Networks, social influence and the choice among competing innovations: Insights from open source software licenses

Param Vir Singh
psidhu@cmu.edu
David A. Tepper School of Business
Carnegie Mellon University,
Pittsburgh, PA 15213

Corey Phelps
phelps@hec.fr
HEC Paris
78351 Jouy-en-Josas
15213 France

Version: 19-June, 2012

Both authors contributed equally to this paper.

Forthcoming, Information Systems Research.

* We would like to thank Henrich Greve and Sonali Shah for comments on an earlier draft of this manuscript.
Networks, social influence and the choice among competing innovations: Insights from open source software licenses

Abstract

Existing research provides little insight into how social influence affects the adoption and diffusion of competing innovative artifacts and how the experiences of organizational members who have worked with particular innovations in their previous employers affect their current organizations’ adoption decision. We adapt and extend the heterogeneous diffusion model from sociology and examine the conditions under which prior adopters of competing OSS licenses socially influence how a new OSS project chooses among such licenses and how the experiences of the project manager of a new OSS project with particular licenses affects its susceptibility to this social influence. We test our predictions using a sample of 5,307 open source projects hosted at SourceForge. Our results suggest the most important factor determining a new project’s license choice is the type of license chosen by existing projects that are socially closer to it in its inter-project social network. Moreover, we find that prior adopters of a particular license are more infectious in their influence on the license choice of a new project as their size and performance rankings increase. We also find that managers of new projects who have been members of more successful prior OSS projects and who have greater depth and diversity of experience in the OSS community are less susceptible to social influence. Finally, we find a project manager is more likely to adopt a particular license type when his or her project occupies a similar social role as other projects that have adopted the same license. These results have implications for research on innovation adoption and diffusion, open source software licensing, and the governance of economic exchange.

Key words: open source software license, social networks, innovation adoption and diffusion, social influence.
1. Introduction

The adoption of innovative artifacts – such as new ideas, products, technologies, processes and practices – by members of a social system and their diffusion through such systems has fascinated scholars from a variety of disciplines and research traditions for decades (Rogers 2003, Phelps et al. 2012). Information Systems (IS) scholars, in particular, have sought to explain the adoption and diffusion of a variety of information technologies (IT) and IT-related practices (e.g. Hardgrave et al. 2003; Jeyaraj et al. 2006; Susarla et al. 2012). While the vast majority of IT adoption research and much of the broader adoption and diffusion literature employs a rational decision-theoretic framework and models adoption as a straightforward cost-benefit analysis (Fichman 2004; Hall 2004; Rogers 2003), substantial research shows adoption decisions often involve profound uncertainty about costs and benefits (Rogers 2003). Facing such uncertainty, potential adopters typically turn to prior adopters as socially influential referents for guidance in determining the appropriate adoption choice (e.g. DiMaggio and Powell 1983).

An emerging theoretical and estimation framework that incorporates spatial and temporal heterogeneity in modeling the social influence of prior adopters is the heterogeneous diffusion model (Strang and Tuma 1993). This framework decomposes social influence in terms of three classes of factors: the susceptibility of a potential adopter to social information about the innovation, the infectiousness of information about the innovation from prior adopters, and the social proximity between prior and potential adopters (Greve 2005). Although research has used this framework to explain, inter alia, the adoption and diffusion of particular corporate governance practices (Davis and Greve 1997), prescription drugs (Strang and Tuma 1993), and business strategies (Greve 1998), its use by IS scholars is rare, despite calls for doing so (Fichman 2004). This research suggests the extent to which prior adopters socially influence potential adopters depends on the number and characteristics of prior adopters as sources of information about an innovation, characteristics of potential adopters as recipients and interpreters of this information and the social proximity between them. Social influence also varies over time as greater or fewer actors adopt the innovation and become more or less socially proximate to potential adopters (Greve 2005).

Despite its contributions to an understanding of how and why innovative artifacts are adopted and diffused, research employing a heterogeneous diffusion framework, as well as the broader innovation adoption and diffusion literature, is limited in two important respects. First, the vast majority of research restricts its attention to a single innovation (Strang and Soule 1998). However, individuals and organizations are often
confronted with multiple, substitutable innovations that compete for adoption. This menu of options increases the complexity of the adoption decision and is therefore likely to increase the uncertainty a potential adopter faces, potentially amplifying the role social influence plays in adoption and diffusion. Extant research provides little insight into how, when or why social influence affects the adoption and diffusion of competing artifacts (Strang and Soule 1998). Moreover, the heterogeneous diffusion model has yet to be applied to explain the adoption and diffusion of multiple, competing innovations.

Second, existing adoption and diffusion research typically assumes that potential organizational adopters are exposed to an innovation through mass media and direct and indirect social ties to prior adopters rather than direct experience with the innovation. However, because organizations are collections of individuals, who often move between organizations as they change jobs, employees can gain experience with and knowledge about particular innovations adopted by their current organization and transfer this experience to other organizations when they move (Song et al. 2003). As a result, the nature of the experience employees have had with innovations adopted by their previous employers may influence how their new organizations evaluate the innovations, making them more or less sensitive to the social influence of prior adopters. Although interfirm employee mobility is a primary mechanism by which organizations learn from and influence one another (Song et al. 2003), research does not consider when or how the experiences of organizational members who have worked with particular innovations in their previous employers affect their current organizations’ adoption of such innovations. Little research has examined the characteristics of potential organizational adopters, such as their employees’ prior exposure to particular innovations, which affect their sensitivity to the choices of prior adopters (Wejnert 2002).

This study seeks to address these important limitations of innovation adoption research employing a heterogeneous diffusion framework and the broader adoption and diffusion literature. We do so by investigating the conditions under which prior adopters of competing innovations socially influence how a potential adopter chooses among such innovations and how the experiences of members of a potential organizational adopter with particular competing innovations affects its susceptibility to this social influence. Of particular importance to IS research, we investigate this question by studying the adoption of particular types of open source software (OSS) licenses by new open source projects.

The OSS context is an ideal setting to investigate our research question for several reasons. First, the OSS licensing framework was an innovative departure from previous legal mechanisms to promote
cooperation and beneficial exchange among actors (i.e., software developers and users) with divergent incentives (Demil and Lecocq 2006). Within the broad OSS framework, there is a variety of specific licenses. Because managers of new OSS projects choose a license at the inception of the project, specific types of OSS licenses represent discrete, competing innovative licensing practices that are at risk of being adopted by new OSS projects. Second, project managers face substantial uncertainty in choosing an appropriate license because of the novelty and large number of licenses (Rosenberg 2000) and the challenge of predicting how their choice of a particular license will affect developers’ incentives to join their projects (Shah 2006). This uncertainty is exacerbated by the fact that the choice of an OSS license is a one-shot, largely irreversible decision made by licensors who are typically software developers with little or no legal expertise (McGowan 2001). Given this uncertainty, managers of new OSS projects are likely to consider existing projects as social referents that guide their license choice (DiMaggio and Powell 1983). Third, new OSS projects are often initiated by developers who have previously worked on other OSS projects (Hahn et al. 2008), allowing them to gain experience with the licenses used by these projects. These developers carry this experience with them to their new projects, which may influence how susceptible their new organizations are to the social influence of the license choices of extant projects. Moreover, in the OSS setting the project manager is the sole member of the project when the license choice is made, allowing for a clear, direct link between employee prior experience and organizational adoption choice. A final reason to study OSS license choice is that while research shows the choice of license by OSS projects can affect their performance (Comino et al. 2007; Stewart et al. 2006), OSS licensing research typically examines the choice between open and closed source licenses rather than the choice among OSS licenses. Only two studies have examined the determinants of OSS project license choice (Lerner and Tirole 2005b; Sen et al. 2009). Both studies, however, ignore the potential social influence of prior adopters of particular licenses on a project manager’s license choice and how this influence may vary over time and by a project manager’s social proximity to established OSS projects, characteristics of these projects and the manager’s previous OSS experience. This is a surprising and substantive limitation of this research given the substantial uncertainty managers of new OSS projects face in choosing a license, which suggests they may be socially influenced by the license choices of existing projects.

We adapt and extend the heterogeneous diffusion model to accommodate multiple, competing innovations. To do so, we investigate each dimension of this model – infectiousness, social proximity and susceptibility – and derive predictions related to each dimension. We incorporate a novel source of influence
on a potential adopter’s susceptibility to the social influence of prior adopters – namely, the experiences of organizational members who have worked with particular innovations in other organizations.

We test our predictions in a sample of 5,307 OSS projects hosted at SourceForge. After controlling for factors shown to affect OSS license choice (Lerner and Tirole 2005b), our results suggest the most important factor determining a new project’s license choice is the type of license chosen by existing projects that are socially closer to it in its inter-project social network. Moreover, we find that prior adopters of a particular license are more infectious in their influence on the license choice of a new project as their size and performance rankings increase. We also find that managers of new projects who have been members of more successful prior OSS projects and who have greater depth and diversity of experience in the OSS community are less susceptible to social influence. Finally, we find a project manager is more likely to adopt a particular license type when his or her project occupies a similar social role as other projects that have adopted the same license.

This study contributes to the innovation adoption and diffusion literature by addressing important limitations of the heterogeneous diffusion framework and to the literature on open source software by being the first study to explore when and how social influence from existing OSS projects affects a new project’s license choice. This study also has substantive implications for understanding the origins and influence of the social institutions that govern economic exchange.

2. Open Source Software

2.1 Open Source Software Development Process

All OSS projects follow a similar process. An “initiating developer” begins a project by working on an idea and then hosts the source code and invites other developers to contribute. Developers volunteer to perform specific tasks and collaborate as a team, incorporating their individual creations into a single body of source code. Once an executable version of the software is developed, it is released for testing and feedback. The software evolves as new features are added, existing features are modified, and bugs get fixed. The process involves the sharing of ideas and joint problem-solving that fosters social bonds among collaborators. Given the small number of developers typically involved in OSS projects (Krishnamurthy 2002) and the frequency and intensity of their interactions over time, the social ties among them can be quite strong (Hahn et al. 2008; Singh et al. 2011a, Singh and Tan 2010, Singh 2010). Because OSS projects stimulate the formation of social ties among teams of developers and because developers often work on multiple projects, a social network is
produced that directly and indirectly connects developers participating in the broader OSS community. While projects create ties among developers, projects also become connected by sharing common developers (Grewal et al. 2006). We examine the influence of this latter inter-project social network in this study.

2.2 Open Source Software Licenses

To be characterized as “open source,” software must be offered under a license that satisfies several conditions. Both the Free Software Foundation and the Open Source Initiative approve OSS licenses. OSS licenses differ in the extent to which they restrict how users may use and modify the software. At one extreme are highly restrictive licenses, such as the GNU general public license (GPL), and at the other extreme are permissive licenses, such as the Berkeley Software Distribution (BSD) license. Highly restrictive licenses differ from permissive licenses in two key ways (de Laat 2005):

1. They require that, when modified versions of the program are distributed, the source code must be made generally available. This provision is called the “copyleft” clause.

2. They prohibit the software to be mingled with other software that does not use the same license. This provision is called the “viral” or the “reciprocal” clause.

The highly restrictive licenses were the first free software licenses. The most famous of these, GPL, was authored by Richard Stallman, an early proponent of OSS and the initiating developer of a free operating system, GNU. The “copyleft” and “viral” clauses were designed to protect the software from being hijacked by proprietary software developers. These clauses require that any modification or derivation of the software has to be offered under GPL, making GPL'd software less attractive to commercial actors. The “viral” clause restricts the software from exploiting complementarities with other software, reducing its appeal to both contributors and users. While the BSD license makes software attractive for commercial use because it allows modified or derivative works to be kept private, it does not protect the software from being hijacked. Stallman authored the Lesser General Public License (LGPL) as a compromise between the highly restrictive GPL and the permissive BSD license (Stallman et al. 2002). The LGPL includes the copyleft but not the viral provision. (Lerner and Tirole 2005b) refer to LGPL-type licenses as restrictive. Any OSS license can be categorized as permissive, restrictive, or highly restrictive.

---

3. Prior Research on the Choice of Open Source License

Few studies have examined the determinants of OSS project license choice (Lerner and Tirole 2005b; Sen et al. 2009). Lerner and Tirole (2005b) model the licensor’s problem as an optimizing balancing act between choosing a more restrictive license to attract more contributing developers and adopting a permissive license to preserve her ability to commercialize the software. Conditional on the exogenous characteristics of the software development project, the manager chooses the license that maximizes her expected benefits from the project given her evaluation of the expected response by potential developers. Consistent with their expectations, Lerner and Tirole (2005b) found that characteristics of an OSS project; such as its intended audience (e.g., end users), application genre (e.g., gaming), operating system (e.g., Linux), and user interface (e.g., GUI); determine its choice of license (Lerner and Tirole 2005b). In contrast to the focus on project characteristics, Sen et al. (2008) examine how the choice of OSS license type can be explained by the intrinsic and extrinsic motivations and the attitudes of the project manager. They showed that project managers who were motivated by the problem-solving challenges of OSS projects preferred moderately restrictive licenses, while those motivated by peer recognition preferred unrestrictive licenses. They also found that, when choosing a license, project managers are more concerned with the ideological principle that all OSS should be able to be freely redistributed than they are with end-users’ rights and that licensors prefer licenses that are aligned with these attitudes.

Despite the insights this research provides into understanding OSS license choice, neither study considers the potential social influence of prior adopters of particular licenses on a new project’s license choice. This represents an important limitation of this research because a manager of a new OSS project is likely to be socially influenced by existing projects’ licensing choices given the substantial uncertainty he or she faces in choosing a license (as described above). Indeed, we visited many on-line forums in which developers queried one another about which license to choose for their projects and observed substantial confusion and uncertainty on this topic. The two postings below are representative of what we observed:

The number of licenses is INSANE, I may as well just write my own and add it to the list (joking...mostly) I have a few open source apps and I think they are probably not licensed as to what I want, but I can’t seem to find a good source to get a suggestion, and I don’t want to read all of them2.

I wish there was just a spreadsheet of "This license has this feature". I’m on the verge of releasing something and it’s very hard to

---

To explore the role social influence plays in the choice of open source license, we build on and extend the heterogeneous diffusion model (Strang and Tuma 1993).

4. Heterogeneous Diffusion Model

The heterogeneous diffusion model was developed to explain how social context influences an actor’s adoption of an innovation and its diffusion within a population of actors (Strang and Tuma 1993). This model allows for spatial heterogeneity in the influence of prior adopters on individual adoption behavior and accommodates temporal heterogeneity by allowing the influence of prior adopters and social proximity to vary over time (Strang and Tuma 1993). In this model, the extent to which an actor’s social context influences its adoption behavior depends on three categories of specific explanatory factors: infectiousness, social proximity and susceptibility (Strang and Tuma 1993). The infectiousness of a prior adopter of an innovation refers to how influential information generated about its actions is for potential adopters, which is a function of characteristics of the prior adopter such as its size, performance or status (Greve 2005). The social proximity of the source to a potential adopter describes how easily information is transmitted between them, based on their social distance from each other (Greve 2005). The susceptibility of a potential adopter refers to how open, receptive or sensitive it is to being influenced by information available about the innovation and depends on inherent attributes that affect its motivation and ability to adopt the particular innovation (Greve 2005). The mechanism by which prior adopters influence potential adopters is the transmission of information between them, either through direct and indirect social ties, or through potential adopters observing prior adopters (Strang and Soule 1998). The model also specifies that an actor has intrinsic propensities to adopt an innovation based on its own inherent characteristics, independent of social context. We consider intrinsic characteristics affecting adoption as control variables that provide a baseline model.

Because the OSS licensing framework represents a novel set of contractual practices for protecting intellectual property (Demil and Lecocq 2006), we characterize types of OSS licenses as discrete and substitutable innovative practices that are at risk of being adopted by new OSS projects. We focus on the adoption of specific categories of licenses rather than individual OSS licenses. While there are over 40 OSS licenses, they are small variations on two underlying themes — whether they contain copyleft and/or viral clauses (de Laat 2005; Lerner and Tirole 2005b; Stewart et al. 2006) — and thus represent three categories

---

(types) of OSS licenses: unrestrictive (neither clause is present), restrictive (only the copyleft clause is present), and highly restrictive (both clauses are present) (Lerner and Tirole 2005b). Because these categories capture the essential differences among OSS licenses (de Laat 2005), each type represents a prototypical OSS license. Given the uncertainty surrounding these complex artifacts, potential licensors will tend to simplify their licensing decisions by focusing on prototypical licenses because these represent the essential and salient differences among the many individual licenses (Kahneman et al. 1982). Indeed, online tools (such as OSS Watch and Three.org) and articles (e.g., Niiranen 2009) designed to help developers choose OSS licenses typically explain and prescribe the licenses in terms of these prototypical characteristics rather than peripheral features. When an innovative practice is complex and uncertain, its prototypical characteristics rather than peripheral features drive its adoption (Ansari et al. 2010). Next, we define the type of social network in the OSS setting through which we expect social influence to operate. We then develop hypotheses linking variables associated with each dimension of the heterogeneous diffusion model to the likelihood a new OSS project will adopt a particular license type.

4.1. Inter-Project Social Network

The nature of the development process in OSS leads to the emergence of affiliation networks (Grewal et al. 2006). An affiliation network is a two-mode network because it consists of actors connected by their participation in common events and events that are connected by common actors (Wasserman and Faust 1994). An affiliation network therefore represents two different types of one-mode networks: an inter-actor network and an inter-event network (Wasserman and Faust 1994). In the OSS setting, the actors are individual developers and the events are projects. Developers have social ties with one another as a result of working together on the same project and projects are connected to one another as a result of sharing common developers (Grewal et al. 2006). Rather than focus on the inter-developer social networks that result from OSS affiliation networks, we focus on inter-project networks as the appropriate social network. An OSS project administrator initiates a project and is the sole developer on the project at the time of project registration and license choice. Thus, at the time of license choice, the project and administrator (i.e., licensor) are one and the same. Our theory suggests an administrator is influenced by his/her social proximity to, and characteristics of, other projects (rather than individual developers). Thus, an administrator’s relevant social network at the time of license choice consists of the social ties the administrator-as-project has to previously established OSS projects.
4.2 Social Proximity

Social proximity refers to the social distance between two actors in a social network and determines how easily information is transmitted between them and the relevance of this information (Coleman et al. 1966). A socially proximate actor provides an influential frame of reference by which a focal actor evaluates and interprets information (Leenders 2002). Two approaches to conceptualizing social proximity in a network exist, each with its own causal mechanism linking proximity with social influence. The first approach – social cohesion – defines proximity in terms of the number, length and strength of the paths that connect actors in a network (Marsden and Friedkin 1993). The second approach – equivalence – defines proximity in terms of the similarity of two actors’ profiles of network relations (Marsden and Friedkin 1993).

4.2.1 Social Cohesion. The social cohesion approach defines social proximity in terms of the number, length and strength of the paths that connect actors in a network (Marsden and Friedkin 1993). We focus on social distance (i.e., path length) as the primary dimension of social cohesion. The simplest form of cohesion is when two actors, such as a potential and prior adopter of an innovation, share a direct social tie. Directly connected actors communicate and share information with each other more frequently and with greater fidelity than indirectly connected actors (Burt 1982). Direct ties are conduits for the communication of rich, personalized information, which tends to be more influential than impersonal information sources (Rogers and Kincaid 1981). The volume and fidelity of information decays as the number of links indirectly connecting actors increases (Shannon 1949), making indirectly connected actors less socially influential on a potential adopter than direct contacts (Burt 1982). Research in social psychology suggests involvement in shared activities provides opportunities for social cohesion to develop and that shared attitudes develop from social cohesion (Homans 1961). Faced with an uncertain situation, such as the adoption of an innovation, individuals discuss it with their proximate peers and develop a consensual normative understanding of the associated costs and benefits (Rogers 2003). Social ties provide detailed, personalized and more persuasive information on costs and benefits of adoption than general information sources (Rogers 2003). Discussions with prior adopters of an innovation build social pressures on the potential adopter to adopt the innovation when faced with an opportunity to do so (Rogers and Kincaid 1981). Social pressure increases with social cohesion and hence a potential adopter is more likely to adopt an innovation that has been adopted by his or her most proximate peers. Research has found that shared attitudes and behavior develop among people or organizations that are connected through direct communication channels (Coleman et al. 1966; Davis and Greve 1997; Haunschild
In the OSS context, projects on which a licensor has worked in the past provide greater opportunities for communication and thus for social cohesion. Through her discussions with developers on prior projects, a focal licensor develops a shared understanding of the costs and benefits associated with the license type chosen for those projects. While direct involvement in a project provides opportunities to observe the consequences of adopting a license, a licensor may also receive useful and persuasive information from projects with which she is not involved but has social ties with developers who are. These ties can provide the licensor with detailed, personalized and persuasive information on the costs and benefits of the particular license types adopted by these projects. Prior adopters have experience with a particular license type and thus understand it better than a potential adopter and may communicate their preferences persuasively via social ties with the focal project administrator. Hence, socially cohesive prior adopters exert social pressure on a potential licensor to adopt the same license type.

HYPOTHESIS 1: *A licensor is more likely to choose a license type that was adopted by other projects to which he or she is more closely socially connected (i.e., socially cohesive).*

4.2.2 Role Equivalence. An alternative conceptualization defines social proximity in terms of equivalence—the similarity of two actors’ profiles of network relations (Marsden and Friedkin 1993). The social network literature initially conceptualized equivalence as structural equivalence. Two actors are structurally equivalent to the extent they have ties to the exact same other actors (Burt 1976). Research findings on the influence of structural equivalence on actor behavior can, however, be interpreted in two ways (Mizruchi 1993). First, competition between actors over the same resources provided by the same alters triggers imitation between socially substitutable actors (Burt 1987). Second, social cohesion due to direct ties with the same set of third-parties induces similar behaviors.

Because of these alternative interpretations, we employ the concept of role equivalence to capture how competition among actors, such as prior and potential adopters, results in imitative behavior. Two actors are role equivalent to the extent they engage in the same kinds of relationships with third parties (Mizruchi 1993). For example, two organizations are more role equivalent when they produce or trade in the same products, and therefore have similar types of upstream and downstream relationships, although not necessarily with the same suppliers or buyers (Guler et al. 2002; Winship and Mandel 1983). In this vein, equivalence among organizations has been defined as the extent to which they produce similar products or produce products using similar technologies and therefore have similar kinds of vertical relationships (Bothner 2003; Davis and Greve
Role equivalence captures the degree to which two actors occupy similar social roles and thus serve as common referents for one another, regardless of whether or not they share direct ties or are structurally equivalent (Mizruchi 1993).

Because role equivalence does not depend on the presence (or absence) of a tie between the actors being compared, the social influence effect of role equivalence on actor behavior is different from the influence associated with direct ties. The causal mechanism linking equivalence and the similarity of actors’ behavior is social influence via a process of observational social comparison. Although role equivalent actors do not necessarily share relations with the same third parties, they compete with one another to retain their existing ties because third parties view such actors as substitutable objects of interaction (Guler et al. 2002). Competition among two equivalent actors increases their incentives to monitor and compare behaviors to ensure neither has an advantage or falls behind (Burt 1987). The more equivalent, the more one actor is likely to adopt an innovation previously adopted by another because it may make the other more attractive as the object or subject of relationships, resulting in some alters abandoning the non-adopter in favor of the actor that adopted (Guler et al. 2002). Non-adopters monitor and imitate the adoption behavior of role equivalent adopters in order to retain existing relationships and the benefits they provide (Guler et al. 2002). Actors look to equivalent others to identify appropriate behavior, particularly in contexts characterized by substantial uncertainty about how to behave (Burt 1987). The equivalence model of social proximity highlights symbolic communication among social substitutes rather than direct communication among contacts, which a social cohesion perspective emphasizes (Leenders 2002). This argument is consistent with the emphasis in neo-institutional research on mimetic isomorphic processes within industries (DiMaggio and Powell 1983). Prior research shows the extent to which organizations produce similar, competitive products or employ the same technologies are more likely to imitate each other’s adoptions of innovations (Bothner 2003; Davis and Greve 1997; Flingstein 1985).

Based on prior research (Guler et al. 2002; Winship and Mandel 1983), we define role equivalence as the extent to which OSS projects engage developers and users in the same technologies. As such, OSS projects are role equivalent to the extent they employ the same technology platform in their development. Indeed, the primary data source we use in this study, SourceForge, organizes all OSS projects it hosts into common domains or “foundries” based on their common usage of a technology platform, such as projects that use the Perl programming language. Projects in the same foundry typically target and compete for similar users and for developers with similar skills (Hahn et al. 2008). The extent to which OSS projects compete to maintain relation-
ships with users and developers increases the extent to which the projects are social substitutes (from the perspective of users and developers). This increasing competition should increase the incentives OSS projects have to monitor and imitate each other’s adoption of practices that can influence the performance of their respective projects, such as the choice of license. Because license choice affects OSS project performance and such choice involves considerable uncertainty, a licensor of a new project should be particularly sensitive to the license(s) previously adopted by equivalent projects. Compatibility issues associated with an OSS project’s choice of license may increase the tendency to mimic the previous licensing decisions of equivalent projects. Complementarities often exist among software products and the exploitation of these complementarities influences a product’s success (Gallaugher and Wang 2002). The license that governs the development and distribution of a software product influences its ability to exploit complementarities with other software products because such licenses influence the extent to which products are legally compatible and can be mingled and mixed by developers and users (Lerner and Tirole 2002). Licensors may fear the opportunity costs of choosing an inappropriate license – one that reduces its compatibility with other projects – and thus look for guidance on the appropriate choice from equivalent projects. In sum, a licensor will be more likely to monitor and imitate the license choices of projects in the same foundry.

**HYPOTHESIS 2:** A licensor is more likely to choose a license type that role equivalent projects have adopted more widely.

### 4.3 Infectiousness of Prior Adopters

The infectiousness of a prior adopter of an innovation describes how influential the information about its actions is for the adoption decision of a potential adopter. Because infectiousness is a property of information generated by prior adopters, its influence depends on whether or not it can reach a potential adopter and the distance it must travel to do so (Strang and Tuma 1993). In other words, the influence of infectiousness depends on the social proximity between a prior and potential adopter. Prior research suggests prior adopters can be more or less influential as social referents for a subsequent potential adopter based on characteristics such as their size, performance or status (Greve 2005). We focus on prior adopter performance as the primary driver of its infectiousness. License adoptions by higher performing OSS projects will be more influential on the choice of license type by new projects for several reasons. First, successful projects attract more attention, which leads to more information being available to potential adopters about them (Greve 2005). Second,

---

4 Although we focus on a prior adopter’s performance as the primary construct affecting infectiousness, we argue that OSS project performance determines its status within the OSS community (Stewart, 2005) and we use the number of developers a project attains (i.e., its size) as a measure of project performance.
potential adopters often attribute the level of success an actor subsequently achieves to their choice of innovation (Greve 2005), which suggests potential OSS license adopters may attribute a prior adopter's success, in part, to its choice of license, making information about such adopters more influential on the adoption decisions of others. However, such attributions may be unnecessary because people often copy the behavior of successful or prestigious referents, even when the focal behavior has nothing to do with the referent’s success or prestige (Henrich and Gil-White 2001). Third, while the adoption of an OSS license is characterized by substantial uncertainty, innovations adopted by high status organizations are viewed as less uncertain and therefore more likely to be imitated by others (DiMaggio and Powell 1983). In this way, an actor’s status provides an uncertainty-reducing signal about its underlying quality and competence, increasing the extent to which others attend to and are influenced by its actions (Podolny 2001). Such status signals are particularly influential when recipients find it difficult to search for other useful information sources to inform their adoption decision (Simcoe and Waguespack 2011). In the OSS community, the success of a project reflects its social status in the community (Stewart 2005). Thus, the adoption of a particular license type by successful projects is likely to be imitated by new projects. Because the social proximity of established projects determines the extent to which they will be used as social referents by a focal project, we expect the license types chosen by socially proximate projects will have a stronger effect on a focal project’s license choice when these projects are more successful.

**HYPOTHESIS 3:** A licensor is more likely to choose a license type that has been adopted by successful projects that are socially proximate.

### 4.4 Susceptibility of Licensor

Susceptibility refers to how much a potential adopter is affected by information about the practices adopted by others (Greve 2005). Holding constant the information about prior adopters that reaches potential adopters, a potential adopter is more susceptible when his/her adoption decision is more sensitive to and influenced by this information (Greve 2005). Little research has examined the characteristics of potential organizational adopters that affect their sensitivity to the choices of prior adopters (Wejnert 2002). In particular, research does not consider when or how the experiences of organizational members who have worked with particular innovations in other organizations affect their current organizations’ adoption of such innovations. Potential adopters differ in their susceptibility to the social influence of prior adopters based on differences in a potential adopter’s motivations to search for new practices and learn from the actions of others (Greve 1998). In addressing an important limitation of extant innovation adoption research, we argue that these motivational
differences will be affected by the nature of the experience key organizational members (i.e., the project manager) have had with the innovations in their work on other projects that had previously adopted.

The success and failure of other OSS projects on which the licensor of a new project has worked will influence her susceptibility. Individuals and organizations are motivated to search for new practices and learn from others’ efforts when they are dissatisfied with their own performance (Cyert and March 1963). Thus, the search for new behaviors or practices is triggered by the need to solve the problem of poor performance (Cyert and March 1963). Actors simplify the evaluation of their performance by assessing it relative to their aspiration level and by dichotomizing actual performance as “success” or “failure” depending on whether it was above or below their aspiration level (March and Simon 1958). An aspiration level is “the smallest outcome that would be deemed satisfactory by a decision maker” (Schneider 1992: 1053). Aspiration levels are determined by a process of social comparison (Cyert and March 1963; Festinger 1954; Greve 1998). Individuals compare themselves and their organizations with referent others who are similar to the focal actor and who therefore serve as a reference group (Greve 2005). A measure of the aggregate performance of the reference group, such as the average, constitutes the social aspiration level for an actor, who compares its own performance to this level to determine success or failure (Cyert and March 1963; Greve 2005). Poor performance triggers the search for new behaviors and practices, while success decreases the likelihood of such search. In the context of innovation adoption, performance below an actor’s social aspiration level increases its willingness to consider adopting innovations adopted by its referent group, while success reduces incentives to change existing behavior and practices (Greve 2005). Poor performing organizations have been found to be more susceptible to social influence when faced with decisions to adopt novel practices (Davis and Greve 1997; Kraatz 1998).

In the OSS context, we expect the extent to which a licensor of a new project has previously worked on successful or unsuccessful projects will influence her susceptibility to the social influence of other socially proximal projects. As discussed previously, socially proximal projects are likely to serve as the referent group for the licensor of a new project. Relatively poor performance in prior projects should increase a licensor’s propensity to search for appropriate practices, which will increase his or her susceptibility to outside influence. Experience working on successful OSS projects, in contrast, will tend to reduce a licensor’s willingness to attend to the adoption decisions of other projects. Prior successful experience also sends a positive signal about the capabilities of the licensor to the community (Podolny 2001; Stewart 2005), which increases the project’s attractiveness to potential contributors, irrespective of the license adopted. Such a licensor is less likely to seek
help from others to determine the appropriate license type for his or her new project. Therefore, we expect that a licensor’s susceptibility to influence from other projects decreases with the success of his or her prior OSS projects.

HYPOTHESIS 4: The effects of social proximity (social cohesion and role equivalence) on the likelihood a licensor will adopt a particular license type will decrease with the relative success of the other OSS projects on which the licensor worked.

The depth and breadth of a licensor’s experience with OSS projects will also influence her susceptibility to social influence. The need to learn from others’ behavior increases with a potential adopter’s unfamiliarity with an innovation or the setting in which it is used (Wejnert 2002). When encountering something previously unknown or out of the ordinary, an individual begins a process of inquiry (Shultz 1964) and relies upon information and interpretation from others to make sense of the situation (Louis 1980). As individuals gain experience in a particular setting, they normally know what to expect of a situation, making them less susceptible to others’ interpretations or behaviors (Louis 1980). While the adoption of an OSS license has uncertain implications, licensors may differ in their familiarity with the OSS community and its practices, owing to differences in their levels of experience. OSS licensors with substantial depth of experience on OSS projects will tend to have more knowledge about OSS licenses than less experienced developers. Greater experience and knowledge increases individuals' sense of self-efficacy and reduces their incentives to alter their behavior (Bandura 1986). In particular, such individuals are less susceptible to social influence in their decision-making processes. Moreover, the accumulation of knowledge by individuals about a particular phenomenon increases their propensity to exploit this knowledge in making decisions and solving problems, rather than search for information and solutions outside their expertise, because local search tends to be less costly and generates less variable results than external, exploratory search (Audia and Goncalo 2007). Thus, licensors with more, or ‘deeper,’ experience will be less susceptible to social influence.

The diversity of a licensor’s experience will also reduce her susceptibility. Experience with a variety of different licenses increases a licensor’s familiarity with more licenses, reducing the likelihood the licensor will need to rely on others for information about a particular license. In general, diverse experience with a particular phenomenon encourages individuals to consider the phenomenon from a variety of perspectives, which challenges their fundamental assumptions and stimulates deeper and richer causal understandings about it (Argyris and Schön 1974). This richer understanding fosters healthy skepticism in decision-making, reduces the potential for decision biases and improves decision quality (Janis 1972). Thus, licensors who have experience
with a greater diversity of OSS licenses will possess richer knowledge about such licenses and will consequently have less incentive seek out or be influenced by external sources of information. In sum, a licensor with deeper and more diverse experience with OSS licenses will be less susceptible to social influence in her licensing decision.

HYPOTHESIS 5: The effects of social proximity (social cohesion and role equivalence) on the likelihood a licensor will adopt a particular license type will decrease with the depth and diversity of the licensor’s prior experience in the OSS community.

5. Data and Methods

5.1 Sample and Social Network Construction

To test our hypotheses we constructed a dataset of OSS projects hosted at SourceForge.net (SF). SF is the largest hosting site for OSS projects and accounts for about 90% of all OSS projects. SF provides web space as well as services such as emailing facility, discussion forums, CVS repository hosting, and download servers for OSS projects to organize and coordinate their development activities. Many studies of OSS have used SF data (e.g., Comino et al. 2007; Grewal et al. 2006; Lerner and Tirole 2005b; Singh et al. 2011a; Singh et al. 2011b).

We consider all projects registered on SF between its inception in November 1999 and December 2003. Data regarding project characteristics and developers were obtained from project websites and the SF database. To ensure the developers whose names are associated with the projects in the SF database actively participated in the projects during the period under consideration, we matched their contribution efforts and time of involvement through CVS log files, project communications, and the project documentation.

The hypotheses focus on a licensor’s social network. To test our hypotheses about the influence of these social networks, we need to observe them. Constructing a social network starts with the identification of the network’s boundary—the collection of actors who, for analytical purposes, is regarded as a bounded social collective (Marsden 2005). Building on prior research that identifies criteria for establishing network boundaries for empirical research (Laumann et al. 1983), we used participation in a foundry as the network boundary. A foundry includes all projects that use a common technology platform, such as projects that share the Perl programming language. Using a foundry as a network boundary criterion is valid for two reasons. First, projects within foundries are technologically similar and, as such, foundries provide a meaningful context for knowledge sharing among network members. Participation in a foundry as a network boundary for OSS projects at SF has

---

5 Concurrent Versioning System (CVS) is a software tool that stores the project source code, tracks changes made to the source code and stores programmers’ comments about the changes they made.
been used in related research (Grewal et al. 2006; Singh 2010; Singh et al. 2011a). Second, social relationships almost always occur among individuals within a foundry and rarely occur across foundries. We analyzed 2000 randomly selected developers who worked on multiple projects and found that only about 4% worked across two or more foundries.

In constructing our sample, we considered all projects that were registered in all 22 foundries at SF between November 1999 and December 2003. We discarded projects that never showed any development activity, leaving 29,995 projects. Although we use all of these projects to construct our networks, we limit our analysis to explaining variation in license adoption to projects that were started between January 2002 and December 2003 (n = 21,220). Because relationships endure over time, constructing networks using only relationships formed at a particular point in time would greatly understate the network’s true connectivity. Data on both pre-sample relationship formation and relationship duration is needed to accurately assess network structure at a point in time. We assume that relationships formed in the pre-sample period endure through the end of the sample period (December 2003). The assumption of individual social ties decaying after four years is consistent with prior research on interpersonal affiliation networks (e.g., Cattani and Ferriani 2008; Uzzi and Spiro 2005). Because our social network measures are meaningful for only those new projects at risk of license adoption that are directly socially connected to other projects, we remove from our final analysis all socially isolated projects – i.e., projects that do not have relationships with other sample projects (n = 15,913), leaving us with a final sample of 5,307 projects. As we explain below, we account for the possibility that this self-selection into our foundry-based social networks biases our results.

5.1 Network Construction

To assess social proximity, we constructed 22 affiliation networks. Each network contains projects assigned to a particular foundry. In these networks developers have social ties with one another as a result of working together on the same project and projects are related to one another as a result of sharing developers. Because a project’s administrator is responsible for choosing the project’s license and is the only member of the project at the time of license choice, the project and administrator are one and the same at the time of license choice. Our theory suggests an administrator is influenced by his/her social proximity to, and characteristics of, other projects (rather than individual developers). Because the unit of analysis is the project, we assessed social

---

6 The beginning of this time period represents the beginning of the SourceForge data. We ended data collection at the end of 2003 because this represents the closest year end to when we began this research project.
proximity among projects by projecting each developers-by-project affiliation network into its respective unipartite (i.e., one mode) inter-project network.

5.2 Dependent Variables

We expect a licensor of a new project will be more likely to adopt the type of prototypical license that projects to which she is socially proximate have adopted. Prototypical licenses are defined by whether they include copyleft and/or viral clauses because these clauses define the essential differences among OSS licenses (de Laat 2005; Lerner and Tirole 2005b; Stewart et al. 2006). These two license clauses have fundamentally different implications for legally copying, modifying and distributing software source code, represent fundamental ideological differences among proponents of OSS, and have important implications for the success of OSS projects (de Laat 2005; Stewart et al. 2006). Accordingly, we categorize licenses using these clauses. Because a project may offer its software under multiple licenses we follow Lerner and Tirole (2005b) and categorize a project under four different categories: (1) all licenses highly restrictive or not (ALLHR) (2) some licenses highly restrictive or not (SHR), (3) all licenses restrictive or not (ALLR), and (4) some licenses restrictive or not (SR). These categories are not mutually exclusive. The requirement for a license to be classified as highly restrictive is the presence of both 'copyleft' and 'viral' clauses, whereas the presence of only the 'copyleft' clause is needed for it to be classified as restrictive. Thus, all highly restrictive licenses also satisfy the restrictive license requirement. Hence, SHR is a subset of SR and ALLHR is a subset of ALLR. To compare our results with those of Lerner and Tirole (2005b), we follow them and treat each of these four categories as separate dichotomous dependent variables.

We estimate four different license choice models. In Model 1, the dependent variable is coded “1” if all licenses chosen by the focal project were highly restrictive and “0” otherwise. The dependent variables for other models were coded similarly. Because our theory specifies that a new project is more likely to adopt a particular type of OSS license when socially proximate projects have previously adopted the same license type, we do not construct a single dependent variable that reflects increasing license restrictiveness and model the

---

7 19% of sample projects use multiple licenses. Consequently, dropping sample projects with multiple licenses would substantially misrepresent the true inter-project social network structure and bias social cohesion measures. For example, assume a focal project A was directly connected to two other projects (B and C). Assume project B has adopted multiple highly restrictive licenses and C has adopted a single unrestricted license. Half of the projects to which project A is socially connected have adopted (some) highly restrictive licenses and therefore the value of social cohesion for A's inter-project network for highly restrictive licenses is 0.5. If projects that use multiple licenses (such as B) are dropped from the sample, then project A's inter-project network would consist of only project C and would erroneously be assigned a social cohesion value for highly restrictive licenses of 0.
likelihood a project will adopt a more restrictive license using, for example, an ordered probit estimator. As robustness checks of our primary results, however, we use two alternative constructions of our dependent variable. We discuss these robustness checks in the online supplement.

5.3 Explanatory Variables

5.3.1 Social Cohesion. Social cohesion defines social proximity in terms of the number, length and strength of the paths that connect actors in a network (Marsden and Friedkin 1993). We focus on social distance as the primary dimension of social cohesion. Our measure of Social Cohesion is based on a measure of social distance in which distance is defined as the number of network links that exist on the shortest path between two connected actors. To compute this measure, we first calculate the social distance among each pair of projects in the foundry-specific inter-project network. The social distance matrix is then used to calculate nearness among each pair of projects using an exponential decay transformation (Burt 1982). Larger values of nearness represent stronger social cohesion between two projects. Social distance is calculated from all projects to the focal project for which there is a path between projects. Because social distance is only meaningful for projects that started at SF prior to the focal project, we do not consider the social distance between a focal project and other projects that started after it. As we explain below, values in this project-specific nearness matrix are used as a measure of social cohesion in the construction of a project-specific social influence variable.

5.3.2 Role Equivalence. In the OSS context, projects are role equivalent to the extent they employ the same technology platform in their development. OSS projects are assigned to foundries based on their common usage of a technology platform. Thus, we construct four versions of the variable, Role Equivalence, each indicating the fraction of projects in the focal project’s foundry that are governed by the particular license type corresponding to the particular dependent variable. Consequently, the construction of this variable does not make use of the foundry-specific affiliation networks in contrast to the social cohesion measure.

5.3.3 Susceptibility. We operationalize susceptibility to social influence with several measures. To capture the relative prior success of the licensor of the focal project (to test hypothesis 4), we use the average size of his or her other projects at the inception of the focal project at SF. The size of a project is the number of developers involved in the project. An OSS project relies on voluntary contributions from developers to survive. A project that is able to attract a large number of developers is more likely to survive and achieve a stable release (Comino et al. 2007; Lerner and Tirole 2005b), which indicates a project’s technical success (Grewal et al. 2006). Many studies have used this measure to capture an OSS project’s success (Comino et al. 2007; Lerner
and Tirole 2005b; Singh et al. 2011a; Stewart et al. 2006). To test the robustness of our results involving this variable, we use the project’s percentile rank on SourceForge as an alternative measure of project success (Lerner and Tirole 2005b, Singh et al. 2011b). This measure captures the ranking of the project based on its activity and popularity among developers and users and has been used in prior research to capture OSS project performance (Lerner and Tirole 2005b). Following (Greve 1998), we computed alternative measures of the social aspiration level for an OSS project as the average size and average rank of all projects in the same foundry at the time of the focal project’s license choice. We then computed the relative success of an OSS project by subtracting the social aspiration level for the project from the project’s performance using average size ($REL SIZE$) and project rank ($REL RANK$), respectively, at the time of the focal project’s license choice. These two variables are used to test H4.

To capture the depth of experience of the licensor of the focal project (hypothesis 5), we use his or her tenure ($TENURE$) in the OSS environment, measured as the number of months the licensor had been active on SF at the time of license choice for the focal project, and experience ($EXP$), measured as the number of projects in which he or she had been involved by the start of the focal project. To measure the diversity of a licensor’s experience with OSS licenses ($DIVERSEXP$) (hypothesis 5), we use (Simpson 1949) index of diversity, also known as Blau’s index of heterogeneity and the Hirschman-Herfindahl index. Experiential diversity was computed as a reverse-scaled Simpson index:  

$$\text{DIVERSEXP} = \frac{1}{\sum p_j^2}$$

where $p_j$ is the proportion of licensor $i$'s total months of experience across all projects spent on projects using license type $j$. We reverse code Simpson’s measure so that higher values indicate greater experiential diversity.

Our theory predicts the susceptibility variables will reduce the influence of social cohesion and role equivalence on a focal project’s license choice. Thus, to estimate the effects of these variables on the susceptibility of the focal project to social influence, we multiply these measures with the two social proximity measures. We include the susceptibility variables alone to ensure proper model specification.

### 5.3.4 Infectiousness.

We measure the infectiousness of a prior adopter by its success (hypothesis 3), where success is measured as the number of developers ($ASIZE$) involved in the project at the start time of the focal project. In the OSS community, the success of a project reflects its social status in the community (Stewart 2005). As an alternative measure of the success of a prior adopter’s project we use the project’s percentile rank on SourceForge. We assess the robustness of our results to the use of this alternative measure. Because infectiousness is a property of information generated by prior adopters, its influence depends on whether or
not it can reach a potential adopter and the distance it must travel to do so (Strang and Tuma 1993). Because the influence of infectiousness depends on the social proximity between a prior and potential adopter, we interact the success measures with the two social proximity measures.\(^8\)

5.4 Control Variables

5.4.1 Structural Equivalence. To ensure the effects of social cohesion and role equivalence are not confounded by unobserved structural equivalence, we control for structural equivalence. Two actors are structurally equivalent to the extent they have ties to the exact same other actors (Burt 1976). Projects in the same foundry are role equivalent, but not necessarily structurally equivalent. Structural equivalence is measured by computing the Euclidean distance between projects \(i\) and \(j\) based on the dissimilarities in the relations of projects \(i\) and \(j\) to all other projects, \(k\), in the same foundry. We follow (Burt 1987) and transform the Euclidean distances to represent similarities in the relations of projects \(i\) and \(j\) by subtracting them from the maximum value of Euclidian distance involving \(i\). We compute structural equivalence only for projects that are indirectly connected by at least one other project. Including projects that share a direct tie would make it impossible to separate the effect of direct communication from the social comparison effect associated with structural equivalence (Leenders 2002).

5.4.2 Focal Project Propensity. To minimize alternative explanations and isolate the marginal effects of the explanatory variables, we include an extensive set of licensor- and project-specific controls.

Licensor Characteristics. Because of behavioral inertia, a licensor who has experience with a particular license type from working on other projects may choose a similar license type for their focal project. To control for this effect, we construct four versions of the variable \(EXPL\), each indicating the fraction of projects to which the focal licensor has contributed that are governed by one of the four particular license types used as our dependent variables.

Project Characteristics. Following Lerner and Tirole (2005b), we construct an extensive list of software project characteristics. Each project is characterized by the following dimensions: Intended Audience

---

\(^8\) Because there are many existing projects in a new project's foundry at the point in time it chooses a license, observations of ASIZE for existing projects need to be aggregated to the foundry level (so that each observation of a new project's license choice has a corresponding observation of ASIZE). In interacting each existing project's ASIZE with its social proximity to a new (focal) project, we aggregate across all existing projects (that have adopted license type \(j\)) by computing a network weight matrix that scales the ASIZE of each existing project (see section 5.5.1). This approach does not control for the main effect of existing projects' ASIZE. The construction of such a main effect variable would require similar aggregation across projects to the foundry level. Extant theory does not suggest a particular approach to aggregation. In unreported results, we used the mean. We did this for both measures of infectiousness. The inclusion of the two infectiousness variables as main effects in all estimated models do not change the results reported in tables 3 and 4.
(e.g., end users, system administrators), Topic (e.g., games, Internet), Operating System (e.g., POSIX, Windows), User Interface (e.g., GUI, text based), and Natural Language (e.g., English, Chinese). In all, we control for 46 characteristics of a focal OSS project. For a list of these controls, see Tables A6-A9 in the online supplement. We control for every variable used by Lerner and Tirole (2005b) except project development stage. Because the license choice is made upon project registration at SourceForge, all of our sample projects have a development stage of 0.

5.5 Model

The basic framework for the license choice model developed here captures both economic and social influence factors that affect a licensor’s choice of a particular license type. Let $y_i$ be an indicator variable that equals “1” if project $i$ has license type L and “0” otherwise. The model is specified as:

$$P(y_i = 1 | X, Z, D, F) = F(\beta X + \gamma Z + \delta D)$$

In the above, $X$ represents the social influence variables, $Z$ represents project characteristic controls, $D$ represents licensor characteristic controls, and $F$ is the logistic function. The variables in $Z$ correspond to the economic incentives argument put forth by Lerner and Tirole. While the construction of $Z$ and $D$ is straightforward, we explain the construction of $X$ in the next sub-section.

5.5.1 Social Influence Variables. Social influence variables account for three factors: social proximity of $j$ to $i$, infectiousness of $j$, and the susceptibility of $i$ (Strang and Tuma 1993). We follow Burt (1987) and assume that concrete social proximity is subjectively perceived by a focal licensor. Thus, projects at greater social distance may be perceived to have less social influence than their actual social distance would imply. In other words, social influence is perceived to exponentially decay with increasing social distance (Burt 1987). Given a concrete measure of social proximity between a focal project, $i$, and the projects in its social network, $j$, the focal project licensor’s subjective perception of that proximity can be represented by the power function:

$$(\text{proximity } i \text{ to } j)^e$$

$e$ is the relative weight (or intensity) of social influence perceived by a licensor from another project is then computed as follows. Let $P$ be the set of all projects in the focal project’s foundry that started before the focal project $i$, and $Y_P=[y_1, y_2, \ldots, y_p]$ be a vector of indicator variables representing license choice of all projects in the set $P$. While social proximity captures the social frame of reference by which a licensor evaluates a type of license, the license type that socially proximal projects have previously adopted
captures the normatively-sanctioned artifact they are influencing the licensor to adopt. In essence, in computing the social influence of a particular license type on a particular licensor, the portion of projects in the licensor’s foundry that previously adopted the license type are weighted by their subjective social proximity to the licensor. Finally, previous adopters of particular license types are also weighted in their social influence on a licensor by their infectiousness. Let $w_j$ capture the weight or intensity by which $j$ influences $i$, relative to all other projects in the network, moderated by the infectiousness of project $j$. These values, commonly referred to as network weights (Burt 1987; Leenders 2002), are calculated as follows:

$$w_i = \left[ w_j \times \text{REL SIZE}_i \times \text{TENURE}_i \times \text{EXP}_i \times \text{DIVERSE EXP}_i \times w_l \right].$$

where $w_i$ varies between zero and one. Higher values of $w_i$ correspond to higher social influence. The exponent $v$ measures the extent to which focal project $i$ is conservative in relying on others. Values of $v$ much larger than one imply that a focal project’s evaluation of a license is affected by its most proximate projects, whereas fractional values indicate that it is affected almost by anyone to which it is connected. Hence, $v$ determines the social frame of reference for the focal project. This construction of $w_j$ allows us to account for multiple competing innovations. Social cohesion and role equivalence are two different measures of social proximity that enter $w_i$. As we explain below in the Results section, we identify and use the value of $v$ that maximizes the estimated likelihood of our model. The susceptibility of a licensor to social influence is estimated by multiplying $w_j$ with the licensor characteristics of project $i$ expected to influence susceptibility:

$$X_i = [w_j \times \text{REL SIZE}_i \times \text{TENURE}_i \times \text{EXP}_i \times \text{DIVERSE EXP}_i \times w_l].$$

### 5.5.2 Accounting for Unobserved Heterogeneity

While we include an extensive number of controls, unobserved heterogeneity may still be a source of endogeneity and bias our parameter estimates. For example, licensors may have unobserved differences in their intrinsic preferences for particular license types or in their susceptibility to social influence. LICensorS also may have unobserved differences in their social relationships that influence their license choice, such as when some licensors have worked on their new projects with developers outside and/or prior to their registration on SF. We allow for unobserved heterogeneity in license choice behavior by allowing coefficients for social influence and licensor characteristics to vary randomly across licensors. Licensor random coefficients are introduced hierarchically. The above specified model is modified as:

$$X_i = [w_j \times \text{REL SIZE}_i \times \text{TENURE}_i \times \text{EXP}_i \times \text{DIVERSE EXP}_i \times w_l].$$

and

, and...
where \( \mu \) is a vector that corresponds to the mean of \( X \), and \( \Sigma \) is the corresponding covariance matrix.

5.5.3 Accounting for the Reflection Problem. Another source of endogeneity bias in estimating the social influence effects is the “reflection problem” (Manski 1993). Reflection poses a problem for causal identification of social influence because: (a) actors typically form social ties with another based on homophily – i.e., when they share similar attributes, beliefs, tastes or interests (McPherson et al. 2001), (b) the variables that cause homophilous tie formation are typically unobserved by analysts investigating the causal effect of one actor socially influencing another, and (c) the actors’ similar, yet unobserved attributes, beliefs, tastes or interests are correlated with (or cause) their similarity of behavior. Consequently, the observed effect of social influence via social ties may be a spurious result of an omitted variables bias. The reflection problem in our context would mean the licensor of a new project chooses the same license type chosen by socially proximate projects in his or her social network not because of the treatment effect of social influence but because he or she is similar in some unobserved ways (e.g., demographic or behavioral) to the developers on those projects and it is this unobserved similarity that also causes these similar actors to self-select into forming the social relationship connecting the projects. We address this source of endogeneity in two ways. First, by construction, only relationships and projects that exist prior to the adoption of a license by the focal project can influence the focal administrator’s license decision. This eliminates the potential for a simultaneity bias because the choice of license by a focal project temporally follows the formation of relationships with other projects that previously adopted their own licenses. Second, we use an approach developed by (Bramoulle et al. 2009) to identify the true effect of social influence by accounting for the reflection problem. The approach involves introducing unobservable effects common to all projects that belong to the same network component. The unobservable effects are treated as component-specific fixed effects and eliminated by using appropriate differencing.

5.5.4 Left-Censoring Bias. The possibility of having social ties with prior adopters of a license depends on the pool of adopters. Thus, the projects that started at SourceForge close to its inception are likely to have fewer ties compared to projects that started later. This creates a potential left-censoring bias. To minimize such a bias we used all projects to construct our social networks, but we used only those projects that were registered from January 2002 to December 2003 to test our hypotheses.

5.5.5 Sample Selection Bias. Of the full sample of 21,220 projects registered at SF during our study, 15,913 projects were social isolates. The social influence variables are undefined for such projects. Not accounting for these projects in the estimation may lead to sample selection bias (Heckman 1976). Accordingly, we estimated a
first-stage selection model for all 21,220 projects using instruments and then entered the inverse Mills ratio computed from this model in the license choice models as a control for unobserved selection into the sample of socially connected new projects. The first stage estimated the selection hazard as a function of all non-social influence variables along with two instruments: technical expertise of the licensor and network density of the focal project's foundry. To measure technical expertise, we used the licensor’s self-reported level of expertise reported to SF. The category indicating the lowest level of expertise is “want to learn,” which we coded as “1” if the developer selected that category and “0” otherwise. The want to learn developers are more likely to be new to OSS and less likely to be part of the existing social network of projects. Network density is the extent to which all the projects within a foundry are connected with each other. Higher density implies a higher level of interconnection. We expect the likelihood a new project is started by a developer from within the network increases with its density.

5.5.6 Estimation Procedure. The estimation procedure consists of two steps. First, we estimated a selection model using Probit regression. Second, we estimated four different license choice models in which the inverse Mills ratio obtained from step one was included as a covariate. Each license choice model is estimated through a standard Markov Chain Monte Carlo (MCMC) hierarchical Bayes Logit estimation procedure, using a Gibbs Sampler and the Metropolis-Hastings algorithm coded in Matlab (Rossi et al. 2005). To reduce the autocorrelation between draws of the Metropolis-Hastings algorithm and to improve the mixing of the MCMC, we used an adaptive Metropolis adjusted Langevin algorithm (Atchade 2006). In the hierarchical Bayes procedure, the first 100,000 observations were used as burn-in, and the last 25,000 were used to calculate the conditional posterior distributions. To assess the convergence of the MCMC, we compared the within- to between-variance for each parameter estimated across multiple chains (Gelman and Rubin 1992).

6. Results

Descriptive statistics for the key variables are reported in Table 1. We log transformed TENURE and EXP since they were heavily skewed. We standardized the component variables (RELSIZE, TENURE, EXP, DIVERSEXP, social cohesion, and role equivalence). The estimated model includes these standardized variables and their interactions. This step reduced correlations to acceptable levels: variance inflation factors (VIFs) for all variables were below 2.

---Insert Table 1 about here---

The selection model results are not reported to conserve space. Both technical experience and network
density significantly predicted the likelihood a new project has ties to other projects of length greater than 1. Moreover, these variables were not correlated with any of the license choice variables.

Before we explain the license choice results, we need to explain the procedure that we followed for estimating the parameter \( \nu \), which is used to construct the social influence variables. Following Burt (1987), we treat this parameter as a constant when calculating the value of social proximity. To identify the optimal value of \( \nu \), we compared likelihoods for different values of \( \nu \) (i.e., 0.1, 0.25, 0.5, 0.75, 1, 1.5, 2, 2.5, and 5) and then chose the value that provided the greatest likelihood. Setting \( \nu = 1 \) for both Social Cohesion and Structural Equivalence provided the best likelihood. This indicates the licensor gives more weight to the most proximate prior adopters. We present results for social proximity measures calculated by using \( \nu = 1 \).

--- Insert Table 2 about here ---

The results of the license choice model are provided in Table 2. Coefficient posterior means and variances are reported. Model 1 assesses the likelihood that all licenses chosen by a focal project are highly restrictive. Model 2 tests the probability that some of the licenses are highly restrictive. Models 3 and 4 test the probabilities for all and some licenses being restrictive, respectively. To save space, software characteristics and time period effects, while estimated, are not reported. Standardized coefficients are reported for all the models for comparison purposes.

Hypotheses 1 and 2 predicted a focal project’s social proximity, in terms of social cohesion and role equivalence, to other OSS projects would increase the likelihood it adopted the same license type as these other projects. Consistent with these hypotheses, the coefficients for social cohesion and role equivalence are positive and significant in all models. Social cohesion effect sizes are substantially larger than those for role equivalence, which suggests that licensors are more strongly influenced by direct communication than symbolic communication in their inter-project social networks. In hypothesis 4, we predicted licensors that had previously worked on relatively successful OSS projects would be less susceptible to social influence in their license choice. Hypothesis 5 predicted licensors with greater depth and diversity of experience on OSS projects would be less susceptible to social influence. Consistent with these two hypotheses, the interactions of the social proximity measures (social cohesion and role equivalence) and susceptibility measures are, with one exception, negative and significant across all models. The estimated coefficients for social cohesion, role equivalence, \( RELSIZE, TENURE, EXP \) and \( DIVERSEXP \) are simple effects rather than main effects because the interaction terms are significant (Jaccard and Turrisi 2003). To assess the net effect of each of these
variables, the main and the interaction effects must be combined. Using the results from Model 1, the effect for social cohesion is:

\[ p = 1.394 - 0.192 \text{SocialCohesion} \times \text{EXP} - 0.160 \text{SocialCohesion} \times \text{RELSIZE} - 0.191 \text{SocialCohesion} \times \text{TENURE} - 0.077 \text{SocialCohesion} \times \text{DIVERSEXP} \]

The coefficient estimate of 1.394 for social cohesion in Model 1 is conditional on RELSIZE, TENURE, EXP and DIVERSEXP taking on the value of zero (thus removing the effects of the interactions involving these variables), while the effect of Social Cohesion when RELSIZE, TENURE, EXP and DIVERSEXP are not zero will depend on their values and the values of their coefficients. For example, when RELSIZE, EXP and DIVERSEXP are at their means, the effect of a one-unit change in social cohesion for one standard deviation below and above the mean TENURE is 1.203 and 1.585, respectively.

---Insert Table 3 about here---

Hypothesis 3 predicted that prior adopters of OSS licenses differ in their infectiousness based on their success and that the influence of socially proximal projects on a focal project’s license choice will increase with the success of these projects. To test this prediction, we computed social influence weights \((w)\) in two ways: (1) non-weighted social proximity (i.e., set infectiousness to 1 for all prior adopters) and (2) ASIZE-weighted social proximity (multiplied proximity by ASIZE of prior adopters). These are non-nested models. Following (Newton and Raftery 1994), we compare these models on log marginal density. The Bayesian Information Criterion (BIC), commonly used for model selection in classical statistical analysis, asymptotically approximates the Bayesian posterior marginal density. Using this measure, the best fitting model is the one that minimizes \(-2\) log marginal density. The models (each with a specific construction of \(w\)) were run separately and their log marginal densities were calculated. The \(-2\) log marginal density values are reported in Table 3. For all four dependent variables, the models that account for infectiousness of prior adopters based on their ASIZE have the lowest \(-2\) log marginal density. This indicates that more successful prior adopters are more infectious. This provides support for Hypothesis 3. The inclusion of foundry-level average ASIZE as a main effect does not change the reported results (see footnote 10 above).

Finally, by comparing the standardized coefficients in Table 3, the major determinants of license choice can be identified. The three most important characteristics that determine the type of license used for governing an OSS project are the type of licenses used by the closest projects in the social network of the licensor, the dominant license type in the project’s foundry, and the licensor’s tenure on OSS projects.
We compared the predictive power of our full model to that of the baseline controls-only model, which corresponds to Lerner and Tirole’s (2005b) economic model, to assess the incremental improvement in fit provided by our full model (see Table 4). We used projects registered from Jan. 2002 to Dec. 2003 as our estimation sample and projects registered from Jan. to Jun. 2004 as a prediction sample. We used the estimated coefficients from the two models to predict the actual choice of license in the prediction sample. Because the adoption (or not) of a particular license type is binary, we assumed a particular license type was correctly predicted when its predicted probability exceeded 50%. For both the baseline and full model, we calculated the overall hit rate, focal license hit rate, and non-focal license hit rate\(^9\). Such hit rates are standard metrics of model fit used in prior research (e.g., Netzer et al. 2008). We also computed the root mean square prediction error (RMSPE) between the two models. The values for these metrics are presented in Table 4. For all metrics our full model substantively outperforms the economic incentives model. Our results are robust to a number of robustness checks, which we report in the online supplement to this paper

---Insert Table 4 about here---

7. Discussion

This study was motivated by gaps in the literature on innovation adoption and diffusion, particularly that which employs the heterogeneous diffusion framework, and the OSS licensing literature. First, existing innovation adoption and diffusion research provides little insight into how, when or why social influence affects the adoption and diffusion of competing artifacts (Strang and Soule 1998). Second, this research does not consider when or how the experiences of organizational members who have worked with particular innovations in their previous employers affect their current organizations’ adoption of such innovations. Finally, research provides an incomplete understanding of OSS license choice because it ignores the potential social influence of prior adopters of particular licenses on a project manager’s license choice and how this influence may vary over time and by a project manager’s social proximity to established OSS projects, characteristics of these projects and

---

\(^9\) The overall hit rate is defined as the percentage of projects for which the predicted license matches the actual outcome (e.g., ALLHR and ‘not ALLHR’). For example, assuming there were 100 projects in the prediction sample, 60 of which chose ALLHR and 40 chose ‘not ALLHR,’ and our model accurately predicted 70 of these licenses, then the overall hit rate is 70%. The focal license hit rate is calculated as the percentage of those projects’ licenses that were correctly predicted in the overall hit rate that were correctly predicted as being a particular type of license (e.g., ALLHR). Continuing the example, if out of the 70 projects accurately predicted the model predicts 50 ALLHR adopter projects correctly, then the focal license hit rate for ALLHR is 83.33% (50/60). The non-focal license hit rate is the percentage of those projects’ licenses that were correctly predicted in the overall hit rate that were correctly predicted as being not a particular type of license (‘not ALLHR’). In the example, of the 40 ‘not ALLHR’ projects in the sample, the model accurately predicts 20 (70-50), which implies a non-focal license hit rate of 50% (20/40).
the manager’s previous OSS experience.

In addressing these limitations, we examined the conditions under which prior adopters of competing OSS licenses socially influence how a new OSS project chooses among such licenses and how the experiences of the project manager of a new OSS project with particular competing licenses affects its susceptibility to this social influence. In investigating this question, we adapted the heterogeneous diffusion model (Strang and Tuma 1993) to accommodate multiple, competing innovations and extended it by identifying a novel source of influence on a potential organizational adopter’s susceptibility to the social influence of prior adopters. This allowed us to examine how the decision to adopt a particular type of OSS license is shaped not only by economic incentives but also by the license choices made by other projects to which a licensor is socially proximate. We tested our predictions in a sample of 5,307 OSS projects hosted at SourceForge.

After controlling for a large variety of factors shown to affect OSS license choice (Lerner and Tirole 2005b), our findings reveal the most important factor determining a new project’s license choice is the type of license chosen by existing projects that are closer to it in its inter-project social network. Moreover, the likelihood that a new OSS project adopts a particular license increases when more role equivalent OSS projects have previously adopted such a license and when these projects are large and successful. We also found that project managers with greater depth and diversity of experience in the OSS community and who have previously been members of successful OSS projects are less susceptible to social influence. Finally, the results indicate that social influence dominates economic incentives in the choice of license type. These results, which are robust to the use of a large number of controls, substantial efforts to control for various sources of endogeneity, alternative measures of the dependent variable and alternative estimation techniques, have important implications for research and practice.

7.1 Implications for Research

This study has substantive implications for research on: innovation adoption and diffusion informed by the heterogeneous diffusion framework, open source software licensing, and the governance of economic exchange. First, the results contribute to innovation adoption and diffusion research by showing how social influence from prior adopters affects a potential adopter’s choice among competing innovations. Although nearly all innovation adoption and diffusion research restricts its focus to a single innovation, potential adopters are frequently confronted with choosing among multiple, competing innovations. We show that the heterogeneous diffusion framework can be readily extended to the context of competing innovations by
focusing on how the prior adoption of one particular innovation affects a potential adopter's choice among all competing innovations, and how this relationship is influenced by characteristics of the prior adopters and potential adopters and the social proximity between them. The results generated from this approach suggest early adoption choices of one particular innovation can stimulate subsequent adoptions of the same innovation, potentially resulting in a bandwagon effect in which one (or a few) innovations diffuse widely and come to dominate while others are largely locked out (Abrahamson and Rosenkopf 1997). Our results suggest OSS licenses are subject to such a positive feedback effect in their adoption and diffusion and such effects are more likely to occur when: (a) early adopters of a particular innovation are more infectious (e.g., when they are viewed as successful or prestigious), (b) prior and potential adopters are more socially proximate, and (c) potential adopters are more susceptible to social influence (e.g., when they lack direct experience with any of the competing innovations). In contrast, adoption and diffusion in a population of actors is less likely to tip in favor of one (or a few) of many competing innovations when these factors are absent. The results of this study provide novel insight into understanding when a population of actors will exhibit more or less heterogeneity in the adoption and diffusion of competing innovations based on social influence.

Our results also contribute to innovation adoption and diffusion research by showing how an organization’s susceptibility to the social influence of prior adopters is shaped by the experiences of organizational members who have worked with particular innovations in other organizations. While existing research assumes potential organizational adopters do not have direct experience with an innovation before adopting it, interorganizational employee mobility makes it possible for employees to gain direct experience with an innovation in one organization and transfer this experience when they move to a new organization. In the OSS context, we showed that new projects founded by managers with little experience with other OSS projects or who had contributed to unsuccessful OSS projects were most influenced by prior adopters in their inter-project social networks. With more, and more diverse, experience and greater success on prior projects licensors became less susceptible to social influence. These results suggest the experiences of a key organizational decision-maker who has worked with particular innovations in other organizations can affect his/her current organization’s susceptibility to the social influence of prior adopters’ choices. These results contribute to an understanding of the micro-foundations of organizational innovation adoption because they show how organizational adoption decisions are affected by the experiences of individual organizational decision makers. Little research has examined the micro-foundations of the organizational susceptibility to the
social influence of prior adopters (Wejnert 2002).

This study also has clear implications for research on OSS license choice. The little research that has examined this topic has not considered the potential for established OSS projects to socially influence the license choice of new projects, which is surprising given the substantial uncertainty and confusion licensor typically face in choosing an appropriate license. Our results suggest social influence from prior license adopters strongly affects the scope of license chosen when licensors are uncertain about the proper license choice. Given that the effect sizes of the social influence variables were substantially greater than all other variables, our results suggest that prior research provides a substantively incomplete understanding of OSS license choice. Our theory and results do not offer an alternative explanation of OSS license choice as we do show economic incentives matter, but we do provide an important, and heretofore missing, complement to existing research by demonstrating the important role social influence plays. Both economic incentives and social influence from socially proximate established projects affect OSS license choice. Our results also indicate that the effect of social influence on license choice declines, allowing for a stronger influence of economic incentives, as a licensor’s depth and diversity of experience on prior OSS projects and the success of these projects increases.

Beyond providing an improved understanding of the OSS license choice, our results have implications for understanding the origins and influence of the social institutions that govern economic exchange. We address a fundamental question at the intersection of economics and sociology of how societies arrive at particular social institutions that govern economic activity (North 1990; Zukin and DiMaggio 1990). Our results indicate the development of social institutions that order economic activity, such as particular contractual governance mechanisms (North 1990), are the result of a social construction process by and among organizations in which social networks play a fundamental role. Under conditions of uncertainty about the appropriate contractual governance practice to adopt, socially proximate and infectious source organizations serve as influential social referents for potential adopters, resulting in a process of selective imitation that has a homogenizing effect on the types of contractual governance adopted. This insight confirms previous research that suggests particular practices diffuse and become widely shared and taken-for-granted (i.e., institutionalized) in particular organizational fields through a process of mimetic isomorphism among socially proximate organizations (Davis and Greve 1997; DiMaggio and Powell 1983; Zucker 1987). Our results also indicate there are limits to this process of mimetic isomorphism in that adopters with more depth and diversity of experience
with particular governance practices are less likely to conform to the governance choices of socially proximate source organizations. In elucidating the role of networks in the diffusion and institutionalization of particular governance practices, we provide a meso-level explanation by bridging macro-level explanations, such as those that highlight the influence of the frequency of prior adoption, with micro-level explanations that focus on individual utility-maximizing behavior. We were able to bridge these perspectives theoretically and methodologically by adapting and extending the heterogeneous diffusion model.

7.2 Implications for Practice

Our results also have implications for practice. Uncertainty about OSS licenses seems to be driving many managers of new projects to seek guidance from license choices made by existing projects. Indeed, prior research (Rosenberg 2000) and our observations of queries posted by open source software developers to Internet-based discussion forums reinforce the insight that managers of new OSS projects face substantial uncertainty when choosing a license. Research that examines the role of social influence in adoption decisions characterized by substantial uncertainty prescribes centralized, third party systems to gather and disseminate information to decision-makers (Bikhchandani et al. 1998). Accordingly, we recommend that centralized educational initiatives should be established to translate the results of research on OSS license choice and disseminate it to would-be project managers to help them choose an appropriate license for their projects. One useful form that such initiatives could take would be the development and implementation of an interactive software tool that would allow potential licensors to specify characteristics of their projects and receive guidance on the best license options. Organizations that might be well-suited for this role include the Free Software Foundation, the Open Source Initiative and SourceForge.net.

The results of this study also provide insight into how to stimulate the adoption and diffusion of a competing innovation. First, given the positive influence of the number of prior adopters on a potential adopter’s choice, innovation promotion strategies should focus on establishing a critical mass of early adopters. These efforts should be targeted at seeding the innovation with highly prestigious or successful actors as our results suggest their adoption decisions are particularly influential in subsequent adoption. Because such actors also tend to be less susceptible to social influence themselves, marketing strategies to encourage adoption will need to be particularly persuasive and attractive. Next, given the positive influence of the social proximity of prior adopters on potential adopters, promotional strategies should also focus on seeding the innovation with actors who have high closeness centrality in their respective networks and who have high role equivalence with
many other actors. Finally, because we show that actors who lack experience with any of the competing innovations are the most susceptible to social influence, promotional strategies should target such actors and emphasize in promotional communications to them the number of prior adopters and those that are particularly prestigious and successful.

7.3 Limitations

The limitations of this study suggest opportunities for future research. First, given the motivation for this study, we focused on aspects of existing and new projects and the network ties among them. Consequently, we did not consider the potential influence of information sources beyond the population of OSS projects. Innovation diffusion research suggests external actors such as the mass media, consultants and professional communities influence innovation adoption and diffusion by shaping how potential adopters in a population interpret innovations as appropriate and legitimate (Strang and Soule 1998; Wejnert 2002). The question of how OSS license choice is influenced by these external actors is thus an important question for future research. Second, although we argued that information transmission, observation and learning are the causal mechanisms underlying social influence, like most network diffusion studies our data do not allow us to directly observe and distinguish among these explanations. Future research should therefore seek to isolate these mechanisms, possibly through longitudinal qualitative study, in order to better specify how and why social influence from prior adopters affects the choices made by new adopters. A final limitation concerns the generalizability of the results as our results may be limited to the time period analyzed. The time period we studied includes the founding of SourceForge and thus an early era in centrally-hosted OSS projects, which may have faced greater uncertainty about license choice than projects started much later. Thus, additional data from more recent time periods may be needed to externally validate our results.

References


<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>S.D.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 All licenses highly restrictive (ALLHR)</td>
<td>0.55</td>
<td>0.5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2 Some licenses highly restrictive (SHR)</td>
<td>0.6</td>
<td>0.49</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Description</td>
<td>Value1</td>
<td>Value2</td>
<td>Value3</td>
</tr>
<tr>
<td>---</td>
<td>--------------------------------------------------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>3</td>
<td>All licenses restrictive (ALLR)</td>
<td>0.73</td>
<td>0.44</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>Some licenses restrictive (SR)</td>
<td>0.76</td>
<td>0.43</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>Social cohesion ALLHR</td>
<td>0.54</td>
<td>0.23</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>Social cohesion SHR</td>
<td>0.64</td>
<td>0.23</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>Social cohesion ALLR</td>
<td>0.54</td>
<td>0.2</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>Social cohesion SR</td>
<td>0.45</td>
<td>0.25</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>Role equivalence ALLHR</td>
<td>0.58</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>Role equivalence SHR</td>
<td>0.74</td>
<td>0.1</td>
<td>0.4</td>
</tr>
<tr>
<td>11</td>
<td>Role equivalence ALLR</td>
<td>0.77</td>
<td>0.06</td>
<td>0.5</td>
</tr>
<tr>
<td>12</td>
<td>Role equivalence SR</td>
<td>0.87</td>
<td>0.05</td>
<td>0.5</td>
</tr>
<tr>
<td>13</td>
<td>Licensor's experience on OSS projects (EXP)</td>
<td>2.94</td>
<td>1.69</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>Average relative size of licensor's prior projects (RELSIZE)</td>
<td>0.13</td>
<td>1.14</td>
<td>-3.14</td>
</tr>
<tr>
<td>15</td>
<td>Licensor's tenure in OSS environment (TENURE)</td>
<td>6.19</td>
<td>6.40</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>Diversity of licensor's experience with OSS licenses (DIVERSEXP)</td>
<td>0.65</td>
<td>0.24</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>Structural equivalence ALLHR</td>
<td>0.18</td>
<td>0.26</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>Structural equivalence SHR</td>
<td>0.2</td>
<td>0.28</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>Structural equivalence ALLR</td>
<td>0.51</td>
<td>0.48</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>Structural equivalence SR</td>
<td>0.52</td>
<td>0.48</td>
<td>0</td>
</tr>
</tbody>
</table>
### Table 2. Hierarchical Bayes Parameter Estimates for Logit License Choice Model

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Variables</th>
<th>Coefficient All Licenses</th>
<th>Coefficient Some Licenses</th>
<th>Coefficient All Licenses</th>
<th>Coefficient Some Licenses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Highly Restrictive?</td>
<td>Highly Restrictive?</td>
<td>Restrictive?</td>
<td>Restrictive?</td>
</tr>
<tr>
<td>H1</td>
<td>Social Cohesion</td>
<td>1.394***</td>
<td>1.739***</td>
<td>2.144***</td>
<td>1.627***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.317]</td>
<td>[0.347]</td>
<td>[0.421]</td>
<td>[0.322]</td>
</tr>
<tr>
<td>H2</td>
<td>Role Equivalence</td>
<td>0.966***</td>
<td>1.122***</td>
<td>1.446***</td>
<td>1.163***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.342]</td>
<td>[0.207]</td>
<td>[0.151]</td>
<td>[0.229]</td>
</tr>
<tr>
<td>H4</td>
<td>Social Cohesion X RELSIZE</td>
<td>-0.160***</td>
<td>-0.134*</td>
<td>-0.138**</td>
<td>-0.097*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.105]</td>
<td>[0.045]</td>
<td>[0.096]</td>
<td>[0.101]</td>
</tr>
<tr>
<td>H5</td>
<td>Social Cohesion X TENURE</td>
<td>-0.291*</td>
<td>-0.216*</td>
<td>-0.383**</td>
<td>-0.192*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.035]</td>
<td>[0.031]</td>
<td>[0.053]</td>
<td>[0.092]</td>
</tr>
<tr>
<td>H5</td>
<td>Social Cohesion X EXP</td>
<td>-0.192**</td>
<td>-0.117***</td>
<td>-0.103***</td>
<td>-0.105***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.222]</td>
<td>[0.209]</td>
<td>[0.324]</td>
<td>[0.219]</td>
</tr>
<tr>
<td>H5</td>
<td>Social Cohesion X DIVERSEXP</td>
<td>-0.077*</td>
<td>-0.065*</td>
<td>-0.123**</td>
<td>-0.111**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.032]</td>
<td>[0.027]</td>
<td>[0.096]</td>
<td>[0.055]</td>
</tr>
<tr>
<td>H4</td>
<td>Role Equivalence X RELSIZE</td>
<td>-0.145**</td>
<td>-0.097**</td>
<td>-0.101**</td>
<td>-0.077*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.026]</td>
<td>[0.005]</td>
<td>[0.017]</td>
<td>[0.026]</td>
</tr>
<tr>
<td>H5</td>
<td>Role Equivalence X TENURE</td>
<td>-0.044*</td>
<td>-0.079**</td>
<td>-0.117***</td>
<td>-0.094**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.013]</td>
<td>[0.012]</td>
<td>[0.023]</td>
<td>[0.003]</td>
</tr>
<tr>
<td>H5</td>
<td>Role Equivalence X EXP</td>
<td>-0.034*</td>
<td>-0.038*</td>
<td>-0.026</td>
<td>-0.029*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.014]</td>
<td>[0.019]</td>
<td>[0.019]</td>
<td>[0.011]</td>
</tr>
<tr>
<td>H5</td>
<td>Role Equivalence X DIVERSEXP</td>
<td>-0.119***</td>
<td>-0.104***</td>
<td>-0.181***</td>
<td>-0.204***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.071]</td>
<td>[0.052]</td>
<td>[0.063]</td>
<td>[0.123]</td>
</tr>
<tr>
<td>Licensor’s experience on OSS projects (EXP)</td>
<td>-0.009*</td>
<td>-0.014*</td>
<td>-0.055**</td>
<td>-0.024**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.007]</td>
<td>[0.024]</td>
<td>[0.009]</td>
<td>[0.011]</td>
</tr>
<tr>
<td>Relative size of Licensor’s prior projects (RELSIZE)</td>
<td>-0.016**</td>
<td>-0.039***</td>
<td>-0.027***</td>
<td>-0.051***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.005]</td>
<td>[0.015]</td>
<td>[0.009]</td>
<td>[0.019]</td>
</tr>
<tr>
<td>Licensor’s tenure at SourceForge (TENURE)</td>
<td>-0.512***</td>
<td>-0.492***</td>
<td>-0.447***</td>
<td>0.455**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.351]</td>
<td>[0.226]</td>
<td>[0.135]</td>
<td>[0.055]</td>
</tr>
<tr>
<td>Diversity of Licensor’s Experience at SourceForge (DIVERSEXP)</td>
<td>-0.012</td>
<td>-0.074</td>
<td>-0.069</td>
<td>-0.023</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.117]</td>
<td>[0.162]</td>
<td>[0.151]</td>
<td>[0.159]</td>
</tr>
<tr>
<td>Structural Equivalence</td>
<td>0.177**</td>
<td>0.284**</td>
<td>0.573**</td>
<td>0.443**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[0.201]</td>
<td>[0.292]</td>
<td>[0.277]</td>
<td>[0.193]</td>
</tr>
</tbody>
</table>

Number of Projects: 5307

Posterior licensor-specific variances are given in square brackets below coefficient means; *** implies the 99% confidence interval does not include zero. ** implies the 95% confidence interval does not include zero; * implies the 90% confidence interval does not include zero. All models include Software characteristics and Time dummy control variables. Social Cohesion and Role Equivalence measures account for infectiousness based on size (ASIZE).
### Table 3. Choosing Competing Models (-2 log marginal density)

<table>
<thead>
<tr>
<th>Role Equivalence</th>
<th>Same Infectiousness</th>
<th>Infectiousness differ by Size</th>
<th>Dependent Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Cohesion</td>
<td>Same infectiousness</td>
<td>12595.06</td>
<td>12494.73</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13691.35</td>
<td>13407.32</td>
</tr>
<tr>
<td></td>
<td>Infectiousness differ by size</td>
<td>14615.29</td>
<td>14419.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14439.50</td>
<td>14259.94</td>
</tr>
</tbody>
</table>

### Table 4: Model Fitness and Predictive Power Analysis

<table>
<thead>
<tr>
<th>Measures</th>
<th>Economic Incentives Model</th>
<th>Social Influence + Economic Incentives Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall License Hit Rate (ALLHR)</td>
<td>42.39%</td>
<td>71.72%</td>
</tr>
<tr>
<td>Focal License Hit Rate (ALLHR)</td>
<td>45.21%</td>
<td>75.59%</td>
</tr>
<tr>
<td>Non-Focal License Hit Rate (ALLHR)</td>
<td>38.94%</td>
<td>66.99%</td>
</tr>
<tr>
<td>RMPSE (ALLHR)</td>
<td>7.42%</td>
<td>3.93%</td>
</tr>
<tr>
<td>Overall License Hit Rate (SHR)</td>
<td>47.84%</td>
<td>73.61%</td>
</tr>
<tr>
<td>Focal License Hit Rate (SHR)</td>
<td>49.05%</td>
<td>78.24%</td>
</tr>
<tr>
<td>Non-Focal License Hit Rate (SHR)</td>
<td>46.02%</td>
<td>66.67%</td>
</tr>
<tr>
<td>RMSPE (SHR)</td>
<td>6.91%</td>
<td>3.66%</td>
</tr>
<tr>
<td>Overall License Hit Rate (ALLR)</td>
<td>63.95%</td>
<td>82.37%</td>
</tr>
<tr>
<td>Focal License Hit Rate (ALLR)</td>
<td>74.44%</td>
<td>84.29%</td>
</tr>
<tr>
<td>Non-Focal License Hit Rate (ALLR)</td>
<td>35.59%</td>
<td>77.18%</td>
</tr>
<tr>
<td>RMSPE (ALLR)</td>
<td>6.58%</td>
<td>3.16%</td>
</tr>
<tr>
<td>Overall License Hit Rate (SR)</td>
<td>65.52%</td>
<td>85.19%</td>
</tr>
<tr>
<td>Focal License Hit Rate (SR)</td>
<td>76.17%</td>
<td>87.92%</td>
</tr>
<tr>
<td>Non-Focal License Hit Rate (SR)</td>
<td>31.79%</td>
<td>76.54%</td>
</tr>
<tr>
<td>RMSPE (SR)</td>
<td>5.96%</td>
<td>2.91%</td>
</tr>
</tbody>
</table>