Between IPRs and Public-Funded Research: Is a Community just a "Fancy" Science?

Lessons from the Free/Open Source Software community case[1]

Francesco Rullani

DRUID & INO, Dept. of Innovation and Organizational Economics

Copenhagen Business School

Kilevej 14A, 2000 Frederiksberg, Denmark.

Ph: +45 3815 2992; Fax: +45 3815 2540. Email: fr.ino@cbs.dk

This version*: May, 2009*

Abstract

The study of social institutions producing and disseminating knowledge has mainly concentrated on two main concepts: *science* and *technology*. This paper examines a recent social institution that seems not to resemble either of the other two; that is, knowledge-intensive communities, where individuals freely exchange knowledge through information and communication technology. Using Free/Libre/Open Source Software (FLOSS) as an example, I develop a model where this phenomenon is compared to science and confronted with technology with respect to their ability to attract researchers. My findings show that knowledge-intensive communities and science have the same *nature*, but a *different cost/benefit trade-off*.

JEL classification*: O31, L86, L88*

Key Words*: free/libre/open source software, science, technology, community, intellectual property rights*

[1] Acknowledgements: I am grateful to Carlo Carraro, and to participants of both the FEEM seminar in Milan, and the December 7–8, 2007 RIPE Conference in Copenhagen.

1. Introduction

This paper presents a model that compares *technology* and *science* (Dasgupta and David, 1987, 1994) with one of the most prominent examples of *knowledge-intensive communities* (David and Foray, 2003)[2], Free/Libre/Open Source Software (FLOSS).

In FLOSS, a large number of individuals spread all over the world (Gonzalez-Barahona et al., 2008) cooperate online to create software, and release it freely and 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; while—thanks to the license scheme adopted by the community—no one can appropriate the software jointly developed. This makes the FLOSS model different from technology, where intellectual property rights are used to extract rents from the distribution or use of the produced knowledge. However, it is unsure whether FLOSS, and knowledge-intensive communities in general, represent a new institutional setting able to provide researchers, managers, and policy makers with new solutions for the trade-offs typical of the innovation activity, or if it simply reproduces the same institutional structures of academic research. This paper investigates this point by developing a model where the FLOSS community, used as an example of a knowledge-intensive community, confronts technology with respect to attracting researchers.

I first discuss the theoretical background upon which this paper can be projected, showing the motivation for the study, and retrieve the specific perspective to be applied in the modeling section. Section 2 presents my research question, and in Section 3, I begin the modeling exercise and present the basic model upon which I will build my analysis (Carraro et al., 2001; Carraro and Siniscalco, 2003). I then draw from the empirical literature on FLOSS and argue that, when the payoff function describing developers' actions is built around motivations like signaling, reputation, and own-use, the FLOSS community can be considered very close to science, precisely because the incentive schemes of the two institutions are similar (Section 3.2). In Section 3.3 I elaborate on the basic model defined to "test" if such a science-like community can be sustainable when it confronts technology for attracting researchers. The finding is that it is extremely difficult for a science-like community to be able to attract enough researchers to survive. Since we observe the exponential growth of the FLOSS community and the birth and success of other knowledge-intensive communities such as user-innovator online communities (Jeppesen and Frederiksen, 2006), the conclusion is that the gap between the reality and the theory can be closed only if we admit that the FLOSS community, as every other knowledge-based community, must be something more than a

[2] The authors use also the term "knowledge-based communities".

simple science-like institution.

In Section 4, I identify other factors used in the analysis, based on the examination of the empirical literature on the FLOSS community. Drawing from contributions from Bagozzi and Dholakia (2006), Lakhani and Wolf (2005), and Shah (2006), I find that the science-like payoff function of the community as defined so far neglects the crucial role played by the social involvement of the community members. This same mechanism also exists in the realm of science— embodied in the scientific community—but the presence of the State and the structural differences between the two institutions create a different distribution of the "weights" of each factor. *In nuce*, science and community have the same nature, but a different cost/benefit structure.

In Section 5 I try to capture the main features of the social processes described in the literature using community of practice theory (Wenger, 1998). On this basis I create an "emended" payoff function able to account for the main social mechanisms at work in the community, and again run the model where this new community form competes with technology to attract researchers. The results, discussed in Section 6, are:

1. A community based on the social mechanisms typical of a community of practice can effectively survive the competition with technology, and, under certain conditions, can "endogenously" become fast and large
2. This "emended" model highlights the specificities, mechanisms, and features of an online community.
3. The analysis of the parameters inside and outside the model makes it possible to derive policy advice that takes into account this particular incentive scheme (Section 7).

2. Is the FLOSS community just a "fancy" science?

2.1 Background and motivation

One of the main challenges the development of what has been called a "knowledge society" imposes to economic theory is the assessment of the changes that occurred in the institutions enabling 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 constructed.

The established literature on institutions connected to knowledge production usually builds on two main contributions. The first is Arrow's (1962) famous account of the peculiarities of the knowledge production and distribution processes.[3] The incentive to invest in innovation is greatly

[3] Arrow's (1962) paper is not focused on knowledge, but on information. The difference is substantial, because

reduced because it is impossible to fully control the access to the knowledge produced and there are degrees of uncertainty and risk associated with the research activity (Nelson, 1990; Rosenberg, 1996). Thus, institutions able to provide innovators with substitutive incentives are needed.

The second contribution is the development of the economic discourse around these institutions.[4] In particular, most of the studies adopting Dasgupta and David's (1987, 1994) point of view recognize mainly two institutional models. The first, *science* (David, 1991) is based on the public funding of knowledge production. The second, *technology*, is based on the market exchange based on intellectual property rights (IPR). Technology and science are not stand-alone institutions. They form a complex web of relations (Rosenberg, 1982) that continually evolves over time.

In the "shaded areas" of this dual system a series of examples of an open model of knowledge production—where agents develop and distribute knowledge without external funding or rents assured by the IPR regime—has been identified. Collective invention (Allen, 1983, Nuvolari, 2004), or communities of user innovators (Jeppesen and Frederiksen, 2006; von Hippel, 1988), are just two examples of the specific forms this model can take, and represent a challenge to the explanatory power of the science/technology dual system. In recent years, this challenge has become particularly important due to the pervasive role of what David and Foray (2003) call *knowledge-intensive communities.* These communities are characterized by "three main elements: a significant number of a community's members combine to produce and reproduce new knowledge (diffuse sources of innovation); the community creates a 'public' space for exchanging and circulating the knowledge; new information and communication technologies are intensively used to codify and transmit the new knowledge" (David and Foray, 2003, p. 5). In such communities, the widespread use of information and communication technologies enabled the overcoming of physical limitations to (codified) knowledge transfer, unlocked the possibility to extend the open model of innovation to other sectors, and resulted in unexpected performances. An investigation on the nature of knowledge-intensive communities, especially with respect to the features of the science-technology system, is then a crucial step to be undertaken by the research in the field.

2.2 The literature and the research question: Comparing science and knowledge-intensive communities

As a starting point it is essential to describe how knowledge-intensive communities relate to the dual system represented by technology and science as defined by Dasgupta and David (1987,

---

knowledge has totally different characteristics: it embodies different degrees of tacitness (routines, for example.) See Coriat and Dosi, 1998, and Nelson and Winter, 1982; it can be acquired only by interpretation, and is subject to learning, ... so that "getting a PhD is not simply acquiring information," Dosi (1988, p. 1130). Nevertheless Arrow's (1962) intuition still applies to knowledge with a high degree of transferability. This characteristic, in fact, embodies the features of information through which Arrow's intuition develops.

[4]On the general role of institutions in economic development and their theoretical status see, for example, Rosenberg

1994). The literature on FLOSS, the case I use as a reference point, allows us to imagine a possible answer to this question. It was, in fact, in this literature that the question was asked relative to whether, and to what extent, the community model seemed to just resemble the academic world. Bezroukov (1999a; 1999b) was among the first authors identifying a possible homomorphism between the two institutions in terms of the produced outcome, the involved incentives, the typology of teamwork and organization of collaboration, and the way in which the activity is financially supported. In particular, Bezroukov stresses the similar role of financial institutions, such as research institutes, universities, or private research labs, in providing the individuals with the funds to undertake their activities in the directions they desire; and the similarity between the rules upon which science is based and the practices typical of the FLOSS community, which are also based on a public debate where priority over solutions and peer review are the crucial mechanisms used to regulate and direct individuals' activities (Dasgupta and David, 1987; Lee and Cole, 2003). Kelty (2001) stresses the same similarities. On the one hand he states that "[…T]he funding that supports many projects (in most cases indirectly) comes from those well-known scientific institutions" (Kelty, 2001, online). On the other hand, he also argues that the structure of incentives and the organization of the collaborative effort of developers and scientists are very close to one another, both based on rules connecting the openness of the results to the individual pursuit of recognition and reputation. Mustonen (2003) shares the same point of view "The essential property of the copyleft licensing scheme [i.e. GPL] is that it creates a particular incentive structure… [that] has properties that are equivalent to the incentive structures of scientific communities" (Mustonen, 2003, p. 104). Following a similar path, Bonaccorsi and Rossi (2003b) recall the origins of FLOSS inside the university labs to claim that "Emerging as it does from the university and research environment, the movement adopts the motivations of scientific research" (Bonaccorsi and Rossi, 2003b, p. 1245). Dalle and David (2003) also share a similar point of view, stressing the parallelism between the FLOSS 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 (1998c), who suggests that the correspondence between the two phenomena is just the outcome of the fact that the scientific and the FLOSS enterprises had simply given the same answer to the same problem of collective knowledge production.

However, similarities between science and FLOSS do not imply that the two systems simply coincide. The history of the FLOSS community[5] can be useful to understand why differences are

---

and Birdzell (1986).

[5]An overview on several aspects of the FLOSS scene and history is presented in Giuri et al. (2002). To have an idea of the cultural environment in which this community developed see Himanen et al. (2001) and Raymond (1998a). The

more relevant than the previous picture may seem to suggest. FLOSS began during the eighties when Richard Stallman founded the Free Software Foundation (FSF). Stallman was a researcher at the Artificial Intelligence Labs at the MIT, but during those years, software development in the scientific environment began to be influenced by the enforcement of IPRs (Williams, 2002). Hardware and software were provided to the Labs, but with temporary *un-disclosure* clauses. In opposition to this practice, Stallman decided to create a new operating system, GNU, which had to be and remain totally open and free. To do that, he decided to leave the MIT to create an organization *outside* that scientific institution. The GNU community grew fast, producing software, creating ideas, values and programming principles. Nowadays the instruments created to guarantee software openness and freedom, as the General Public License (GPL), are widely spread also in the sciences environment, but the community is still something different form the academic structure. Had the FLOSS experience being just a new scientific enterprise, there would have been no need for the GPL and for an external organization such as the FSF. In fact, in the FLOSS-EU survey (Ghosh *et al*., 2002) only a bit more than 30% of the surveyed developers is composed by students or universities employees, and only 6% of the projects undertaken in the frame of *SourceForge* (www.sourceforge.net, January 2003) is intended for "Science/Research" or "Education" audiences. It is then true that, as Kelty (2001) argues, the FLOSS community was born in the scientific environment and culture; but the perspective offered here suggests that it moved out of it and grew autonomously, relying only partly onto the academic structure.

Given this uncertainty about the nature of FLOSS, and by extension of knowledge-intensive communities, we can ask: *Is the nature of the FLOSS community the same as that of science?* The same conclusion is reached by David et al. (2001), who state: "This analogy with open science research networks […] calls [for the] understanding [of] the conditions under which voluntary, open source software development can co-exist in productive balance with proprietary software development" (David et al., 2001, online). Also Giuri et al. (2002) apply a similar perspective arguing that we need to understand "whether the [FL]OSS is different from proprietary software because it is closer to open science" (Giuri et al., 2002, p. 82).

Following this lead, I create a model where the FLOSS community is initially defined as a simple science-like structure (i.e. I assume FLOSS is just the same as science, *mutatis mutandis*) and then confront it with technology for the attraction of researchers. The resulting equilibrium of the model, however, suggests that the definition has limitations that the real FLOSS community

---

"core" of the FLOSS community model can be found in Raymond's articles, as Raymond (1998b), while some insight on the individual motivations and on the "micro" level of the community can be read in Torvalds and Diamond (2001) and Williams (2002). More documentation can be found directly on the Internet, at the Open Source Initiative (www.opensource.org/) and Free Software Foundation (http://www.fsf.org) Web sites.

does not experience. I then propose a wider definition of community, where the empirical literature on FLOSS developers' motivations is used to obtain a new set of equilibria compatible with the development observed in reality. Finally, I use this emended model to discuss possible policies aimed at supporting the creation and growth of knowledge-intensive communities.

3. Investigating science-like community sustainability

3.1 The basic model

Many models have been produced to describe the competition of the FLOSS community model and the IPR-based system, not only in attracting developers (e.g. Gambardella and Hall, 2006; Johnson 2006), but also along a wide variety of economic dimensions, such as market share or sales (e.g. Casadesus-Masanell and Ghemawat, 2006; Economides and Katsamakas, 2006; Lanzi, 2009). However, the perspective outlined by the previous quotes by David et al. (2001) and Giuri et al. (2002) call for a specific modeling structure. FLOSS should be compared to science in terms of its capabilities to confront technology. The work presented in Carraro et al. (2001) and Carraro and Siniscalco (2003) can be used as a starting point for the discussion.

In these papers the authors develop a model in which science and technology incentive schemes are defined by specific payoff functions and then compared on their "capability" to attract researchers. Thus, the field of confrontation between the two institutions is characterized by a representation of the incentives they provide their members, and on their relative capabilities to attract individuals in their ranks. The payoff functions agent *i* is facing are:

$$Science: \qquad \Pi_i^S = F_i(n) + P_r^S(N, x_i^S, X_{-i}^S, X^T)k_i^S - c_i^S(x_i^S) \tag{1}$$

$$Technology: \qquad \Pi_i^T = P_r^T(N, x_i^T, X_{-i}^T, X^C)p_i^T(x_i^T)\gamma_i(x_i^T) - c_i^T(x_i^T) \tag{2}$$

where *N* is the number of total researchers, *n* the number of them who have chosen technology (and thus *(N-n)* is the number of scientists); F_i is the wage of a scientist[6]; P_r is the probability of innovating (which is a positive function of *i*'s own effort x_i and of the externalities from her own group X_{-i} and from the other group *X*[7]); k_i is the "prize" obtained by the success in innovating; c_i is the cost of the research; $\gamma(x_i^T)$ is the knowledge production function; and $p_i^T(x_i^T) = D[\gamma(x_i^T)]$ is

[6]The wage of a scientist is supposed to decrease with the number of scientists, i.e. increase in *n*.
[7]The externalities from science are supposed to always be positive, whatever the group to which the agent belongs, while the externalities from technology are supposed to always be negative.

the inverse demand function. Notice that k_i contains different dimensions, from the increased reputation among peers and in the job market, to the possibility to use the produced knowledge for further research.

The researchers are assumed to have identical preferences, and the payoff that science and technology can provide to their members is a function of the number of each institution's members. Thus, the decision of each agent to be in science or technology is given by a two-stage game where:

1^{st} stage: each researcher compares the payoffs she can garner by entering science or technology, given the decision of the others;

2^{nd} stage: each agent decides 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 and, provided the number of researchers *N* is high enough, the optimal value of *n*, *n**, can be found using the usual conditions (Laffont and Tirole, 1993):

$$\Pi_i^T(n^*) \cong \Pi_i^S(n^*)$$

In graphical terms, then, the equilibrium can be found where the two payoff curves, described as functions of *n*, cross each other.

This model can be used to see what new equilibrium is established when, instead of confronting technology with science, the former is confronted with a knowledge-intensive community such as the FLOSS one. In the next section, this exercise will be undertaken assuming first that the FLOSS community can be considered an institution based on the same mechanisms as science.

3.2 The incentive structure of a science-like community

Many studies, the majority of them inspired by von Hippel's (1988) work, have highlighted 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 the tools to develop software can easily customized the software she uses and even produce the one she needs (von Hippel, 2001). As Bessen (2006) showed, in fact, software is a complex good that can be personalized much more effectively by skilled users than by manufacturers. Once produced, the software is very inexpensive to exchange through the Internet, so that even a very small reward can push developers to exchange the codes they have written. This effect can be achieved by a second set of incentives: developers may aim at producing code and distributing it freely in order to signal their capabilities in the job market (Ghosh, 1998; Lerner and Tirole, 2002; Raymond, 1998c) and acquire a strong reputation among their peers (Dalle and David, 2003).

If own-use, signaling, and reputation are the main incentives in action, the FLOSS community can actually be seen as a "special kind of science" because the two institutions' incentives schemes are very close to each other. As Dasgupta and David (1987, 1994) argued, science is mainly based on the reputation mechanism. Recognition by peers, career concerns, and expectations relative to future research opportunities are the basic leverages acting on scientists' incentives. So in this case, Dalle and David's (2003) and Lerner and Tirole's (2002) intuition relative to the FLOSS community is by and large consistent with the Dasgupta and David (1987, 1994) definition of science.

Moreover, own-use has a relevant role in the scientific environment, at least relative to software development. In the research fields where software is a fundamental instrument, as happens in econometrics, for example, scientists often decide to develop the tools they need, and sometimes they decide to distribute their work widely and freely (Gambardella and Hall, 2006). Thus, the FLOSS community and science also "overlap" with respect to the own-use incentive. If the FLOSS community emerges just from reputation-seeking behaviors, signaling and own-use, this community can be conceptualized as a special kind of science.

3.3 Model I: The science-like community model

Now consider the payoff function[8] assigned to scientist *i*, i.e. equation 1. If a community is nothing more than a group of individuals following an incentive scheme very close to the science one, than the incentive structure of the community must embody the same "forces" of the science scheme. So what is needed now is to build up a payoff function of a community member following the track of equation 1. Notice that equation 1 is, in fact, able to express both own-use and signaling incentives, so it can gather all the claims of the literature on FLOSS based on these incentives. In science, the term $P_r^S k_i^S$ represents the payoff connected with the success in the innovative activity. If the scientist innovates, she's able to signal her capabilities and effort as well as use the software she has just developed. The gain k_i^S represents both of these two dimensions.

To be able to capture the real effect of those dimensions, and thus to describe properly the relationship between k_i^S and k_i^C, we can refer to a wide range of surveys and empirical studies. In particular, we can focus on two of the main surveys, FLOSS-EU (Ghosh et al., 2002) and Boston

[8]Since here *n* represents the number of individuals in technology, the number of scientists or, alternatively, of

Consulting Group (Lakhani et al., 2002), which offer a good representation of the results of all the other empirical surveys.[9]

[Table 1 about here]

As shown in Table 1, own-use related incentives are ranked among the most important motivations, while reputation and signaling do not appear to have such an important role. Given this, we can state that $k_i^C < k_i^S$ because:

1. In a software production environment, the own-use activity gives the same reward in both groups of individuals, so that the level of k_i reached through own-use in science is the same as the level reached in the FLOSS community;
2. Most of the empirical literatures agree that signaling plays a role, but the empirical analysis showed that this kind of incentive is not very important with respect to the others. This is not the case in science, where one's future work opportunities depend heavily on the possibility to signal to the scientific community and other social subjects (the government, the society) one's own effort and capabilities. This central role of the signaling mechanism is proven by the fact that funds and other working opportunities heavily and explicitly depend on the signal each scientist is able to "produce."[10] This means that the *k*-level determined by the signaling incentive is quantitatively less effective in the community, and thus $k_i^C < k_i^S$.

Considering additionally that at the community level there is no government intervention, and thus that the term F_i is equal to zero, we can write the payoff function the member of science-like community faces as:

$$\Pi_i^C = P_r^C(N, x_i^C, X_{-i}^C, X^T)k_i^C - c_i^C(x_i^C) \tag{3}$$

community members is given by *(N-n)*.

[9] These surveys have been chosen as they are among the most representative of the empirical results as a) they rely on a broad sample, b) their formulation is such that they can account for a wide range of different phenomena, and c) their findings are not so qualitatively different from the results of other empirical analysis.

[10]Notice that while a community is *totally* a self-organizing mechanism, and thus it has "transient shape," science tends to be much more stable and structured, despite its degree of self-organization (Giuri et al., 2002). This is because, on the one hand, the relationships with two central actors of the social environment (namely, state and public opinion) force it to be reliable and transparent; on the other hand state and firms also rely on reputation to allocate resources for further research. In science there is not only the *possibility*, as in the community, but rather the *necessity* to realize an articulated structure of reputational gains (for example, providing a series of prizes at different levels) and to spread the relevant information as much as possible to efficiently allocate those gains and magnify the signaling incentive.

Let's recalculate our basic model substituting science with the science-like community described above and considering the comparison between this new institution and technology. Even if there are differences in the two payoff functions, nothing qualitatively changes in the computation of the equilibrium value of the second stage of the game. $F_i(n)$ does not affect the marginal conditions derived with respect to x_i^C, while the inequality $k_i^C < k_i^S$ does not change a non-quantitative discussion. So we just have to focus on the first stage of the game, which describes how the payoff functions of the agents change with respect to *n*, so that it is possible to define the equilibrium value of *n*, i.e. the value of *n* at which no one wishes to move to the other group. In order to move from the model by Carraro and Siniscalco (2001) to the present model, the scientist *i*'s payoff function:

$$\Pi_i^S(n) = F_i(n) + P_r^S[N, \hat{x}_i^S(n), \hat{X}_{-i}^S(n), \hat{X}^T(n)]k_i^S - c_i^S(\hat{x}_i^S(n)) \tag{4}$$

must be substituted by the community member *i*'s payoff function:

$$\Pi_i^C(n) = P_r^C[N, \hat{x}_i^C(n), \hat{X}_{-i}^C(n), \hat{X}^T(n)]k_i^C - c_i^C(\hat{x}_i^C(n)) \tag{5}$$

while $\Pi_i^T(n)$, the payoff function of the technology group member *i*, remains as in equation 2 in both the models.

Now consider the difference between the two payoff functions 4 and 5. To clearly recognize the changes in passing from science to the community, focus on the shifts A and B in Figure 1, where technology (*T*) faces science (*S*).

[Figure 1 about here]

The first change is in the slope of the curves (the A shift). In the case of science, the derivative with respect to *n* of the payoff function is:

$$\frac{\delta\Pi_i^S(n)}{\delta n} = \left\{\frac{\delta P_r^S}{\delta x_i^S}\frac{\delta \hat{x}_i^S}{\delta n} + \frac{\delta P_r^S}{\delta X_{-i}^S}\left[(N-n-1)\frac{\delta \hat{x}_i^S}{\delta n} - \hat{x}_i^S\right] + \frac{\delta P_r^S}{\delta X^T}\left[n\frac{\delta \hat{x}_i^T}{\delta n} + \hat{x}_i^T\right]\right\}k^S +$$

$$-\frac{\delta c_i^S}{\delta x_i^S}\frac{\delta \hat{x}_i^S}{\delta n} + \frac{\delta F_i}{\delta n} \tag{6}$$

In the community case, the slope of the payoff function is the same as in equation 6, but without the $\frac{\delta F_i}{\delta n}$ term and with the term k_i^C instead of k_i^S, where $k_i^C < k_i^S$. This leads directly[11] to the announced inequality:

[11] The fact that $x_i^C(0) \neq x_i^S(0)$ does not generate problems is explained in Appendix A.1.

$$\frac{\delta\Pi_i^C(n)}{\delta n} < \frac{\delta\Pi_i^S(n)}{\delta n}$$

The second shift, B, derives from the change in the intercept, i.e. when $n = 0$. Since there is no reason to assume relevant differences in the shape of P_r and of c_i, in fact, the absence of $F_i(0)$ in $\Pi_i^C(0)$ and $k_i^C < k_i^S$ gives directly $\Pi_i^C(0) < \Pi_i^S(0)$ (see proof in Appendix A.1).

The result is that the number of researchers entering the science-like community shrinks, in the extreme case moving the equilibrium level of *n* from $n*$ to $N-1$, that is, no community at all. Thus, looking at Figure 1, the weakness of a science-like community emerges clearly. The lack of public funding, which is represented by the absence of F_i, and the scarce importance of the signaling mechanism, that leads to $k_i^C < k_i^S$, make the community mechanism unstable. Technology, in fact, is able to completely rule out the community, as in the depicted case, or has the ability to at least limit its size and growth.

This result is consistent with a view that highlights the crucial role that in modern science is played by the State (represented by the non-zero F_i) and by the professionalization of the scientific career (captured in the model by the term k_i^S, whose higher values can be connected to the more structured "reputational game" science relies on). Consider the argument developed by Dalle and David (2007) and David et al. (2001) in the following quotes:

"…one should observe that that the parallel [between science and FLOSS] is by no means exact: formal professional accreditation and institutional affiliation are salient de facto requirements for active participation in modern academic and public sector research communities, yet the computer programming and other software development tasks—whether in the commercial or the free and open-source spheres—remain activities that have resisted becoming 'professionalized.'" (Dalle and David, 2007, p. 393n4)

"[T]he aspect of voluntarism … find[s] interesting historical parallels in the activities of the West's 'amateur' and 'gentlemen scientists' of the late eighteenth and nineteenth centuries. The eventual professionalization of scientific careers might therefore suggest a different fate for the voluntarism that the Open Source movement has thus far managed to harness" (David et al., 2001, online).

In both cases the authors stress the analogy between the FLOSS community and science, but also highlight that the current structure of the FLOSS community resembles the initial stages of the development of open science, where, use of the parameters of the model, F_i and k_i^S were much lower. The modern development of the academic profession, where the role of the State has become crucial in determining fund allocation and the careers of individuals have been systematized into a precise setting of requirements, has moved away from the original embryo of science and gives more importance to F_i and k_i^S. What better resembles the FLOSS community is not the modern structure of science, but rather its original form, where the two incentives represented by F_i and k_i^S were much less effective in attracting researchers. This supports the results derived from Figure 1.

4. The importance of the social dimension

4.1 Empirical evidence

The previous model predicts we should observe no or very limited (in number and size) knowledge-intensive communities. In particular, we should have observed a very small and unstable FLOSS community. However, none of these predictions has did happen: the FLOSS community, as many other knowledge-intensive communities, is growing at a high rate. We this need to modify the model to make it more realistic. In order to do that, the first step is empirical: I need to take into account the empirical evidence relative to the FLOSS case and retrieve from there what factors are missing in the picture.

Lakhani and Wolf (2005) use cluster analysis to aggregate data collected through SourceForge.net (http://sourceforge.net/) and obtain four clusters that can be considered categories or "archetypes" of the FLOSS community members. Each one of the archetypes represents roughly one-fourth of the number of developers, and seems to mix different kind of motivations and lifestyles. Consider, however, how they describe the last cluster: "Cluster 4 (19 percent of the sample) consists of people motivated primarily by obligation/community-based intrinsic motivations. A majority of this cluster report group-identity-centric motivations derived from a sense of obligation to the community and a normative belief that code should be open" (Lakhani and Wolf (2005, p. 16). Also a second look at Table 1 focusing on the motivations marked with "O" (other) gives a similar picture. The role of the "social" side of the community is central, and sometimes even more important than own-use, reputation, and signaling. This point of view is consistent with the vast majority of the literature that actually underlines the importance of socially determined motivations in the definition of the FLOSS model (Bonaccorsi and Rossi, 2003a; David and Shapiro, 2008; Elliott and Scacchi, 2003; Ghosh et al., 2002). So it is possible to start from this point of view in order to search for a more realistic description of the FLOSS community.

4.2 The social side of a community

In the context of the FLOSS community, the social dimension of motivations has been analyzed with respect to such theories as gift economy (Raymond, 1998c), communities of practice, or epistemic communities (Cohendet et al., 2001; Edwards, 2001; Lin, 2003a, 2003b, 2004a, 2004b; Mateos-Garcia and Steinmueller, 2008), and general reciprocity (Luo, 2002). In particular, the community-of-practice perspective (Wenger, 1998) can be particularly useful to describe in detail the passages of the social processes at work in the FLOSS community, and thus can be used to emend the description of the community incentive scheme adopted in the model.

Applying this perspective to the FLOSS community means recognizing the central role of the process of "negotiation" (Lin, 2004a) that developers are involved in when creating software. Developing a common project together makes people continuously "renegotiate" the meanings connected with their own actions with the others, giving sense to the common action and to the social context where it takes place. This negotiation of meanings leads to a continuous reshaping of the participants' vision of the world to adapt their *identities* to the social circumstances they are embedded in (Wenger, 1998). But changing individuals' identities means configuring in a new way the principles guiding their actions and their priorities, precisely those principles represented by their payoff functions. In other words, the interaction between the community members' leads to change in their identities that ultimately results in a structural modification of their payoff functions, that now 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." (Shah, 2006, p. 1000). Among 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 [community of FOSS users] members, the community metamorphosizes into a friendship group and a social entity with which one identifies." (Bagozzi and Dholakia, 2006, p. 1111).

4.3 Science is a community-like institution

Taking seriously the social side of the community means *also* recognizing its role in the scientific institution. Like community, in fact, science can also be considered a structure internally organized around a system of meanings debated and shaped by scientists. The scientific community—which is in fact a community—is the real protagonist of the scientific organization, allowing internalization of the rules, learning, production, and allocation of reputation, and providing space in which the produced knowledge is composed into a meaningful system. Thus, the social mechanisms above also act inside the scientific community, making science closer to the communitarian institution. So what is the difference between the two institutions, if any?

Science and community are similar *ab origine*,[12] they have the same *nature*: signaling and reputation, own-use, and social interaction are, in both cases, the main factors determining the payoff function. What is different is the costs/benefit structure. While science can rely on State intervention assuring high levels of F_i and k_i^S, a community has a bottom-up growth, and must be able to generate its own reproduction process without employing these mechanisms. That is why in the FLOSS community, *social* motivations appear to have a central role both theoretically and empirically.

5. Emending the community payoff function

5.1 Accounting for the social dimension: a payoff function for an identity-based community

The social mechanisms described above act along the nexus of the communitarian ties, modifying the structure of the payoff function and the relative importance of its constituting elements. To adequately represent this new payoff structure, we must take into account five factors taken mainly from Wenger's (1998) idea of community of practice:

1. *Communitarian activity*: What makes a group of people becoming a community is the construction of a social environment where identities are defined together with the

[12] Recall that Dalle and David (2003) argue that there is little difference between the FLOSS community and the organization of science in its early stages. This confirms that the two institutions share the same nature. However, there is an important difference: "patrons" are key actors in the evolution of science, as is the State today, while individual and/or state patronage is only a subsequent phenomenon in the FLOSS history. Especially at the beginning, patronage could be considered as just one among a wide range of factors at work in the production of FLOSS. The FLOSS-US survey (David et al., 2003), for example, highlights the fact that 75% of the participants are not paid for developing FLOSS, 14.6% are paid for developing FLOSS, and 10.2% are paid for developing FLOSS in professional work. From other answers in the survey, we find that monetary earnings through FLOSS are zero for the 56.8% of the participants in the survey, and, while participants paid for supporting FLOSS are 13.0%, 7.0% are paid for administrating FLOSS, 6.1% landed a job because of FLOSS experience, and 13.2% refer to other reasons. But the presence of external founding—especially with the entry of the State—is precisely what has determined the deep structure of science and its cost/benefit trade-off. To a certain extent then, the reasoning could be reversed: the "archetype" is not the "open science" structure, but rather the "community structure," which is the same for the early stages of science and for the modern day structure of the FLOSS community. Science developed toward another cost/benefit trade-off thanks to the role of patronage and the State.

activity of the community (e.g. in the FLOSS case, producing software). All the processes take place *in* that environment and *thanks to* that environment. The common activity must have a central role in the payoff function and must depend on the effort of all the members of the community;

2. *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 her payoff function is greater. For example, in the FLOSS case, the GNU/Linux development has a greater effect on the payoff of a person who "believes" in the GNU/Linux project, greater than the payoff ensured by the simple usage of GNU/Linux. This translates in the model of Wenger's (1998) idea of *engagement*, where individuals are involved into a process of reciprocal influence between them and the social environment of the community. The more a member invests in the common activity undertaken in the community to define her identity, the higher is the payoff she gets from that activity;

3. *Endogeneity of personal involvement*: In interacting, the members of the community start the "renegotiation" of their vision of the world. The process develops and becomes stronger in terms of affecting members' behaviors as the "volume" of the interaction grows. So, when a community grows, it not only becomes "quantitatively" stronger (e.g. it produces more software), but also "qualitatively," determined by a higher rate of the personal involvement of its members. Thus, personal involvement must be a function of the community size.

4. *The cooperation costs*: A group of people who collaborate is subject to free riding episodes. The group must then create some rules and enforcing mechanisms to sustain the cooperation and avoid free riding. There are costs, in terms of monitoring others' behavior, spreading the information about it, discovering the break of a rule, and punishing the free rider. These costs are obviously exponentially increasing with the number of the community members, but are decreasing in the personal involvement of the members. This second effect is obtained by two mechanisms. On the one hand, the internalization of the community rules through the personal involvement reduces the "appeal" of free riding behaviors. On the other hand, we can observe something that we can call a "free-riding exclusion" effect;

5. *The "free-riding exclusion" effect*: When member *i* starts to engage in the

communitarian activity and to believe in the common enterprise, she begins to *perceive* the community as a trustworthy environment. Thus, the simple fact that *j* also belongs to the community is taken by *i* as a signal of *j*'s trustworthiness. *j*'s potential free-riding behavior is perceived by *i* as an almost-irrelevant exception, and *i* reduces her monitoring and punishing activities, decreasing the cooperation 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" (Bagozzi and Dholakia, 2006: 1111). So, the cooperation cost each member has to sustain in order to "trust" the others is reduced by the involvement level through the "free-riding exclusion" effect.

The following payoff function tries to express in mathematical terms the five factors analyzed above:

$$\Pi_i^C = P_r^C(N, x_i^C, X_{-i}^C, X^T)k_i^C - c_i^C(x_i^C) + e(n)\cdot Y(N, x_i^C, X_{-i}^C, X^T) - C(n; e(n)) \tag{7}$$

where $e(n)$ is the personal involvement, $Y(x_i^C, X_{-i}^C, X^T)$ is the communitarian activity, and $C(n;e(n))$ are the cooperation costs expressed as a function of *n* and of the involvement $e(n)$. To take into account the previous five points, we must assume that:

$$\frac{\delta e(n)}{\delta n} < 0 \qquad \frac{\delta C(.)}{\delta n} < 0 \qquad \frac{\delta C(.)}{\delta e} < 0 \qquad \frac{\delta Y(.)}{\delta x_i^C} > 0 \qquad \frac{\delta Y(.)}{\delta X_{-i}^C} > 0 \qquad \frac{\delta Y(.)}{\delta X^T} < 0$$

$$C(N-1, e(N-1)) = 0 \qquad \frac{\delta C(.)}{\delta(N-n)} > 0 \qquad \frac{\delta^2 C(.)}{\delta(N-n)^2} > 0 \tag{8}$$

In the case of technology, we can use the same payoff function used by Carraro and Siniscalco (2003) simply by slightly changing the effect of the communitarian externalities. In science, and in a science-like community, the produced knowledge can be easily adapted and translated into an IPR regime by someone other than the innovator. In the FLOSS community, however, the GPL protects the produced knowledge, thus avoiding this risk. This does not cancel out the benefit that technology has from the community production of software (ideas can be reused because GPL is not a patent), but GPL does limit the effects of communitarian externalities on the technology payoff function. This effect is represented with the factor $g < 1$, which multiplies the externalities of the community, X^C. Thus, the payoff function in the case of technology is:

$$\Pi_i^T = P_r^T(N, x_i^T, X_{-i}^T, g \cdot X^C) p_i^T(x_i^T) \gamma_i(x_i^T) - c_i^T(x_i^T)$$

5.2 Model II: The optimal effort and the new communitarian dimension

To derive the equilibrium of the game as done for the previous model we must compute the FOCs with respect to the effort x_i and get the reaction function of each player. The system of these functions together with the symmetry hypothesis gives us the Nash equilibrium and the optimal effort levels. What is necessary to take into account is how $\hat{x}_i^C$ is determined in the present model with respect to the science-like case. Consider the following derivative:

$$\frac{\delta \Pi_i^C}{\delta x_i^C} = \underbrace{\frac{\delta P_r^C(N, x_i^C, X_{-i}^C, X^T)}{\delta x_i^C} k_i^C - \frac{\delta c_i^C}{\delta x_i^C}}_{A} + \underbrace{e(n) \cdot \frac{\delta Y(N, x_i^C, X_{-i}^C, X^T)}{\delta x_i^C}}_{B}$$

While the terms marked with A induce the same reaction as in the science-like case, the term B moves the optimal effort level to a different level. To describe the result of this movement we need to consider how *Y(.)* is construct. Despite the similarities between P_r^C and *Y(.)*, in fact, the two functions are substantially different. While in P_r^C the exchanged knowledge are kind of perfect substitutes, here they reciprocally interfere at the level of the production process in such a way that they become kind of complementary factors.

As in Saint-Paul (2003), the knowledge that A transmits to B enhances B's productivity. Since this effect is related to *Y(.)*, it means that A's transfer enhances B's ability to extract value from the communitarian activity. This happens because *Y(.)* does not represent an intermediate activity, an instrument to reach a different goal. The essence of the communitarian activity, in fact, is realizing an artifact but also cooperating, sharing and working together. Participating in a collective work gives then more value the more has being produced on the *Y(.)* side, and this boosts the optimal effort.

What this implies in mathematical terms is a situation in which the shape of *Y(.)* can increase the agent's optimal effort with respect to the science-like case, as described in figure 2.

[Figure 2 about here]

5.3 The final equilibrium of the game: a identity-based community faces technology

By reproducing the same analysis undertaken by Carraro and Siniscalco (2003) described in section 2.4, we can derive the optimal effort value of each agent as a function of *n*. By replacing the optimal effort value in the payoff functions we obtain:

$$\Pi_i^C(n) = P_r^C[\hat{x}_i^C(n),(N-n-1)\hat{x}_i^C(n),n\hat{x}_i^T(n)]k_i^C - c_i^C(\hat{x}_i^C(n)) +$$

$$+e(n)\cdot Y[\hat{x}_i^C(n),(N-n-1)\hat{x}_i^C(n),n\hat{x}_i^T(n)] - C(n;e(n))$$

$$\Pi_i^T(n) = P_r^T[\hat{x}_i^T(n), g\cdot(N-n)\hat{x}_i^C(n),(n-1)\hat{x}_i^T(n)]p_i^T(\hat{x}_i^T(n))\gamma_i(\hat{x}_i^T(n)) - c_i^T(\hat{x}_i^T(n))$$

To understand the differences between the payoff functions of the present model and the functions used in the previous one (and in Carraro and Siniscalco, 2003) consider the changes in the intercept and in the derivatives with respect to *n*.

Let's start from the **intercepts**.

In the *technology* the intercept in $n = N$ is the same as in the Carraro and Siniscalco model, because there are no people in the community and thus the only value that is different between that model and the present one, the communitarian externalities, is equal to zero.

In the *community*, the intercept in $n = N-1$ is the same as in the science-like case, apart from the term $e(n)\cdot Y[\hat{x}_i^C(n),(N-1)\hat{x}_i^T(n)] > 0$. The shift upward, in any case, is very small, even negligible, given the fact that the term expresses the community environment. Such an environment is obviously associated with at least two members, since one member does not make a community. In this case, the term expresses just the interest of the individual for a project that can be *potentially* undertaken by a community. Moreover, X^T is also very close to the science-like community case, because the GPL effect represented by g acts on few externalities (actually $X^C = x_i^C$), leaving the technology production of knowledge substantially unchanged and thus X^T. So, there is no change in the impact of technology on the community side, and the intercept of the community payoff function is very close to the Carraro and Siniscalco (2003) case.

Computing now the **derivatives** with respect to *n* we have:

The *technology* case:

$$\frac{\delta\Pi_i^T}{\delta n} = \left[\frac{\delta P_r^T}{\delta x_i^T} - \frac{\delta P_r^T}{\delta X_{-i}^T}\right]\frac{\delta x_i^T}{\delta n}R_i^T - g\cdot\underbrace{\frac{\delta P_r^T}{\delta X^C}(1-\delta_i^C)\hat{x}_i^C R_i^T}_{A} - \left[-\frac{\delta P_r^T}{\delta X_{-i}^T}\right](1-\delta_i^T)\hat{x}_i^T R_i^T +$$

$$-P_r^T(.)(1-\varepsilon)p_i^T\frac{\delta\gamma_i}{\delta x_i^T}\left(-\frac{\delta x_i^T}{\delta n}\right)+\frac{\delta c_i^T}{\delta x_i^T}\left(-\frac{\delta x_i^T}{\delta n}\right)$$

where $R_i^T = p_i^T\gamma_i(x_i^T)$ is the revenue, $\varepsilon<1$ is the inverse of demand elasticity, and δ_i^C and δ_i^T are the elasticities of the optimal effort with respect to the group size of community and technology respectively.

The only difference with the Carraro and Siniscalco(2003) model is the factor " g ," which multiplies the term A.[13] Since $g<1$ and the factor A is positive, the derivative in the present model must be less, in absolute terms, than the derivative in that model. So the function is flatter.

The *community* case*:*

$$\frac{\delta\Pi_i^C(n)}{\delta n}=\left\{\frac{\delta P_r^C}{\delta x_i^C}\frac{\delta\hat{x}_i^C}{\delta n}+\frac{\delta P_r^C}{\delta X_{-i}^C}\left[(N-n-1)\frac{\delta\hat{x}_i^C}{\delta n}-\hat{x}_i^C\right]+\frac{\delta P_r^C}{\delta X^T}\left[n\frac{\delta\hat{x}_i^T}{\delta n}+\hat{x}_i^T\right]\right\}k^C+$$

$$-\frac{\delta c_i^C}{\delta x_i^C}\frac{\delta\hat{x}_i^C}{\delta n}+e(n)\left\{\frac{\delta Y}{\delta x_i^C}\frac{\delta\hat{x}_i^C}{\delta n}+\frac{\delta Y}{\delta X_{-i}^C}\left[(N-n-1)\frac{\delta\hat{x}_i^C}{\delta n}-\hat{x}_i^C\right]+\frac{\delta Y}{\delta X^T}\left[n\frac{\delta\hat{x}_i^T}{\delta n}+\hat{x}_i^T\right]\right\}+$$

$$+\frac{\delta e(n)}{\delta n}Y(.)-\frac{\delta C(.)}{\delta n}-\frac{\delta C(.)}{\delta e}\frac{\delta e(n)}{\delta n}$$

To understand the changes in the derivative, first of all we need to know how $\hat{x}_i^C$ reacts to the changes in *n*. As noted above, the shift of *Y(.)* can counterbalance the movement of P_r^C, as in Figure 2. The net effect is then unknown. By the way, we can derive some observations from the theory to state the specific behaviors of the curves.

A community is a network, and thus, as far as terms showing complementarity as *Y(.)* are concerned, Metcalfe's law holds. This means that the communitarian environment shows a kind of increasing return to scale: the more the members, the more the success of the community and the

[13]There is no difference in the influence of the community/science between the two models. Here, as in the previous model, the only effect of the community/science on the technology payoff function is through its externalities, which in both cases follow the rule $\frac{\delta X^C}{\delta n}<0$.

more the marginal benefit of increasing members' own efforts. The counterpart of this explosive behavior is a slow and difficult beginning, when *(N-n)* is small and the members are few. Moreover Metcalfe's law faces the obstacles of coordination costs.[14] In big communities, i.e. for the highest values of *(N-n)*, coordination costs overcome the increasing returns of Metcalfe's law, determining overall decreasing returns. Up to that point, by the way, it is realistic assuming $\frac{\delta^2 Y(.)}{\delta(N-n)^2} > 0$, that is, the more the *(N-n)*, the more the effect of a change in *n* on the *Y(.)*.

If this is the pattern, we can realistically assume that when *n* is high (i.e. *(N-n)* is small) the effect of its changes on the optimal effort level is dominated by P_r^C, decreasing $\hat{x}_i^C$. When *n* is instead in a kind of "middle region," *Y(.)* dominates the movements of the whole payoff curve, determining an increases in $\hat{x}_i^C$. After a certain threshold, both P_r^C and *Y(.)* begin to show decreasing returns and to slow down the increase in $\hat{x}_i^C$. Figure 3 shows the different phases of the process.

[Figure 3 about here]

Mathematically:

During the α phase $\rightarrow \frac{\delta \hat{x}_i^C}{\delta n} > 0$

During the β phase $\rightarrow \frac{\delta \hat{x}_i^C}{\delta n} < 0$

During the χ phase $\rightarrow \frac{\delta \hat{x}_i^C}{\delta n} = 0$

A second question regards the effect of the change in *n* and in the optimal effort on the externalities of the community.

In the β phase of Figure 3, this is straightforward: as *n* decreases (*(N-n)* grows), the optimal effort increases. The increase in the number of community members and in the effort increases the communitarian externalities. This increment acts-back on the optimal effort level, on the one hand, containing its increase through P_r^C, on the other hand, expanding it through *Y(.)*. Since *Y(.)* is stronger in this phase, the optimal effort increases again, even if less than before because of the P_r^C effect.

[14] Notice that here we do not refer to *cooperation* costs, which have been defined above as something much more related to the agents' strategic behaviors and expectations rather than to division of labor and *coordination* of operations.

In the α phase, however, the reduction of *n* (the growth of $N-n$) is counterbalanced by the decrease in $\hat{x}_i^C$. In the Carraro and Siniscalco (2003) model, the same ambiguity is solved by an assumption (assumption 3, p.582), which assigns a greater effect to the size dimension than to the effort dimension. In mathematical terms this means:

$$\tilde{n}_i^C = -\left[\frac{\delta \hat{x}_i^C}{\delta(N-n)} \frac{N-n}{\hat{x}_i^C}\right] < 1 \qquad \tilde{n}_i^T = -\left[\frac{\delta \hat{x}_i^T}{\delta n} \frac{n}{\hat{x}_i^T}\right] < 1 \tag{9}$$

where $\tilde{n}_i^C$ and $\tilde{n}_i^T$ are the elasticities of the optimal effort with respect to the group size. This hypothesis, then, induces again an increase in $\hat{X}^C$, which acts-back on $\hat{x}_i^C$ inducing a greater contraction of the effort. In both cases, then, we have $\frac{\delta X^C}{\delta n} < 0$[15].

Consider now that the reaction of technology to these movements is in both cases qualitatively the same as in Carraro and Siniscalco (2003), so that $\frac{\delta X^T}{\delta n} > 0$. A reduction of *n* then diminishes X^T, and this enhances P_r^C, which tends to reduce the effort, and also *Y(.)*, which instead increases it. Since in the β phase the *Y(.)* effect is greater, the optimal effort increases, while in the α phase it is exactly the opposite.

Eventually, in the χ phase, when the community is huge (*n* is close to zero), a movement in *n* implies negligible variations of the optimal effort and also of the externalities perceived by each member, so that $\frac{\delta x_i^C}{\delta n} \cong 0$ and $\frac{\delta X^C}{\delta n} \cong 0$. On the technology side, a huge community means a very small technology, so we can state that $\frac{\delta X^T}{\delta n} \cong 0$. In this phase, then, nothing moves in P_r^C, in *Y(.),* or in the optimal efforts.

All we have seen up to now is consistent with Figure 3, so the three phases characterize the optimal effort even when all the variables are considered.

We are now ready to describe the behavior of $\frac{\delta \Pi_i^C}{\delta n}$. In particular, the differences between the science-like community (*Slc*) and the identity-based community (*Idc*); that is, the difference:

$$\frac{\delta \Pi_i^{Slc}(n)}{\delta n} - \frac{\delta \Pi_i^{Idc}(n)}{\delta n} =$$

[15]This is the result anticipated in Footnote 15.

$$\underbrace{\left\{\frac{\delta P_r^{Slc}}{\delta x_i^{Slc}}\frac{\delta \hat{x}_i^{Slc}}{\delta n}+\frac{\delta P_r^{Slc}}{\delta X_{-i}^{Slc}}\left[(N-n-1)\frac{\delta \hat{x}_i^{Slc}}{\delta n}-\hat{x}_i^{Slc}\right]+\frac{\delta P_r^{Slc}}{\delta X^T}\left[n\frac{\delta \hat{x}_i^T}{\delta n}+\hat{x}_i^T\right]\right\}k^{Slc}}_{A}+$$

$$\underbrace{-\frac{\delta c_i^{Slc}}{\delta x_i^{Slc}}\frac{\delta \hat{x}_i^{Slc}}{\delta n}}_{B}+$$

$$\underbrace{-\left\{\frac{\delta P_r^{Idc}}{\delta x_i^{Idc}}\frac{\delta \hat{x}_i^{Idc}}{\delta n}+\frac{\delta P_r^{Idc}}{\delta X_{-i}^{Idc}}\left[(N-n-1)\frac{\delta \hat{x}_i^{Idc}}{\delta n}-\hat{x}_i^{Idc}\right]+\frac{\delta P_r^{Idc}}{\delta X^T}\left[n\frac{\delta \hat{x}_i^T}{\delta n}+\hat{x}_i^T\right]\right\}k^{Idc}}_{C}+$$

$$+\underbrace{\frac{\delta c_i^{Idc}}{\delta x_i^{Idc}}\frac{\delta \hat{x}_i^{Idc}}{\delta n}}_{D}-e(n)\underbrace{\left\{\frac{\delta Y}{\delta x_i^{Idc}}\frac{\delta \hat{x}_i^{Idc}}{\delta n}+\frac{\delta Y}{\delta X_{-i}^{Idc}}\left[(N-n-1)\frac{\delta \hat{x}_i^{Idc}}{\delta n}-\hat{x}_i^{Idc}\right]+\frac{\delta Y}{\delta X^T}\left[n\frac{\delta \hat{x}_i^T}{\delta n}+\hat{x}_i^T\right]\right\}}_{E}+$$

$$\underbrace{-\frac{\delta e(n)}{\delta n}Y(.)}_{F}+\underbrace{\frac{\delta C(.)}{\delta n}}_{G}+\underbrace{\frac{\delta C(.)}{\delta e}\frac{\delta e(n)}{\delta n}}_{H} \tag{10}$$

The analysis of the difference must be divided in the three cases, α, β, and , χ corresponding respectively to the three phases in Figure 3.

The α case:

In this case *n* is too big ($N-n$ too small) to produce optimal effort levels and communitarian and technological externalities that are significantly different between the two cases, so:

$$\hat{x}_i^{Slc} \cong \hat{x}_i^{Idc} \qquad \frac{\delta \hat{x}_i^{Slc}}{\delta n} \cong \frac{\delta \hat{x}_i^{Idc}}{\delta n}$$

Given the isomorphism of P_r and c_i and the fact that $k^{Idc}=k^{Slc}$, in equation 10, B cancels out D, and A cancels out C.

From the hypothesis described in equation 8, we can directly derive that term F is positive, G is negative, and H is positive. Notice that in this phase the G term must be very small, because when the community is small, increasing the number of members by one unit does not significantly change the cost that each member has to bear to cooperate with the others, including the new member. The same increase by one unit, however, has a greater effect on *e(.)*, because in a small community, one more member means one more person to interact with and to cooperate with directly, almost “one to one.” This implies a significant level of $\frac{\delta e(n)}{\delta n}$, and thus an increase both in terms F and H. In this phase, then, we should assume that F+G+H > 0, even if this sum can also be low.

The remaining term E is also very small, since in the α phase the personal involvement $e(n)$ and the action of $\frac{\delta Y(.)}{\delta n}$ are not very significant because of the small size of the community. Moreover, *Y(.)* is the net value of two opposite forces, which enforces the previous argument. To see this,[16] recall equation 9 and rewrite term E as:

$$e(n)\frac{\delta Y(.)}{\delta n}=\underbrace{\left[\frac{\delta Y(.)}{\delta x_i^{Idc}}-\frac{\delta Y(.)}{\delta X_{-i}^{Idc}}\right]\frac{\delta \hat{x}_i^{Idc}}{\delta n}e(n)}_{I}+$$

$$-\underbrace{\left[\frac{\delta Y(.)}{\delta X_{-i}^{Idc}}(1-\tilde{n}_i^{C})\hat{x}_i^{Idc}e(n)+\left(-\frac{\delta Y(.)}{\delta X_{T}^{Idc}}\right)(1-\tilde{n}_i^{T})\hat{x}_i^{T}e(n)\right]}_{II} \qquad (11)$$

If we assume realistically that a change in i's own effort x_i^{Idc} has an effect on *Y(.)* greater than a change in the externalities X_{-i}^{Idc}, then the I term is positive, because in this phase $\frac{\delta \hat{x}_i^{Idc}}{\delta n}>0$. For this last reason, and because of equation 9, the II term is also positive. Since term E is given by (I-II), it is the net value of two opposite forces, which partly compensate each other, reducing E. The result of the whole discussion then is that E can be reasonably assumed to be very small. So, in the α phase, *the derivative of the Idc curve is slightly less than the Slc one*:

The β case

In this case it is possible to reproduce the same decomposition applied in equation 11 to what in equation 10 is the term A (relative to $P_r^{Slc}(.)k^{Slc}$). This is possible because also in this phase $\frac{\delta \hat{x}_i^{Slc}}{\delta n}>0$ and equation 9 still holds. So, $P_r^{Slc}(.)k^{Slc}$ is just the net value of two opposing forces. Moreover, on the Slc side, decreasing returns start to have a significant effect, so that $\frac{\delta P_r^{Slc}}{\delta n}$ is now small. The result is that in equation 10, term A can also be considered very small.

Applying the same decomposition in equation 11 to terms C and E in equation 10, and noticing that in this phase $\frac{\delta \hat{x}_i^{Idc}}{\delta n}<0$, equation 9 directly leads to C < 0 and E < 0 (that is: $-C-E>0$). Consider also that here the community is big enough to develop a significant level of $e(n)$ and

$\frac{\delta Y(.)}{\delta n}$, so that E is significantly high.

The part of equation 10 defined as (F+G+H) shows the same behavior in the α phase, even if all the forces are moving toward a balance. The derivative $\frac{\delta e(n)}{\delta n}$ is still high, because the community is still able to produce consistent levels of personal involvement, but the marginal effect of *n* starts to decrease as the community grows. In any case, in this phase, *Y(.)* effect begins to be very important, both in absolute and marginal terms, so that its effect on F is strong and positive. So, (F+H), is still high enough to balance the negative effect of term G. In this phase, then, the costs of cooperation become important, but the community is still able to manage the mix of trust and punishment that is at the basis of its activity. So, assuming (F+G+H) are very small, or even slightly positive, is realistic.

Eventually, B is positive and D is negative, because the optimal effort in the *Slc* case still follows the *n* movements, while in the *Idc* case it now moves in the opposite direction.
At the end, the sign of the whole difference is given by:

$$\underbrace{-C}_{+}\underbrace{-E}_{+}\underbrace{-B}_{-}\underbrace{+D}_{-}$$

To make it positive, the following inequality must hold:

$$\frac{-C-E+D}{B}>1 \tag{12}$$

The previous inequality clearly expresses an effect on the ability of a community mechanism to create incentives as the number of members grows. Inequality 12, in fact, can be read as follows: as *n* decreases, the ratio between the net gain coming from the *Idc* community mechanism[17] and the gain—in terms of decrease in the costs of effort—from the *Slc* side must be greater than one. In other words, an *Idc* community must produce a greater gain than the cost saving in the *Slc* case. This seems a plausible condition, since the involved gains on the *Slc* side are relative just to the costs-saving mechanism, while term C, and especially term E, are significantly positive. To support this idea, consider the fact that in this phase there are many members in the community. This means that:

1. the decreasing returns of the optimal effort curve had already decreased $\frac{\delta \hat{x}_i^{Slc}}{\delta(N-n)}$, and thus $\frac{\delta \hat{x}_i^{Slc}}{\delta n}$ is also small;
2. the marginal cost of effort can be assumed to be increasing with the effort, and thus the derivative $\frac{\delta c_i^{Slc}(\hat{x}_i^{Slc})}{\delta x_i^{Slc}}$ decreases as the effort decreases. So, since an increase in *(N-n)*

---

[16]See Carraro et al. (2001) for a similar discussion about the P_r^S term.
[17]That is, the sum of (-C), the gains coming from the higher probability to innovate, plus E, the community project

determines a contraction in the effort, the previous derivative is small.

Given this, the idea that inequality 12 holds can be considered realistic.

In any case we don't need to consider it valid for all the β phase. As *(N-n)* grows, it is possible that the inequality 12 becomes false. This is not a problem in the model, it simply implies that in the β phase the community is able to provide members with a gain higher than in the Slc case, just below a certain threshold, a certain community size. Thus, in the β phase, *the derivative of the Idc curve is much smaller than the Slc one for the highest values of n. This can be true also for the lowest values of n, but not necessarily.*

The χ case

In this phase the decreasing returns are the dominant force, so externalities and efforts are not influenced by movements of *n*, and neither is the term $e(n)$. All the derivatives are then equal to zero. The only derivative that is non-zero is the term G, because the marginal cost of cooperation increases in *(N-n)*. In the χ phase, then, *the derivative of the Idc curve becomes greater than the Slc one*.

Summarizing, the *Idc* curve and the *Slc* function are in the relationship described by Figure 4:

[Figure 4 about here]

6. Analysis of the results

6.1 The equilibria of the game

Let's now put all the pieces of the model together. The final result is depicted in Figure 5.

[Figure 5 about here]

As it is easy to see, the equilibria of the model in the case of a identity-based community are:[18]

a, which is a stable equilibrium associated with a community strong enough to resist the technology competition. The difference between the equilibria associated with the

implementation, minus the increase in costs due to the higher effort boosted by *Y(.)*.

[18]This is not the only possible outcome. If in $n=(N-1)$ the "community" payoff is less than the technology one, the equilibria ***a'***, a_{noGPL} and even ***a*** disappear. In this case no small community is possible, and if a big enough community is created, it grows endogenously until ***c***. The same discussion applies if the curves are such that equilibria ***a*** and ***b*** coincide or do not exist at all. Of course, opposite arguments can be stated if ***c*** does not exists or if it coincides with ***b***. Eventually, if the community (technology) payoff is always greater than the technology (community) one, then there is

science-like community, ***a***, and the one associated with the identity-based comunity, ***a'***, represents exactly this point. In particular, ***a*** is the equilibrium in which the model expresses the possibility of the *existence* of a community thanks to the introduction of the social side of the community;

b, which is an unstable equilibrium, a threshold. Such a threshold, widely recognized in the literature about communities (Bonaccorsi and Rossi, 2003b), divides the realm of small communities from the set of communities that is able to grow fast and large.[19] In each one of those spaces, the dynamics of the model shows a bandwagon effect consistent with what other authors have found (e.g. Bitzer and Schröder, 2005);

c, which is the second stable equilibrium, associated with a large community. This equilibrium expresses the potential of a community much better than ***a***. If the community grows enough and overcomes the limit ***b***, then it grows endogenously until the equilibrium ***c***. This is the case of the FLOSS community.[20]

Several forces contribute to determining the selected equilibrium of the model. Some of them are parameters determining the "landscape" of the payoffs, others are "engines" able to push agents along one particular path of this landscape.

6.2 The protective structure of the community

The effect of the GPL is fundamental in enhancing the community sustainability, but it is just the precondition, it is not the engine of the community growth. In the model, the passage from ***a'*** to a_{noGPL} is negligible, showing that a community in which the social side is important, but where any protection mechanism is lacking, can only reproduce a science-based community. Without GPL, even this community is too weak.[21] However, the GPL, and in general all the rules communities set-up in order to protect themselves from "hostile" behaviors, must be conceptualized as necessary structures to give the community the *possibility* to survive and grow, *but they cannot trigger that growth*. To grow a community needs an engine, which must be found in other parameters pushing the model toward higher-level equilibria as ***c***.

just a corner solution: all agents in community (technology).

[19] On the related process of growth see, for example, Bonaccorsi and Rossi, 2003b.

[20] Notice that this approach takes into account the quantitative aspect of the FLOSS 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 put 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).

[21] It is easy to see from Figure 5 that the shift GPL induces in the technology payoff has a significant effect in the science-like community case also.

6.3 Exploring the space of the model's parameters

Some parameters have been included in the model such as the personal involvement $e(n)$ or the communitarian activity *Y(.)*, while others, which I introduce now, have not been discussed, such as the level of protection granted by intellectual property rights.

Moreover, if we consider the model in a dynamic framework, the process of the equilibrium selection is substantially composed by two kinds of forces. The first includes all the parameters that can affect the initial *n* level. If the initial *n* is the region (***b***; N], at the end the equilibrium will be $n = n_a$, while for initial values belonging to [0; ***b***), the selected equilibrium will be $n = n_c$.[22] The second is comprised of forces shifting and reshaping the payoff functions, and thus changing the relative position of the equilibria. In this case, if ***b*** is shifted to the right, the interval (***b***; N] shrinks. The parameters setting the initial condition of *n*, then, can assume a wider range of values to which the ***c*** equilibrium is associated, making it more likely.

6.4 The parameters outside the model

6.4.1 The interest toward the community activity

Considering the FLOSS case we can observe that communities become economically relevant when they fill an unfilled market, creating one *ex novo* or providing the conditions to fill an established one. 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 FLOSS case allowing users to be part of the process—becomes crucial. If this is true, the simple existence of a community can attract all the individuals interested in that market (Green, 1999). The more the community responds to unfilled gaps, the more attractive it becomes to interested individuals.

Moreover, communities, as other institutions, cover a particular space of social interaction. Their role is to 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 FLOSS 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 free-riding exclusion to recognition by peers), meanings (what 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

[22]The probability of an initial value of *n* such that the equilibrium corresponds exactly to a payoff=***b*** is zero, because it has the same probability to draw one specific value from a continuous distribution.

attracted to the community, and may become eventually members. The FLOSS community history seems to have walked along both paths.

In our static model, the more this dynamical mechanism is strong and reflects the capability of a community to attract new developers, the smaller has to be the initial chosen value of n.

6.4.2 Opportunities: the IPR regime and community space

Another force that can contain or expand the initial n is the ability of the community to offset the expected costs connected with its own activity. By definition, communities are built on voluntary contributions. If members can undertake the communitarian activity, which gives "life" to the community, by only bearing a very high cost, their initial motivation must be high enough to overcome the expected costs, and vice versa. These costs have to be computed taking into account the expectations relative to future opportunities of the community, and may rise when future opportunities are constrained.

Inside the broad category of parameters defining these opportunities, the IPR regime has a central role. Communities, in fact, are devices to produce knowledge and artifacts, and thus future opportunities to develop and expand are also shaped by the proprietary regime into which their knowledge is placed. In the case of software, the possibility to patent software code, and in general the capability of IPR to enhance software production, is still a challenging debate. But in the FLOSS community case, the same debate puts forward questions on the community survival (Stallman and Garfinkle, 1992). In this case, the role of GPL is fundamental in creating a space in which the FLOSS community can develop *inside* the property rights structure (Gambardella and Hall, 2006); while the enforcement of IPR, for example expanding the right of software producers to patent their code, can have exactly the opposite effect, depressing the development of the community software and limiting its growth. As Linus Torvalds and Alan Cox put it: "Software patents are also the utmost threat to the development of Linux and other free software products, as we are forced to see every day while we work with the Linux development" (Torvalds and Cox, 2003).

But property rights are not the only factor determining the opportunities of communities. The second factor is the availability of "spaces" in which the community can act. In the FLOSS case, all the literature agrees on the fundamental role of the Internet in enabling the GNU/Linux development (Cammozzo, 2007). With such a space, where codified knowledge exchange can be done at no cost and can spread among all the users, the cost of exchanging knowledge is negligible. In this case, limits to Internet access (because of legal constraints or because of the digital divide) decrease the opportunities of the FLOSS community development.

6.4.3 Agents' heterogeneity

In our static model, agents are considered symmetric. However, in reality there is a structural idiosyncrasy among agents. If we drop the hypothesis of symmetry, the model becomes much more complicated. However we can use it to give an intuition of what happens if the specificities of the game are not the same for every agent.

Consider the problem in dynamical terms, as done before. The community is initially set up by people with a high interest in the activity that the community is going to undertake and in the "vision" it embodies, an interest high enough to make them bear the costs connected with the small size of the community. In the model, this means selecting equilibrium ***a***. The community, then, could be created and developed, even if linking only a few individuals. But when the community reaches ***a***, it is now able to produce artifacts. In ***a***, moreover, the community starts to develop a structured identity. Following what has been said in section 6.4.1 As the artifacts produced start to fill a new market, and as the identity becomes more precise and well known, other agents could find it desirable to join the communitarian project, even if, at the beginning, it was too costly for them. At this point, if the individuals joining the community number enough to overcome the threshold ***b***, the communitarian terms of the payoff function, namely $e(n)$ and *Y(.)*, will trigger the increasing returns of the function, making the community grow toward ***c***. Notice that in the previous description it is easy to recognize the FLOSS community, even if the process is just sketched out (Bonaccorsi and Rossi, 2003b; Bitzer and Schröder, 2005).

6.5 The parameters inside the model

6.5.1 The terms $e(n)$ and Y(.)

Both $e(n)$ and *Y(.)* express the core of the communitarian dimension. To see how the shape of these functions affects the selection of the equilibria, consider their mechanisms. Cooperation determines endogenously a space in which the individuals interact along the structure of the production, creating social ties. The interaction, in a sense, becomes "thick," and the exchange of meanings, values, and visions are gradually associated with the exchange of artifacts. This acts on the payoff function of the individuals, through a collective term, *Y(.)*, and an individual term, *e(.)*, which produces the effects discussed in the introduction of the model. The shape of the two functions, then, is a "property" of the community, because it is given by the social processes that constitutes its social structure. So, the shapes of the curves, and thus the relative positions of the equilibria, depend on all the features of these processes: the "vision" embodied in the communitarian activity, the characteristics of the produced artifacts, the spaces of the interaction, the kind and level of debate developed inside the social nexus, and so on. If these elements change, the "distance" between ***a***, ***b*** and ***c*** also changes, making more (or less) plausible the birth of a community and/or its growth toward ***c***.

6.5.2 The IPR protection level

The level of protection granted by the IPR is not included in the model. However, it is useful to consider their effect on the payoff functions. IPR limit the scope of the innovative activity (constraining the "field" in which research could be exploited without violating them), in both community and technology cases. Moreover, they constrain the communitarian activity while enhancing the profits of the inventor in the technology regime. Thus, if we "plug" the IPR protection level into P_r, *Y(.),* and R_i^T, it is possible to see how the equilibria change if the IPR protection level changes. We do not need to recalculate the equilibria to understand that both of the payoff functions move in such a way that ***a*** and ***c*** are pushed to the right, while the effect on ***b*** is ambiguous. This is because, on the one hand, the higher level of IPR protection has a negative impact on the community payoff, constraining its activity and moving downward the payoff function. On the other hand, researchers in technology also face a stronger limitation of their research, but they can rely on higher revenues,[23] so that their payoff function moves upward. Considered all together, this reasoning results in a movement of the equilibria, such that if a community grows much less than before, if it is created t all.

7. Conclusion

7.1 Policy implications

The policy question I want to assess is: *what are the instruments that policy makers can use to enhance knowledge-intensive communities?* The previous discussion about $e(n)$ and *Y(.)* terms clearly shows that the main problem connected with the policy design for a knowledge-intensive community is the impossibility to act *directly* on the community engine. This is because a community has a lot to do with intrinsic motivation and social communication, which by definition are not affected directly by external incentive structures. If we look back at the analyzed parameters, it is easy to see that the policy maker's role is indirect, based on construction of spaces and opportunities rather than on the promotion of the community mechanism *per se* (Garzarelli, 2003). In a "lighter" IPR structure, the creation of spaces in which individuals can easily communicate, where the diffusion of the skills and of the infrastructures need to access the

---

[23]Notice that if one looks at firms' behaviors in the software market, it is easy to consider this plausible: $\frac{\delta R_i^T}{\delta IPR} > \frac{\delta P_r^T}{\delta IPR}$.

community, the possibility for the communities to spread their ideas and to create an "osmosis" with the society as a whole are the key parameters.

However this means that policy makers "give up" their direct control on the communitarian production of knowledge. Communities are by definition voluntary phenomena, based on the needs and desires of their members. As Garzelli and Galoppini (2003) note "the specific assets in FS/OSS production are mostly knowledge (different capabilities or skill levels) and spare time. But, in the main, these specificities are both self-selective: knowledge 'enters' its most valued use spontaneously when it desires, and it 'exits' in exactly the same fashion, 'marginalizing' lock-in scenarios" (Garzarelli and Galoppini, 2003, pp. 23–24). Moreover, even if the communitarian activity is based on rules, these rules are able to enhance the communitarian activity only when they are internalized by the members, making the common activity easier and less costly. Putting all this in a radical way: there is no direct way to control the community activity, because if the community is not able to create endogenously the interest for the production of a certain artifact or "piece of knowledge," that artifact or knowledge will not be produced. If the community endogenously reaches a breaking point when it becomes unstable or even disappears, there are few things a policy maker can do without distorting or even disrupting the inner functioning of the community. This means that in the long run, communities are, from a policy making point of view, "unreliable." Other more "reliable" institutions, such as science and technology, must be promoted around the community to fill the gaps it cannot fill.

7.2 Science, technology and communities

If this is true, the difference between a community and science and technology mechanisms emerges clearly. The role of the State is fundamental in science, as well as in technology, in making it a reliable mechanism. The extrinsic motivation provided with the IPR structure and the management of research funds assigns to policy makers the power to control the direction and the volume of knowledge production. Of course, the control is based on a dialog between the scientific community and the economic agents, but the point is that policy makers have a fundamental role in *actively* determining the rules of the game. This role is much less effective in the communitarian environment, because the social dimension, which is much less sensitive to policy makers' usual tools, is the fundamental engine of the institution.

From a policy making point of view, then, the coexistence of the three kinds of institutions is a central issue, because communities can cover only part of the knowledge production space. The FLOSS case shows this clearly: the community develops autonomously, but economic agents act in all the profitable regions in which community cannot and/or does not want to be present[24]. A similar argument can be made for universities researching in the areas where a community cannot move because of the huge research costs. The role of these two institutions in filling the gaps created by the community is then crucial.

[24] Firms are participating in different ways in the FLOSS arena, experimenting a wide variety of business models. See among others Fosfuri et al., 2008; Henkel, 2006; and Dahlander and Wallin, 2006.

References

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

Arrow, K., 1962. Economic welfare and the allocation of resource for inventions. In Nelson, R.R. (Ed.), The Rate and the Direction of Innovation Activity. Economic and social factors. Princeton University Press, Princeton, NJ, pp. 609–626.

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

Bessen, J., 2006. Open source software: Free provision of complex public goods. In Bitzer J., Schröder, P.J.H., (Eds.), The Economics of Open Source Software Development, Elsevier, Amsterdam, The Netherlands, pp. 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. Schröder, J.H.P., 2005. Bug-fixing and code-writing: The private provision of open source software. Information Economics and Policy 17, 389–406.

Bonaccorsi, A., Rossi, C., 2003a. Comparing motivations of individual programmers and firms to take part in the open source movement: From community to business. Social Science Research Network, at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=460861#PaperDownload.

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

Cammozzo, A., 2007. Enabling environments: Internet, Web, GNU/Linux, Debian. Presented at the Workshop OSSEMP2007, Limerick, June.

Carraro, C., Pomè, A., Siniscalco, D., 2001. Science versus profit in research: Lessons from the human genome project. Working Paper #60.2001, FEEM and Working Paper #2890 CEPR.

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

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

Cohendet, P., Creplet, F., Dupouët, O., 2001. Communities of practice and epistemic communities: a renewed approach of organizational learning within the firm. Paper presented at the Workshop on Economics and Heterogeneous Interacting Agents, June 2001 Marseille, France.

Coriat, B., Dosi, G., 1998. The institutional embeddedness of economic change: An appraisal of the 'revolutionary' and 'regulationist' research programmes. In Nielsen, K., Johnson, B. (Eds.), Institutions and Economic Change: New Perspectives on Markets, Firms and Technology. Edward Elgar: Cheltenham, UK, pp. 3–32.

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

Dalle, J.-M., David, P.A., 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., David, P.A., 2007. Simulating code growth in Libre (open source) mode. In Brousseau, E., Curien, N. (Eds.), Internet and Digital Economics. Cambridge University Press, New York, pp. 391–422.

Dasgupta, P., David, P.A., 1987. Information disclosure and the economics of science and technology. In Feiwel, G. (Ed.), Arrow and the Ascendant of Economic Theory. New York

University Press, New York, pp. 519–542.

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

David, P.A., 1991. Reputation and agency in the historical emergence of institutions of 'open science.' Center for Economic Policy Research Publication # 261, Stanford University.

David, P.A., Arora, S., Steinmueller, W.E., 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, Foray, D., 2003. Economic fundaments of the knowledge society. Policy Futures in Education 1 (1), 20–49.

David, P.A., Shapiro, J.S., 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.

Dosi, G., 1988. Source, procedures, and microeconomic effects of innovation. Journal of Economic Literature XXVI (September), 1120–1171.

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

Edwards, K. (2001), Epistemic communities, situated learning and open source software development. Working paper, Department of Manufacturing Engineering and Management, Technical University of Denmark, at http://opensource.mit.edu/papers/kasperedwards-ec.pdf

Elliott, M.S., Scacchi, W., 2003. Free software: A case study of software development in a virtual organizational culture. Working paper, Institute for Software Research, University of California– Irvine, April.

Fosfuri, A., Giarratana, M.S., Luzzi, A., 2008. The Penguin Has Entered the Building: The

Commercialization of Open Source Software Products. Organization Science, 19(2), 292-305

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

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

Garzarelli, G., 2003. Open source software and the economics of organization. Industrial Organization, Economics Working paper #0304003, at http://opensource.mit.edu/papers/garzarelli.pdf

Garzarelli, G., Galoppini, R., 2003. Capability coordination in modular organization: Voluntary FS/OSS production and the case of Debian GNU/Linux. Industrial Organization, Economics Working paper #0312005, at http://opensource.mit.edu/papers/garzarelligaloppini.pdf

Ghosh, R.A., 1998. Cooking pot markets: An economic model for the trade in free goods and services on the Internet. First Monday 3 (3), March, at http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/580/501

Ghosh, R.A., 2003. Understanding free software developers: Findings from the FOSS study. Presented at the Conference "New Models of Software Development," Harvard Business School/MIT Sloan Free/Open Source Software, 19–20 June.

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

Giuri, P., Rocchetti, G., Torrisi, S., 2002. Open source software: From open science to new marketing models an enquiry into the economics and management of open source software. Working paper, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies (LEM), Pisa, Italy, at: http://www.lem.sssup.it/WPLem/files/2002-23.pdf

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

Green, E.L., 1999. Economics of open source software. at http://badtux.org/home/eric/editorial/economics.php

Henkel J., 2006. Selective revealing in open innovation processes: The case of embedded Linux. Research Policy 35, 953–969

Himanen, P., Torvalds, L., Castells, M., 2001. The Hacker Ethic and the Spirit of the Information Age. Secker & Warburg, London, UK.

Jeppesen, L.B., 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., 2006. Collaboration, peer review and open source software. Information Economics and Policy 18, 477–497.

Kelty, C.M., 2001. Free software/free science, First Monday 6 (12) at http://firstmonday.org/issues/issue6_12/kelty/index.html

Laffont, J.J., Tirole, J., 1993. The Theory of Incentives in Procurement and Regulation. MIT Press, Cambridge, MA.

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

Lakhani, K.R., Wolf, R., Bates, J., Di Bona, C., 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, DOI: 10.1016/j.infoecopol.2008.11.004.

Lee, G.K., Cole, R.E., 2003. From a firm-based to a community-based model of knowledge creation: The case of the Linux Kernel development. *Organization Science* 14 (6), 633–649.

Lerner, J., Tirole, J., 2002. Some simple economics of open source. Journal of Industrial Economics

50 (2), 197–234.

Lin, Y., 2003a. Culture, creativity, and innovation: Practices of locally-crafted software. Presented at the joint seminar SATSU and HCI group at computer science and psychology departments on 20 May, University of York, UK. Available at http://www.york.lug.org.uk/talks/yuwei_satsu/satsu_paper.html (accessed May 5th, 2005).

Lin, Y., 2003b. The institutionalisation of hacking practices. Ubiquity 4 (4), 2, 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

Luo, J., 2002. Why people are willing to share knowledge: an open source account. Queen's KBE Centre for Knowledge-Based Enterprises, Doctoral Consortium 2002 Papers, June.

Mateos-Garcia, J., Steinmueller, W.E., 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 & Governance 10, 381–400.

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

Nelson, R., Winter, S., 1982. An Evolutionary Theory of Economic Change. Harvard University Press, Cambridge, MA.

Nelson, R.R., 1990. Capitalism as an engine of progress. Research Policy 19 (3), 193–214.

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.

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

Raymond, E., 1998a. A brief history of hackerdom. At http://earthspace.net/ esr/faqs/hacker-hist.html

Raymond, E., 1998b. The cathedral and the bazaar. First Monday 3 (3), at http://www.firstmonday.dk/issues/issue3_3/raymond/index.html

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

Rosenberg, N., 1982. Inside the Black Box. Cambridge University Press, Cambridge,UK.

Rosenberg, N., 1996. Uncertainty and technological change. In Landau, R., Taylor, T., Wright, G., (Eds.) The Mosaic of Economic Growth. Stanford University Press, Stanford, CA, pp. 334–353.

Rosenberg, N., Birdzell Jr., L., 1986. How the West Grew Rich. Basic Books, New York.

Rullani, F., 2006. Dragging developers towards the core. How the Free/Libre/Open Source Software community enhances developers' contribution. Presented at the CCC Thirteenth Annual Colloquium for Doctoral Student Research, 19–21 May, Lausanne, Switzerland.

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

Saint-Paul, G., 2003. Information sharing and cumulative innovation in business networks. GREMAQ-IDEI, Université des Sciences Sociales deToulouse and CEPR, October.

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

Stallman, R., Garfinkle, S., 1992. Against software patents. Communications of the ACM 35 (1), 17–22, 121.

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

Torvalds, L., Diamond, D., 2001. Just for Fun: The Story of an Accidental Revolutionary. Texere, New York.

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.

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

Williams, S., 2002. Free as in Freedom: Richard Stallman's Crusade for Free Software. O'Reilly, Sebastopol, CA.

Appendix

A.1 Proof of $\Pi_i^C(0) < \Pi_i^S(0)$

Since ${x_i}^C(0) \neq {x_i}^S(0)$, it could be the case that ${x_i}^C(0) > {x_i}^S(0)$ such that $P_r^C(0) > P_r^S(0)$ is enough to say $\Pi_i^C(0) > \Pi_i^S(0)$, falsifying my claim that $\Pi_i^C(0) < \Pi_i^S(0)$. Indeed, it is possible to demonstrate that this is not the case.

Without considering any change in the level of externalities, since P_r is concave and c_i convex, if k_i decreases, the optimal level of x_i decreases and $\hat{x}_i^C < \hat{x}_i^S$. This situation corresponds to less externalities from the community, and this limits the contraction of $\hat{x}_i^C$, but at the end the equilibrium level of the effort will be less than the initial level, because the initial level is too costly for the actual k_i^C. The result is then a decrease of the optimal effort level of community members with respect to scientists, and a contraction of their externalities. Thus $P_r^C(0) < P_r^S(0)$ must be true, leading to $\Pi_i^C(0) < \Pi_i^S(0)$.

To move from here to the overall model, consider that if $n = 0$ all the agents are in science or community and, thus, there is no need to take into account the optimal effort level of the agents in technology. □

Tables and Figures

Table 1. Developers' motivations from FLOSS-EU and Boston Consulting Group surveys.

FLOSS-EU

"For what reason(s) do you go on with developing and/or distributing OS/FS (maximal four answers)?"	***Type***	***%***
to learn and develop new skills	OM/**OU**	70.47
to share my knowledge and skills	OM	67.22
to improve OS/FS products of other developers	**OU**	39.83
because I think that software should not be a proprietary product	OM	37.85
to participate in new forms of cooperation	OM	37.23
to participate in the OS/FS scene	**SR**	35.54
to improve my job opportunities	**SR**	29.82
to solve a problem that could not be done by proprietary software	**OU**	29.61
to limit the power of large software companies	OM	28.87
to get help in realizing a good idea for a software product	**OU**	26.98
to make money	**SR**	12.31
to get a reputation in the OS/FS developers' scene	**SR**	12.03
to distribute not marketable software	OM	10.01

Boston Consulting Group

"Please indicate your top 3 reasons for contributing to this project"	***Type***	***%***
the code for this project is intellectually stimulating to write	OM	44.90
my activity on this project improves my programming skill	OM	41.30
my contribution creates specific functionality in the code needed for my work	**OU**	33.80
I believe source code should be open	OM	33.10
my contribution creates specific functionality in the code needed for my non-work life	**OU**	29.70
I feel a personal obligation to contribute since I use free/open source software	OM	28.50
I like working with the development team on this project	OM	20.30
my contributions will enhance my professional status	**SR**	17.50
other	OM	16.30
I dislike proprietary software or the companies that produce it and want to help the free/open source community defeat them	OM	11.10
my contributions will enhance my reputation in the free/open source software community	**SR**	11.00
the license for this project forces me to contribute my changes	OM	0.20

*Notation: **OS**=Own-use; **SR**=Signaling and Reputation; **OM**=Other motivations. The classification is ours*[25].

[25] The classification given by the FLOSS-EU authors is different from the present one. As Ghosh explains (2003, p. 14) proposing a categorization of the FLOSS-EU data: "There are many ways of forming these categories and how one forms them would change how they correlate to other variables." See Rullani (2006) for an account of this.

Figure 1. The community payoff function is always below and steeper than the science one.

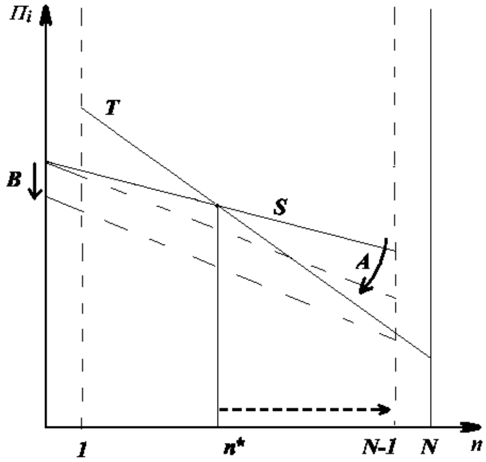

Figure 2. Components of the community payoff function.

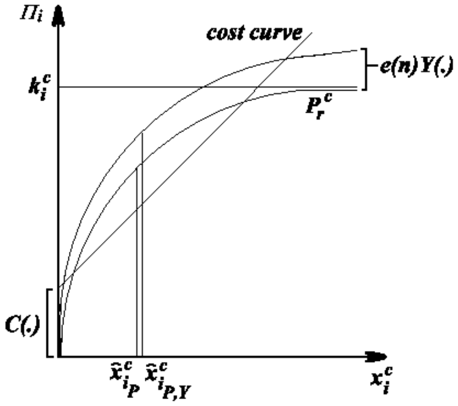

Figure 3. The optimal effort path.

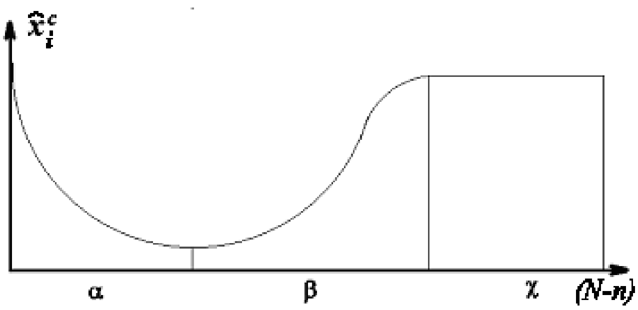

Figure 4. Identity-based community (Idc) and science-like community (Slc) payoff functions.

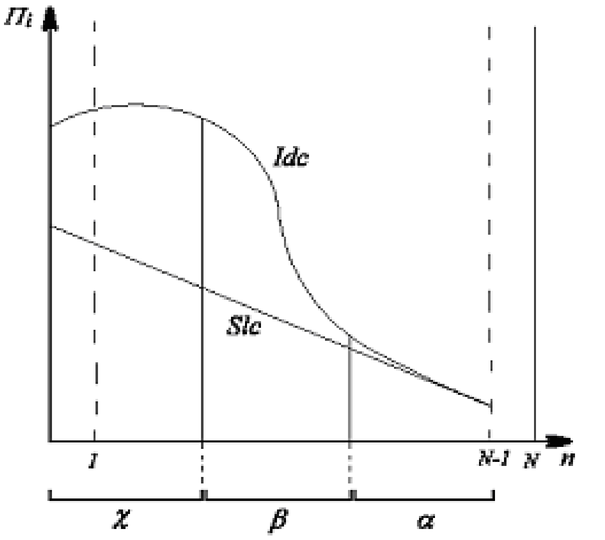

Figure 5. Equilibria of the game with a science-like community (*Slc*) and a identity-based community (*Idc*).