Paper to be presented at the

35th DRUID Celebration Conference 2013, Barcelona, Spain, June 17-19

Information, control, and small worlds: studying returns to individual network positions under different global structures

Antoine Vernet
Imperial College London
Business School
a.vernet@imperial.ac.uk

Martin Kilduff
University College London
Management
m.kilduff@ucl.ac.uk

Ammon Salter
Imperial College London
Business School
a.salter@imperial.ac.uk

Abstract

Why are some people able to anticipate the future direction of a community, spotting trends and breakthroughs? Building on network theory, we investigate how structural positions ? individual-level closure and project-level brokerage - enable individuals to identify and realize new opportunities in fluid, dynamic technical communities. Using the context of open source software projects, we argue ? and empirically demonstrate ? that closure enables individuals to anticipate which projects capture the attention of the community, leading to their eventual release. However, we suggest that individuals who connect disconnected projects are in a better position to identify those initiatives that later become innovations. We show that the benefits of these structural positions are moderated by the dynamism and small-worldedness of the fields in which these individuals are active. We hypothesize that in dynamic fields the effect of closure is weakened, whereas the effect of brokerage is heightened. In turn, field-level small-worldedness moderate the effects of brokerage. We find partial support for these moderators. We explore the implications of these findings for community-based organizing, opportunity identification and network theory.

Jelcodes:O32,D23

INFORMATION, CONTROL, AND SMALL WORLDS: STUDYING RETURNS TO INDIVIDUAL NETWORK POSITIONS UNDER DIFFERENT GLOBAL STRUCTURES

ABSTRACT

Prior research showed that both global network structures and individual network positions affect individual innovative outcome. These two approaches theorize about related processes to explain their effects. The two levels are not usually theoretically integrated which leads us to think that researchers might be unknowingly looking at the expression of the same process at different levels. We develop a theoretical framework for the interplay of individual level and network level processes. Building on network theory, we investigate how structural positions – between individual brokerage, between project brokerage and small-world properties of the community – enable individuals to identify and realize new opportunities in fluid, dynamic technical communities. Using the context of open source software projects, we show that different types of brokerage enable individuals to anticipate which projects are ultimately completed. We demonstrate that individuals benefit from their network position when certain conditions in the structure of the whole community are met. We explore the implications of these findings for opportunity identification and network theory.

Keywords: Small World, Social Capital, Brokerage, Closure, Innovation, Open Source Software

Relationships are resources individuals can mobilize to make sense of their environment. Prior research explored how individual position influences individuals' decisions. However, little work exists on the interaction of local and global structures. Are the benefits of one's position contingent on the global structure of the network or are networks only structuring local processes? Strong evidence exists showing that the individual's position in a network shapes individual's success (Burt, 1992 ; Coleman, 1988 ; Obstfeld, 2005 ; Shipilov & Li, 2008 ; Gargiulo et al. 2009). Similarly strong evidence shows that global network configuration influence individual or team success (Watts, 2004 ; Uzzi & Spiro, 2005 ; Uzzi et al. 2007 ; Schilling & Phelps, 2007). Studies which examine individuals' position usually do not develop a theoretical framework for the effect of global network characteristics. Similarly, theoretical endeavours about global network structures do not consider the effect of individual level characteristics. This yields two questions. First, it suggests that conjoint examination of the effect of global level characteristics and individual level characteristics will help assess whether the two are different expressions of the same underlying process: indeed, some past research suggests that the global properties of the network arise from local behaviour (e.g. preferential attachement leading to the creation of hubs) (Watts & Strogatz, 1998; Newman, 2003; Robins et al. 2005). If we follow the logic of those studies strictly, there is no theoretical reason to expect to observe any effect of global characteristics of the network when controlling for individual level positions in the network. However, we argue that even if global structures are emerging properties of individual level tie creation decisions in the past, at every point in time global structure is still exogenous to individuals and therefore have an influence on network flows at the global level. In turn these can hinder or enhance the benefits individuals capture from their position in the network. This leads to our second question: if there is distinct influence from the global network structure and the individual level position, we need to examine the interplay of global characteristic of the network and individual position. If characteristics of the whole network predict that overall performance of individuals in the network is better, the question of the distribution of this added performance between individuals remains. We argue that based on their position, individuals will experience different outcomes for specific states of the whole network. By doing this, we further prior attempts at examining contingency returns of social capital (Burt, 1997). We explore two dimensions of individuals' position in the network to further our understanding of brokerage. Brokerage has been showed to bring benefits to individuals through two mechanisms: access to diverse information and control of this information flow. We take advantage of the unique setting of our study to partially disentangle "information brokerage" from "control brokerage".

To test our theory we study a large open source software community. We investigate the *effects of individual network position and global network structure on individuals' choices.* We also explore whether or not the ability of individuals to capture returns from their network position is contingent on global network structure.

Open-source software (OSS) development has been described as resembling 'a great babbling bazaar of different agendas and approaches' (Raymond, 1999) (p.30) in which large numbers of individuals self-organize into a variety of different projects. Each individual programmer in this bazaar, however, is likely to make rational decisions by taking into account personal interests, effort/reward trade-offs and the likelihood of project success. Self-organization is likely to be underpinned by careful calculation on the part of all those engaged in scanning for potential projects to join. The result of all these individual decisions tends to be a system in which programmers cluster around a relatively small number of projects. In addition, the innovative nature of the enterprise (developing original software) leads to high uncertainty in the outcome and therefore many unsuccessful projects. In this context, the goal of an OSS programmer is to choose projects that have the greatest potential for achieving a viable outcome (i.e., a released piece of software) (Hertel et al. 2003; Lakhani & Wolf, 2005). Yet, it is difficult to know ex-ante which new projects are likely to be completed. Programmers are at risk of joining and contributing to projects that fail to develop. If programmers commit significant effort to failing projects, they may find that they garner little in the way of rewards (Lerner & Tirole, 2002). Join the wrong projects and your efforts may be lost. Join too late and you may become lost in the crowd, lowering the potential that you will be recognized for your efforts.

We know that programmers achieve informal leadership in OSS communities through the familiar processes of technological contributions and social networking (Fleming & Waguespack, 2007) but there is little work exploring how social networks affect individual programmers' choice of projects (but see Singh et al. (2011) , for analysis of how team social networks contribute to OSS team projects' success). To help explore this area, we examine how the individual's network position influences their ability to join the right projects and how the global structure of the network, because it transforms global information flows, shapes returns individuals can capture through their network position.

The following section develops our theoretical framework in more depth. Then the three following sections present our data, analytical strategy and results in turn. In a final section, we discuss our results from a theoretical perspective and develop their implications for future research, and for managers.

THEORETICAL AND EMPIRICAL BACKGROUND

Actors who bridge social structures rich with structural holes potentially benefit from information and control advantages (Burt, 1992). Information advantages derive from brokers' access to diverse sources of timely news and knowledge, whereas control advantages derive from brokers' ability to strategically manipulate this flow of resources between disconnected contacts. A common difficulty is to disentangle the two type of brokerage activity. Both brokerage activities potentially provide benefits to the focal actor, but attributing the benefits of a brokerage position to one or the other has often been impossible.

Interest in individual's network position has grown in parallel but independently from the interest in network searchability. This interest originated with formulation of the small world problem (Milgram, 1967). Following the development of a formal model for search in large networks (Watts & Strogatz, 1998), several studies have shown how smallworldedness of a network improves performance outcomes for teams or firms within it (Uzzi & Spiro, 2005; Schilling & Phelps, 2007). Both studies show positive effects for firms or teams in small-world networks, holding individuals' characteristics constant. However, they do not explore the interplay of global network structure and individual level position. Following Kilduff et al. (2006) , p. 1043 we argue that individuals with bounded rationality will have difficulties in grasping the global network structure and where they sit in it, and will remain highly biased toward their local network. However, in a small world, it is possible that the ability to extract value from one's position will differ that in a non-small world network.

Our study explores both the effect of individuals' network position and network global structure on individual behavior. We further contribute to the literature by examining the global structure influence on how individuals' can take advantage of their position in the network. Our final contribution is to consider two types of brokerage activity to disentangle between information and control advantages. Control for a broker implies the ability to allow information to flow or not between two of his/her unconnected alters and to act as an intermediary that can transform the information while transfering it. The other advantages associated with brokerage presume ease of access to diverse information often referred to in the literature as a 'vision advantage' (Burt, 2005 , p. 59). Therefore, we develop two measures to capture both separately, theorizing and measuring their effects at different levels.

In this study, combining individual structural position and global network structure, we put forward that the interplay of global structure and individual position is a promising avenue for deepening our understanding of social networks and the ability of individuals to reap benefits from their structural positions.

Between-project brokerage: information access

Prior research has associated brokerage with the ability to access original information and control the flux of this information (Burt, 1992; Burt, 2005; Shipilov & Li, 2008). In our settings, access to information about competing projects is critical because most projects will not be completed. Developers want to join projects that are likely to be released, to avoid expending efforts for no returns. Therefore, programmers scan the community for indications of projects' potential to identify the ones worth joining. However, individuals have limited cognitive resources and the size of the community makes it impossible for them to evaluate every project.

Projects can be seen as repositories of knowledge as they keep records of the technical history of the project through version control, bug tracking systems and discussion forums. Therefore, between-project brokers (members of otherwise disconnected projects) are likely to recognize opportunities to transfer ideas or features across projects. Those individuals are in an ideal position to bring together and recombine partial solutions to a problem found in two disconnected projects. This is likely to give them an edge when they are considering joining a project: they might be able to envision how to solve some technical hurdles in that new project and therefore have a more accurate evaluation of the potential of the project. This is a vision advantage. It is also likely that between-project brokers access more diverse information because they connect separated parts of the community. In doing so, they are exposed to more diverse collaboration styles in different projects. In summary, project brokers gain two types of advantages from connecting disconnected projects: an ability to transfer technical solutions between projects and an access to more diverse information and diverse styles of collaboration making them more versatile.

Accordingly, individuals in between-project brokerage positions are, we suggest, in a better position to recognize which early stage projects have potential. Thus, we hypothesize:

Hypothesis 1. Between-project brokerage enables the individual to increase chances of joining ultimately-successful projects.

Between-individual brokerage: control

By individual level brokerage, we refer to the classical definition of brokerage as the extent to which one's contacts are disconnected. In an OSS community, this translates into whether the collaborators of a developer collaborate together.

In an open source software community, projects record a lot of information on the software development and on exchanges between project members through version control of the code, bug tracking systems and forums. However, more informal knowledge, like a developers current engagement with a project, is not easy to evaluate without direct exchange with the individual or one of his/her collaborators. As a result individuals have to rely on information relayed by their contacts. Between-individual brokers gain information on their contacts and their collaborators that others who do not bridge disconnected components of the community are not aware of. Those between-individual brokers also enjoy control advantages. These take two forms. First, because they connect disconnected developers, they control the local flux of information between those individuals. This means that they can choose whether or not to transmit information from one to the other. They also can use information from one developer to gain an advantage when dealing with the other. The other type of control advantage enjoyed by between-individual brokers is the possibility to bring together disconnected individuals after spotting potential synergies (Obstfeld, 2005). Having collaborated with two individuals on two different projects, a between-individual broker might become aware of complementarity in their skills. Therefore, this individual would realize the potential of a project in which those two individuals are involved while others would not. More important, this between-individual broker can take it upon himself to convince those individuals to join and realize the synergy.

Between individual brokers enjoy two types of advantages: a better knowledge of developers' skills and engagement and the ability to control the local flux of knowledge and to foster potential synergies between developers by bringing them together. Thus we hypothesize:

Hypothesis 2. Between-individual brokerage enables the focal individual to increase chances of joining ultimately-successful projects.

Small world

A network is a small world if it exhibits two properties: high clustering and short path length (Watts & Strogatz, 1998). Clustering measures the average local density of ties in the network. A high clustering is associated with high transmission of information within the cluster (Schilling & Phelps, 2007). Path length measures the average distance between any two nodes of a network. It is associated with access to high quantity and diversity of information (Burt, 2004; Schilling & Phelps, 2007). This has led to an interest for small-world networks as a desirable structure for efficient information transmission.

Several studies have taken on the task of exploring the effect of the global structures of the network on individual or team level outcomes following the renewed interest for 'small worlds' spurred by the formalization of the phenomenon first studied by Milgram (Milgram, 1967; Watts & Strogatz, 1998). Results from those studies have been somewhat mixed. Most studies have found a positive association of small world and performance (Uzzi et al. 2007). However, in some cases, the association is linear (Schilling & Phelps, 2007), in other it is an inverted U (Uzzi & Spiro, 2005; Guimera et al. 2005). Yet, other studies fail to find any association between small world and innovation altogether, but find a positive effect of shorter path length on innovative outcome (Fleming et al. 2007).

To further our understanding of the effect of small world on innovative output, we want to examine its effect on innovation in a setting requiring intense knowlegde exchange. We argue that because information flows better and more uniformely through a small world, those are likely to show better innovative performance than non small-world networks. Consequently, we hypothesize:

Hypothesis 3. Small-*world characteristics of the individual's community increases the* chances of the project to be ultimately successful

Interplay between individual network positions and network state

Small world studies sometime accounted for individual position in the network, but they have overlooked the potential interplay of the two levels. However, because small worlds imply the presence of hubs between clusters (Baum et al. 2003; Robins et al. 2005) and that those hubs closely resemble brokers (brokers have to be hubs, but not all hubs are brokers), it begs the question of how the benefits of small world are distributed between individuals in the network.

Prior studies called for future research to explore the multilevel effect of small world (Uzzi & Spiro, 2005; Uzzi et al. 2007). Burt (2004) suggested that brokers are extracting value from being embedded in a small world, because they are in a position to extract and recombine diverse knowledge drawn from various homogenous clusters.

Adopting this logic, we expect individuals in brokerage position to benefit more from a network exhibiting small world characteristics than individuals not in this position, because they will be able to monitor and partly control the intense flow of information in small world networks (Schilling & Phelps, 2007). Thus, in the context of open source software, to the extent that the network exhibits small world properties, the advantages enjoyed by brokers are likely to be increased because brokers will be able to draw from highly-clustered sources of distinctive knowledge. However, we expect the 2 types of brokerage positions we have defined to be differently affected by being embedded in a small world community.

For between-project brokers, being embedded in a small world means that information flows are more intense. This is likely to lead to a more homogeneous spread of information. However, between-project brokers sit between projects but not necessarily between individuals, as a result, their local control on who has access to the knowledge flowing in the network is limited. They have better information than average, but they cannot control its flux. As a result, in a small world those brokers should retain an advantage but this advantage should be smaller because information is more heavenly spread in the community.

For between-individual brokers, in a small world, they control information circulation between individual at the local level, so they will retain the ability to block or allow the transmission of a piece of knowledge. Therefore, they should be able to gain a larger advantage from a small world. Accordingly, we expect:

Hypothesis 4. Small-world characteristics will negatively moderate the effect of between project brokerage on the ability of individuals to anticipate projects that go to release

Hypothesis 5. Small-world characteristics will positively moderate the effect of between individual brokerage on the ability of individuals to anticipate projects that go to release

METHODS

Our study is based on an analysis of the SourceForge Research Data Archive (SRDA) (http://zerlot.cse.nd.edu) (Gao et al. 2007; Van Antwerp & Madey, 2008). SourceForge is the largest online repository of open source software projects, covering a wide range of individuals and types of projects. The SRDA is available through the University of Notre Dame, which sublicenses the data they receive directly from SourceForge to other researchers. Although SourceForge has already received some attention in the literature on open source software (Madey et al. 2002; Xu et al. 2005; David & Rullani, 2008; Fershtman & Gandal, 2011), there are few studies on the effects of networks on open source development using this data (see Singh et al. (2011) for an example of one such studies). The SourceForge dataset has a number of advantages for the study of networks and project outcomes. First, it is comprehensive as it covers all the projects hosted by the website, as well as all individual developers. This allows us to observe both successful and unsuccessful projects, helping to correct the tendency, common in the literature, to focus on successful OSS projects. The data also covers a wide range of programs, allowing us to avoid focusing on a specific project or subset of projects. Moreover, since OSS project communication and project work takes place online, we are able to reconstruct almost complete records of the nature of communication and exchange in the project team. We use detailed and time-stamped information on 297,110 individuals working on 234,147 open source projects. The database we constructed from the SourceForge records contains monthly data over the period January 2006 to June 2008. Using this information, we can observe individuals, projects and ties over time. Exploiting the time dimension and richness of this data, we address some concerns about causality by using lagged network variables to explore how individuals' closure and brokerage positions and field layout shape their perspective and therefore influence the type of projects they choose. In addition, the database provides information at multiple levels. The data file contains the possibility to build links between the projects and the individuals over time. As a result, we can access three levels of analysis: individual-level, project-level and dyad-level (characteristics of the link between the individual and the project). In sum, by using SourceForge, we are able to develop a fine-grained and dynamic view of the activities and affiliations of individuals to different projects over time in the community. For each month, our sample is composed of new ties to projects created in the previous three months and to projects that have not issued a release at the time of observation. The lag structure built into our dependent variables shortens our sample of the initial three months of observations that are used to calculate the dependent variables. Finally, measuring the dependent variable a year after the observation period cuts a year from the end of our sample. As a result, our estimations are based on observations starting in March 2006 and ending in June 2007.

Dependent variable

Our dependent variable records whether or not the software project has been released a year after the individual has joined. This variable is equal to one if the project has had a release and 0 otherwise. A release means the project led to a working software program, however rough it may be. A release is a significant milestone among OSS projects, as the vast majority of these projects do not lead to working programs. The mean time before software release prior to January 2009 (when our data is censored) for all the projects started in 2006 is 78 days and after 250 days, 90% of those that would be released, had been (the mean is 59 days and the 90th percentile is 198 days for 2007). 40% of the projects started in 2006 were released before January 2009, as well as 36% of those started in 2007.

We allow for a year to pass before measuring whether or not the software has been released to avoid miscategorizing software that was slow to launch.

Independent variables

Usually brokerage positions are determined by looking at links between individuals. However, we have 2-mode data, therefore we can use the information contained in the 2-mode network to define brokerage measures that account for different types of brokerage positions. This approach allows us to make use of the richness of this 2-mode data exploiting the person-to-project network to assess individuals position in the person-to-project as well as the person-to-person network. In so doing, we take into account that individuals in an affiliation network can be in a different structural position with regards to individuals and projects (Everett & Borgatti, 2012).

Between-project brokerage — We measure project brokerage as the proportion of projects of the individual between which that individual is the only link. Here, individuals only on one project have an indefinite value for our brokerage measure. However, this means that such individual do not have any project brokerage potential. As a result, we replace indefinite values with 0 and we create a dummy variable that takes the value 1 if the project brokerage value was indefinite and 0 otherwise. This allows us to disentangle individuals who are on only 1 project and individuals on several projects which all have some overlap in membership. Measuring brokerage in this way is an attempt to show how many non-redundant sources of information an individual developer has access to, considering projects as repositories of knowledge. Because it is a proportion that varies between 0 and 1, the higher the score, the more disjointed the projects the individual is involved in, and the more that individual is in a brokerage position. This measure is taken one month prior to the measure of our dependent variable to reduce concerns about reverse causality and captures the state of the individual's network before the joining event (which is our unit of analysis). We use all past affiliations in our data to build the network. To facilitate the interpretation of our interaction term, the variable is mean-centered.

Between-individual brokerage — Clustering coefficient has been used extensively in the literature since it was first conceptualized by Luce & Perry (1949). For example, it has been used in studies of team performance on Broadway (Uzzi & Spiro, 2005), of firms' innovative capabilities (Schilling & Phelps, 2007), on how people's cognitive limitations influence the way they perceive their friendship ties and ties between their friends (Kilduff et al. 2008). To measure individual level brokerage, we use the 2-mode clustering coefficient (Opsahl, 2011). This measure is similar to other clustering measures defined for 2-mode networks (Latapy et al. 2008) and extends the definition of Watts & Strogatz (1998) for 2-mode networks. Projecting the 2-mode network on the developers' mode would not be satisfactory because the resulting 1-mode network would have a density that mainly arises as an artefact of the projection (Robins & Alexander, 2004; Latapy et al. 2008).

The 2-mode clustering coefficient allows us to avoid those pitfalls. The local version of the clustering coefficient that we use to calculate the score of each individual counts the number of closed 4-paths centered on the individual divided by the total number of 4-paths centered on the individual. We use 4-path because their equivalent in a projected 1-mode network is a triad. This triad is open for an open 4-path and closed for a closed 4-path.

Insert figure 1 about here

The score varies between 0 and 1; 0 indicates that the individual is present only on open four-paths; 1, that all 4-paths are closed. To make the interpretation easier, we transform this score by taking 1-clustering coefficient. Therefore, a low value indicates that the focal individual's collaborators are often connected and brokerage opportunities are scarce; a high value indicates that the individual has few collaborators that are connected, indicating more brokerage opportunities. However, since the clustering coefficient relies on 4-paths, individuals that do not lie on a 4-path have an undefined clustering coefficient.

One approach to overcome this problem consists of dropping all individuals without a definite clustering coefficient. The drawback of this approach is that in a large, sparse network, the number of individuals without a clustering coefficient value can be expected to be large. In our case, it would significantly reduce our sample for no clear theoretical reason, because most individuals only participate in one project and therefore will not have a clustering coefficient value. In our setting, an undefined clustering coefficient is equivalent to a special case of not having any brokerage opportunity: that is having no collaborators, or having collaborators on only one project. This consideration leads us to retain a strategy that consists in replacing the missing clustering coefficient values. We replace the missing values with zeros to reflect that individuals have no brokerage opportunity if they are not in contact with multiple individuals through 4-paths.

However, we recognize that the absence of collaborators on different projects is qualitatively different from the situation in which one's collaborators are all connected through multiple paths (which also yields a score of 0 closure with our measure). To account for this difference, we construct a dummy variable, no individual level brokerage, to indicate whether or not the value has been replaced for an individual. The measurement is taken a month prior to the current observation's month to capture the state of individuals' networks before their current month activity. This helps reduce concerns about reverse causality. We use all the past affiliation information to construct the network on which we do the measurement. To help interpret our interaction term, the closure score is mean-centered.

Small world — A network exhibits small world properties if it has high clustering and short path length (Watts & Strogatz, 1998). To assess what is 'high' and 'short', it has become classic to compare the observed network to similar random networks (Uzzi & Spiro, 2005). Following this literature, we compute the small world Q as the ratio of two ratios. First the ratio of the clustering coefficient observed in our network over the clustering coefficient of a similar random network. Second the ratio of the path length in our network over the path length in a similar random network. The randomization procedure we use is called tie rewiring (Opsahl et al. 2008) and produces random networks that are a better representation of real world network than the classic Erdos-Renyi random networks (Erdos & Renyi, 1959). This makes our estimation of whether a network is a small world or not more conservative than it would be with classic random networks. This reduces concerns that an effect of small-world Q will be due to measurement errors.

We compute our small world measure every month on sub-communities of SourceForge to capture how individuals fit in the larger community. To identify the subcommunities, we build a 2-mode network of projects linked to their type of software. We do a weighted projection of that network to obtain a weighted one mode network of software linked by their shared categories. A link between two software represents the sharing of categories, the weight of the link is the number of shared categories. On the resulting network, we detect communities using the Louvain method (Blondel et al. 2008). This method groups software according to their proximity in the network of shared categories. To reflect the fact that the SourceForge platform is a dynamic environment, we do not freeze topical communities over time. Therefore we repeat the above procedure for every month of our sample. This yields between 5 and 10 communities each month. We observe that the small-world score is mainly driven by the clustering coefficient ratio, as in the study by Uzzi & Spiro (2005) . Schilling & Phelps (2007) argued this is a sign the networks are mature and relatively stable. We give projects that do not belong to a community (because they are not listed as being of any type) a score of 0 for that variable (2,145 projects-month among 35,850 unique pairs per project-month, or 6% of the sample).

Controls

Factors that could influence project outcome are numerous. To account for alternative explanations of our results, we include a wide range of control variables at three different levels: field, project and individual-level.

Field maturity — In mature fields, the number of competing projects is higher, making it more difficult for individuals to assess the prospects of a project. It also controls for the possibility that projects in mature areas are better defined and therefore more likely to go to release. The variable measures the mean number of projects in each category of software in which the focal project is registered.

Dynamism of the field — measures the change in the maturity of field, as the proportion of the global maturity score that was acquired by the field over the 6 months prior to the focal month. A high value of dynamism indicates an area that is currently getting more attention. This controls for the effects on projects of a rapidly changing environment. In such an area, the high number of newly created projects might make it more difficult for developers to evaluate prospects of a project.

We also seek to control for project-level effects.

Breadth of project — captures how many categories of software are used to describe the project. A project that is registered as a combination of many different categories is probably ill defined and thus more likely to fail.

License — records which Open Source Software licence the software uses. Prior research has shown that license choice influences development activity on a project (Stewart et al. 2006). Some license are likely to both attract more contributors and influence their engagement with the project, thus influencing its likelihood of going to release.

Contributors$_{t-1}$ — controls for the number of project contributors one month prior to the focal month. This helps control for the fact that bigger projects in terms of team size are more likely to be released.

Finally, we control for individual-level effects.

Breadth of skills$_{t-1}$ — To capture the breadth of an individual's skills in software development, we measure the diversity in types of software projects an individual has actively participated in. This measure is the number of unique categories of software (e.g. database management, scheduling tool, text editor) listed on projects the individual is involved in, prior to joining the focal project. It is likely that individual with broader skills will enjoy some of the vision advantage we attribute to between-project brokers. This control helps make sure that the effects of between-project brokerage are not due to the characteristics of the broker rather than to their structural position.

Number of projects joined$_{t-1}$ — records the number of projects joined by the individual one month prior to the focal month. It controls for individual strategies that on joining many projects in order to increase the likelihood of joining a successful one, helping us to ensure that our findings are not simply driven by individuals being polygamous.

We also control for an individual's level of activity on OSS projects using two measures: bug solving and posting$_{t-1}$ and message posting$_{t-1}$. We think controlling for past activity is essential as previous studies have shown that activity in an OSS community strongly correlates to obtaining positions of leadership in communities (Dahlander & O'Mahony, 2011) which might influence network positions.

Bug solving and posting$_{t-1}$ — this measure of activity counts the numbers of bugs submitted or solved by each individual in the six months prior to the observation period. We use a lagged measure to ensure that activity does not simply reflect current levels of effort on the project. Instead, it measures technical commitment to the OSS community. As well as controls for the technical expertise of the individual. This expertise is likely to influence their evaluation of projects they could join and ultimately their affiliation decisions.

Message posting$_{t-1}$ — measures activity as the number of messages posted by the individual on forums in all the projects they are involved in, in the six months prior to the observation period. The information gathered through reading and answering posts is likely to shape the individual's perspective. Voicing opinions and concerns is also likely to shape the type of opportunity an individual will have access to. As such, the overall activity of an individual on forums is likely to influence his/her assesment of projects and therefore his choices.

We control for an individual's activity on the focal project, to control for the fact that the ability to choose successful projects could be largely driven by that individual' s contribution to the projects they are involved with. We use two measures to capture this personal effect: bug solving and posting on project and message posting on project.

Bug solving and posting on project — measures the number of bugs solved and posted between the time they join and a year after, proxying an individual's technical activity on the project.

Message posting on project — records the number of message posted by the focal individual on the project's forums between the time they join and a year after.

Tenure — controls for how long an individual has been a member of the OSS community. This is measured by the number of days of tenure on the Sourceforge website starting on the day the individual joins. This controls for differences in point of view and opportunity access between seasoned and more junior community members.

Timezone — controls for country-effects and synchronous communication advantages. It is a dummy variable that takes value 1 if the individual is in the US (the most represented country in terms of membership) and 0 otherwise. Time zones are registered at the time an individual joins the website, which means that there is some noise in the variable, especially for long-time members.

Number of projects founded$_{t-1}$ — Count of the number of projects created by the individual over the six months prior to the observation period. This controls for the effect of an entrepreneurial experience of the individual and how such an experience shapes how individuals' evaluate projects they might join.

Number of administrative roles$_{t-1}$ — records the number of adminitrator roles the individual has in different projects in the six months before the observation period. Administrators have a leadership role on projects and are generally more experienced and also regarded as technical authorities by other developers (Dahlander & O'Mahony, 2011). Thus controlling for administrator roles is a proxy for unobserved quality of individuals' contributions.

Past successes — Due to the structure of our data and the limitation in the time period it spans, we cannot construct an exact duplicate of our dependent variable. In order to try to address the influence of past success on future success, we measure the number of affiliations to young unreleased project in the six months prior to the observation that were released before the observation.

Table 1 gives an overview of our controls.

Insert Table 1 about here

Estimation strategy

We estimate logistic regression models using release as a dependent variable. All models include individual and month random intercepts.

Table 2 reports descriptive statistics and correlations between the variables. Overall, we do not observe high correlations among our variables. However, project brokerage correlates highly with breadth of skills (0.60), no between individual brokerage (-0.48) and no between project brokerage (-0.57). Similarly, breadth of skills correlates highly with no between individual brokerage (-0.64) and no between project brokerage (-0.62) and Small world ratio (0.60). Small world ratio also correlates highly with no between project brokerage (-0.98). We estimate separate models with no between project brokerage and small world ratio to avoid multicollinearity issues. The results of both models are similar. This leads us to estimate models with only small world ratio. Our models with random intercepts are robust to correlations within individuals and within months. This leads us to believe multicollinearity is not a major issue here. Observation of the values reported in Table 2 led us to transform some of the variables prior to inclusion in our models. We took a log transformation of the following variables: bug solving and posting$_{t-1}$, message posting$_{t-1}$, bug solving and posting on project, message posting on project, maturity and tenure. The small-world Q is logged and then centered to facilitate the interpretation of interactions (Gelman & Hill, 2007).

Insert Table 2 about here

RESULTS

Table 3 reports the estimated random intercept logistic models for the effect of between project brokerage, between individual brokerage and small-world Q on the ability of a developer to choose projects that go to release.

Insert Table 3 about here

Our first hypothesis stated that the more individuals were in a between-project brokerage position the greater their likelihood of joining a project that goes to release. We report results for the model in column 4 of Table 3 as it is the model including small-world ratio that fits the data the best. We found support for this hypothesis. Between project brokerage is positive and significant when the interaction with small world is included. In Table 3, column 4, the coefficient is positive and significant (ß=1.29, $p<0.01$). The modification in the odds ratio of an increase of 1 unit in the between project brokerage score was 265% ($e^{1.29}=3.65$).

What about between individual brokerage? Recall that hypothesis 2 suggested that between individual brokerage enabled individuals to better identify projects likely to achieve release. This hypothesis is not supported as shown by the results in Table 3, column 4: the coefficient for between individual brokerage is negative and significant (ß=-1.90, p<0.01). The modification in the odds ratio of an increase of 1 unit in the between individual brokerage score is a reduction of 86% ($e^{-1.90}$=0.14).

Our third hypothesis stated that an individual in a small-world community had higher chances of joining a project that would go to release. We find support for this hypothesis. The coefficient for small world is positive and significant (ß=0.027, p<0.01). The modification in the odds ratio of an increase of 1 unit on the logged small world score is 3% ($e^{0.027}$=1.03).

This gives us a picture of how individuals in the open-source software community succeeded and failed depending upon whether their networks exhibited between project brokerage or between individual brokerage and whether the smallworldedness of the community of the individual influenced their likelihood of joining projects that go to release. We now want to focus on the effect of the interaction between small world and the individual level variables. Hypothesis 4 stated that between project brokers would get extra benefits from being embedded in a small world. We find no support for this hypothesis. To the contrary, we find a significant negative effect of the interaction of project brokerage and small-world ratio. Figure 2 shows how small-world ratio moderates the effect of between project brokerage. We see that in a community with a lower small-world ratio, a higher brokerage score is highly beneficial. In a community with a higher small-world ratio, a higher brokerage score is still beneficial but less so.

Insert Figure 2 about here

Finally, hypothesis 5 stated that individual in a between individual brokerage would benefit from being embedded in a small world community. We find support for this hypothesis. We find a positive and significant effect for this interaction. Figure 3 shows how small-world ratio moderates the effect of between individual brokerage. In communities with both high and low small world, the effect of a higher between individual brokerage score is a decrease in the likelihood of going to release. However, in a community with a low small world score, the decrease in probability is steeper for low to medium score of between individual brokerage.

Insert Figure 3 about here

DISCUSSION

In this paper, we have focused on how individual choices are influenced by social network position as well as by the structure of the network as a whole. In our setting, in which time is a scarce resource, programmers have to choose carefully which projects to join. We hypothesized that the network of affiliation to projects, and to other programmers, influences programmers' perspectives and therefore their ability to allocate their time to specific projects. The main result of our study is to emphasize the importance of the global network structure for individuals to take advantage of their network position. More specifically, the study suggests that being in a between-project brokerage position — connecting otherwise unconnected projects — helps individuals gain advantages in identifying projects that get released. However, being able to take advantage of this individual level position is dependent on the structure of the community: in environment with low smallworldedness, between-project brokers gain much more from their position than between project brokers in community with a high small world ratio. This suggests that in a small world community, the information advantages enjoyed by between project brokers are not fully retained by them. Indeed, their advantage lies in better and more diverse information access. In a small world in which information is spread more evenly, their advantage is eroded because they do not control the local flux of information and therefore cannot maintain their informational advantage toward other members of the community.

Being in a between-individual brokerage position is detrimental to identifying projects that subsequently go on to be released irrespective of whether or not the individuals are embedded in a small world. But in a small-world community, the negative effect of a between-individual brokerage position is less steep. This suggest that in a small world, these between-individual brokers can use their ability to control the flux of information to their advantage. The evidence we presented suggests that the role of brokers at each level in small worlds necessitates further research to better understand the conditions in which brokers can benefit from their individual position.

The difference between the advantages accruing to between project brokers and the disadvantages experienced by between individual brokers suggest that in our setting, information advantages are more critical than control advantages. However, this might be partially due to the fact that control in our setting is harder to enforce than in more traditional organization in which formal boundaries reinforce barriers between individuals, making it easier for brokers to maintain and exploit their structural advantage.

Our study extends the social network literature in two important ways. First, we study networks at two-levels and the interactions between them. Second, we further the brokerage literature by opening the black box of brokerage to try to assess the independent effects of vision and control. Our results suggest that brokers can only exploit the full potential of their position in specific settings. It seems that smallworldedness of the community acts as an equalizer, limiting the downsides of between individual brokerage and the advantages of between-project brokerage.

We break away from recent studies of brokerage in the management literature (Burt, 1992; Burt, 1997; Shipilov & Li, 2008) that use one structural measure for brokerage. Numerous studies use the network constraint to measure brokerage (Burt, 1992). We see this measure as inherently limited as it cannot separate the effects of vision and control.

We take full advantage of our access to accurate affiliation data to define our network measures. We define between-project brokerage to assess individuals' structural position towards projects and between-individual brokerage to capture presence of structural holes among an individual's collaborators.

Because our measures capture two distinct constructs (they are only correlated at 0.02), we argue that brokerage needs to be measured at 2 levels to disentangle the advantages accruing to brokers from information and from control. This opens an avenue for future research to help define more appropriate measures for brokerage, especially for 2-mode networks, that both have a strong theoretical basis but are also rooted in the specificities of the data analyzed.

Our study shows information advantages associated with between project brokerage are not straightforward and necessitate a specific community structure to be efficient. Between individual brokerage did not provide any advantage, but here again, the extent to which this position is detrimental depends on the community structure at large.

Limitations and future research

Our study has several limitations. The main limitation lies in our setting. Although OSS is a growing and important area, it has specific routines and institutional norms that may make it difficult to generalize from this area to other domains of organizational life. To confirm that there are multi-level effects of networks on individual choice, further study in other settings is necessary.

Second, we use position in the network as a given at every time point. However, network position at a certain point in time is likely to be a function of past network position as well as individual choices. We address this limitation by controlling for the individual's past activity (as number of projects joined and past activity in the community). Nonetheless, further research could examine the antecedents of network position in open source communities.

Third, by exploring the interplay between global network structure and individual's position on this structure, we answer in part the question posed by Kilduff et al. (2006:1043), on the effect of networks at different levels. Brokers do experience different returns from their network position based on the structure of their community. However, we do not explore network cognition and therefore do not fully answer their call. Again, further research should examine interaction between network cognition, individual position and global network structure to fully address this.

Fourth, our analysis is also limited by the nature of the dataset. Although Sourceforge is an impressive dataset, our sample covers a short range of months and some of the communication that takes places between OSS developers does not appear in the data (e.g., private e-mail communication). We proxy global communication through the use of the public communication available to us (message on forum and bug posting and solving trackers).

Fifth, our sampling method excludes inactive projects and developers. Some individuals might come in (out) of our analysis as they join (leave) their only project. However, we think this is not a strong limitation because we focus on how individual choices are shaped and thus individuals who fail to participate in any projects and therefore do not make joining decisions are outside of the scope of this study.

Sixth, this study has focused on individuals's choices rather than teams'. It is clear that much of OSS involves the engagement of teams working together on common problems. Our approach accounts for the project and individual networks, but it does not observe the movement of groups over time and how these migrations shape OSS outcomes. There is an opportunity here for further research on the movement of people between projects and how different forms of social capital lead them to join particular projects.

Managerial Implications

Our study has two types of implications for management. Implications that are specific for software development platform and more general insights on how managers can use knowledge about social network topology.

Our results should encourage platforms to shift their focus from project to individual when trying to identify promising projects. Indeed it seems that individuals are better able to identify promising projects when occupying a certain network positions. This should lead a platform to identify those individuals at any point in time and carefully monitor their decisions to uncover the most promising projects.

At a more general level, our study highlights the importance of networks in decision-making and should lead managers to carefully think about mapping their collaborators and organization networks and using the map to identify individuals who are likely to be the best at evaluating uncertain ventures. Consequently managers will rely more heavily on those individuals' opinions when making a decision.

Conclusions

Most creative efforts in OSS end in failure, increasing the chances than individuals will expend their time and energy on abortive projects. In order to confront this uncertainty, individuals draw upon their personal and professional networks to mobilize resources, support and advice. Our paper shows that an individual's network position — measured by their levels of between-individual brokerage in the network of open source software developers and by their levels of between-project brokerage — can significantly shape their ability to choose successful projects. In addition, we show that the structure of the overall community in which an individual evolves also shapes their choices. Finally, we show that community structure can hinder or enhance advantages individuals draw from their position.

The advantages of between-project brokerage are vision advantages, but it seems these are less efficient in a small world that exhibits stronger flow of information. Between-individual brokers are at a disadvantage in our setting, but it seems that control becomes more crucial when information is spread more evenly. Our study provides a first foray into contingencies of global network structure and individual level position, more research is needed to draw a more precise picture of the necessary conditions for small worlds to impact the benefits individuals are able to draw from their network position.

REFERENCES

Baum, J. A. C., & Shipilov, A. V., & Rowley, T. J. 2003. Where do small worlds come from? **Industrial and Corporate Change,** 12(4): 697-725

Blondel, V. D., & Guillaume, J., & Lambiotte, R., & Lefebvre, E. 2008. Fast unfolding of communities in large networks. **Journal of Statistical Mechanics: Theory and Experiments,** 2008(10): P10008

Burt, R. S. 1992. **Structural hole. The social structure of competition.** Cambridge: Harvard University Press

Burt, R. S. 1997. The Contingent Value of Social Capital. **Administrative science quarterly,** 42(2): 339-365

Burt, R. S. 2004. Structural Holes and Good Ideas. **American Journal of Sociology,** 110(2): 349-399

Burt, R. S. 2005. **Brokerage and Closure, An Introduction to Social Capital.** Oxford: Oxford University Press

Coleman, J. S. 1988. Social Capital in the Creation of Human Capital. **American Journal of Sociology,** 94(Supplement): S95-S120 UChicago Press

Dahlander, L., & O'Mahony, S. 2011. Progressing to the Center: Coordinating Project Work. **Organization Science,** 22(4): 961-979

David, P. A., & Rullani, F. 2008. Dynamics of innovation in an "open source" collaboration environment: lurking, laboring, and launching FLOSS projects on SourceForge. **Industrial and Corporate Change,** 17(4): 647-710

Erdos, P., & Renyi, A. 1959. On Random Graphs I. **Publicationes Mathematicae Debrecen,** 6: 290-297

Everett, M., & Borgatti, S. 2012. The dual-projection approach for two-mode networks. **Social Networks** (forthcoming).

Fershtman, C., & Gandal, N. 2011. Direct and indirect knowledge spillovers : the "social network" of open-source projects. **Rand Journal of Economics,** 42(1): 70-91

Fleming, L., & King, C., & Juda, A. I. 2007. Small Worlds and Regional Innovation. **Organization Science,** 18(6): 938-954

Fleming, L., & Waguespack, D. M. 2007. Brokerage, Boundary Spanning, and Leadership in Open Innovation Communities. **Organization Science,** 18(2): 165-180

Gao, Y., & Antwerp, M. V., & Christley, S., & Madey, G. 2007. A Research Collaboratory for Open Source Software Research. **International Workshop on Emerging Trends in FLOSS Research and Development (FLOSS'07: ICSE Workshops 2007),**: 4-4 Ieee

Gargiulo, M., & Ertug, G., & Galunic, C. 2009. The Two Faces of Control: Network Closure and Individual Performance among Knowledge Workers. **Administrative Science Quarterly,** 54(2): 299-333

Gelman, A., & Hill, J. 2007. **Data Analysis Using Regression and Multilevel/Hierarchical Models.** Cambridge: Cambridge University Press

Guimera, R., & Uzzi, B., & Spiro, J., & Amaral, L. A. N. 2005. Team assembly mechanisms determine collaboration network structure and team performance.. **Science,** 308(5722): 697-

702

Hertel, G., & Niedner, S., & Herrmann, S. 2003. Motivation of software developers in Open Source projects: an Internet-based survey of contributors to the Linux kernel. **Research Policy,** 32(7): 1159-1177

Kilduff, M., & Tsai, W., & Hanke, R. 2006. A paradigm too far? A dynamic stability reconsideration of the social network research program. **Academy of Management Review,** 31(4): 1031-1048

Kilduff, M., & Crossland, C., & Tsai, W., & Krackhardt, D. 2008. Organizational network perceptions versus reality: A small world after all? **Organizational Behavior and Human Decision Processes,** 107(1): 15-28

Lakhani, K. R., & Wolf, R. G. 2005. Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects. **Perspectives on Free and Open Source Software** MIT Press

Latapy, M., & Magnien, C., & Del Vecchio, N. 2008. Basic notions for the analysis of large two-mode networks. **Social Networks,** 30: 31-48

Lerner, J., & Tirole, J. 2002. Some Simple Economics of Open Source. **The Journal of Industrial Economics,** L(2)

Luce, R., & Perry, A. 1949. A method of matrix analysis of group structure. **Psychometrika,** 14(1): 95-116

Madey, G., & Freeh, V., & Tynan, R. 2002. The Open Source Software Development Phenomenon: an Analysis Based on Social Network Theory.: 1806-1813 **Americas conf. on Information Systems (AMCIS2002)**

Milgram, S. 1967. The Small-World Problem. **Psychology Today,** 1(1): 61-67

Newman, M. E. J. 2003. The structure and function of complex networks. **SIAM review,** 45(2): 167-256

Obstfeld, D. 2005. Social Networks, the Tertius lungens Orientation, and Involvement in Innovation. **Administrative Science Quarterly,** 50(1): 100-130

Opsahl, T. 2011. Triadic closure in two-mode networks: Redefining the global and local clustering coefficients. **Social Networks,** (forthcoming)

Opsahl, T., & Colizza, V., & Panzarasa, P., & Ramasco, J. 2008. Prominence and Control: The Weighted Rich-Club Effect. **Physical Review Letters,** 101(16): 1-4

Raymond, E. 1999. The cathedral and the bazaar. **Knowledge, Technology & Policy,** 12(3): 23-49

Robins, G., & Alexander, M. 2004. Small worlds among interlocking directors: Network structure and distance in bipartite graphs. **Computational & Mathematical Organization Theory,** 10(1): 69-94

Robins, G., & Pattison, P., & Woolcock, J. 2005. Small and Other Worlds: Global Network Structures from Local Processes. **American Journal of Sociology,** 110(4): 894-936

Schilling, M. A., & Phelps, C. C. 2007. Interfirm Collaboration Networks: The Impact of Large-Scale Network Structure on Firm Innovation. **Management Science,** 53(7): 1113-1126

Shipilov, A. V., & Li, S. X. 2008. Can you have your cake and eat it too? Structural holes' influence on status accumulation and market performance in collaborative networks. **Administrative Science Quarterly,** 53(1): 73-108

Singh, P. V., & Tan, Y., & Mookerjee, V. 2011. Network effects: The influence of structural social capital on open source project success. **MIS Quarterly,** 35(4): 813-829

Stewart, K. J., & Ammeter, A. P., & Maruping, L. M. 2006. Impacts of License Choice and Organizational Sponsorship on User Interest and Development Activity in Open Source Software Projects. **Information Systems Research,** 17(2): 126-144

Uzzi, B., & Amaral, L. A. N., & Reed-Tsochas, F. 2007. Small-world networks and management science research : a review. **European Management Review,** 4(2): 77-91

Uzzi, B., & Spiro, J. 2005. Collaboration and Creativity: The Small World Problem. **American Journal of Sociology,** 111(2): 447-504

Van Antwerp, M., & Madey, G. 2008. Advances in the sourceforge research data archive. **Workshop on Public Data about Software Development (WoPDaSD) at The 4th International Conference on Open Source Systems**

Watts, D. J. 2004. The "New" Science of Networks. **Annual Review of Sociology,** 30(1): 243-270

Watts, D. J., & Strogatz, S. H. 1998. Collective Dynamics of 'Small-World' Networks. **Nature,** 393(4): 440-442

Xu, J., & Gao, Y., & Christley, S., & Madey, G. 2005. A Topological Analysis of the Open Source Software Development Community. 00(C): 1-10 **Proceedings of the 38th Hawaii International Conference on System Sciences**

TABLES

Table 1: Overview of controls

Control	Measure	Theoretical rationale
Field-level controls		
Field maturity	Mean number of projects in all software categories the focal project is registered	Controls for the possibility that small world effects would only be an effect of being located in a mature area in which software are better defined and therefore more likely to go to release. It also controls for competition faced by the new project
Dynamism of the field	Proportion of the maturity score that was acquired in the previous 6 months	Controls for effects specific to rapidly changing areas
Project-level controls		
Breadth of project	Number of software categories used to describe the software	Controls for effects of ambition of projects. A less ambitious projects are likely to be better defined and go to release more often
License	Type of license of the software (13 types)	Certain open source software licenses attract more contributors and also influence their engagement, thus influence the likelihood those projects will go to release
$\text{Contributors}_{t-1}$	Number of project contributors	The larger the team the more likely it is that the project is going to see high level of activity and the more likely the software is to be released
Individual-level controls		
Breadth of skills_{t-1}	Number of unique software categories listed on projects the individual is involved in	Individual with more breadth in their skills are likely to enjoy some of the vision advantage project-level brokers enjoy. With this control we ensure that effect of project-level brokerage are not linked to individual characteristics instead of their structural position
Number of projects joined_{t-1}	Number of projects joined by the individual in the previous month	Controls for the possibility that individuals might choose projects that go to release by simply joining many projects
Bug solving and posting_{t-1}	Number of actions (code posted, messages) on the bug tracking systems of all the softwares the individual is involved with over the previous 6 months	Measure technical activity of the individual. Proxy for the technical experience of the individual. This experience is likely to shape how individual evaluate projects
Message posting_{t-1}	Number of messages posted on forums of all the software the individual is involved with over the previous 6 months	Proxy for the "political" experience of the individual and his/her participation in general discussions when involved on a project
Bug solving and posting on project	Number of actions by the individual on the bug tracking system of the project	Controls for the activity of the person on the project
Message posting on project	Number of messages on the forums of the project posted by the individual	Controls for the activity of the person on the project
Tenure	Number of days since the individual registered on the platform	Controls for the experience of the individual with the platform
Timezone	Whether or not the individual is in the US	US developers are overrepresented and the platform is run from the US, this controls for potential benefits of proximity to the core of developers
Number of project founded_{t-1}	Number of projects founded by the individual in the previous 6 months	This controls for effect of an entrepreneurial orientation of the individual and the potential effect that might have on the ability to evaluate projects
Number of administrator roles_{t-1}	Number of administrator roles of the individual in the previous 6 months	Controls for leadership experience on projects
Past successes	Number of affiliations to young projects that went to release in the six months prior to the observation period	Controls for the influence of past success on future success

Table 2: Correlation between independent variables and controls

	Mean	S.d	1	2	3	4	5	6	7	8	9
1. past success	0.03	0.21									
2. Individual closure	0.00	0.04	0.05***								
3. Project brokerage	-0.00	0.31	0.20***	0.02***							
4. Number of projects joined (t-1)	0.01	0.14	0.24***	0.01*	0.13***						
5. Bug solving and posting (previous)	2.28	48.55	0.07***	0.06***	0.05***	0.02***					
6. Message posting (previous)	0.65	11.26	0.03***	0.01**	0.05***	0.00	0.13***				
7. Bug solving and posting on project	2.08	46.42	0.00	0.00	0.00	0.00	0.00	0.01			
8. Messsage posting on project	0.51	10.27	0.00	0.00	0.00	0.00	0.00	0.00	0.03***		
9. Tenure	308.62	582.35	0.05***	0.09***	0.37***	0.01	0.05***	0.05***	0.00	0.00	
10. Number of project founded	0.85	1.02	0.28***	0.02***	0.21***	0.15***	0.04***	0.00	0.00	0.01*	0.07***
11. Number of administrator roles	1.20	1.45	0.32***	0.10***	0.38***	0.14***	0.12***	0.04***	0.00	0.01	0.25***
12. Breadth of project	1.61	1.12	0.00	-0.01	-0.01*	0.00	-0.02***	-0.01	0.01*	0.01**	0.02***
13. Number of contributors (t-1)	2.60	4.08	-0.03***	0.00	-0.06***	-0.01	-0.01**	0.00	0.00	0.00	-0.06***
14. Past successes	0.05	0.30	0.56***	0.04***	0.20***	0.15***	0.06***	0.03***	0.00	0.00	0.04***
15. Breadth of skills	1.51	3.75	0.24***	0.22***	0.60***	0.13***	0.16***	0.11***	0.00	0.00	0.41***
16. Maturity	2359.01	2383.72	0.00	-0.01*	-0.02***	-0.01*	-0.01**	-0.01**	0.00	0.00	-0.02***
17. Change in maturity	0.25	0.14	0.01	-0.01*	-0.03***	-0.02***	-0.02***	-0.02***	0.00	0.01*	0.03***
18. Closure dummy	0.92	0.27	-0.14***	-0.32***	-0.48***	-0.08***	-0.10***	-0.08***	0.00	0.01	-0.32***
19. Brokerage dummy	0.71	0.46	-0.23***	-0.14***	-0.57***	-0.15***	-0.07***	-0.07***	0.01	0.00	-0.46***
20. Small world ratio (centered)	-0.00	3.15	0.22***	0.14***	0.57***	0.14***	0.07***	0.07***	-0.01	0.00	0.47***

	10	11	12	13	14	15	16	17	18	19
11. Number of administrator roles	0.71***									
12. Breadth of project	-0.04***	-0.02***								
13. Number of contributors (t-1)	-0.26***	-0.19***	0.12***							
14. Past successes	0.28***	0.33***	0.02***	0.01						
15. Breadth of skills	0.16***	0.40***	0.00	-0.04***	0.23***					
16. Maturity	0.01*	0.02***	-0.04***	-0.02***	0.01	-0.04***				
17. Change in maturity	0.00	0.03***	0.07***	0.04***	0.00	-0.04***	-0.09***			
18. No closure	-0.06***	-0.26***	0.02***	0.01	-0.15***	-0.64***	0.04***	0.04***		
19. No brokerage	-0.20***	-0.37***	0.01**	0.05***	-0.21***	-0.62***	0.03***	0.03***	0.45***	
20. Small world ratio (centered)	0.19***	0.36***	-0.01**	-0.05***	0.20***	0.60***	-0.02***	-0.03***	-0.44***	-0.98***

Table 3: Results from mixed effects logistic regression on release

	Controls	Main effects (no brokerage)	Main effects (Small world)	Interactions
Constant	-2.5658***	-1.7852***	-1.9999***	-1.9211***
	(0.0962)	(0.3034)	(0.3018)	(0.3048)
Field maturity	-0.0382***	-0.0393***	-0.0400***	-0.0392***
	(0.0127)	(0.0127)	(0.0127)	(0.0127)
Dynamism of the field	0.1572*	0.1620*	0.1597*	0.1598*
	(0.0912)	(0.0913)	(0.0912)	(0.0913)
Breadth of project	0.2165***	0.2178***	0.2177***	0.2173***
	(0.0103)	(0.0103)	(0.0103)	(0.0103)
Number of contributors (t-1)	-0.0103***	-0.0112***	-0.0111***	-0.0112***
	(0.0033)	(0.0034)	(0.0033)	(0.0034)
Breadth of skills	0.0202***	0.0034	0.0060	0.0054
	(0.0043)	(0.0057)	(0.0056)	(0.0057)
Number of projects joined (t-1)	-0.1298	-0.1811**	-0.1726**	-0.1746**
	(0.0873)	(0.0872)	(0.0872)	(0.0871)
Bug solving and posting (previous)	0.0288	0.0247	0.0270	0.0247
	(0.0227)	(0.0227)	(0.0227)	(0.0227)
Message posting (previous)	0.0103	0.0085	0.0095	0.0103
	(0.0314)	(0.0315)	(0.0314)	(0.0315)
Bug solving and posting on project	0.6580***	0.6602***	0.6598***	0.6593***
	(0.0201)	(0.0201)	(0.0201)	(0.0201)
Messsage posting on project	0.9689***	0.9748***	0.9732***	0.9736***
	(0.0371)	(0.0372)	(0.0371)	(0.0372)
Tenure	0.0185***	-0.0011	0.0014	0.0009
	(0.0055)	(0.0061)	(0.0061)	(0.0061)
Time Zone	-0.0482	-0.0470	-0.0471	-0.0478
	(0.0314)	(0.0314)	(0.0314)	(0.0314)
Number of project founded	-0.0634***	-0.0768***	-0.0745***	-0.0768***
	(0.0202)	(0.0205)	(0.0205)	(0.0205)
Number of administrator roles	0.0239*	0.0181	0.0182	0.0182
	(0.0131)	(0.0132)	(0.0132)	(0.0133)
Past successes	0.1992***	0.1628***	0.1720***	0.1650***
	(0.0594)	(0.0588)	(0.0589)	(0.0588)
Independent variables				
Between project brokerage		0.0685	0.0808	1.2939***
		(0.0527)	(0.0527)	(0.3019)
Between individual brokerage		-0.5815*	-0.5885*	-1.9005***
		(0.3185)	(0.3183)	(0.5452)
No between project brokerage		-0.2951***		
		(0.0391)		
No between individual brokerage		-0.4947	-0.5032	-0.5306*
		(0.3079)	(0.3077)	(0.3075)
Small world ratio			0.0354***	0.0273***
			(0.0056)	(0.0086)
Small world ratio x between project brokerage				-0.2478***
				(0.0610)
Small world ratio x between individual brokerage				0.2635***
				(0.0913)
License dummies	Included	Included	Included	Included
Month random intercept	Included	Included	Included	Included
Individual random intercept	Included	Included	Included	Included
AIC	50342	50284	50300	50287
BIC	50606	50583	50599	50603
Log Likelihood	-25141	-25108	-25116	-25108
Deviance	50282	50216	50232	50215
N	48423	48423	48423	48423

*** p < 0.01, ** p < 0.05, * p < 0.1

FIGURES

Figure 1: 4-path in a 2-mode network and corresponding triads in a 1-mode network

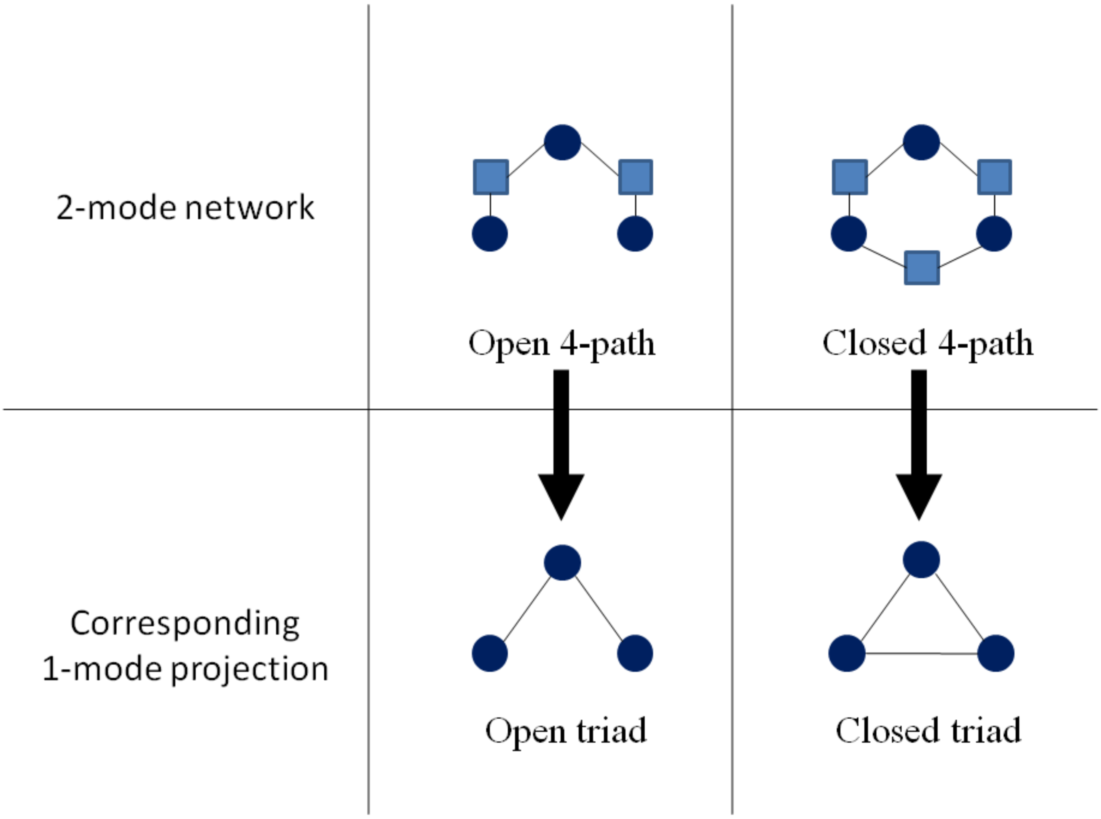

Figure 2: Interaction between small world ratio and between project brokerage and 90% confidence intervals

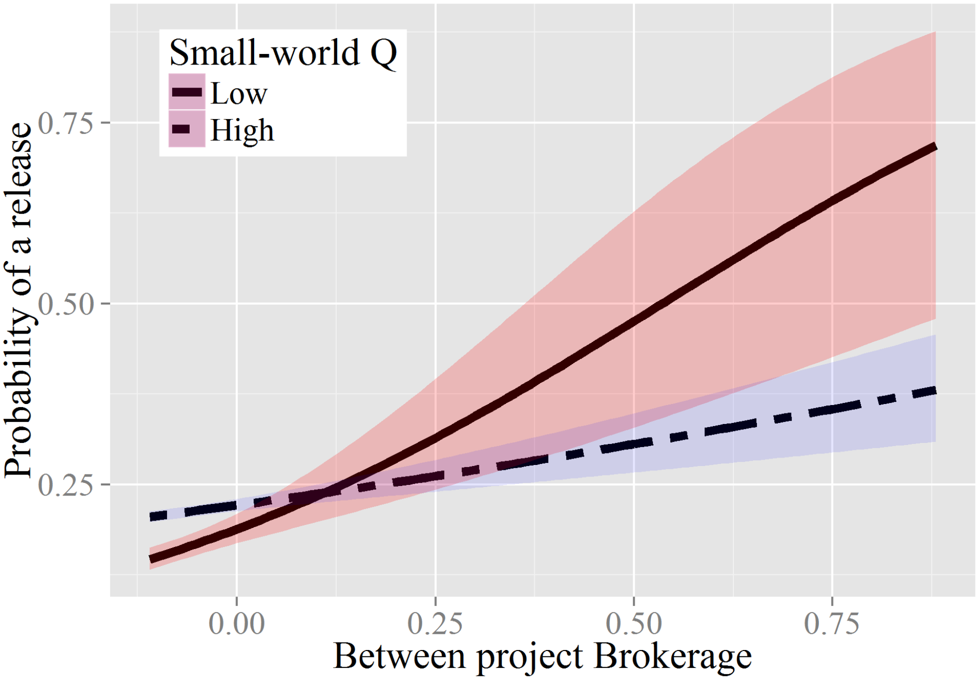

Figure 3: Interaction between small world ratio and between individual brokerage and 90% confidence intervals