Even in this circumstance, firms have created business models able to leverage the capabilities of the community and create a positive coexistence with it (Dahlander and Magnusson, 2005; Dahlander and Wallin, 2006; Dahlander, 2007; Bonaccorsi et al., 2006; Fosfuri et al., 2008). Other examples of the specific forms this model can take include Wikipedia, collective invention (Allen, 1983, Nuvolari, 2004), or the communities of user innovators (Jeppesen and Frederiksen, 2006; von Hippel, 1988).

This paper develops a formal analysis where an open collaborative innovation community (simply “Community” hereafter) confronts Technology with respect to the ability of attracting researchers. Our model has two distinctive features. First, we represent the broad set of motivations that affects the functioning of Community and the individual choices between two institutions. In addition to the individual motives put forth by the early literature on open source (Raymond, 1998a; von Hippel, 2001; Lerner and Tirole, 2002; Torvalds and Cox, 2003), i.e., signalling and reputation, fun, and own-use, we pay particular attention to the social dimension of communities, introducing explicitly in our model insights from the community of practice literature (Wenger, 1998). This literature is here used as a framework to deal with those social processes recently identified as crucial to generating community members’ incentives (e.g., Bagozzi and Dholakia, 2006).

Second, our model explicitly takes into account the role of each researcher’s externalities (“spillovers”) towards those who work in the same institution and in other institutions, a mechanism usually left in the background in the literature or captured simply as the possibility of all agents enjoying the openly distributed content (Gambardella and Hall, 2006). A key point is that Community and Technology differ in the sign of such externalities. While the openness of Community generates positive spillovers both within the institution and towards Technology, the appropriability strategies prevailing in Technology remove a larger proportion of knowledge from the inputs other actors can openly access and thus lead to negative externalities in both cases.

Our results refer both to the "size" of each institution (i.e., the number of researchers who choose it) and to its performance, as measured by degree of innovativeness. In terms of group size, we find that when the social processes taking place in Community generate important incentives for its members, a threshold (in terms of the number of individuals initially choosing each institution) divides the realm of communities doomed to remain small from the set of communities that are able to grow *endogenously* fast and large. Path dependence (David, 1985) is observed in the growth of Communities. This threshold has been widely recognised in the literature about communities (e.g., Bonaccorsi and Rossi, 2003); what is new in our argument is that the threshold is not based on demand factors, but on the structure of developer motivation, i.e., on supply-side factors. This peculiar feature of the model is particularly useful to inform managers and project leaders on the strategies they should use to create communities around their projects. Initial steps are much more important for further efforts, possibly the only important ones. Being able to overcome threshold size at the beginning is what distinguishes a community capable of future growth from those doomed to remain small.

In this context, Communities whose protection mechanisms, such as licenses, are effective at preserving the openness of the produced knowledge and which prevent Technology from appropriating it are more likely to be established and grow by reducing Technology attractiveness and thus triggering the endogenous mechanism of Community growth.

As for innovativeness, the higher importance of motives triggered by the social processes taking place in Community directly affects the performances of both institutions. Contrary to what one can expect, from the Community perspective what matters most is the level of individuals' overall investments in Technology rather than in Community: if this increases, then the negative spillover from Technology towards Community may offset the positive effect of larger intra-group spillovers due to the larger Community. From a Technology perspective, innovativeness may instead increase due to the larger positive spillovers from Community.

Community's inability to protect openness, and thus the larger amount of knowledge that Technology can appropriate (i.e., larger spillovers form Community to Technology), not only leads to smaller communities, but when they also trigger negative spillovers from Technology, unambiguously reduces their innovativeness. At the same time, Technology innovativeness may be reduced if researchers in Community significantly reduce their efforts. Again, what determines the overall level of innovativeness is how Technology is structured. Policymakers need to keep this in mind when trying to foster knowledge production that favours the diffusion of open collaborative innovation communities.

The remainder of this paper is organised as follows. Section 2 elaborates on the appreciative theory of communities as knowledge-related institutions upon which the model is based, by focusing in particular on researchers' motivations. Section 3 describes the model and put it the perspective of the existing, formal literature on free and open source communities. Section 4 derives the results, and Section 5 discusses their properties in light of the discussion in Section 2. Finally, Section 6 presents managerial and policy implications and concludes.

1 The open collaborative innovation community as a knowledge-related institution

This section develops an appreciative theory of the institutional status of open collaborative innovation communities, which is then formalised in Section 3. In particular, we focus on the motivations that prevail in this institution in order to characterise the main determinants of the payoff of the individuals acting in Community. As for the empirical literature to support our argument, we mainly refer to free and open source communities, which have been the most studied example of such communities. In Section 3, we confront them with the motivations that prevail in Technology, whose characteristics are well known since the description of the dual system by Dasgupta and David (1987, 1994).

1.1 *Established motivations in Community*

The study of individual motivations for participation has been a central concern in the literature on free and open source communities (e.g., Nuvolari and Rullani, 2007), which, as we said, is a leading example of open collaborative innovation communities. Bezroukov (1999a, 1999b) argues that the structure of incentives and the organisation of the collaborative efforts of developers and scientists are very close to one another, given that they are both based on rules that connect the openness of the results to the individual pursuit of recognition and reputation (see also Lerner and Tirole, 2002). Dalle and David (2003) also share a similar point of view, stressing the parallelism between the free and open source institutional setting and the rules of "open science", where "the norm of openness is incentive compatible with a collegiate reputational reward system based upon accepted claims to priority" (Dalle and David, 2003, pp. 3–4). A similar point is made by Raymond (1998a), who suggests that the correspondence between the two phenomena is just the outcome of the fact that scientific and free and open source enterprises have simply given the same answer to the same problem of collective knowledge production.

Individuals may also seek peers' regard (Dalle and David, 2005) for more instrumental reasons, such as a means to reach a financial reward. For example, they may want to signal their ability as programmers to the job market in the hope that a company may hire them (Lerner and Tirole, 2002). An empirical account of the role of monetary incentives of this kind is offered by Roberts et al. (2006), who show that when developers are moved by status-related motivations, being paid increases contribution.

In addition to reputation-based incentives that relate peer-judgment to possible psychological and financial rewards (Lerner and Tirole, 2002), own-use has also been underlined as an important motivation in free and open source communities. This relates to the literature inspired by von Hippel (1988), which highlights the role of users as a source of innovation in a wide range of fields (e.g., sports equipment, as in Franke and Shah, 2003). In the software case, an individual who has the knowledge and tools to develop software can easily customise the software he or she uses and even produce what he or she needs (von Hippel, 2001). As Bessen (2006) shows, in fact, software is a complex good that can be personalised much more effectively by skilled users than it can by manufacturers. Once produced, software is inexpensive to exchange through the Internet, so that even a small reward can push developers to exchange the codes they have written (von Hippel and von Krogh, 2003). A final important motivation was originally put forward by software developers (Raymond, 1998a; Torvalds and Diamond, 2001) and it was only later empirically proven and theoretically discussed by scholars (Lakhani and Wolf, 2005): fun. Indeed, many software developers create code in their spare time as a leisure activity (David et al., 2001) because they consider it to be a way to "scratch a personal itch" (Raymond, 1998b). When developers find a coding challenge that matches their interest and skills, they enter a state of "flow" (Lakhani and Wolf, 2005) where enjoyment is maximised by devoting attention only to the code and the development activity.

Therefore, signalling one's talent, reputation, fun, and own-use are the main individual incentives for action in free and open source communities. Surveys and empirical studies such as the FOSS-EU survey (Ghosh et al., 2002), Boston Consulting Group survey (Lakhani et al., 2002), and many others (Bonaccorsi and Rossi, 2006; David and Shapiro, 2008) confirm that own-use-related incentives and psychological motivations such as fun are among the most important drivers, while reputation and signalling are less crucial (Lakhani and Wolf, 2005). In addition, they find that the desire to learn from others is also a fundamental incentive to join the collective production of code (von Hippel and von Krogh, 2003; Ghosh et al., 2002).

1.2 Social motivations in Community

In the context of the free and open source community, the social dimension has been analysed with respect to theories such as gift economy (Raymond, 1998a), communities of practice, and epistemic communities (Amin and Cohendet, 2004; Lin, 2003, 2004a, 2004b; Mateos-Garcia and Steinmueller, 2008). In particular, the community of practice perspective (Wenger, 1998) can be particularly useful to describe in detail the passages of social processes at work in the free and open source community based on mutual learning (another crucial incentive).

Applying this perspective to the free and open source community means recognising the central role of a specific learning process of "negotiation" (Lin, 2004a) that developers are involved in when creating software. Developing a common project together forces people to interact and compare their visions of the problems, possible solutions, and actions. This collective sense-making activity can be thought of as a continuous “renegotiation” of the meanings connected with developers’ own actions, giving sense to the common activity and to the social context in which it takes place. This negotiation of meanings leads to a continuous reshaping of participants’ visions of the world to adapt their *identities* to the social circumstances and opinions they are exposed to in the community (Wenger, 1998). Changing individuals’ identities means configuring in a new way the principles that guide their actions and priorities, namely those principles normally represented in economics by their payoff functions. In other words, the interaction between community members leads to changes in their identities that ultimately result in a modification of the importance of the elements of their payoff functions to take into account the priorities and rules shared by the whole community (Muller, 2006).

An empirical account of this process can be found in Bagozzi and Dholakia (2006) and Shah (2006). Shah (2006) describes the evolution in developers’ motivations as follows: “... a need for software-related improvements drives initial participation. The majority of participants leave the community once their needs are met, however, a small subset remains involved. For this set of developers, motives evolve over time and participation becomes a hobby” (p. 1000). Among the possible explanations for this process, the author also identifies the hypothesis that the “interaction with the community leads to a shift in the individual’s identity and self-perception” (Shah, 2006, p. 1011). This is the perspective taken by Bagozzi and Dholakia (2006), who write: “Initial participation by novice users is driven by specific task-oriented goals .... But over time, as the user comes to form deeper relationships with other [free and open source community] members, the community metamorphosizes into a friendship group and a social entity with which one identifies” (p. 1111).

If the free and open source community is conceived as a community of practice following Wenger’s (1998) intuition, the social mechanisms described above should act along the nexus of communitarian ties, influencing the structure of the payoff function and the relative importance of its constituting elements. This payoff structure can be adequately represented including and giving importance to three factors taken directly from Wenger's (1998) idea of community of practice.

The first aspect is related to the *communitarian activity*. What makes a group of people become a community is the construction of a social environment where identities are defined through a process that is interwoven with the activity of the community (e.g., in the free and open source case, producing software). All processes take place *in* that social environment and *thanks to* that social environment. Thus, the common activity has a central role in the payoff function and depends on the effort of all members of the community.

The second aspect is *personal involvement*: if a member's identity is strongly tied to the common activity (i.e., the project undertaken by the community), the effect of that activity on his or her payoff function is greater. For example, in the free and open source case, the development of GNU/Linux (the most famous open source operating system) has a greater effect on the payoff of a person who "believes" in the GNU/Linux project compared with the payoff ensured by the simple usage of GNU/Linux. This translates as the model of Wenger's (1998) idea of *engagement*, where individuals are involved in a process of the "renegotiation" of their visions of the world and reciprocal influence between them and the social environment of the community. The more a member invests in – and counts on – the common activity undertaken in the community to define his or her identity, the higher is the psychological payoff he or she gets from that activity. Thus, personal involvement is intended as endogenous to the development of the community, as the previous quote from Bagozzi and Dholakia (2006) suggests: it develops and becomes stronger (in terms of affecting members' behaviours) as the "volume" of the interaction grows. Therefore, when a community grows, it not only becomes "quantitatively" stronger (e.g., it produces more software), but also "qualitatively", determining a higher average rate of the personal involvement of its core members. Personal involvement is thus considered to be a function of community size.

The third aspect refers to *coordination costs*: in general, a group of people who collaborate is expected to be subject to free riding episodes. The group must then create some rules and enforcing mechanisms to sustain cooperation and avoid free riding (O'Mahony, 2003). These costs originate from activities such as monitoring others' behaviour, spreading information about it, discovering the breaking of a rule, and punishing the free rider. Further, they increase with the number of community members, suggesting that they limit the growth opportunities of the community.

2 The model

2.1 The structure of the model

Our paper builds on the work presented in Carraro and Siniscalco (2003), and can be summarised in the following terms. A population of N researchers is active in a given field of research. Researchers are identical in terms of both preferences and productivity. They exert effort to produce knowledge, and they can do this in two institutional settings: Technology and Community. It is assumed that researchers, before choosing how much effort to exert, choose which institution they intend to belong, based on an expected payoff comparison. Participation in one institution is exclusive.[1]

Technology and Community differ both in their payoff structures, capturing different motivations, and in the nature of externalities within and between institutions. As for payoff structure, in the field of Technology economic return constitutes the main source motivation: new knowledge is kept secret, embodied in patents, or protected by copyright law. In Community, where new knowledge is disclosed, benefits are associated to the own-use, signalling, and reputation motives (simplified in the model by a single parameter k^C), while the social dimension is linked to the degree of personal involvement and to the communitarian activity. We explicitly introduce a parameter α to capture the relative importance of social motivation. In addition to effort costs, participation in Community involves coordination costs.[2]

Technology and Community differ in terms of both "intra-group" and "inter-group" spillovers from other researchers. In Technology, the knowledge produced by a researcher that chooses this institution has a negative impact on the probability of any others' success in knowledge creation, since the limits imposed

[1] As an alternative interpretation, researchers belong to both institutions, and the relevant decision is the allocation of resources (e.g., time and effort) to each organisation.

[2] Of course, in reality these differences are much more blurred. A researcher that works for a firm but embedded in the scientific debate with his or her colleagues from other firms can reach the same social motivation as an open source developer. Likewise, the latter can find a job in an open source-based company and receive a monetary incentive similar to that of the former. However, we seek to grasp the inner difference between the two institutions, and thus we magnify the difference in the payoff functions they offer to their researchers.

by property rights or secrecy reduce the space for further innovations. Community's externalities, instead, have a positive effect, because openness favours recombination and cumulativeness. However, the magnitude of such externalities may depend on the specific institutional setting. Open collaborative innovation communities may find ways to protect the knowledge they produce from appropriation, protecting in this way their openness (O'Mahony, 2003). In open source, for instance, licenses may limit the rights of others to use the code as inputs for their productions of "closed software", thus reducing spillovers from Community to Technology. In creative industries, Creative Commons licenses may have the same effect. We explicitly introduce a parameter β to model this aspect.

Formally, the payoff from participating in Technology and Community are respectively:

$$\Pi_i^T = \Pr^T(x_i^T, X_{-i}^T, \beta X^C)R^T - c^T(x_i^T) \tag{1}$$

$$\Pi_i^C = \Pr^C(x_i^C, X_{-i}^C, X^T)k^C - c^C(x_i^C) + \alpha e(n)Y(x_i^C, X_{-i}^C, X^T) - C(n) \tag{2}$$

In equations (1) and (2), n $(N\text{-}n)$ denotes the number of researchers in Technology (Community). We define $I = T, C$ as the index for institution and x_i^I as the individual effort of researcher i if active in institution I. X_{-i}^I and X^I represent the sum of efforts of all researchers in i's institution (excluding i) and in the other institution, respectively, $\Pr^I(\circ)$ is the probability of innovation (successful production of knowledge). It is assumed that $\frac{\partial \Pr_i^I}{\partial x_i^I} > 0$ (the higher the effort of the researcher, the higher the probability of innovation is); $\frac{\partial \Pr_i^I}{\partial X_{-i}^T}(\frac{\partial \Pr_i^I}{\partial X_i^T}) < 0$ (externalities from Technology are always negative); and $\frac{\partial \Pr_i^I}{\partial X_{-i}^C}(\frac{\partial \Pr_i^I}{\partial X_i^C}) > 0$ (externalities from Community are always positive).[3] $c_i^I(\circ)$ represents the cost of

[3] We assume, and we use these assumptions in the proof of Proposition 1, that the probability of innovation is strictly concave in x_i^I and that $\frac{\partial^2 \Pr_i^I}{\partial x_i \partial X_{-i}^T}(\frac{\partial^2 \Pr_i^I}{\partial x_i \partial X_i^T}) < 0$ and $\frac{\partial^2 \Pr_i^I}{\partial x_i \partial X_{-i}^C}(\frac{\partial^2 \Pr_i^I}{\partial x_i \partial X_i^C}) > 0$, namely individual efforts are strategic complements with total efforts in Community and strategic substitutes with total efforts in Technology.

effort, with $\frac{\partial c_i^I}{\partial x_i^I} > 0$ and $\frac{\partial^2 c_i^I}{\left(\partial x_i^I\right)^2} > 0$.

In equation (1), R^T represents the economic return from innovation in Technology in the case of success. This can be represented by profits from entrepreneurial activity; alternatively, $\Pr^T(x_i^T, X_{-i}^T, X^C)R^T$ may represent the expected wage for employed software developers. β is a spillover parameter, which takes values in the interval [0;1]. In equation (2), k^C is the "prize" obtained from successful innovating. k^C captures the different motivational dimensions at the individual level, from the increased reputation among peers and in the job market (Lerner and Tirole, 2002) to the possibility of using the produced knowledge (von Hippel, 2001) to the intrinsic enjoyment of the development activity (Lakhani and Wolf, 2005). The other terms in (2) instead unfold the social dimension as defined in the previous section. $e(n)$ is the personal involvement in Community (with $\frac{\partial e(n)}{\partial n} < 0$); $Y(x_i^C, X_{-i}^C, X^T)$ is the value of the communitarian activity, with $\frac{\partial Y}{\partial x_i^C} > 0$, $\frac{\partial Y}{\partial X^T} < 0$ and $\frac{\partial Y}{\partial X_{-i}^C} < 0$[4]; α>0 measures the relative weight of social motivations with respect to individual motivations; and $C(n)$ are the coordination costs expressed as a function of *n*, with $\frac{\partial C(n)}{\partial n} < 0$. $\frac{\partial^2 C(n)}{\partial n^2} > 0$ and $C(N) = 0$.

Researchers' interaction is represented by a two-stage non-cooperative game: in the first stage, each researcher decides whether to enter Technology or Community, while in the second stage, after observing *n*, each agent decides simultaneously his or her effort level. The game is solved backward, computing the optimal effort of each researcher given *N* and *n*. Then, the analysis moves to the first stage, where

---

[4] We also assume that Y is strictly concave in x_i^I and that $\frac{\partial^2 Y_i^I}{\partial x_i \partial X_{-i}^T} < 0$ and $\frac{\partial^2 Y_i^I}{\partial x_i \partial X_{-i}^C} > 0$, namely individual efforts are strategic complements with total efforts in Community and strategic substitutes with total efforts in Technology.

researchers choose the institution that predicts correctly the outcome, and then the payoff, for any value of n. We restrict our attention to pure strategy Nash equilibria in which n^* researchers choose Technology in equilibrium and N-n^* choose Community. Furthermore, we consider only symmetric equilibria in terms of efforts within each institution. Consequently, we define $\Pi^T(n)$ and $\Pi^C(n)$, the reduced-form payoff in the first stage for a researcher choosing Technology and Community, as a function of the number of researchers in Technology. If N is large enough, the determination of an interior equilibrium n^* is well approximated by the condition:

$$\Pi^T(n^*) = \Pi^C(n^*) \tag{3}$$

so that n is treated as a continuous variable. An equilibrium n^* is (locally) stable if:

$$\frac{d\Pi^T(n^*)}{dn} - \frac{d\Pi^C(n^*)}{dn} < 0 \tag{4}$$

which implies that there is a neighbourhood of n^* such that for any n in such a neighbourhood the myopic (with respect to the choice of institution) best response dynamic adjustment process converges to n^*. Informally, an allocation of researchers between Technology and Community is stable if (sufficiently small) exogenous shocks in institution size do not move the equilibrium away (permanently) from the initial configuration.

In addition to their size, Section 4 also assesses institution performance, measured by the expected number of innovations (institution *innovativeness* hereafter). Therefore, innovativeness in Technology is defined as:

$$n^* * \Pr^T(\hat{x}^T(n^*), (n^*-1)\hat{x}^T(n^*), (N-n^*)\hat{x}^C(n^*)) \tag{5}$$

while innovativeness in Community is defined as:

$$(N-n^*) * \Pr^C(\hat{x}^C(n^*), (N-n^*-1)\hat{x}^C(n^*), (n^*)\hat{x}^T(n^*)) \tag{6}$$

As previously mentioned, our framework is based on the model developed by Carraro and Siniscalco (2003), who describe the choice of researchers between Science and Technology. While we represent

Technology along their same lines, two differences exist between our payoff function for participation in Community and their payoff function for participation in Science, which affect the equilibria. First, the role of the State in Science leads to the presence of a fixed salary, which reduces the number of researchers that choose this institution. Second, Carraro and Siniscalco (2003) do not consider the social dimension of the institution, which in our model is captured by the positive value attached to personal involvement and communitarian activity and by the existence of coordination costs. In addition, we model explicitly the magnitude of externalities (from Community to Technology), which allows us to perform comparative statics exercises. Finally, we also discuss the stability properties of equilibria, whereas all equilibria are stable in Carraro and Siniscalco's set-up.

2.2 The model in light of the formal literature on free and open source software

Our paper directly relates to the fast-growing literature that is developing formal models on diverse aspects of free and open source community because of the attention it attracted as an exemplary form of the open collaborative innovation model. A first stream of this literature has looked at the conditions for developers (or user-developers) to contribute to free and open source communities, emphasising supply-side and motivation issues. Johnson (2002), Baldwin and Clark (2006), and Bitzer and Schroder (2005), for instance, consider open source software as a public good and develop a game-theoretic model of contribution by self-interested individuals, while Gambardella and Hall (2006) and Johnson (2006) consider how the competition of the free and open source community and IPR-based system attracts developers. Our work relates to this literature, but it moves some steps away from it. Specifically, on one hand, it builds on the community of practice literature to explicitly analyse the role of social motivations in explaining the relative attractiveness of the community model; on the other hand, it does that by taking into account how spillovers link not only members of the same institutions, but also those part of competing ones.[5]

A second stream of the literature on the formal models of open source has looked at competition between

[5] Baldwin and von Hippel (2011) also analyze the conditions for the economic viability of the open collaborative innovation model, compared with producer and user innovation models, but they do not try to determine their relative importance when more than one model is viable.

proprietary and open source software from consumers' points of view. Among others, Casadesus-Masanell and Ghemawat (2006) develop a dynamic duopoly model between a profit-oriented firm and an open source community. Economides and Katsamakas (2006) consider the two-sided competition between proprietary and open source platforms, with particular attention paid to the incentives for complimentary good production. Lanzi (2009) jointly considers product differentiation, lock-in and network externalities, and consumers' experiences in software use and implementation. Finally, Dalle and Jullien (2003) and Bonaccorsi and Rossi (2003) take a technology diffusion perspective to study the conditions under which open source software can overcome an existing and dominant proprietary software. Our model also draws from this literature and expands it thanks to its linkage with the supply-side stream of literature presented above, thus explaining the coexistence of models of software production, and their competition, on the basis of the structure of developers' motivations.

Another model that aims to bridge the demand and supply sides of the literature is Mustonen (2003), but it assumes that developers are heterogeneous in their productivity and that more productive developers choose the open source model, whereas our model shows that the two institutions can coexist even when the heterogeneity of individuals is ruled out.[6]

3 Results

This section presents and discusses the results from our model. The model is solved in subsection 4.1, where we first determine the equilibrium efforts in the second stage of the game for a given allocation of researchers in Technology and Community, and then we proceed backward by analysing the first stage decision to determine institution size equilibria and their stability properties. Section 4.2 discusses these results, which also highlight the contribution of our analysis to the existing literature on open collaborative innovation communities. Finally, subsection 4.3 analyses the impact of α (importance of social motivation) and β (Community's capabilities of protecting openness, i.e., magnitude of spillovers from Community to Technology) on the level of the innovativeness of each institution.

[6] In a recent contribution, Kumar et al. (2011) develop a model where two firms marketing commercial open source software compete both in the product market and in a developer market in which firms compete for developers. In their model, however, consumers do not interact directly with the community.

3.1 Equilibrium in the second stage

In the second stage of the game, each researcher, both in Technology and Community, chooses the effort that maximises his or her payoff given n and the effort choices of the other researchers. The first order conditions for payoff maximisation in Technology and Community are given by:

$$\frac{\partial \Pi_i^T}{\partial x_i^T} = \frac{\partial \Pr^T(x_i^T, X_{-i}^T, \beta X^C)}{\partial x_i^T} R^T - \frac{\partial c^T(x_i^T)}{\partial x_i^T} = 0 \tag{7}$$

$$\frac{\partial \Pi_i^C}{\partial x_i^C} = \frac{\partial \Pr^C(x_i^C, X_{-i}^C, X^T)}{\partial x_i^T} k^C - \frac{\partial c^C(x_i^C)}{\partial x_i^C} + \alpha e(n) \frac{\partial Y(x_i^C, X_{-i}^C, X^T)}{\partial x_i^C} = 0 \tag{8}$$

Since we are interested in symmetric Nash equilibria, the equilibrium efforts in Technology and Community (as a function of n), denoted by $\hat{x}^T(n)$ and $\hat{x}^C(n)$, are implicitly defined by:

$$\frac{\partial \Pr^T(\hat{x}^T(n),(n-1)\hat{x}^T(n),(N-n)\hat{x}^C(n))}{\partial x_i^T} R^T - \frac{\partial c^T(\hat{x}^T(n))}{\partial x_i^T} = 0 \tag{9}$$

$$\frac{\partial \Pr^C(\hat{x}^C(n),(N-n-1)\hat{x}^C(n),n\hat{x}^T(n))}{\partial x_i^T} k^C - \frac{\partial c^C(\hat{x}^C(n))}{\partial x_i^C} + e(n)\frac{\partial Y(\hat{x}^C(n),\hat{x}^C(n),\hat{x}^T(n))}{\partial x_i^C} = 0 \tag{10}$$

Proposition 1, proven in the Appendix, characterises the effect of n on the effort exerted by each researcher in Technology and Community.

Proposition 1 *An increase in group size reduces individual effort in Technology and increases it in Community. i.e.,* $\frac{\partial \hat{x}^T(n)}{\partial n} < 0$ *and* $\frac{\partial \hat{x}^C(n)}{\partial n} < 0$.

The intuition of this result is straightforward. In Technology, an increase in group size increases competition within the group and reduces spillovers from Community, both effects being detrimental to the productivity of individual effort.[7] In Community, an increase in size leads to more efforts because of the complementarity among researchers' investments and because of the lower negative externalities from Technology.

[7] Notice that here "competition" refers to technological competition, namely that for discovering innovation opportunities in a given technological space. The model does not include product market competition, which also affects innovation efforts, but in different ways. See Section 4.3 for more on this point.

When we look at total efforts in each institution, i.e., $\hat{X}^T(n) = n\hat{x}^T(n)$ and $\hat{X}^C(n) = (N-n)\hat{x}^C(n)$, it is clear that total efforts in Community are decreasing in *n*, i.e., increasing in size. For Technology, by contrast, the effect is ambiguous. Following Carraro and Siniscalco (2003), we solve this ambiguity by assuming that total effort is also always increasing in group size in Technology, i.e., $\frac{d\hat{X}^T(n)}{dn} > 0$.

Plugging the equilibrium efforts into the payoff functions provides the reduced-form payoff used for comparison in the first stage:

$$\Pi_i^T(n) = \Pr{}^T(\hat{x}^T(n), (n-1)\hat{x}^T(n), \beta(N-n)\hat{x}^C(n))R^T - c^T(\hat{x}^T) \tag{11}$$

$$\Pi_i^C(n) = \Pr{}^C(\hat{x}^C(n), (N-n)\hat{x}^C, n\bar{x}^T)k^C - c^C(\hat{x}^C) + \alpha e(n)Y(\hat{x}^C(n), (N-n)\hat{x}^C, n\bar{x}^T) - C(n) \tag{12}$$

In order to identify the equilibria and their stability properties it is useful to derive the first derivatives of $\Pi_i^T(n)$ and $\Pi_i^C(n)$. By the use of the envelope theorem, we obtain:

$$\frac{d\Pi_i^T(n)}{dn} = \left\{ \frac{\partial \Pr^T}{\partial X_{-i}^T} \frac{dX_{-i}^T}{dn} + \frac{\partial \Pr^T}{\partial X^C} \beta \frac{dX_{-i}^C}{dn} \right\} R^T \tag{13}$$

$$\frac{d\Pi_i^C(n)}{dn} = \left\{ \frac{\partial \Pr^C}{\partial X_{-i}^C} \frac{dX_{-i}^C}{dn} + \frac{\partial \Pr^C}{\partial X^T} \frac{dX^T}{dn} \right\} k^C + \alpha \left\{ \frac{de}{dn} Y + e \left[\frac{\partial Y}{\partial X_{-i}^C} \frac{dX_{-i}^C}{dn} + \frac{\partial Y}{\partial X^T} \frac{dX^T}{dn} \right] \right\} - \frac{dC}{dn} \tag{14}$$

In order to simplify the proofs, but without affecting the qualitative discussion that follows, we assume that $\frac{d^2\Pi_i^T(n)}{dn^2} > 0$ (which is satisfied if the effort cost function is sufficiently convex) and $\frac{d^2\Pi_i^C(n)}{dn^2} < 0$ (which is satisfied whenever coordination costs are sufficiently convex.). These assumptions guarantee the existence of at most three equilibria, reducing in this way the number of cases to be considered. The next Proposition is proven in the Appendix.

Proposition 2 *Payoffs from Technology always decrease the number of researchers in the group, i.e.,* $\frac{d\Pi_i^T(n)}{dn}$ *is always positive. By contrast, payoffs from Community always increase, always decrease, or* *increase and then decrease group size, i.e.,* $\frac{d\Pi_i^C(n)}{dn}$ *can be i) always negative ii) always positive, or iii) first positive and then negative.*

The intuition of the results in Proposition 2 closely mimics that in Proposition 1. Payoffs in Technology decrease the size of this group because, first, more researchers in Technology implies more competition within the institution and, second, it implies fewer researchers active in Community, and thus lower positive spillovers. In Community, group size has a positive effect on researchers' payoff for three reasons: i) the larger positive spillover within the group; ii) a positive impact on the communitarian activity; and iii) lower negative externalities from Technology. However, large communities incur large coordination costs. This negative effect of group size can easily prevail for large groups, hence case iii).

3.2 Equilibrium in the first stage

Given the reduced-form payoffs (11) and (12) and the notions of equilibrium and stability (3) and (4), we are now ready to examine the equilibria.[8]

Proposition 3 *The equilibria of the game are characterised as follows:*

(Scenario I*) If* $\Pi_i^T(0) > \Pi_i^C(0)$ *and* $\Pi_i^T(N) > \Pi_i^C(N)$*, but* $\Pi_i^T(n) < \Pi_i^C(n)$ *for some values of n, then there are two stable equilibria* $n_1{*} \in (0, N)$ *(coexistence of Technology and Community) and* $n{*} = N$ *(all researchers in Technology), and one unstable equilibria* $n_2^* \in (0, N)$ *(coexistence of Technology and Community), with* $n_1^* < n_2^*$.

(Scenario II*) If* $\Pi_i^T(0) < \Pi_i^C(0)$ *and* $\Pi_i^T(N) > \Pi_i^C(N)$*, then there are three equilibria: two stable equilibria,* $n{*} = 0$ *(all researchers in Community), and* $n{*} = N$ *(all researchers in Technology), and one unstable equilibria* $n{*} \in (0, N)$ $n_2^* \in (0, N)$ *(coexistence of Technology and Community).*

(Scenario III*) If* $\Pi_i^T(0) > \Pi_i^C(0)$ *and* $\Pi_i^T(N) < \Pi_i^C(N)$*, then the equilibrium value*

[8] From Proposition 3 we exclude the trivial cases in which $\Pi^T(n) > \Pi^C(n)\ \forall n \in [0; N]$, so that all researchers are in Technology as a unique equilibrium, and $\Pi^T(n) < \Pi^C(n)\ \forall n \in [0; N]$, in which all researchers are in Community as a unique equilibrium.

$n^* \in (0, N)$ *(coexistence of Technology and Community) is unique and stable.*

A graphical representation of the equilibria determined in Proposition 3 is shown in Figure 1.

INSERT FIGURE 1 ABOUT HERE

4 Discussion

4.1 Institution size and the social dimension of Community

In this section, we comment upon the different scenarios described in Proposition 3, in particular relating them to the social dimension of Community. Before entering into a discussion, we first argue how variations in the relevance of personal involvement, the value of communitarian activity, and the coordination costs affect the payoff function in Community.

An increased importance of personal involvement and of the value of communitarian activity is captured by an increase in the parameter α. Formally, $\frac{d\Pi_i^C(n)}{d\alpha} = e(n)Y(\hat{x}^C(n),(N-n)\hat{x}^C, n\bar{x}^T) > 0$. Therefore, an increase in α has the effect of moving the payoff from Community upwards, making it more attractive for any *n*. This effect is also more significant for large Communities, since both *e(n)* and $Y(\hat{x}^C(n),(N-n)\hat{x}^C, n\bar{x}^T)$ are decreasing in *n*.[9] It follows that an increase in α leads to a higher sensitiveness of the payoff of a researcher belonging to Community to his or her group size, i.e., it makes $\frac{d\Pi_i^C(n)}{dn}$ larger in absolute value.

As for coordination costs, their increase has the effect of reducing the Community payoff for all *n*. In addition, the assumptions that $C(N) = 0$ and $\frac{\partial^2 C(n)}{\partial n^2} > 0$ imply that any increase in coordination costs has a greater impact the smaller is *n* (i.e., the larger is the community). Therefore, a high level of coordination costs leads to a low value for $\Pi_i^C(0)$ and possibly $\frac{d\Pi_i^C(n)}{dn}$ to increase for small values of *n* (i.e., for large communities).

[9] *Y* is decreasing in *n* since we assume (subsection 4.1.1) that total effort in technology in increasing in *n*.

In Scenario I, two stable equilibria exist, one in which all researchers choose Technology and one in which Community is "large" (while Technology is "small"); on the contrary, the equilibrium with a "small" Community is unstable. In this scenario, $\Pi_i^T(0) > \Pi_i^C(0)$ and $\Pi_i^T(N) > \Pi_i^C(N)$: $\Pi^C(n)$ must have an inverted U shape. This is associated with a situation where both coordination costs and α (the importance of communitarian activity and personal involvement) are high, i.e., factors that the literature on the communities of practice has identified as their characterising elements. High coordination costs lead to $\Pi_i^T(0) > \Pi_i^C(0)$ and to $\frac{d\Pi_i^C(n)}{dn}$ increasing for low n ; a high value of α, making Community payoff highly (and positively) dependent on group size, makes $\Pi_i^C(n)$ strongly decreasing for high values of *n*, inducing an inverted U relationship in the Community payoff and $\Pi_i^T(N) > \Pi_i^C(N)$.

As a first remark, we notice that in this Scenario there exists one stable equilibrium in which Technology and Community coexist, with group size depending on parameter values. From an empirical point of view, this equilibrium is consistent with what is observed in software, where similar competing products are offered under proprietary and open regimes.[10]

While a large community is stable, the equilibrium where the community is small is an unstable equilibrium. As suggested in the previous section, the model allows a dynamic interpretation, where individuals play a best response strategy to the current allocation of researchers between institutions. In this case, the unstable equilibrium constitutes a threshold that divides the realm of small communities from the set of communities that are able to grow fast and large. In each one of those spaces, the dynamics of the model shows a sort of bandwagon effect. If a community, for whatever reason, is able to grow enough and overcomes the threshold, then it grows endogenously until the large equilibrium**,** which in a sense expresses the full potential of a community. This seems to be the case of the free and open source community, as widely recognised in the literature (e.g., Bonaccorsi and Rossi, 2003; Bitzer and

[10] A notable result is that this has been obtained with ex-ante symmetric researchers and it is the outcome of the endogenous mechanisms within each institution.

Schröder, 2005).[11] What is new in our “Critical Mass” argument for free and open source development is that it is not based on demand factors, (such as, for instance, in Bonaccorsi and Rossi, 2003), but instead on the structure on developers’ motivation,. In particular, the shape of the social forces we described and rooted in Wenger’s (1998) community of practice determines the behaviour we observe in the model.

Consider now Scenario II. This occurs when $\Pi_i^T(0) < \Pi_i^C(0)$ and $\Pi_i^T(N) > \Pi_i^C(N)$. In this case, the stable equilibria correspond to the corner solutions, while an unstable interior equilibrium separates the two “basins of attraction”. This scenario corresponds to a situation where α is high, but coordination costs are low. In a sense, this scenario is a special case of Scenario I: low coordination costs lead to $\Pi_i^T(0) < \Pi_i^C(0)$, instead of $\Pi_i^T(0) > \Pi_i^C(0)$, posing no limit to Community growth. What this scenario shows with clarity is that strong communitarian activity may create a large community, but this is not necessarily so: to be fully expressed, the self-reinforcing growth process needs a critical mass at the beginning.

Finally, consider Scenario III, in which the unique and stable equilibrium is the coexistence of Technology and Community. This case requires $\Pi_i^T(0) > \Pi_i^C(0)$ and $\Pi_i^T(N) < \Pi_i^C(N)$, and consequently the “absolute” value of $\frac{d\Pi_i^C(n)}{dn}$ is “small” (compared with $\frac{d\Pi_i^T(n)}{dn}$). Therefore, this case is consistent with a situation where the value of communitarian activity, the degree of personal involvement, and coordination costs are not significant. Low values of α tend to induce low values of $\Pi_i^C(0)$ [12]; low values of α and coordination costs tend to make $\Pi_i^C(0)$ relatively insensitive to *n*, i.e., $\frac{d\Pi_i^C(n)}{dn}$ is “small”. This scenario coincides with one of those identified in Carraro and Siniscalco

[11] Notice that this approach takes into account the quantitative aspect of free and open source community growth, but not its qualitative side. When communities grow, their social space becomes more complex and their forms of participation and governance structures are placed under pressure. The case of Debian is a clear example of the radical transformation needed to make a growing project able to bear the challenges determined by its own growth (Mateos-Garcia and Steinmueller, 2008; O’Mahony and Ferraro, 2007; Sadowski et al., 2008).

[12] Since low coordination costs tend to increase $\Pi_i^C(0)$, we assume that this effect is dominated by the other. If this is not case, we could expect the situation where all researchers choose Community to prevail.

(2003),[13] where, as we said, the social side of Science (the institution “competing” with Technology in their model) is not considered. Therefore, social motivations are crucial to generate the threshold level between small unstable communities and large stable communities, which is observed in the free and open source case, but not in Carraro and Siniscalco’s competition between Technology and Science.

4.2 Path dependence and the growth of Community

In this section, we focus on Scenario I, where the social dimension of Community is the most relevant. In the dynamic interpretation of the model, the basin of attraction of the two stable equilibria (in terms of the initial condition for n) is determined by the unstable equilibrium, whose values depend on the parameters of the model. Path dependence (David, 1985), which is shown by the importance of the initial condition for n in determining, first, which equilibrium is selected, and then the size of Community, points out the fundamental role that the initial ability of attracting researchers has for the establishment and growth of this institutional mode. Community can be initially attractive through a series of different processes.

First, communities may become economically relevant filling an unfilled market, either creating one *ex novo* or providing the conditions to fill an established one (Bonaccorsi and Rossi, 2003). The definition of “market”, of course, must be interpreted in a wide sense, so that not only the product is important, but also the model of production—in the free and open source case allowing users to be part of the process. The simple existence of a community attracts all the individuals interested in that market. Thus, the more the community responds to such unfilled gaps, the more attractive it becomes to interested individuals.

Moreover, communities, as other institutions, cover a particular space of social interaction. They provide members with a specific interaction environment, ruled by implicit laws and grounded in peculiar identities, i.e., structures of meanings, principles, and values. One of the debates around which the free and open source community is structured concerns the concepts of free and open source software (Dahlander, 2007). This debate shapes the environment in which developers act, defining rules (from rules against free-riding to recognition by peers (O’Mahony, 2003) as well as the meanings (what

[13] The other two scenarios identified by Carraro and Siniscalco involve all individuals choosing a single institution, a case which we excluded from the analysis in the present paper.

“openness” means), values (whether software should be always free), and visions of the world (whether all the produced knowledge should be free). Such interaction contributes then to building the “identity” of the community. Non-members interested in this debate and sensitive to such an identity are then attracted to the community, and may eventually become members.[14]

Another mechanism can also be activated by trust building, which in small communities can lead to a common language, established rules, and improved efficiency at the organisational level. Community of practice theory (Lave and Wenger, 1991; Wenger, 1998) suggests that even in their initial phases communities become structured in a series of concentric circles. Inner circles connect “senior members”, possibly the founders of the community, who perform a great share of the work and lead the community. Other outer circles group together individuals who are less involved in the community, namely those that participate to a lower degree the further is the circle from the centre (Crowston and Howison, 2006). The passage from outer circles to inner circles is a learning process (von Krogh et al., 2003): new members joining the community engage in a series of activities with senior members, getting to know about the community in more depth and letting the community acquire knowledge about them. This process of reciprocal learning, termed Legitimate Peripheral Participation by Lave and Wenger (1991), triggers the process of the negotiation of meanings discussed earlier, affecting the identities of the new member, of the other members involved in the activity, and of the community as a whole. The gradual acceptance of new members into the community increases the level of trust, where belonging to inner circles also means being recognised as more trustworthy. This implies that when member i starts to engage in the communitarian activity, increasingly interacting with the inner circle and acquiring more trust, he or she

[14] Our model excluded heterogeneity among agents, which was instead adopted by other authors in the same field (Mustonen, 2003). Researcher heterogeneity would indeed influence the dynamic interpretation of the model in that, for any given characteristics of the “average” researcher, higher heterogeneity would favour the constitution of a community. Let us focus on Scenario I. Initially, the community is set up by people with a high interest in the activity that the community is going to undertake (captured by a high value of “their” k) and in the “vision” it embodies (captured by a high value of “their” α), an interest high enough to make them bear the costs connected with the small size of the community. The community then can be created and developed, even if it links only a few individuals. The simple existence of a community would make it more rewarding for other individuals to join, thus triggering community growth and making the threshold more likely to be overcome. In this description, it is easy to recognise the actual evolution of the free and open source community (Bonaccorsi and Rossi, 2003; Bitzer and Schröder, 2005). Notice that this argument follows the logic of so-called threshold models (Granovetter, 1978).

begins to *perceive* the community as a trustworthy environment. Thus, the fact that *j* also belongs to the inner community is taken by *i* as a signal of *j*'s trustworthiness. Therefore, *j*'s potential free-riding behaviour is perceived by *i* as an almost irrelevant exception, and *i* reduces his or her monitoring and punishing activities, decreasing coordination costs. This maps the results obtained by Bagozzi and Dholakia (2006), who, as already noted, find that "the community metamorphosizes into a friendship group and a social entity with which one identifies" (p. 1111). Legitimate Peripheral Participation creates trust, and this in turn increases the payoff for participating in Community and makes it more attractive for potential members (handling them without increasing coordination costs), thereby fuelling community growth.

Filling an unfilled market, identity, and trust building can all attract new potential members and trigger the self-reinforcing growth described in the model as a movement from a community below the threshold to one above the threshold that is able to grow endogenously.

4.3 IPR and externalities across institutions

Proposition 3 also provides insights into the role of externalities across institutions in the development of communities. In the case of open source, for instance, the role of licences has been at the centre of a lively debate since the beginning (e.g., Lerner and Tirole, 2005; Comino et al., 2007). In the model, tools that protect openness and prevent appropriation such as strict licenses (e.g., the GPL) are captured by lower values of β, since they reduce the positive externalities from Community to Technology without affecting intra-group externalities. This has the consequence of decreasing the payoff from Technology ($\frac{d\Pi_i^T(n)}{d\beta} = R^T \frac{\partial \Pr^T}{\partial X^C} X^C > 0$) and leading towards equilibrium in which Community is larger and/or enlarges the basin of attractions of these equilibria. Therefore, instruments that protect openness are fundamental for enhancing community sustainability, creating the condition of community growth (Gambardella and Hall, 2006), although they are not part of the engine that fosters it. In the specific case of open source, furthermore, the GPL can help create the critical mass at the initial stage of community development by attracting individuals that share the ideological component at the basis of the GPL.[15]

4.4 Institution innovativeness

Thus far, we have limited our attention to the size of groups that choose each institution at the equilibrium. However, from a social point of view, it is the performances of institutions that matter. As we anticipated in Section 3, we now assess performance in terms of the expected number of innovations in the institution, i.e., its innovativeness. In particular, we perform some comparative statics exercises with respect to the two main parameters of our model, i.e., α (importance of social motivations, in line with the literature on communities of practices) and β (spillovers from Community to Technology, which controls for how much openness is protected). The results of our analysis, whose proofs are in the Appendix, are summarised in the following Proposition.

Proposition 4 *(i) An increase in β has an ambiguous effect on Technology innovativeness, while it reduces Community innovativeness if total efforts in Technology increase (otherwise the effect is ambiguous).*

(ii) An increase in α has an ambiguous effect on Technology innovativeness, while increases Community innovativeness if total efforts in Technology decrease (otherwise the effect is ambiguous).

The intuition that underlies Proposition 4 is as follows. Consider (i) first. An increase in β allows individuals in Technology to take more advantage of the positive externalities from Community, making Technology more attractive. From a Technology perspective, even if the number of researchers increases,

[15] Although a fully-fledged analysis of this point is outside the scope of this paper, our model also hints at the effects of patenting software code on the viability of open source communities. As a first effect, an increase in the strength of IPR in Technology implies an increase in the economic return *R*, which increases in the payoff from Technology. Second, stronger IPR limit the scope of the innovative activity (constraining the "field" in which research could be exploited without violating them), in both Technology and Community cases. Formally, this is captured by a more negative effect on the probability of innovation for any value of the total effort in Technology. While in Technology this effect is most likely to be dominated by the increase in *R*, for Community the negative effect is the only one. This unambiguously leads to equilibria in which a community grows much less than before, if it is created at all. This result is consistent with the concerns about extending the rights of software producers to patent their code in Europe (Torvalds and Cox, 2003).

the impact on innovativeness is ambiguous for two reasons. First, each individual in this institution may increase or decrease his or her effort, where the latter case is possible if, with an increase in n^*, the intra-group negative externalities prevails. Second, a smaller Community reduces its total effort in a way that can reduce, or even offset, the effect of the increase in β. From the point of Community, the number of individuals choosing it is smaller, each individual exerts lower effort, and intra-group positive externalities are lower. If total effort in Technology increases, then the expected number of innovations in Community unambiguously decreases; otherwise, the effect is also ambiguous in the Community case.

Consider now (ii). An increase in α corresponds to a greater importance of social motivations in Community. This makes Community more attractive: a larger institution increases positive intra-group externalities, with a positive effect on innovativeness. On the Technology side, the size reduction has an ambiguous effect on innovativeness because of the symmetric argument put forth before: on one side, there are fewer researchers; on the other side, the effort exerted in Technology by each individual increases. Therefore, the effect on total effort is ambiguous. Furthermore, a larger Community increases the positive inter-group externalities towards Technology. From the point of view of Community, if the total effort in Technology decreases, this reinforces the direct effect of an increase in α on Community innovativeness, making it positive. Otherwise, the impact of α on Community innovativeness is ambiguous.

The results summarised in Proposition 4 show that the endogeneity of institution size and nature of externalities influence variations in β and α that are difficult to predict. However, making assumptions on the effect of the variation of those two parameters (which concern the functioning of Community) on total effort in Technology can resolve the ambiguity in Community. In other words, in order to know the sign of the impact of those two institutional characteristics on innovativeness in Community, we need to know what happens in Technology. This result is interesting per se: what really needs to be determined if Community can actually increase the innovation activity of the whole system is how Technology is shaped, even when the characteristics of the Community environment play a prominent role in determining the payoff for its members. Thus, Technology determines the overall level of innovativeness.

Both in cases (i) and (ii), a crucial role in resolving the ambiguity in Community is played by the impact that an increase in the number of researchers in Technology has on the individual research effort ($\frac{\partial \hat{x}^T(n)}{\partial n}$). This suggests a connection with the literature on the relationship between (product market) competition and innovation. While early empirical work on this issue obtained mixed results, more recent contributions seem to suggest a positive (or nonlinear) relationship between the two (Aghion and Griffith, 2005). In our model, which considers only technological competition, we proved (Proposition 1) that $\frac{\partial \hat{x}^T(n)}{\partial n} < 0$. If, following the direction of this empirical evidence, we assume that $\frac{\partial \hat{x}^T(n)}{\partial n}$ is sufficiently small in absolute value, then total effort in Technology is increasing in β and decreasing in α, thus allowing clear predictions on the effect of these parameters on Community innovativeness. The Community's innovativeness unambiguously increases when β decreases (i.e., when there is a higher capability to protect openness) and α increases (i.e., a higher importance of communitarian social processes).

5 Implications and conclusion

In this paper, we developed a model where open collaborative innovation communities are confronted with Technology (Dasgupta and David, 1994) in their ability to attract researchers. In particular, attention is paid to the social nature of the Community institution, which following the literature on communities of practice (Wenger, 1998) is captured by the degree of personal involvement, the value of communitarian activity, and coordination costs.

As a main result, we confirm the presence of a threshold size for Community, below which it can only remain small or disappear and above which it is pushed by internal forces to grow large. However, in contrast to the previous literature that focused on the final market for knowledge products (demand side), we do that by highlighting the forces at work on the supply side (the input market where institutions compete for knowledge workers). This point of view adds another perspective to the competition dynamics between Community and Technology and unfolds a series of new mechanisms that the literature should take into account.

5.1 Managerial implications

This conclusion is also important for firms. It suggests that when firms decide to initiate a community around their innovation processes, they cannot adopt a step-by-step procedure. Such a "real option-like" approach, where firms gradually increase their investments by deciding in each subsequent step whether and how to foster community growth and solidity, does not square with our findings that emphasize the importance of the initial conditions and of overcoming the initial threshold size. Instead, we describe the subsequent growth as an almost automatic process driven by endogenous forces. The lesson is that firms need to invest a lot in planning and realising the initial phase of community development in order to gather a committed enough group of initial members that is large enough to place the community beyond the threshold size. If a firm succeeds in this, the simple size of the community will trigger an endogenous growth process, attracting researchers from outside and enlarging the payoffs for those already in the community. Firms can compensate for the larger expenses this strategy calls for in the initial phase with lower control and support costs in the growth phase, as according to our model fostering community expansion is intrinsic to the community mechanism itself. This intuition, we believe, has wide implications for managers and project leaders because it speaks against the diffused wisdom that community growth 1) can be treated as a gradual process and 2) should be closely attended by the firm. We claim here instead that an important and careful investment at the beginning would be enough to generate endogenous (and thus almost costless) growth later on. For instance, Spaeth et al. (2010) show that, in the case of the Eclipse development process, not only did IBM release the source code, but its employed contributors played a fundamental role in fuelling the growth of the community in its starting phase. This sort of "preemptive generosity", as the authors call it, is the strategy our model indicates as the best one.

This approach has two main corollaries. First, it means that firms need to detect the threshold size for the sector in which they operate. Having done that and distinguished the initial phase from the growth phase, they need to move resource utilisation from later stages to initial stages. In this case, investments that are apparently oversized with respect to the small initial group of people (e.g., large sponsorship or diverting a lot of internal human resources to community-related tasks) is a perfectly rational strategy. Second, managers and project leaders can use the leverages described above to act upon the initial conditions of community formation and favour the reach of threshold size. In planning the community environment and the tasks the community is called to perform (e.g., suggesting innovative solutions to product-related problems, participating in the brand construction process, and providing actual inputs (such as code) for the firm's project), the firm needs to make sure that the community fills an "unfilled market" (Bonaccorsi and Rossi, 2003). As explained in the text, communities that respond to the social and/or economic needs that are currently unanswered have higher probabilities of attracting those that manifest those needs (Green, 1999), easing the path towards a larger initial community. The other two levers firms can use are centred around the social processes intrinsic to community life. As seen in the model, social processes play a fundamental role in determining researchers' payoffs and thus their motivations. A lively community environment, favouring peer-to-peer interaction and fostering debates and the emergence of a shared identity among members, increases the attractiveness of the community and the rewards for those already a part of it. Firms should thus 1) design the space of members' interaction such that they can communicate directly in an engaging manner (e.g., through discussion groups, social networks, sand-boxes for idea generation, online as well as face-to-face) and 2) provide the initial community with a clear and easily communicable identity, suddenly shared and continually rebuilt with those initial members that choose to gather around the firm's project. The last lever a firm can use is a direct consequence of the dynamics of the communitarian social processes described by community of practice theory (Wenger, 1998). A community, as explained, is organised in a concentric structure, with inner circles populated by the most committed members, while would-be members move from outer circles inward in the Legitimate Peripheral Participation process (Lave and Wenger, 1991). During this process, trust is formed and increased the more new members enter the inner circles of the community. This allows participants in the community activity to suddenly detect those who can be trusted (because they belong to inner circles) and those who, by now, can be trusted only to a certain extent (because they are not yet fully part of the community). This eases community governance and management and decreases coordination costs, in our model making Community more attractive. Firms can exploit this property of communities that favours the development of Legitimate Peripheral Participation processes, for example by providing an infrastructure for members' interaction that organizes the community into concentric circles. Providing instruments to clearly show the seniority of each member, allowing more senior members to access part of the community's online space that are not available to younger members, and giving the former rights on community activities and products that they can delegate to the latter are all tools for moving in that direction.

5.2 Policy implications

Further results concern the innovative performances of the two institutional models. In particular, we focus on how the two main drivers of Community influence the overall level of institution innovativeness: establishing and/or strengthening instruments aimed at protecting the openness of the knowledge produced by the Community (O'Mahony, 2003) and the degree to which the social processes typical of a social body such as Community generate motivations and incentives for its members. It is shown that the endogeneity of institution size and nature of externalities influence ambiguous variations in these two parameters. However, if research effort in Technology is not too sensitive to the number of individuals in the institution, i.e., to the level of competition, then protecting openness (generating thus less spillover from Community to Technology) or the higher importance of social motivations both increase innovativeness in Community, while we still obtain ambiguous results for Technology. The policy implication we can derive from this is counterintuitive: even when focusing the analysis on Community and its institutional structure, what really needs to be determined for the actual level of the innovativeness of both institutions is how the space of Technology is designed by policymakers. If the regulatory background that describes IPR and related markets for technology (Arora et al., 2001) are designed in ways that protect the effort of researchers from high-level competition, this generates a positive effect on Community, where protecting openness and assigning more importance to social motivations means increasing its innovativeness. The possibility for Community to avoid ambiguity in the effects of its determinants relies entirely on the way Technology is shaped by policymakers. This paper does not aim at investigating how Technology should be designed to solve its own ambiguity, which persists in the previous case, and to foster innovativeness in the overall system. However, it clearly indicates to policymakers that the interrelations between Community and Technology are tight and subtle, and to obtain positive results on the one side it is likely that action is needed on the other side. Designing markets affects non-market social bodies as well.

5.3 Further research

While the model is suggestive of several forms of interaction between the two institutional modes of Community and Technology, its stylised form gives several opportunities for potentially useful extensions. First, product market competition (including the issue of pricing and product differentiation) could be modelled explicitly, both in Technology and in Community. Second, the role of firms in Community could be considered, removing explicitly the assumption that participation in one institution is exclusive. Finally, the value of innovation (rather than the probability of innovation) could be made endogenous.

6 References

Aghion, P., Griffith, R., 2005. *Competition and Growth: Reconciling Theory and Evidence*. The MIT Press, Cambridge MA.

Allen, R.C. 1983. Collective invention. *Journal of Economic Behavior and Organization*, **4** (1), 1–24.

Amin, A., Cohendet, P., 2004. *Architectures of Knowledge: Firms, Capabilities and Communities*. Oxford University Press: Oxford, UK.

Arora, A., A. Fosfuri, A. Gambardella. 2001. *Markets for Technology: The Economics of Innovation and Corporate Strategy*. The MIT Press, Cambridge MA.

Bagozzi, R.P., U.M. Dholakia. 2006. Open source software user communities: a study of participation in Linux user groups. *Management Science,* **52** (7), 1099–1115.

Baldwin, C., Clarke, K. 2006. The architecture of participation: does code architecture mitigate free riding in the open source development model? *Management Science*, **52** (7), 1116–1127.

Baldwin, C., von Hippel, E. 2011. Modeling a paradigm shift: from producer innovation to user and open collaborative innovation. *Organization Science*, **22** (6) 1399–1417.

Bessen, J. 2006. Open source software: free provision of complex public goods. J. Bitzer, P.J.H. Schröder, eds. *The Economics of Open Source Software Development*, Elsevier, Amsterdam, The Netherlands, 57–81.

Bezroukov, N. 1999a. Open source software development as a special type of academic research (critique of Vulgar Raymondism). *First Monday* **4** (10) at: http://firstmonday.org/issues/issue4_10/bezroukov/index.html

Bezroukov, N. 1999b. A second look at the cathedral and Bazaar by Nikolai Bezroukov. *First Monday,* **4** (12) at: http://firstmonday.org/issues/issue4_12/bezroukov/index.html

Bitzer, J., J.H.P Schröder. 2005. Bug-fixing and code-writing: the private provision of open source software. *Information Economics and Policy*, **17,** 389–406.

Bonaccorsi, A., S. Giannangeli, C. Rossi. 2006. Entry strategies under competing standards: hybrid business models in the open source software industry. *Management Science*, **52,** 1085-1098.

Bonaccorsi, A., C. Rossi. 2006. Comparing motivations of individual programmers and firms to take part in the open source movement: from community to business. *Knowledge, Technology and Policy*, **18** (4), 40–64.

Bonaccorsi, A., C. Rossi. 2003. Why open source software can succeed. *Research Policy,* **32** (7), 1243–1258.

Carraro, C., Siniscalco, D. 2003. Science versus profit in research. *Journal of the European Economics Association*, **1** (2–3), 576–590.

Casadesus-Masanell, R., P. Ghemawat, P. 2006. Dynamic mixed duopoly. *Management Science,* **52** (7), 1072–1084.

Comino S., Manenti F.M., Parisi M.L. 2007. From planning to mature: On the success of open source projects. *Research Policy*, **36** (10), 1575–1586.

Crowston K., Howison J., 2006. Hierarchy and centralization in free and open source software team communications. *Knowledge, Technology, and Policy*, **18** (4), 65–85.

Dahlander, L. 2007. Penguin in a new suit: a tale of how de novo entrants emerged to harness free and open source software communities. *Industrial and Corporate Change*, **16** (5), 913–943.

Dahlander, L., M. G. Magnusson. 2005. Relationships between open source software companies and communities: observations from Nordic firms. *Research Policy*, **34**, 481–493.

Dahlander, L., M.W. Wallin. 2006. A man on the inside: unlocking communities as complementary assets. *Research Policy*, **35**, 1243–1259.

Dalle, J.-M., P.A. David. 2003. The allocation of software development resources in 'open source' production mode. SIEPR Discussion Paper #02-27, Stanford University at http://opensource.mit.edu/papers/dalledavid.pdf

Dalle, J.-M., P.A. David 2005. The allocation of software development resources in 'open source' production mode. In: Feller, J., Fitzgerald, B., Hissam, S., Lakhani, KR. (Eds.), Making Sense of the Bazaar: Perspectives on Open Source and Free Software. MIT Press, Cambridge, MA, 297–328.

Dalle, J.-M., N. Jullien 2003. "Libre" software: turning fads into institutions? *Research Policy*, **32**(1), 1-11.

Dasgupta, P., P.A. David. 1987. Information disclosure and the economics of science and technology. G. Feiwel, ed. *Arrow and the Ascendant of Economic Theory*. New York University Press, New York, 519–542.

Dasgupta, P., P.A. David. 1994. Towards a new economy of science. *Research Policy*, **23** (5), 487–521.

David, P.A., S. Arora, W.E. Steinmueller. 2001. Economic organization and viability of f/oss software: a proposal to the national science foundation. SIEPR, Stanford University, 22 January, at http://www-siepr.stanford.edu/programs/OpenSoftware_David/NSF-CISE-OpSourFinal.htm

David, P.A., D. Foray. 2003. Economic fundaments of the knowledge society. *Policy Futures in*

Education, **1** (1), 20–49.

David, P.A., Rullani, F., 2008. Dynamics of innovation in an "open source" collaboration environment: Lurking, laboring and launching floss projects on Sourceforge. *Industrial and Corporate Change* **17** (4), 647–710.

David, P.A., J.S. Shapiro. 2008. Community-based production of open source software: what do we know about the developers who participate? *Information Economics and Policy*, **20** (4), 364–398.

David, P.A. 1985. Clio and the Economics of QWERTY. *American Economic Review*, **75** (2), 332–337.

Economides, N., E. Katsamakas. 2006. Two-sided competition of proprietary vs. open source technology platforms and the implications for the software industry. *Management Science*, **52** (7), 1057–1071.

Fosfuri, A., M.S. Giarratana, A. Luzzi. 2008. The penguin has entered the building: the commercialization of open source software products. *Organization Science*, **19** (2), 292–305.

Franke, N., S. Shah. 2003. How communities support innovative activities: an exploration of assistance and sharing among end-users. *Research Policy*, **32** (1), 157–178.

Gambardella, A., B. Hall. 2006. Proprietary vs public domain licensing of software and research products. *Research Policy*, **35** (6), 875–892.

Ghosh, R.A., R. Glott, B. Kriege, G. Robles. 2002. Free/libre and open source software. Part iv: Survey of developers. International Institute of Infonomics, Berlecom Research GmbH, at http://www.infonomics.nl/free and open source/report/Final4.pdf.

Gonzalez-Barahona, J.M., G. Robles, R. Andradas-Izquierdo, R.A. Ghosh, R.A. 2008. Geographic origin of libre software developers. *Information Economics and Policy*, **20**, 356–363.

Granovetter, M. 1978. Threshold models of collective behavior. *American Journal of Sociology*, **83** (6), 1420–1443.

Green, L. G. (1999). Economics of Open Source Software, at http://www.badtux.org/eric/editorial/economics.html.

Jeppesen, L.B., L. Frederiksen, L., 2006. Why do users contribute to firm-hosted user communities? The case of computer-controlled music instruments. *Organization Science*, **17** (1), 45–63.

Johnson, J.P., 2002. Open source software: private provision of a public good. *Journal of Economics & Management Strategy*, **11** (4), 637–662.

Johnson, J.P., 2006. Collaboration, peer review and open source software. *Information Economics and Policy*, **18**, 477–497.

Kumar, V., B.R. Gordon, K. Srinivasan, 2011. Competitive strategy for open source software. Marketing Science, **30** (6), 1066-1078.

Lakhani, K.R., R. G. Wolf. 2005. Why hackers do what they do: understanding motivations and effort in free/open source software projects. J. Feller, B. Fitzgerald, S. Hissam, K.R. Lakhani, eds. *Perspectives on Free and Open Source Software*. MIT Press, Cambridge MA.

Lakhani, K.R., R. Wolf, J. Bates, C. Di Bona. 2002. Hacker survey (release 0.73), the Boston Consulting Group. Presented at the O'Reilly Open Source Software Conference, 24 July, San Diego, CA. Available at http://www.osdn.com/bcg.

Lanzi, D., 2009. Competition and open source with perfect software compatibility. *Information Economics and Policy*, **21** (3), 192–200.

Lave, J., E. Wenger 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press, Cambridge UK.

Lerner, J., J. Tirole. 2002. Some simple economics of open source. *Journal of Industrial Economics*, **50** (2), 197–234.

Lerner J., Tirole J., 2005. The scope of open source licensing. *Journal of Law, Economics, and Organization*, **21**, 20–56.

Lin, Y., 2003. The institutionalisation of hacking practices. *Ubiquity* **4** (4) at: http://www.acm.org/ubiquity/views/y_lin_1.html

Lin, Y., 2004a. Contextualising knowledge–making in Linux user groups. *First Monday* **9** (11) at: http://firstmonday.org/issues/issue9_11/lin/index.html

Lin, Y., 2004b. Epistemologically multiple actor-centered systems: or, EMACS at work! *Ubiquity* **5** (1) at: http://www.acm.org/ubiquity/views/v5i1_lin.html

Mateos-Garcia, J., W.E. Steinmueller 2008. The institutions of open source software: examining the Debian community. *Information Economics and Policy* **20**, 333–344.

Muller, P., 2006. Reputation, trust and the dynamics of leadership in communities of practice. *Journal of Management and Governance* **10**, 381–400.

Mustonen, M., 2003. Copyleft—the economics of Linux and other open source software. *Information Economics and Policy*, **15**, 99–121.

Nuvolari, A., 2004. Collective invention during the British Industrial Revolution: The case of the Cornish pumping engine. *Cambridge Journal of Economics*, **28** (3), 347–363.

Nuvolari A. and Rullani F. (2007) Curious exceptions? Open Source Software and 'Open' Technology. In St. Amant, K., Still, B. (Eds.), *Handbook of Research on Open Source Software: Technological, Economic, and Social Perspectives*. Idea Group Publishing, Hershey, PA.

O'Mahony, S., 2003. Guarding the commons. *Research Policy*, **32** (7), 1179–1198.

O'Mahony, S., F. Ferraro. 2007. The emergence of governance in an open source community. *Academy of Management Journal*, **50** (5), 1079–1106.

Raymond, E., 1998a. Homesteading the noosphere. *First Monday* 3 (10) at http://www.firstmonday.dk/issues/issue3_10/raymond/#d12

Raymond, E. S. 1998b: *The Cathedral & the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*. Sebastopol, CA: O'Reilly Associates.

Roberts, J.A., I. Hann, S.A. Slaughter. 2006. Understanding the motivations, participation, and performance of open source software developers: a longitudinal study of the apache projects. *Management Science*, **52** (7), 984–999.

Sadowski, B.M., G. Sadowski-Rasters, G. Duysters. 2008. Transition of governance in a mature open software source community: evidence from the Debian case. *Information Economics and Policy*, **20**, 323–332.

Shah, S.K. 2006. Motivation, governance, and the viability of hybrid forms in open source software development. *Management Science*, **52** (7), 1000–1014.

Spaeth, S., Stuermer, M., von Krogh, G. 2010. Enabling knowledge creation through outsiders: towards a push model of open innovation. *International Journal of Technology Management*, **52** (3/4), 411–431.

Torvalds L., Diamond D. 2001. *Just for fun: the story of an accidental revolutionary*. Harper Business.

Torvalds, L., A. Cox. 2003. Open letter on software patents from Linux developers, 21 September 2003, http://www.effi.org/patentit/patents_torvalds_cox.html

von Krogh, G., S. Spaeth, K. R. Lakhani, 2003. Community joining, and specialization in open source software innovation: A case study, *Research Policy*, **32**, 1227–1241.

von Hippel, E. 1988. *The Sources of Innovation*. Oxford University Press, New York.

von Hippel, E. 2001. Innovation by user communities: learning from open source software. *Sloan Management Review*, **42** (4), 82–86.

von Hippel, E., G. von Krogh. 2003. Open source software and the 'private-collective' innovation model: Issues for Organization Science. *Organization Science* **14** (2), 209–223.

Wenger, E., 1998. *Community of Practice: Learning Meaning and Identity*. Cambridge University Press, Cambridge UK.

7 Appendix

Proof of Proposition 1

By the use of the implicit function theorem, we get:

$$\frac{\partial \hat{x}^T}{\partial n} = -\frac{\left\{\frac{\partial^2 \Pr^T}{\partial x^{\mathrm{T}} \partial X_{-i}^{T}} \hat{x}^T - \frac{\partial^2 \Pr^T}{\partial x^{\mathrm{T}} \partial X^{C}} \hat{x}^C\right\} R^T}{\frac{\partial^2 \Pr^T}{\left(\partial x^T\right)^2} R^T - \frac{\partial^2 c^T}{\left(\partial x^T\right)^2}}$$

$$\frac{\partial \hat{x}^C}{\partial n} = -\frac{\left\{\frac{\partial^2 \Pr^C}{\partial x^C \partial X_{-i}^T}\hat{x}^T - .\frac{\partial^2 \Pr^C}{\partial x^C \partial X^C}\hat{x}^C\right\}k^C + \frac{\partial e}{\partial n}\frac{\partial Y}{\partial x_i^C} + e(n)\left\{\frac{\partial^2 Y}{\partial x^C \partial X_{-i}^T}\hat{x}^T - \frac{\partial^2 Y}{\partial x^C \partial X^C}\hat{x}^C\right\}}{\frac{\partial^2 \Pr^C}{\left(\partial x_i^C\right)^2}k^C - \frac{\partial^2 c^C}{\left(\partial x_i^C\right)^2} + e(n)\frac{\partial^2 Y}{\left(\partial x_i^C\right)^2}}$$

Concerning $\frac{\partial \hat{x}^T}{\partial n}$, we notice that the denominator is negative since $\frac{\partial^2 \Pr^T}{\left(\partial x^T\right)^2}$ ($\frac{\partial^2 c^T}{\left(\partial x^T\right)^2}$) is negative (positive) by assumption. As for the numerator, $\frac{\partial^2 \Pr^T}{\partial x^T \partial X_{-i}^T}$ ($\frac{\partial^2 \Pr^T}{\partial x^T \partial X^C}$) is negative by assumption (footnote 4). Therefore $\frac{\partial \hat{x}^T}{\partial n} < 0$.

Concerning $\frac{\partial \hat{x}^C}{\partial n}$, the denominator is negative since $\frac{\partial^2 \Pr^C}{\left(\partial x_i^C\right)^2}$ and $\frac{\partial^2 Y}{\left(\partial x_i^C\right)^2}$ are negative by assumption, while $\frac{\partial^2 c^C}{\left(\partial x_i^C\right)^2}$ is positive. As for the numerator, $\frac{\partial e}{\partial n}$ is negative, while $\frac{\partial Y}{\partial x_i^C}$ is positive; $\frac{\partial^2 \Pr^C}{\partial x^C \partial X_{-i}^T}$ and $\frac{\partial^2 Y}{\partial x^C \partial X_{-i}^T}$ are negative by assumption, while $\frac{\partial^2 \Pr^C}{\partial x^C \partial X^C}$ and $\frac{\partial^2 Y}{\partial x^C \partial X^C}$ are positive. Therefore $\frac{\partial \hat{x}^C}{\partial n} < 0.$

Proof of Proposition 2

The sign of (13) comes from $\frac{\partial \Pr^T}{\partial X_{-i}^T} < 0$, $\frac{\partial \Pr^T}{\partial X^C} > 0$, $\frac{dX^T}{dn} > 0$ by assumption, while we proved that

$$\frac{dX_{-i}^C}{dn} < 0.$$

From Equation (14), we notice that given the assumptions made in the paper:

$$\left\{\frac{\partial \Pr^C}{\partial X_{-i}^C}\frac{dX_{-i}^C}{dn}+\frac{\partial \Pr^C}{\partial X^T}\frac{dX^T}{dn}\right\}k^C<0$$

$$\frac{\partial e}{\partial n}\left[\frac{\partial Y}{\partial X_{-i}^C}\frac{dX_{-i}^C}{dn}+\frac{\partial Y}{\partial X^T}\frac{dX^T}{dn}\right]<0$$

since $\frac{\partial \Pr^C}{\partial X_{-i}^C},\frac{\partial Y}{\partial X_{-i}^C}$>0, and $\frac{\partial \Pr^C}{\partial X^T},\frac{\partial Y}{\partial X^T}$<0 by assumption. However, $-\frac{\partial C}{\partial n}$ is positive by assumption. The overall sign is then ambiguous. If

$$\left\{\frac{\partial \Pr^C}{\partial X_{-i}^C}\frac{dX_{-i}^C}{dn}+\frac{\partial \Pr^C}{\partial X^T}\frac{dX^T}{dn}\right\}k^C+\frac{\partial e}{\partial n}\left[\frac{\partial Y}{\partial X_{-i}^C}\frac{dX_{-i}^C}{dn}+\frac{\partial Y}{\partial X^T}\frac{dX^T}{dn}\right]<-\frac{\partial C}{\partial n}$$

for any *n*, then it is always $\frac{d\Pi_i^C(n)}{dn}>0$. If

$$\left\{\frac{\partial \Pr^C}{\partial X_{-i}^C}\frac{dX_{-i}^C}{dn}+\frac{\partial \Pr^C}{\partial X^T}\frac{dX^T}{dn}\right\}k^C+\frac{\partial e}{\partial n}\left[\frac{\partial Y}{\partial X_{-i}^C}\frac{dX_{-i}^C}{dn}+\frac{\partial Y}{\partial X^T}\frac{dX^T}{dn}\right]>-\frac{\partial C}{\partial n}$$

for any *n*, then it is always $\frac{d\Pi_i^C(n)}{dn}<0$. Suppose now that there are some values $\tilde{n}$ for which $\frac{d\Pi_i^C(\tilde{n})}{dn}=0$. Since we assumed $\frac{d^2\Pi_i^C(n)}{dn^2}<0$, $\tilde{n}$ is unique, and it is the global maximiser of $\Pi_i^C(n)$ in the relevant interval. Consequently, $\Pi_i^C(n)$ is increasing in *n* until $\tilde{n}$, and then decreasing.

Proof of Proposition 3

Consider Scenario I. If $\Pi_i^T(0)>\Pi_i^C(0)$ and $\Pi_i^T(N)>\Pi_i^C(N)$, with $\Pi_i^T(n)<\Pi_i^C(n)$ for some values of *n*, then $\Pi_i^C(n)$ must be first increasing and then decreasing in *n*. Consequently $\Pi_i^T(n)$ crosses $\Pi_i^C(n)$ twice, in $n_1{*}\in(0,N)$ and $n_2^*\in(0,N)$, with $n_1^*<n_2^*$. Since $\Pi_i^T(0)>\Pi_i^C(0)$

n_1* is stable ($\Pi_i^T(n)$ "cuts" $\Pi_i^C(n)$ from below), while $\Pi_i^C(n)$ cuts $\Pi_i^T(n)$ from above in n_2^*, and then n_2^* is unstable. Since n_2^* is unstable, also *n=N* is a stable equilibrium.

Consider Scenario II. If $\Pi_i^T(0) < \Pi_i^C(0)$ and $\Pi_i^T(N) > \Pi_i^C(N)$, then $\Pi_i^T(n)$ and $\Pi_i^C(n)$ cross only once given our assumption. Since $\Pi_i^T(0) < \Pi_i^C(0)$, in $n* \in (0, N)$ where $\Pi_i^T(n) = \Pi_i^C(n)$ $\Pi_i^C(n)$ cuts $\Pi_i^T(n)$ from above, and then the equilibrium in unstable. Consequently $n* = 0$ and $n* = N$ are stable equilibria.

Consider finally Scenario III. If $\Pi_i^T(0) > \Pi_i^C(0)$ and $\Pi_i^T(N) < \Pi_i^C(N)$, $\Pi_i^T(n)$ and $\Pi_i^C(n)$, which are continuous, must cross at least once. Given $\frac{d^2\Pi_i^C(n)}{dn^2} < 0$ and $\frac{d^2\Pi_i^T(n)}{dn^2} > 0$, the value $n* \in (0, N)$ where $\Pi_i^T(n) = \Pi_i^C(n)$ must be unique. Since $\Pi_i^T(0) > \Pi_i^C(0)$, then $\Pi_i^T(n)$ "cuts" $\Pi_i^C(n)$ from above, which guarantees stability.

Proof of Proposition 4

First, we observe that in order to determine the effect of the variation of a parameter on the expected number of innovations we need to proceed as follows:

i) Determine the derivative of *n** with respect to the parameter.

ii) Determine the derivative of the individual and total efforts in each institution with respect to the parameter.

iii) Determine the derivative of the expected number of innovations with respect to the parameter.

Furthermore, we observe that, since we focus on stable equilibria, it must be:

$$\frac{d\Pi^T(n^*)}{dn} - \frac{d\Pi^C(n^*)}{dn} < 0$$

Consider case now (i). Applying the implicit function theorem to the equilibrium condition:

$$\Pi_i^T(n^*) - \Pi_i^C(n^*) = 0$$

we obtain the effect of β on the equilibrium size of the Technology group:

$$\frac{dn^*}{d\beta} = -\frac{R^T X^C \dfrac{\partial \Pr^T}{\partial \beta}}{\dfrac{d\Pi^T(n^*)}{dn} - \dfrac{d\Pi^C(n^*)}{dn}} > 0$$

since the denominator is negative in a stable equilibrium, while the numerator is positive since $\frac{\partial \Pr^T}{\partial \beta} \equiv \frac{\partial \Pr^T}{\partial X^C} > 0$ by assumption.

As for the effect of β on $\hat{x}^T$, this is obtained as the sum of the direct impact of the parameter variation through the first order condition, and the indirect impact due to the variation in the numbers of individuals in the institution. Therefore, by applying the implicit function theorem to equation (9), we obtain:

$$\frac{d\hat{x}^T}{d\beta} = \frac{\partial \hat{x}^T}{\partial \beta} + \frac{\partial \hat{x}^T}{\partial n^*}\frac{dn^*}{d\beta} = \frac{R^T X^C \dfrac{\partial^2 \Pr^T}{\partial \hat{x}^T \partial \beta}}{-\dfrac{\partial^2 \Pr^T}{(\partial x^T)^2} + \dfrac{\partial^2 c^T}{(\partial x^T)}} + \frac{\partial \hat{x}^T}{\partial n^*}\frac{dn^*}{d\beta}$$

In the first term (direct effect), the denominator is positive since, by assumption, $\frac{\partial^2 \Pr^T}{(\partial x^T)^2}$ is negative and $\frac{\partial^2 c^T}{(\partial x^T)}$ is positive. $RX^C \frac{\partial \Pr^T}{\partial \hat{x}^T \beta}$ is positive since $\frac{\partial^2 \Pr^T}{\partial \hat{x}^T \partial \beta} \equiv \frac{\partial^2 \Pr^T}{\partial \hat{x}^T \partial \hat{X}^C}$ is positive by assumption (see footnote 5). $\frac{\partial \hat{x}^T}{\partial n^*} \frac{dn^*}{d\beta}$ (the indirect effect via n^*) is negative, since $\frac{dn^*}{d\beta} > 0$ and $\frac{\partial \hat{x}^T}{\partial n} < 0$ by Proposition 1. Therefore, the overall sign is ambiguous, with $\frac{d\hat{x}^T}{d\beta}$ being positive if the direct effect prevails.

Computing $\frac{d\hat{X}^T}{d\beta}$ we obtain:

$$\frac{d\hat{X}^T}{d\beta} = n^* \frac{d\hat{x}^T}{d\beta} + \frac{dn^*}{d\beta} \hat{x}^T$$

which is positive if $\frac{d\hat{x}^T}{d\beta} > 0$, while it has an ambiguous sign in the opposite case.

On the Community side, we note that only an indirect effect exists. Applying the implicit function theorem on (10), and computing afterwards the derivative of total effort in Community, yields:

$$\frac{d\hat{x}^C}{d\beta} = -\frac{\frac{\partial \hat{x}^C}{\partial n} \frac{dn^*}{d\beta}}{\frac{\partial^2 \Pr^C}{(\partial x^C)^2} - \frac{\partial^2 c^C}{(\partial x^C)}} < 0$$

$$\frac{d\hat{X}^C}{d\beta} = (N - n^*) \frac{d\hat{x}^C}{d\beta} - \frac{dn^*}{d\beta} \hat{x}^C < 0$$

since $\frac{\partial \hat{x}^C}{\partial n} < 0$ by Proposition 1, and $\frac{\partial^2 \Pr^C}{(\partial x^C)^2} < 0$ and $\frac{\partial^2 c^C}{(\partial x^C)} > 0$ by assumption.

Considering the results so far, we finally get the impact of β on the expected number of innovations in Technology:

$$\frac{dn^* \Pr^T(\hat{x}^T(n^*), \hat{X}^T(n^*), \hat{X}^C(n^*))}{d\beta} = \frac{dn^*}{d\beta}\Pr^T + n^* \frac{d\Pr^T}{d\beta} =$$
$$\frac{dn^*}{d\beta}\Pr^T + n^*\left[\frac{\partial \Pr^T}{\partial x_i{}^T}\frac{dx^T}{d\beta} + \frac{\partial \Pr^T}{\partial X_{-i}{}^T}\frac{dX^T}{d\beta} + \beta\frac{\partial \Pr^T}{\partial X^C}\frac{dX^C}{d\beta} + X^C\frac{\partial \Pr^T}{\partial \beta^C}\right]$$

which has an ambiguous sign. The first term is positive. Within the square brackets, the first two terms, capturing the effect of β variation on Technology, have ambiguous signs, which turn out to be positive if the direct effect prevails. The third term is positive and the fourth negative, capturing that an increase in β increases the spillovers towards Technology for given total effort in Community, but also reduces such effort via a reduction in *N*-*n**.

As for the impact of β on the expected number of innovations in Community, we obtain:

$$\frac{d(N-n^*)\Pr^C}{d\beta} = -\frac{dn}{d\beta}\Pr^C + (N-n)^*\frac{d\Pr^C}{d\beta} =$$
$$-\frac{dn^*}{d\beta}\Pr^C + n\left[\frac{\partial \Pr^C}{\partial x_i^C}\frac{dx^C}{d\beta} + \frac{\partial \Pr^C}{\partial X_{-i}{}^C}\frac{dX^C}{d\beta} + \frac{\partial \Pr^C}{\partial X^T}\frac{dX^T}{d\beta}\right] < 0$$

which is ambiguous. However, if $\frac{dX^T}{d\beta}$ is positive, then $\frac{dX^T}{d\beta}$ is unambiguously negative since all the addends are negative.

Consider now case (ii). Regarding the effect of α on the equilibrium size of the Technology group, applying the implicit function theorem to the equilibrium condition yields:

$$\frac{dn^*}{d\alpha} = -\frac{-e(n)V}{\dfrac{d\Pi^T(n^*)}{dn} - \dfrac{d\Pi^C(n^*)}{dn}} < 0$$

The effect of α on $\hat{x}^C$ is obtained as the sum of the direct impact of the parameter variation through the first order condition, and the indirect impact due to the variation in the numbers of individuals in the institution. Then, by applying the implicit function theorem to equation (10), we obtain:

$$\frac{d\hat{x}^C}{d\alpha} = -\frac{e(n)\dfrac{\partial V}{\partial x^C}}{\dfrac{\partial^2 \Pr^C}{(\partial x^C)^2} - \dfrac{\partial^2 c^C}{(\partial x^C)}} + \frac{\partial \hat{x}^C}{\partial n}\frac{dn^*}{d\alpha} > 0$$

In the first term (the direct effect), the denominator is negative and the numerator is positive because of the assumptions made in this paper. The second term (the indirect effect) is negative due to Proposition 1. The effect of total investment in Community is given by:

$$\frac{d\hat{X}^C}{d\alpha} = (N - n^*)\frac{d\hat{x}^C}{d\alpha} - \frac{dn^*}{d\alpha}\hat{x}^C > 0$$

As for the effect on the Technology side, we note that only an indirect effect exists. By applying the implicit function theorem to (9), we obtain:

$$\frac{d\hat{x}^T}{d\alpha} = -\frac{\frac{\partial \hat{x}^T}{\partial n}\frac{dn^*}{d\alpha}}{\frac{\partial^2 \Pr^T}{(\partial x^T)^2} - \frac{\partial^2 c^T}{(\partial x^T)^2}}$$

which is positive since the denominator is positive and $\frac{\partial \hat{x}^T}{\partial n}$ is negative by Proposition 1. In terms of total effort, we obtain:

$$\frac{d\hat{X}^T}{d\alpha} = n^* \frac{d\hat{x}^T}{d\alpha} + \frac{dn^*}{d\alpha}\hat{x}^T$$

whose sign is ambiguous since $\frac{d\hat{x}^T}{d\alpha}$ is positive, while $\frac{dn^*}{d\alpha}$ is negative.

Considering the results so far, we finally get:

$$\frac{d(N-n^*)\Pr^C}{d\alpha} = -\frac{dn^*}{d\alpha}\Pr^C + (N-n)^* \frac{d\Pr^C}{d\alpha} =$$
$$-\frac{dn^*}{d\alpha}\Pr^C + (N-n^*)\left[\frac{\partial \Pr^C}{\partial x_i^C}\frac{dx^C}{d\alpha} + \frac{\partial \Pr^C}{\partial X_{-i}^C}\frac{dX^C}{d\alpha} + \frac{\partial \Pr^C}{\partial X^T}\frac{dX^T}{d\alpha}\right]$$

While the first term is positive (since the Community group is larger), the quantity within the square brackets has an ambiguous sign since $\frac{d\hat{X}^T}{d\alpha}$ has an ambiguous sign (the other terms are positive). Then, the effect of α on the innovativeness in the Community group is overall ambiguous, unless $\frac{d\hat{X}^T}{d\alpha}$ is negative, which would imply $\frac{d(N-n^*)\Pr^C}{d\alpha} > 0$ since $\frac{\partial \Pr^C}{\partial X^T}$<0 . As for the effect on Technology, we obtain:

$$\frac{dn\,\mathrm{Pr}^T}{d\alpha} = \frac{dn^*}{d\alpha}\mathrm{Pr}^T + n^*\frac{d\,\mathrm{Pr}^T}{d\alpha} =$$
$$\frac{dn^*}{d\alpha}\mathrm{Pr}^T + n^*\left[\frac{\partial\,\mathrm{Pr}^T}{\partial x_i^T}\frac{dx^T}{d\alpha} + \frac{\partial\,\mathrm{Pr}^T}{\partial X_{-i}^T}\frac{dX^T}{d\alpha} + \frac{\partial\,\mathrm{Pr}^T}{\partial X^C}\frac{dX^C}{d\alpha}\right]$$

which has ambiguous sign since $\frac{dn^*}{d\alpha} < 0$ and, within the square brackets, the first and third terms in the square brackets are positive, while the second term is ambiguous.

Figure 1

(Scenario I)

(Scenario II)

(Scenario III)