J Evol Econ (2009) 19:277–295
DOI 10.1007/s00191-008-0125-5

REGULAR ARTICLE

Returns from social capital in open source software networks

Rebeca Méndez-Durón · Clara E. García

Published online: 18 September 2008

Abstract Open Source Software projects base their operation on a collaborative structure for knowledge exchange in the form of provision or reception of information, expertise and feedback on the creation of source code. Here, we address the direction of these knowledge flows among projects throughout social networks and their impact on project success. We identify the roles of membership or contribution that individuals play within projects. We found that connections through contributors who bring their knowledge to the project, improve project success, and that connection through members, who transfer their knowledge towards other projects, enhance project success. Finally, we found that ties through shared membership and contributions hamper project success. The analysis of knowledge flows and their impact on project success imply a translation of returns from investment in social capital, where investment takes the shape of knowledge flows and the returns mean the projects' diffusion over the network.

Research funded by the *Spanish Ministry of Science and Education* grant SEJ-2005-06655/ECON on "Challenges and opportunities of new organizational forms: Lessons from Open Source Software projects". We thank Marco Giarratana, Pablo Ruiz-Verdu, two anonymous journal referees, and EMAEE 2007 conference participants for their comments on earlier versions of this paper.

R. Méndez-Durón (✉) · C. E. García
Departamento de Economía de la Empresa, Universidad Carlos III de Madrid, C/Madrid 126, Getafe, Madrid 28903, Spain
e-mail: rebeca.mendez@uc3m.es

C. E. García
e-mail: claragar@emp.uc3m.es

Keywords Knowledge · Networks · Social capital · Software industry

JEL Classification A14 · L14 · L86 · O32 · Z13

1 Introduction

The Open Source Software (OSS) movement has recently received an enormous amount of attention as the OSS development process, product commercialization, and diffusion differ dramatically from those of proprietary software solutions. Individual programmers contribute to the development of core tasks, the debugging, and the improvement of programs for different reasons (Raymond 1999): First, self-interested behavior—programmers need the software and its improvements for their own purposes (Baldwin and Clark 2006; Bergquist and Ljungberg 2001; DiBona et al. 1999; Kollock 1994); Second, reputation, and associated rents gained by those who make high quality contributions (Lerner and Tirole 2001, 2002; Lakhani and Wolf 2005); Third, philanthropic behavior linked to the enjoyment of the contribution itself (Raymond 1999).

Current research on OSS focuses on the mechanisms and foundations underlying the network formation, since they represent a valuable setting for understanding basic aspects of social organization, such as cooperation and the emergence of open networks as enduring forms of governance (Lakhani and von Hippel 2003). In OSS networks, cooperation takes place among a large, heterogeneous and ever changing number of developers working simultaneously on different projects where, typically, participants are strangers and, there are no long-term commitments.[1] However, without active contributions and participation, collaborative knowledge will not succeed. The OSS gains prominence as projects, independently of their domain and scale, are organized according to the principles of a *community of practice* (Brown and Duguid 2001): open and self-organizing networks in which shared values, norms and beliefs rest upon generalized exchange mechanisms and expectations of reciprocity. Then, *pure generalized exchange* might emerge in the presence of altruism, a collective sense of fairness that allows for unilateral giving while preventing free riding behavior (Ekeh 1974). This is especially true when gift givers may choose the recipients (*'fairness-based selective giving'*) (Takahashi 2000).[2] This strategy materializes as a strong reason why

[1] Programmers are geographically distant; they come from different cultures, languages, and traditions, and differ in personal, professional and social features.

[2] The fairness-based selective strategy is characterized by actors endowed with options for leaving the current relation and forming a new relation [...] but it cannot really explain generalized exchange patterns when actors are strangers (Takahashi 2000).

individuals freely reveal innovation-related information within innovation communities (Franke and Shah 2003).

From an empirical perspective, at least three questions stand out: First, why individuals participate in generalized exchange mechanisms in the absence of hierarchical structures and rules:[3] a refined understanding of the mechanisms governing exchange relationships between community members (fairness, trust, generalised and reciprocal exchange); second, the relative importance of skill level versus other relationships in determining an individual position in the network; third, the differences in the emergence of collaborative structure types and their relation to the success of open source projects.

This paper addresses how differences in collaborative structures influence open source projects' success using a panel data of 2,962 software game projects from SourceForge.net. Our data allow us to identify flows of knowledge and information among projects and the collaborative structures that emerge, based on software developers' contributions to those projects.

The architecture of network ties and of interpersonal relations will differ among projects as a result of the projects' heterogeneity in both human and social capital. Programmer skills and knowledge refer to human capital, while the capabilities to access resources embedded in their social and professional networks are termed "social capital." Programmers working across different projects may have a wide network of relationships favoring knowledge and information flows and affecting the projects' differential success rates. Previous research on social capital and networks indicates that *strong ties*, while providing relational benefits, are less likely to provide non-redundant knowledge. Moreover, projects requiring the reception of mostly explicit knowledge, as is the case of OSS, benefit from having a network of *weak ties* (Granovetter 1973; Hansen 1999) while avoiding the costs of developing trustworthy collaborations, or by exploiting access to unique resources (Burt 1992). However, open source projects may also benefit from knowledge sharing outside their boundaries, as previous research has shown in the case of work groups (Ancona and Caldwell 1992). In such a context, work groups' structural diversity contributes to strong effects on performance associated with knowledge sharing. A structurally diverse work group is one in which the members, by virtue of their different organizational affiliations, roles, or positions, can expose the group to unique sources of knowledge (Cummings 2004).

The OSS projects' collaborative structure is based on knowledge sharing, defined here as the provision or reception of information, expertise, and feedback regarding the development and/or modification of a source code. Consequently, open source knowledge sharing represents a generalized

[3] Unilateral resource giving within social and economic exchanges may emerge because of: (1) pure altruistic behavior; (2) collective norms that punish any form of free riding, and (3) rational choice under game theoretical frameworks in which the existence of incentive structures to solve social dilemmas predominates (Olson 1965).

exchange supported by the expectation that if a community member provides assistance today (*provision*), someone else will provide him with assistance when he needs it (*reception*). However, research on generalized exchange is more concerned about free-riding issues and individual behavior/incentives than the structural characteristics of work groups and its effects on success.

Our analysis focuses on the effects of knowledge sharing on projects performance accounting from the effects of both knowledge/information reception and provision. Consequently, the research questions addressed are: (1) Does a focal project that receives knowledge from other actors improves its performance? (2) Are these effects comparable to those obtained from providing knowledge to other actors? In addition (3) what are the effects on performance of a focal project where its members engage in generalized exchange towards other projects? To answer these questions, we base our analysis on social capital theories and on groups' structural diversity.

In this paper, we consider social capital as the ties between programmers that provide access to information and, consequently, between projects on which they work. However, there is empirical evidence to suggest that the nature of ties explains additional variance in outcomes beyond that explained by structure alone (Lin 1999). While research on social capital and social networks has emphasized the *strong/weak* nature of those ties, we account for the directionality of knowledge flows—inwards and outwards—with respect to a focal project as a structural dimension of diversity, since they relate to the role performed by individual programmers in different projects.

We use the project as the unit of analysis; each project represents an event and programmers or actors relate to each other through events, and events relate to other events because of common actors. Though the data analyzed represents a two-mode affiliation networks (Waserman and Faust 1994; Faust 1997), our theoretical analysis and modelling depart from previous works; we project a two-mode network into a one-mode event by event network (Sorenson and Stuart 2008). The modelling of events/projects collaborative structure also introduces some attributes of the actors/programmers on whom the ties form.

Our analysis indicates that knowledge flows among projects are asymmetric as result of both members and contributors' roles. We found that connections through *contributors* who bring their programming expertise to the project, besides being a source of new knowledge, improve project success. Additionally, our findings reveal that connections through members who sign up in other projects hinder the *focal* project success, unless they carry on programming activities transferring their knowledge towards other projects. Finally, we found that ties through shared membership and contribution hamper project success.

This paper presents our analysis as follows. First, we introduce the theoretical framework. In Section 3, we present the data and variables used, and in Section 4, we show the main results. Finally, we conclude by discussing the findings, limitations and further research implications.

2 Theoretical framework and research hypotheses

Previous studies on large OSS projects such as Linux and Apache support the presence of social capital (Bergquist and Ljungberg 2001; McKelvey 2001): as big projects are nurtured from individual contributions, thus the overall effect is straightforward as soon as they capture new contributors. More recently, empirical evidence suggests that social capital, measured by the density of ties, positively influences collective activity and productivity (Long 2006). When referring to OSS networks, Zhang (2007) identifies previous ties among a group of programmers as a powerful predictor of further members' affiliations to specific projects. The existence and density of prior ties between the initiator of the project and developers positively influence the probability of a project to attract more individuals. Consequently, in an analogous idea to Putnam's (1995), in OSS networks social capital seems likely to be a substitute for the absence of organizational structures and hierarchical authority devoted to promoting communication, coordination, resource pooling and knowledge sharing, and contributing to improving efficiency and performance.

External knowledge in the open network and the inflow of ideas comes through programmers involved in two or more projects. While programmers differ in their skills, they also differ in their access to resources embedded in their social and professional networks, and those differences are critical in the success of OSS projects.

Social capital entails the underlying structure for exchange, where relations among individuals create and support a common understanding that promotes the generation of a public good exploitable by all individuals within the structure (Coleman 1988). Social capital is rooted in ties among agents that exchange information and provide access to resources leading to a major emphasis on the density of connections and centrality (Ahuja 2000; Zajac and Westphal 1996).

Research on social networks has an established and long tradition in social disciplines because of both interest in patterns of exchange and the important implications of networks in the spread of information and knowledge (Burt 1992; Granovetter 1973; Larson 1992; Powell 1990; Zaheer and Bell 2005; Zajac and Westphal 1996).

Scholars interpret social capital as a metaphor about the gains individuals or groups may obtain by belonging and interacting in social network structures. The result is an increasing and dense body of literature that uses social capital as an independent variable to explain a wide range of social phenomena, concluding that individuals or firms interacting within a network perform better than when they are alone (Burt 2000). The structural dimension of social capital emphasizes some network properties, including the strength of ties (Granovetter 1973), structural holes (Burt 1992) and embeddedness (Gulati and Gargiulo 1999). Direct ties define strong interactions among projects representing channels for the transfer of useful knowledge (Uzzi 1996, 1997; Levin and Cross 2004). The definition of indirect or weak ties establishes that distant and infrequent interactions play an instrumental role in the diffusion of ideas and information (Granovetter 1985; Uzzi and Lancaster 2003). Hence, prior studies found that knowledge-sharing positively relates to factors such as strong [intra-group] ties (Wellman and Wortley 1990), co-location (Allen 1977), demographic similarity (Pelled 1996), status similarity (Cohen and Zhou 1991), and a history of prior relationships (Krackhardt 1992).

These ties differ in their effects on knowledge-sharing; direct/strong ties based on trust enhance the transfer of both tacit and explicit knowledge, while indirect/weak ties allow the transfer of codified knowledge more efficiently (Hansen 1999; Ahuja 2000). Previous research indicates that *strong ties*, while providing relational benefits, are less likely to provide non-redundant knowledge. Moreover, projects requiring the receipt of mostly explicit knowledge, as is the case of OSS, benefit from having a network of *weak ties* (Granovetter 1973; Hansen 1999), while avoiding the costs of developing trustworthy collaborations, or by exploiting access to unique resources (Burt 1992).

Social capital benefits are the return to investment strategies oriented to institutionalizing group relations into a social network. Portes (1998) considers a *socio-relational* dimension of social capital, where agents gain access to resources, and a *stock* dimension comprising the quantity and quality of those resources. OSS projects reflect a common stock of knowledge. Software projects in general, and OSS projects in particular, exhibit at least two properties: (1) they have a modular architecture and (2) outcomes may vary along the design process allowing for the introduction of new modules and creating an option value for development. These two properties suggest that projects themselves become a stock of social capital supplied and nurtured by members and contributors.

This paper deals with the value of [external] *knowledge sharing* and its effects on OSS projects' success. Knowledge sharing is defined here as the provision or reception of information, expertise, and feedback regarding the development and/or modification of a source code. Thus, we expect that flows of knowledge incoming to a focal project will contribute to its stock of social capital and to its success.

Hypothesis 1 *Knowledge-sharing will positively influence a focal project's success, while the effects of reception will be stronger than the effects of provision.*

However, instead of focusing on the strength of ties—the *closeness* and *frequency* of interactions in a dyadic relationship—and structural embeddedness (Grewal et al. 2006) we focus on the groups' structural diversity. Cummings (2004) has hypothesized that external knowledge-sharing will be more strongly associated with performance when work groups are more structurally diverse. A structurally diverse work group is one in which the members, depending on their different organizational affiliations, roles, or positions, can expose the group to unique sources of knowledge. Structural diversity within OSS projects refers to differential access to external sources of knowledge by virtue of programmers' affiliations and roles. Although structural diversity may also relate to previous collaborative experiences and expertise, we do not deal with this issue here.

Thus, we expect OSS projects' members to have a differential impact on project success accordingly to their role, their function and their social network. For the purpose of this paper, we label the team workforce appointed by the project leader as *members* of the project; these members of the project may perform different activities, from project administrator to language translator. We denote all persons who contribute to software development as *contributors* (though not all *members* of the project are *contributors* and not all *contributors* are *members* of the project). However, some are *members* and *contributors* at the same time (*duality*).

Since knowledge sharing includes the provision and the reception of information, the effects will be stronger if programmers gain unique insight on both projects (i.e. contributor–contributor). Then,

Hypothesis 2 *External knowledge sharing among projects through individuals contributing to software development in more than one project, will positively influence [a focal] project success.*

Projects with individuals representing different functions (i.e. contributors or members) can access diverse social networks they have established in their respective domains and have an effect on a focal project's success from the perspective of social capital. Therefore, we do not expect those gains to differ when individuals in a focal project represent a dual role (contributor/member).

Hypothesis 3 *External knowledge sharing among projects through members contributing to software development in other projects will positively influence the focal project success, independently of the type of role they play in the focal project.*

Woolcock (1998) classified ties as bridging, bonding, and linking social capital. Bridging refers to relations between more distant and heterogeneous members; bonding indicates higher levels of homogeneity, while linking relates to the ability to leverage resources and information beyond the community. We identify 'bonding' ties among OSS projects when members contribute to more than one project, since they imply higher levels of knowledge homogeneity. On the other hand, we consider that the relationships between a focal project member/contributor who is solely a member in other projects represent 'bridging' ties. Knowledge homogeneity will result in lower effects on project performance because of a lack of access to new knowledge resources, and

Hypothesis 4 *External knowledge-sharing among projects through members contributing to software development in the focal project and other projects will have a lower effect on the focal project's success than when members contributing to the focal project are sole members of other projects.*

The analysis of interaction among projects becomes relevant as the emergence of the OSS movement leads to the creation of repositories for software development. Repositories usually are internet-based sites that provide an organizational infrastructure to allow asynchronous exchanges among programmers, mobilize contribution of people otherwise hard to reach, and screen the contributions from redundant knowledge. A repository integrates a community that benefits from the exchange of knowledge among its members.

If repositories track all artifacts and records, we can identify the contributors and the mere members of the project. Moreover, we can identify whether an individual who is a member of a particular project is simultaneously a member of another project. Thus, depending on the affiliation and the role played by individuals somehow involved in the development in each project, the social network for knowledge exchange within the repository comprises nine forms of ties among each pair of projects: 'member–member', 'contributor–member', 'member–contributor', 'contributor–contributor', 'duality–member', 'duality–contributor', 'member–duality', 'contributor–duality' and 'duality–duality'.

The suitability of OSS repositories for identifying the structural dimension of interactions comes from the structure that a repository provides through knowledge exchange. OSS programmers interact by contributing their knowledge to their own and others' projects, but they do not necessarily receive the same amount or quality of contributions from other network actors. Still, they generate expectations about others' behavior for providing their knowledge. Therefore, it is interesting to analyze the knowledge flows throughout the community and their communication paths that sustain and reinforce social capital.

The present study contributes to social capital theory and its empirical analysis in several ways. First, we claim there are different individual strategies for investing in social capital, and these strategies have a relation to structural diversity (role and affiliation). Second, we support the fact that social capital influences project's success only if *contributors* provide new valuable knowledge whether inside or outside the *focal* project.

3 Data and analysis

The analysis of social capital in OSS repositories requires a thorough database that, on the one hand, provides intricate details of project formation, organization and operation, and, on the other hand, provides details about the implication of individuals in the software development. Furthermore, the analysis itself requires disentangling the relationships occurring through the social network.

3.1 Data setting

The data we use in this analysis come from the SourceForge.net Research Data (Department of Computer Science and Engineering, University of Notre

Dame). SourceForge.net is the largest repository of open source software; it hosts over 140,000 projects and gathers over 1.5 million registered users. SourceForge.net belongs to OSTG, Inc, which has shared activity data with the University of Notre Dame for the purpose of academic research on OSS, under the condition of sharing the data with other researchers interested in open source software phenomenon (Christley and Madey 2005).

Sourceforge.net as a repository has several characteristics that promote network exchanges. The purpose of the repository is to provide a platform for software development over a worldwide web infrastructure for knowledge exchange. They host projects and provide tools that allow asynchronous communication, stock people's contributions, and screen software for redundant knowledge. To host a project, an initiator should register in the network, and post a message to the platform indicating the type of software, its purpose and target public, the intellectual property regime, the programming language, the user interface, the phase of development, and the team workforce that will be responsible for developing and controlling the software. Then Sourceforge.net administrators authorize the creation of a space for hosting the project if it satisfies the main premises for the repository. Once the repository hosts the project, then it registers and controls all its activity. Every time there is a movement in the project, the repository electronically archives the information; these movements may include communications among members, forum posts, and, most importantly, the *artifacts* (modules of software code) produced by the people engaged in programming activities. An important feature of a repository is open access to projects to everybody on the worldwide web—every single person, registered or not on Sourceforge.net, may see the project, look at the code, and contribute. The repository records changes in the team workforce, so it is possible to know if there are new members or if some of them have already quit the project. As repository tracks all artifacts records, we can identify who are *contributors*, who are *members*, and who constitute the *duality* at each project. Hence, we can observe how projects connect within the repository through individuals. These individuals establish a tie among the available pool of knowledge for each project.

Particularly, our dataset is a sub-sample of 2,962 valid observations over 12 months of projects aimed at developing game software. To get this sample, we look at the monthly dumps of data and select all projects that belong to the category of GNU Public License (GPL).[4] We restrict the sample to such a license to guarantee that projects were not subject to any restriction for copying, adoption and distribution, so all software is equally prone to reach the same audience. Then, we limit the sample to games in a general category. This selection serves the purpose of tapping an appropriate measure of project success. Although we will further discuss performance and success measures, we should note that we are interested in measuring performance as *market*

[4] GPL grants the programmers and users of software the privileges or freedom to distribute and to modify copies of the software, and transfers those privileges to further developments.

penetration of software; thus, we should pick a category of projects the target markets of which were not constrained by the programming skills and abilities of end users. Furthermore, we restrict the sample to the projects that were alive during the whole sample; we take this decision because dataset classifies a project as ‘alive’ when the repository still hosts the project and ‘dead’ when it quit the repository. Unfortunately, it is impossible to distinguish whether the project has quit the repository independently, or if it moved to another repository, or just ended; thus, their disappearance does not necessarily relate to market demands, or product lifecycles, but to managerial premises. Finally, we look at those projects the information on downloading activity of which is available. As the source of information comes from relational database, some observations could be missing in the joining process. We follow this procedure over twelve months running from February 2005 to January 2006, to build an unbalanced panel with 25,722 total observations.

3.2 Variables and analysis

3.2.1 Dependent variable. Project success

We present *project success* as a dependent variable; for that we use number of downloads per month as a measure of project success. A download means one user retrieves the executable files of the software for private use; thus downloads are an output measure of success or popularity among a mass of users. Number of downloads is consistent with the nature of a software game as a digital good through the Internet, because downloading is the sole way of access; furthermore, it is a common measure of performance in OSS (Crowston et al. 2003; Grewal et al. 2006).

3.2.2 Independent variable. Social network as a source of social capital

Social capital measures are largely heterogeneous within the academic literature, thus they vary as to references to different social capital perspectives.[5] In this paper, our measure of social capital lies in the structure of knowledge flows and refers to the sum of [complementary] resources attainable through a social network of projects which include both unidirectional and bidirectional relations. In our analysis, we test whether direction in knowledge flows in the social network of projects generate different outcomes in terms of the projects’ success.

Here we measure directionality of ties among projects through identifying roles and affiliations, that is, contributors, members and duality roles. The

[5]For example, in Putnam’s tradition, social capital measures membership in voluntary organizations as an indicator of networks together with norms and social trust (Putnam 1995; Tsai and Ghoshal 1998). In Granovetter’s and Burt’s tradition, social capital measures, to some extent, the information flows through exchanges among partners (Glaeser et al. 2000; Koka and Prescott 2002; Oh et al. 2004).

presence of such ties indicates knowledge flows that enhance performance of the focal project.

We focus on dyadic relationships since they represent ties between pairs of projects linked through at least one of their members, or contributors. Nonetheless, if projects share more than one individual, we count just one tie if they are the same typology of relationships. In doing so, we are measuring the relevance of structural diversity instead of the density of relationships. This practice also helps to diminish a plausible effect of excess downloading just for having more individuals involved in both projects simultaneously. For example, if the number of individual programmers from project A contributing to project B is three, it counts as one tie between projects A and B. Thus, our measure for ties attempts to equate a linkage between the knowledge of the focal project and the knowledge of any other project within SourceForge.net.

The important feature we want to test is the effect of knowledge flow direction on project success. We distinguish between those ties that receive *inbound* knowledge—ties through individuals who play the 'contributor' or a 'duality' role in the focal project—from those ties that provide *outbound* knowledge to other projects. All categories of ties measure knowledge flows from the focal project's point of view, therefore, although above we define nine forms of ties, we can group those that represent the same directions of knowledge-sharing for the focal project. In addition, we observe that a tie through individuals who simultaneously play as members and contributors in both projects does not actually represent a share of knowledge, but the common stock of knowledge for both projects, and thus we disregard this category. The independent variables are:

1. *Active-inbound ties*: We consider the reception of knowledge from other projects toward the focal project; nonetheless, other projects receive some share of knowledge. That means ties through non-member individual that contribute to a focal project, and simultaneously contribute to other projects. This measures focal project ties through the forms 'contributor–contributor' and 'contributor–duality'.
2. *Inactive-inbound ties*: We consider the reception of knowledge from other projects toward the focal project, but other projects do not receive any share of knowledge. This means ties through contributors to focal projects that are sole members in other projects. They measure focal project ties through the form 'contributor–member'.
3. *Active-outbound ties:* We consider the provision of knowledge from the focal project toward other projects. This means ties through individuals members of the focal project that actively contribute to other projects. They measure focal project ties through the forms 'member–contributor' and 'member–dual' roles.
4. *Inactive-outbound ties*: We consider potential knowledge flow for both focal and other projects, but none of them actually receives a share of knowledge. This means ties from individuals that are sole members in both

focal and other projects. They measure focal project ties through the form 'member–member'.

5. *Duality-inbound ties*: We consider the reception of knowledge from other projects toward the focal project, where individuals play a duality role and other projects do not receive any share of knowledge. This means ties through individuals that play a dual role as members and contributors in a focal project and are sole members in another. They indicate focal project ties through the form of 'duality–member' ties.
6. *Duality-outbound ties:* We consider the provision of knowledge flow from the focal project, where individuals play a duality role, while other projects receive some share of knowledge. This means ties through individuals that play a duality role as members and contributors in a focal project and contributors in another. They indicate focal project ties through the form of 'duality–contributor' ties.

3.2.3 Control variables

In the present study, we have included the project's characteristics, such as control variables. On the one hand, these control variables reflect characteristics that may induce differences in the dependent variable because of demographic issues, such as size and age of a project, and also those referring to their technical peculiarities, which could demand an increase in knowledge transfer. Our control variables include:

1. *Project SIZE*: Number of members of focal project.
2. *Project AGE*: Age of the project in days.
3. *STAGE*: OSS projects included in our database are at different stages of development, going from planning, pre-alpha, alpha, beta, production, mature and inactive. The status of the project and its ability to attract programmers, and therefore the number of ties and contributions from outside, are strongly associated. At its first stage, the core process is the creation of an initial system that will evolve over time. At the last stages, the core process is the diffusion of the product. Thus, we control for the evolution as the project may require different knowledge along its stages.
4. *Characteristics*: We control for characteristics such as programming language, operating system, user interface, intended audience, and speaking language of the game.
5. *Previous project success*: We control for the number of downloads at the previous period to take into account the possibility that successful projects attract the attention of more programmers. This is lagged dependent variable.

In order to test our hypotheses in a dynamic framework, we use a differences-in-differences fixed effects estimator, taking differences in both dependent and independent variables on a monthly basis. Therefore, we condition a variation in success at month $t + 1$ with respect to month t, to

Table 1 Equations and expected results

	The effect on project success of:	is	than the effect of
H1.	Active-Inbound + Inactive-Inbound	>	Active-Outbound + Inactive-Outbound
H2.	Active-Inbound	>	Inactive-Inbound
H3.	Duality-Outbound	=	Active-Outbound
H4.	Duality-Inbound	>	Duality-Outbound

a variation on ties at month $t + 1$ with respect to month t, controlling for the project's fixed effects over the period of study. We want to measure how a variation on a pattern of knowledge flows influences the monthly rate of downloading. The use of this type of regression method obeys our intention of testing the way in which differential rates of knowledge exchange strategies condition project success. A variation in a particular type of tie means there is knowledge flow from or toward a different project, such as discovering a new source of knowledge, or a new channel for diffusion. The fixed effect estimator allows us to isolate the effect of a particular project, the characteristics of which make it more attractive for programmers. The differences-in-differences fixed effects estimator allows us to isolate variations in dependent variable due to changes in independent variables conditional to the particular project's characteristics, which make them more attractive for programmers.

Restating our hypotheses in terms of our independent variables, we obtain a set of equations to test, as expressed in Table 1.

4 Results

We have presented our data in a descriptive way, and furthermore we performed differences-in-differences fixed effects regressions to test whether our theoretical arguments hold. Table 2 reports basic descriptive statistics for the dependent and independent variables. The initial descriptive analysis indicates that projects, on average, report an increase of 20 downloads every month. Nonetheless, we observe that there is huge dispersion on this variable, as the minimum and maximum increase in downloads is approximately 400,000.

Table 2 Basic descriptive statistics

Variation in:	Mean	Std. dev.	Min.	Max.
Project success:				
Downloads	20.819	5,004.585	−468, 651.500	463,422.500
Active-inbound	0.005	0.095	−3.995	7.277
Inactive-inbound	0.010	0.209	−8.444	14.738
Active-outbound	0.003	0.063	−3.270	3.003
Inactive-outbound	0.000	0.000	0.000	0.000
Duality-inbound	0.002	0.159	−3.089	16.547
Duality-outbound	0.004	0.073	−1.178	5.458
Size	0.007	0.175	−8.618	10.279

Concerning the different ties in our study, we also observe an extremely low mean, but an important dispersion, except for the *Inactive-outbound* ties that do not present any variation at all, and these are not included in the regression analysis.

Table 3 reports the results from differences-in-differences fixed effects on downloads. We addressed our empirical analysis of identifying knowledge flows among projects by studying diverse typologies of ties relating the possible 'member' and 'contribution' combinations. Here we show two basic regressions. Model 1 presents the basic model with controls. The control variables do not significantly explain variation in rates of downloads, except for the previous period difference in downloads, which has a negative effect on current period difference in downloads. This result is an indicator that projects present a decreasing trend in their success. Model 2 contains all our measures for ties as explanatory variables. We note that *active-inbound* ties have a positive and significant effect on the rate of downloads. This means there is a benefit from receiving assistance from individuals who play the 'contributor' role in a focal project, while they also contribute to others. We also find that *inactive-inbound* ties have a negative although non-significant effect. This implies there is no gain in just receiving knowledge from outside. The *active-outbound* ties show a surprising positive and significant effect, which means the focal project benefits from individuals who play the 'member' role while they are 'contributors' outside. *Duality-inbound* ties have a negative and significant

Table 3 Results from differences-in-differences fixed effects regressions

Dependent variable: project success	(1)	(2)
Active-inbound		1517.380**
		0.000
Inactive-inbound		−97.452
		0.455
Active-outbound		1217.611**
		0.001
Duality-inbound		−337.702*
		0.032
Duality-outbound		−765.607*
		0.031
Project success$_{t-1}$	−0.490**	−0.489**
	0.000	0.000
Size	−108.597	−225.611
	0.502	0.167
AGE	−36.543	−35.458
	0.440	0.453
STAGE	−159.904	247.215
	0.935	0.899
Project characteristics controls included		
F	0.831	629.98
$P > F$	0.000	0.000

*Significant at 5%
**significant at 1%

effect on project success; meanwhile, *duality-outbound* ties have a significant and strong negative effect on rate of downloads.

Viewed through our hypotheses, we observe that this study confirms hypothesis 1, because the aggregate effects of *active* and *inactive inbound* ties (1,419.48) are greater than the aggregate effects of *active* and *inactive outbound* ties (1,217.611). Therefore, reception of knowledge is more favorable than provision for project success.

In hypothesis 2, we expect to have a greater effect from *active-inbound* ties than from *inactive-inbound* ties. Our result is consistent with hypothesis 2, and confirms that it is better for project success if individuals have unique insight during all the projects.

Hypothesis 3 establishes no [significant] difference among *duality-outbound* and *active-outbound* ties' effects on project success. Nonetheless, we find a stronger, positive effect of *active-outbound* ties compared to *duality-outbound* ties. Our results do not support hypothesis 3. Despite the fact that this result contradicts the theory and our expectations, it may imply that those individuals serving as *active-outbound* ties attract the attention of other individuals because they open new channels for project diffusion or establish a good reputation on other projects.

We found that *duality-inbound* ties have a greater effect than *duality-outbound* ties, although both are negative; therefore, our results confirm hypothesis 4. This may mean that individuals that play the 'duality' role on the focal project, while they contribute to other projects, distract resources and attention from the focal project.

Overall results demonstrate that knowledge flows are significant for project success, and they build channels for flow on the extent that there is structural diversity.

5 Discussion and implications for future research

Social capital, generally defined as the actual and potential resources embedded in relationships among actors, is an important predictor of group and organizational performance (Adler and Kwon 2002; Bourdieu 1983; Leana and Van Buren 1999; Nahapiet and Ghoshal 1998). The structural dimension of social capital focuses on the nature and strength of relationships, and the communication flows embedded in networks of individuals and organizations. The advantages ascribed to social capital include better group communication, efficient collective action, enhanced stocks and use of intellectual capital, and better access to resources. Social capital also applies when answering questions about individual motivations; programmers consider it useful to maintain their social network and nurture the relation by providing their knowledge. This assertion is consistent with Nahapiet and Ghoshal (1998), who postulate that the relational dimension of social capital induces actors to formulate an expectation on the value of the resources they supply to the structure.[6]

A network of OSS projects is fertile ground for testing and supporting the sources of social capital. Contributions to projects are the sole means for development and success, but their patterns differ widely across OSS projects; some projects attract a large number of contributors while others do not. There are projects in which most of the advances come from the voluntary contributions of their own members, while others rely on contributions from actors initially assigned to other projects. Ties between OSS projects represent the network structure, and therefore they are able to indicate differential levels of social capital based on individuals' affiliations and roles. We identified social capital in terms of the social network of the OSS as the number of ties generated through members and contributors and the role they play by solving programming gaps in both focal and other projects.

In this research, the straightforward outcome concerns differences in social capital depending on the structural diversity of project members by identifying paths of communication across different projects. Moreover, empirical results also indicate that some social capital arguments do not hold, as the effects of the role played by individuals is a significant source of variance in the access to network resources and project success. Explicitly, we find that project members playing a contributor role in the focal project and in other projects do not contribute to success, but instead hamper it. We also find that knowledge and skills received through inbound ties lead to superior success, whereas providing knowledge and skills to other projects is beneficial to one's own success only if individuals play an active role as contributors in other projects beyond pure membership.

The core of our contribution relies on the measurement and identification of the collaborative structure—by identifying programmers' roles and affiliations—for knowledge sharing and the differential effects on project success related to knowledge reception and provision. As our own results show, this measurement supports the generalized reciprocity exchange theory. Therefore, with the aim of finding significant sources of social capital, the necessity emerges for accounting differences in investment strategies or knowledge exchanges that complement the stock of capital.

However, one of the main limitations of the present study is that we do not measure the quality of the exchange, but only its direction. While we know inbound ties provide new knowledge and improved performance, we do not know anything about the quality of that knowledge. Thus, a project may need

[6]Nahapiet and Ghoshal (1998) identify three dimensions of social capital that inhere in knowledge exchange and recombination: (1) structural, which reflects the impersonal pattern of ties; (2) relational, as a sense of trust, norms, obligations and expectations that actors develop along connections; and (3) cognitive, as the bonding force, such as shared understanding and identification that hold the group together. These three aspects of social capital combine to improve information transmission and absorption among organizational members, thus enhancing overall organizational performance.

to tie to a broad collection of projects because of the poor quality of inbound contributions. Equally, in the case of outbound knowledge, we could assume the contributions of project members make them visible to the network, but we do not know how valuable their contribution is for others. Neither do we account for modularity of the project; a project with plenty of modules is more attractive to programmers because they clearly identify tasks and goals (Baldwin and Clark 2006).

Another drawback involves our inability to link the different functions played by the contributors. We affirmed the presence of skill complementarities [that should be positively associated to project performance] in networks of cooperation across OSS projects; nevertheless, we do not measure complementarities but assume they occur through ties. Different participants perform different functions that facilitate the rapid change and creation of stable releases, including testing, contributing new changes, coordinating releases, and maintaining documentation. An important aspect of the collaborative approach is to help individuals find tasks in which they can better apply and exploit their talents. Moreover, self-selection and voluntary participation shape projects to reach the required level of complementary skills, which we were unable to include in our analysis.

We foresee a promising vein of research about social capital and OSS networks, as they present a symbiotic environment. Research on social capital will nurture findings derived from the study of open source software networks. Meanwhile, our understanding of the open source software phenomenon enriches us with a deep consciousness regarding the exchange processes for the creation of a valuable stock of social capital.

References

Adler PS, Kwon S (2002) Social capital: prospects for a new concept. Acad Manage Rev 21(1): 17–40

Ahuja G (2000) Collaboration networks, structural holes, and innovation: a longitudinal study. Adm Sci Q 45(3):425–455

Allen T (1977) Managing the flow of technology. Technology transfer and the dissemination of technological information within the R&D organization. MIT Press, Cambridge

Ancona DG, Caldwell DF (1992) Bridging the boundary: external activity and performance in organizational teams. Adm Sci Q 37:634–665

Baldwin CY, Clark KB (2006) The architecture of participation: does code architecture mitigate free riding in the open source development model? Manage Sci 52(7):1116–1127

Bergquist M, Ljungberg J (2001) The power of gifts: organizing social relationships in open source communities. Inf Syst J 11(4):305–320

Bourdieu P (1983) Forms of capital. In: Richardson JG (ed) Handbook for theory and research for the sociology of education. New York, Greenwood, pp 241–258

Brown JS, Duguid P (2001) Organizational learning and communities-of-practice: toward a unified view of working, learning and innovation. Organ Sci 21:40–57

Burt RS (1992) Structural holes: the social structure of competition. Harvard University Press, Cambridge

Burt RS (2000) The network structure of social capital. In: Staw BM, Sutton RI (eds) Research on organizational behavior. Elsevier, Amsterdam, pp 345–423

Cohen BP, Zhou X (1991) Status processes in enduring work groups. Am Sociol Rev 56(2): 179–188
Coleman JS (1988) Social capital in the creation of human capital. Am J Sociol 94:S95–S120
Christley S, Madey G (2005) Collection of activity data for sourceforge projects. Technical report: TR-2005-15, Dept. of Computer Science and Engineering, University of Notre Dame, Notre Dame, IN
Crowston K, Annabi H, Howison J (2003) "Defining open source software project success" 24th international conference on information systems. http://opensource.mit.edu/papers/crowstonannabihowison.pdf
Cummings JS (2004) Work groups, structural diversity, and knowledge sharing in a global organization. Manage Sci 50(3):352–364
DiBona C, Ockman S, Stone M (eds) (1999) Open sources: voices from the open source revolution. O'Reilly & Associates, Sebastopol, CA
Ekeh PP (1974) Social exchange theory: the two traditions. Harvard University Press, Cambridge
Faust D (1997) Centrality in affiliation networks. Soc Netw 19(2):157–191
Franke N, Shah S (2003) How communities support innovative activities: an exploration of assistance and sharing among end users. Res Policy 32(1):157–178
Glaeser EL, Laibson D, Sacerdote B (2000) The economic approach to social capital. Working Paper 7728, National Bureau of Economic Research
Granovetter M (1973) The strength of weak ties. Am J Sociol 78(6):1360–1380
Granovetter M (1985) Economic action and social structure: the problem of embeddedness. Am J Sociol 91(3):481–510
Grewal R, Lilien GL, Mallapragada G (2006) Location, location, location: how network embeddedness affects project success in open source systems. Manage Sci 52(7):1043–1056
Gulati R, Gargiulo M (1999) Where do interorganizational networks come from? Am J Sociol 104(5):1439–1455
Hansen M (1999) The search-transfer problem: ther role of weak ties in sharing knowledge across organization subunits. Adm Sci Q 44(1):82–119
Koka B, Prescott J (2002) Strategic alliances as social capital: a multidimensional view. Strateg Manage J 23(9):795–816
Kollock P (1994) The emergence of exchange structures: an experimental study of uncertainty, comitment, and trust. Am J Sociol 100(2):313–346
Krackhardt D (1992) The strength of strong ties: the importance of philos in organizations. In: Nohria N, Eccles R (eds) Networks and organizations: structure, form and action. Harvard Business School Press, Cambridge, MA
Lakhani KR, von Hippel E (2003) How open source software works: "free" user-to-user assistance. Res Policy 32(6):923–943
Lakhani KR, Wolf B (2005) Why hackers do what they do: understanding motivation and effort in free/ open source software projects. In: Feller J, Fitzgerald B, Hissam S, Lakhani KR (eds) Perspectives on free and open source software. MIT Press, Cambridge, pp 3–21
Larson A (1992) Network dyads in entrepreneurial settings: a study of the governance of exchange relationships. Adm Sci Q 37(1):76–104
Leana C, Van Buren HJ III (1999) Organizational social capital and employment practices. Acad Manage Rev 24(3):538–555
Lerner J, Tirole J (2001) The open source movement: key research questions. Eur Econ Rev 45(4):819–826
Lerner J, Tirole J (2002) Some simple economics of open source. J Ind Econ 50(2):197–234
Levin DZ, Cross R (2004) The strength of weak ties you can trust: the mediating role of trust in effective knowledge transfer. Manage Sci 50(11):1477–1490
Lin N (1999) Building a network theory of social capital. Connections 22(1):28–51
Long Y (2006) Social structure, knowledge sharing, and project performance in open source software development. Dissertation project, The University of Nebraska-Lincoln
McKelvey M (2001) Internet entrepreneurship: linux and the dynamics of open source software, centre for research on innovation and competition, Discussion paper no. 44, Manchester
Nahapiet J, Ghoshal S (1998) Social capital, intellectual capital, and the organizational advantage. Acad Manage Rev 23(2):242–266
Oh H, Chung M, Labianca G (2004) Group social capital and group effectiveness: the role of informal socializing ties. Acad Manage J 47(6):860–875

Olson M (1965) The logic of collective action: public goods and the theory of groups. Harvard University Press, Cambridge, MA
Pelled LH (1996) Demographic diversity, conflict and work group outcomes: an intervening process theory. Organ Sci 7(6):615–632
Portes A (1998) Social capital: its origins and applications in modern sociology. Annu Rev Sociology 24:1–24
Powell WW (1990) Neither market nor hierarchy: network forms of organization. Res Organ Behav 12:295–336
Putnam RD (1995) Bowling alone: America's declining social capital. J Democr 6(1):65–78
Raymond ES (1999) The cathedral & the bazaar: musings on Linux and open source by an accidental revolutionary. O'Reilly & Associates, Sebastopol
Sorenson O, Stuart TE (2008) Bringing the context back in: settings and the search for syndicate partners in venture capital investing. Adm Sci Q 53(2):266–294
Takahashi N (2000) The emergence of generalized exchange. Am J Sociol 105(4):1105–1135
Tsai W, Ghoshal S (1998) Social capital and value creation: the role of intrafirm networks. Acad Manage J 41(4):464–476
Uzzi B (1996) The sources and consequences of embeddedness for the economic performance of organizations: the network effect. Am Sociol Rev 61(4):674–699
Uzzi B (1997) Social structure and competition in interfirm networks: the paradox of embeddedness. Adm Sci Q 42(1):35–68
Uzzi B, Lancaster R (2003) Relational embeddedness and learning: the case of bank loans managers and their clients. Manage Sci 49(4):383–399
Waserman S, Faust D (1994) Social network analysis: methods and applications. Cambridge University Press, Cambridge
Wellman B, Wortley S (1990) Different strokes from different folks: community ties and social support. Am J Sociol 96(3):558–588
Woolcock M (1998) Social capital and economic development: toward a theoretical synthesis and policy framework. Theory Soc 27(2):151–208
Zaheer A, Bell G (2005) Benefiting from network position: firm capabilities, structural holes, and performance. Strateg Manage J 26(9):809–825
Zajac E, Westphal JD (1996) Director reputation, CEO-board power, and the dynamics of board interlocks. Adm Sci Q 41(3):507–529
Zhang C (2007) Emergence of new project teams from open source software developer networks: impact of prior collaboration ties. www.mgmt.purdue.edu/academics/mis/workshop/cz_01907.pdf