Can Free Resources Create Economic Value? The Impact of Crowd Contributors on Venture Capital Investment to Open Source Technologies

FRANCISCO POLIDORO JR., The University of Texas at Austin
WEI YANG, George Mason University

1. INTRODUCTION

Open source has become increasingly proliferated with entrepreneurial ventures (Chesbrough, 2003; Von Krogh & Von Hippel, 2006; Alexy & Reitzig, 2013). Although forfeiting the proprietary rights of technologies, open source can also bring unique benefits to new ventures' innovation and growth (Colombo, Piva, & Rossi-Lamastra, 2014). One of such benefits of open source is the extensive collaboration with external contributors in their corresponding online communities (Lee & Cole, 2003; Garriga, Von Krogh, & Spaeth, 2013; Dahlander & Piezunka, 2014; Belenzon & Schankerman, 2015). New ventures can access the human capital and knowledge inputs for free through collaborating with the crowd contributors in open source communities, as those external contributors volunteer to improve ventures' open source technologies without monetary compensation (Lakhani & Von Hippel, 2003; Fleming & Waguespack, 2007; Boudreau, 2012).

However, the economic implications of collaborating with crowd contributors in the open source communities remain unclear for new ventures. The existing literature on entrepreneurship through open source has not addressed the puzzle, whether and how the collaborations with external contributors can translate into financial benefits that ultimately sustain the growth of open source-based new ventures. In this study, we seek to explore this question by investigating on the impact of collaboration with contributors on ventures' acquisition of venture capital investment, an important milestone that reflects a venture's economic value (Baum & Locke, 2004). We propose that collaboration with the crowd increases the likelihood of venture capital investment to open source-based ventures, as the crowd functions as signals valuable market resources. In contrast with the existing literature that regards crowdsourcing collaboration as a process of knowledge creation in the upstream innovation process, our theory development the value of the crowds in the competition of the downstream product market. Crowdsourcing allows external contributors, who are oftentimes lead users in the diffusion and adoption of new technologies (Suarez, 2005; Rogers, 2010), to develop familiarity with the functionalities, fundamental logic, and communication pattern, which makes it easier for firms to lock-in critical users at the emergence of the downstream product market. Hence, crowd collaboration enables ventures to cultivate loyal users as rare and valuable market resources at an early stage, the economic value of which will then be reflected through venture capital investment. Empirically, we test our hypotheses using open source repositories on GitHub.com under an instrumental variable design, in which we use the time taken for crowd collaborators to respond to the venture as an instrument for collaboration completion. Preliminarily analyses show strong support for our arguments.

2. THEORY AND HYPOTHESES

2.1 Collaboration with the Crowd in Open Source Communities

Collaboration with external contributors through crowdsourcing has gained increasing attention in the literature on innovation and strategic management. Crowdsourcing generally refers to the firms' behavior of soliciting suggestions and solutions to problems from the population of individual external contributors (Afuah & Tucci, 2012; Piezunka & Dahlander, 2015). Crowdsourcing is particularly relevant to open source technologies, as they are often based on online communities for external developers and users to interact with each other and with the inventing venture. Such online communities become a platform for crowdsourcing collaborations, where actors jointly discover and solve technological problems while seeking to improve the innovation (Lee & Cole, 2003; O'Mahony, 2003; Ebner, Leimeister, & Krcmar, 2009). In those communities, the collaboration with external contributors in those communities oftentimes driven by the intrinsic interests of the crowd without monetary reward nor proprietary rights (Bagozzi & Dholakia, 2006; Waguespack & Fleming, 2009). Such external contributors in communities constitute free resources that firms can access through crowd collaboration.

Most of the current studies view the crowdsourcing collaboration with external contributors as a form of knowledge sourcing mode, which allows firms to search for distant knowledge and facilitates knowledge creation

and firms' adaption in the face of new technologies (e.g., Hippel & Krogh, 2003; Afuah & Tucci, 2012; Alexy, George, & Salter, 2013; Piezunka & Dahlander, 2015). Recent literature, however, also uncovers the downsides of crowdsourcing in the process of knowledge creation. Even though firms may access the distant knowledge from crowd collaboration, they may not be able to assimilate and utilize such knowledge as the filtering and learning process of such external knowledge is still subject to firms' own path dependencies and limited absorptive capacities (Piezunka & Dahlander, 2015). Those disadvantages cast doubts on how crowd collaboration can create value for entrepreneurial firms developing open source technologies.

2.2. Crowd Collaboration and the Financial Value of Open Source Technologies

In this study, we shift the focus from its role in upstream innovation to its impacts downstream on the product market, to address the puzzle whether and how crowd collaboration creates value for open source based ventures. In investigating the mechanisms that allow crowd collaboration to generate economic value, we highlight the ability of crowd collaboration to gain users and expand the market base, which constitutes critical prerequisites for ventures to subsequently profit from their open source technologies.

In the process of profiting from innovation, new ventures with open source technologies face a unique challenge. The absence of prosperity rights makes it difficult to gauge the potential economic value of the technology, even when its technical quality can be verified. Open source-based ventures often need to resort to business model innovation or launching subsequent priced technologies, to capture the value of their open source technologies (Fosfuri, Giarratana, & Luzzi, 2008; Alexy & Reitzig, 2013). However, despite the diversity of business models, such value capture process all requires an extended and loyal user base, with the potential to pay for service or priced alternatives (Teece, 2007). Hence, the fundamental role of the user base in profiting from open source innovations makes the competition of user and market share particularly pertinent in open source entrepreneurship.

We argue that collaborations with crowd contributors constitute an effective strategy to develop the user base as a valuable resource for open source technologies. The first market resource is the external contributors who directly take part in the crowd collaboration. Research on open source communities has noticed that crowd contributors are often users of the open source technologies (Von Krogh, Spaeth, & Lakhani, 2003; Bagozzi & Dholakia, 2006; Baldwin & von Hippel, 2011; Chatterji & Fabrizio, 2014). For those users, the collaboration process deepens their understanding of a focal technology. As they seek to contribute and improve the innovation, external contributors need to comprehend the principals, as well as the development logic and the underlying knowledge, above and beyond simply using the technology. Collaborators are more likely lock-in to the technology, not only because the technology is modified by themselves to better suit their own needs, but also because they become increasingly familiarly with the technical details. At the same time, the crowd's familiarity with the focal technologies increases their switching cost to competing technologies, which also reinforce the unique user base of the technology. Moreover, the collaboration also reinforces the social and emotional attachment of those users. As the community-based collaboration often requires frequent communication (Dahlander & Magnusson, 2005; Foss, Frederiksen, & Rullani, 2016), during which the crowd develops shared routines, languages and experience with the firm and within the community (Shah, 2006; O'Mahony & Ferraro, 2007), which helps firms retain those collaborators as the user base.

Secondly, the crowd also facilitates the diffusion of technology to other users, further strengthening the firm's advantage in the market competition for users. As mentioned earlier, the crowd of users who seek collaboration are often lead users with considerable technological knowledge. Those lead users play a critical role in the diffusion of technologies, as other users tend to rely on their recommendation and education when deciding which technology to adopt (Von Hippel, 1986; Rogers, 2010). Hence, retaining the lead users through crowd collaboration allows the inventing firm to generate direct network effects among users to win over a larger proportion of market base (Cabral, 1990; Cennamo & Santalo, 2013). In summary, the crowd collaboration enables more expanded user base of ventures' open source technologies, creating the potential to profit such user base. Hence, crowd collaboration for open source technologies will positively affect the ventures' financial value as reflected in venture capital investment.

H1: Collaboration with external contributors increases the likelihood of receiving venture capital investment for an open source-based venture.

Although the literature often portraits the crowd as "free" and "unpaid" (Lakhani & Von Hippel, 2003; Boudreau & Jeppesen, 2015), attracting crowd collaboration is not without cost. We argue the major cost involved in the crowd collaboration is the opportunity cost of disclosing the technological components as open source– knowledge that ventures could have otherwise made proprietary. If crowd collaboration requires

extensive disclosure of ventures' own knowledge, the acquisition of potential user base becomes more costly, as ventures may more lose opportunities to directly profit from their own knowledge that can otherwise become proprietary and protected by intellectual property rights. In other words, the increasing knowledge disclosed creates higher costs to compete for user base through crowd collaboration. In turn, the positive effect of collaborating with the crowd on firm value would be attenuated during venture capital investment, as it increases the opportunity cost for ventures in profiting from their open source technologies.

> *H2: The venture's knowledge disclosure in the open source communities weakens the positive effect of collaborating with external contributors on the likelihood of receiving venture capital investment for an open source-based venture.*

3. EMPIRICAL CONTEXT, RESEARCH DESIGN, AND PRELIMINARY RESULTS

We test the proposed hypotheses in the context of the open source software, the industry with the most active and developed open source innovation. Our primary data source is GitHub.com, currently the world's largest host of computer source codes for open source software programs (Dabbish *et al.*, 2012; Octoverse, 2018). We obtain venture capital investment through the VentureXpert database. The initial data contains longitudinal information on over the open source projects by 40,000 ventures, with more than 4 million records of collaboration from 2013 to 2017, mapped to the investment data based on the company website. Figure 1 presents venture capital investment to ventures with open source projects.

Our dependent variable *Venture capital investment* is the dollar amount of venture capital investment in the focal quarter. To measure ventures' collaboration with external contributors, we take advantage of a unique feature provided by GitHub to the open source technology, the "pull-request". Through pull requests, external contributors can make requests to the owner of open source technology to incorporate contributor's changes to the source codes, and owners or administers can review the changes submitted through pull request before deciding on whether to accept or reject changes. Our *Collaboration* is measured is the number of pull requests from external contributors completed within the focal month. In testing our predictions, an empirical challenge is to account for the possibility unobserved factors may exist that at the same time affect whether a venture collaborates with external contributors and also help explain venture capital investment. To do so, we adopt an instrumental variable method. We use the average time the external contributor taken to respond to the venture during collaboration as an instrument for collaboration. While affecting the collaboration (completion of pull requests), this measure is not related to other factors that can impact on venture capital investment because it is determined by individual collaborators exogenous to the venture.

We measure venture's *Knowledge disclosure* through the kilobytes amount of codes contained in "pushes", changes made by the venture to its source code repository directly, up the focal quarter. Ventures' knowledge breath is measured by *Programming language concentration*, a Herfindahl index based on the programming language used in all the repositories released by the venture up to the focal month. Models also control for venture experience and other knowledge related factors that may affect both venture capital investment and collaboration. Table 1 presents the preliminary results using Radom effects and instrumental variable models, showing support to our proposed hypotheses.

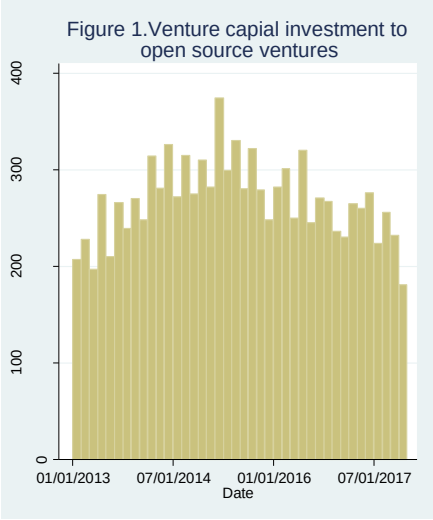

Table 1. Effects of crowd collaboration on venture capital investment

	(1) Random effects model	(2) Instrumental variable - (collaborator time to respond in pull requests)
Pull requests (H1: ß > 0)	0.009***	0.028***
	(0.001)	(0.003)
Pull requests X Push code size (H2: ß < 0)	-0.001***	-0.003***
	(0.000)	(0.000)
Push code size	-0.000	-0.000
	(0.000)	(0.000)
Constant	-0.012***	-0.021***
	(0.001)	(0.001)
Observations	4,300,309	3,873,006
Control, firm & month FE	YES	YES

Robust standard errors in parentheses, *** p<0.001, ** p<0.01, * p<0.05, + p<0.1

REFERENCES

Afuah A, Tucci CL. 2012. Crowdsourcing as a solution to distant search. *Academy of management review* **37** (3): 355-375.

Alexy O, George G, Salter AJ. 2013. Cui bono? The selective revealing of knowledge and its implications for innovative activity. *Academy of Management Review* **38** (2): 270-291.

Alexy O, Reitzig M. 2013. Private–collective innovation, competition, and firms' counterintuitive appropriation strategies. *Research Policy* **42** (4): 895-913.

Bagozzi RP, Dholakia UM. 2006. Open source software user communities: A study of participation in Linux user groups. *Management science* **52** (7): 1099-1115.

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

Baum JR, Locke EA. 2004. The relationship of entrepreneurial traits, skill, and motivation to subsequent venture growth. *Journal of applied psychology* **89** (4): 587-598.

Belenzon S, Schankerman M. 2015. Motivation and sorting of human capital in open innovation. *Strategic Management Journal* **36** (6): 795-820.

Boudreau KJ. 2012. Let a thousand flowers bloom? An early look at large numbers of software app developers and patterns of innovation. *Organization Science* **23** (5): 1409-1427.

Boudreau KJ, Jeppesen LB. 2015. Unpaid crowd complementors: The platform network effect mirage. *Strategic Management Journal* **36** (12): 1761-1777.

Cabral LMB. 1990. On the adoption of innovations with 'network' externalities. *Mathematical Social Sciences* **19** (3): 299-308.

Cennamo C, Santalo J. 2013. Platform competition: Strategic trade-offs in platform markets. *Strategic Management Journal* **34** (11): 1331-1350.

Chatterji AK, Fabrizio KR. 2014. Using users: When does external knowledge enhance corporate product innovation? *Strategic Management Journal* **35** (10): 1427-1445.

Chesbrough HW. 2003. *Open innovation: The new imperative for creating and profiting from technology*. Harvard Business Press.

Colombo MG, Piva E, Rossi-Lamastra C. 2014. Open innovation and within-industry diversification in small and medium enterprises: The case of open source software firms. *Research Policy* **43** (5): 891-902.

Dabbish L, Stuart C, Tsay J, Herbsleb J. Year. Social coding in GitHub: transparency and collaboration in an open software repository. In Proceedings of the Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work.

Dahlander L, Magnusson MG. 2005. Relationships between open source software companies and communities: Observations from Nordic firms. *Research policy* **34** (4): 481-493.

Dahlander L, Piezunka H. 2014. Open to suggestions: How organizations elicit suggestions through proactive and reactive attention. *Research Policy* **43** (5): 812-827.

Ebner W, Leimeister JM, Krcmar H. 2009. Community engineering for innovations: the ideas competition as a method to nurture a virtual community for innovations. *R&d Management* **39** (4): 342-356.

Fleming L, Waguespack DM. 2007. Brokerage, boundary spanning, and leadership in open innovation communities. *Organization science* **18** (2): 165-180.

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

Foss NJ, Frederiksen L, Rullani F. 2016. Problem-formulation and problem-solving in self-organized communities: How modes of communication shape project behaviors in the free open-source software community. *Strategic management journal* **37** (13): 2589-2610.

Garriga H, Von Krogh G, Spaeth S. 2013. How constraints and knowledge impact open innovation. *Strategic Management Journal* **34** (9): 1134-1144.

Hippel Ev, Krogh Gv. 2003. Open source software and the "private-collective" innovation model: Issues for organization science. *Organization science* **14** (2): 209-223.

Lakhani KR, Von Hippel E. 2003. How open source software works:"free" user-to-user assistance. *Research policy* **32** (6): 923-943.

Lee GK, Cole RE. 2003. From a firm-based to a community-based model of knowledge creation: The case of the Linux kernel development. *Organization science* **14** (6): 633-649.

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

O'Mahony S. 2003. Guarding the commons: how community managed software projects protect their work. *Research Policy* **32** (7): 1179-1198.

Piezunka H, Dahlander L. 2015. Distant search, narrow attention: How crowding alters organizations' filtering of suggestions in crowdsourcing. *Academy of Management Journal* **58** (3): 856-880.

Rogers EM. 2010. *Diffusion of innovations*. Simon and Schuster.

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

Suarez FF. 2005. Network effects revisited: The role of strong ties in technology selection. *Academy of Management Journal* **48** (4): 710-720.

Teece DJ. 2007. Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. *Strategic management journal* **28** (13): 1319-1350.

Von Hippel E. 1986. Lead users: a source of novel product concepts. *Management science* **32** (7): 791-805.

Von Krogh G, Spaeth S, Lakhani KR. 2003. Community, joining, and specialization in open source software innovation: A case study. *Research Policy* **32** (7): 1217-1241.

Von Krogh G, Von Hippel E. 2006. The promise of research on open source software. *Management science* **52** (7): 975-983.

Waguespack DM, Fleming L. 2009. Scanning the commons? Evidence on the benefits to startups participating in open standards development. *Management Science* **55** (2): 210-223.