



The Impact of Helping Others in Coopetitive Crowdsourcing Communities

Indika Dissanayake¹, Sridhar Nerur², Jingguo Wang³, Mahmut Yasar⁴, Jie Zhang⁵

¹University of North Carolina at Greensboro, USA, i_dissan@uncg.edu

²University of Texas at Arlington, USA, snerur@uta.edu

³University of Texas at Arlington, USA, jwang@uta.edu

⁴University of Texas at Arlington, USA, myasar@uta.edu

⁵University of Texas at Arlington, USA, jiezhong@uta.edu

Abstract

Organizations are increasingly engaging the community through crowdsourcing platforms to evolve innovative solutions to challenging business problems. Participants on such platforms often simultaneously cooperate and compete with one another to earn top honors. This paper addresses the imperative to understand the dynamics of knowledge sharing in such a coopetitive environment. Specifically, our study relies on the conceptual foundations of social exchange and social capital theories to investigate how help rendered (e.g., exchanging ideas or sharing knowledge) by participants in an online coopetitive crowdsourcing setting affects their performance. Furthermore, the study examines the moderating effects of the intensity of competition. Results of our econometrics analyses suggest that help given in a highly competitive contest, as opposed to a less competitive one, is more likely to be reciprocated, but less likely to improve the contributor's contest performance. In addition, our study found that help received by participants positively impacts their contest performance, and partially mediates the relationship between help rendered and contest performance. This research also provides insight into what motivates participants to share knowledge under conditions of coopetition. The findings of our study have strong implications for both theory and practice.

Keywords: Crowdsourcing, Online Communities, Contest-Based Communities, Knowledge Sharing, Social Exchange, Coopetitive Knowledge Sharing

Rajiv Sabherwal was the accepting senior editor. This research article was submitted on April 21, 2018 and underwent three revisions.

1 Introduction

At a time when information technology is growing at a phenomenal rate and product lifecycles are rapidly shrinking, organizations that fail to learn, adapt, and innovate are doomed to obscurity. Companies are continually striving to gain actionable insights that will help them stay ahead of their competition. However, the lack of in-house talent that is adequately trained in emerging technologies severely hampers their efforts.

Furthermore, there is growing evidence to suggest that the “wisdom of crowds” can be exploited to evolve novel solutions to complex problems (for example, see Howard, 2010). Given this backdrop, it is not surprising that organizations and researchers seek new ideas and high-quality solutions to their real-world problems from crowdsourcing platforms. For example, ZTE crowdsourced its smartphone design (Lukyanenko et al., 2017), and Lego has been able to create value by engaging its consumers in the design of their Lego sets (Kohler, 2015). A case study by Schlagwein and Bjorn-

Andersen (2014) suggests that crowdsourcing also contributes to organizational learning.

Formally, crowdsourcing refers to “a sourcing model in which organizations use predominately advanced internet technologies to harness the effort of the virtual crowd to perform specific organizational tasks” (Saxton & Kishore, 2013 p.3). According to Crowdsourcing.org, ten out of the eleven top global brands use crowdsourcing to seek innovative solutions to their business problems (Crowdsourcing.org, 2015). Furthermore, a survey conducted by the Marketing Executive Networking Group in 2009 found that 75% of company executives think crowdsourcing is highly effective for new product and service development (Sullivan, 2010). Examples of online platforms that facilitate crowdsourcing include Kaggle.com (for predictive modeling projects), IdeaStorm.com (for idea generation), and InnoCentive.com (for R&D).

With the growing popularity of online communities (OCs), the topic of knowledge sharing among members of these communities has attracted increasing attention in the IS literature (e.g., Faraj, Kudaravalli, & Wasko, 2015). Online communities not only provide a space for social interactions but also facilitate integration and knowledge sharing among geographically distributed individuals. Unlike in organizational settings, these collaborations tend to be self-organizing because of the voluntary and informal nature of interactions (Faraj et al., 2015). Crowdsourcing communities are similar to many of these online communities, for they are self-organizing, involve voluntary participation, provide forums for active community engagement, and have members who share common interests. However, the competitive nature of crowdsourcing platforms differentiates them from other online communities (Füller et al., 2014). In a typical crowdsourcing contest, participants, often working in teams, compete with one another to evolve creative solutions to business problems posted by organizations and research institutions. Such contests are characterized by intense competition, as only the best solution is rewarded. Indeed, the members who participate in such crowdsourcing platforms, as opposed to many other online communities, may be motivated by the reward as well as the opportunity to enhance their reputations.

Interestingly, crowdsourcing platforms generally facilitate interactions among competitors, allowing them to exchange ideas and share knowledge related to the contest. Thus, the social structure of cooperation underlying crowdsourcing platforms is akin to multiple units within organizations cooperating while simultaneously competing with one another for resources (for example, see Tsai, 2002). It is, therefore, reasonable to expect knowledge sharing in such cooperative environments to have different consequences compared to online communities that are collaborative. The question then is: *Why would*

participants jeopardize their chances of winning by sharing knowledge and ideas with competitors in an intensely competitive environment?

It is apparent from the preceding discussions that the relationships between knowledge sharing (i.e., rendering and receiving help) and performance in a cooperative setting are yet to be subjected to rigorous empirical scrutiny (for example, see (Hutter et al., 2011). Researchers have viewed knowledge sharing behavior in online communities as a form of social exchange (e.g., Yan et al., 2016). The benefits (i.e., social capital) that members may gain by virtue of establishing structural ties with other participants (Nahapiet, 2008) may be yet another reason for knowledge sharing in cooperative environments. Therefore, to fill the void in the literature, our study draws on the theoretical underpinnings of social exchange and social capital theories (Aronson, Wilson, & Akert, 2006; Blau, 1964) to explore the following research questions:

1. How does the active engagement of participants in knowledge sharing—either in terms of rendering or receiving help—impact their performance? Furthermore, how does the type of knowledge being shared (e.g., procedural versus general inquiries) differentially affect performance? Are social exchanges in cooperative crowdsourcing communities reciprocal? Specifically, do competitors who contribute more knowledge receive more help and perform better? In other words, does the extent of help received by participants mediate the relationship between the extent of help rendered and their performance?
2. Competition makes crowdsourcing platforms such as Kaggle different from other online communities. How does the level of competition impact knowledge sharing behavior in crowdsourcing communities? Specifically, does the intensity of competition moderate the relationships between knowledge sharing and performance?

Our study uses data from Kaggle.com, a crowdsourcing platform in which competition and cooperation among participants coexist, to explore the research questions. Thus, our study makes a unique contribution to both theory and practice by exploring knowledge sharing behaviors in a cooperative environment. While prior empirical studies have focused on collaborative communities (Ardichvili, Page, & Wentling, 2003; Wasko & Faraj, 2005), ours is among the few that have examined *help seeking* and *help providing* in a crowdsourcing community characterized by the simultaneous presence of collaboration and competition (Füller et al., 2014; Hutter et al., 2011). Furthermore, our study reaffirms the findings of previous empirical works (e.g., Faraj et al., 2015) with regard to the positive influence of structural social capital (e.g., the indegree

and betweenness centrality of a social network) on performance. For example, our paper demonstrates that rendering help enhances one's indegree and betweenness centrality, which, in turn, leads to superior performance. Our study also clarifies the types of knowledge shared, namely technical and context-related, in a knowledge-intensive crowdsourcing context and investigates how each impacts performance. Finally, the insights provided by this study may be used by crowdsourcing platform designers to create a collaborative environment that motivates participants to actively contribute knowledge that will eventually lead to more innovative and effective solutions.

The remainder of this paper is organized as follows. The ensuing section provides a review of the pertinent literature, followed by an illustration of the research model and the hypotheses resulting from it. Subsequently, we describe the dataset and explain how the variables used in our study were operationalized. Next, we present our analysis and results, followed by a discussion of the implications of our findings. The concluding section sums up the paper and suggests directions for future research.

2 Theoretical Background and Hypotheses

2.1 Knowledge Sharing in Online and Crowdsourcing Communities

An online community (OC) is a virtual space where people voluntarily come together and exchange resources and information (Faraj et al., 2016). Participants in these communities often share common interests and experiences. Such communities foster a climate of knowledge sharing because they are "boundaryless," highly visible, and continually change as new members with diverse backgrounds and skillsets join them. In the special section introduction of "online communities as a space for knowledge sharing," Faraj et al. offer the following definition: "OCs are collective spaces of knowledge flows characterized by a continuous morphing and are mutually constituted by digital technologies and participants" (Faraj et al., 2016, p.669). As mentioned earlier, a crowdsourcing community is a special type of online community that shares many similarities with other OCs, but the spirit of competition they instill in their members distinguishes them from other online communities investigated in prior studies (Bullinger et al., 2010). For example, OCs related to open-source software (OSS) development entail collective effort. In contrast, crowdsourcing largely relies on independent work (Seltzer, 2012). Paradoxically, crowdsourcing participants often voluntarily share their contest-related knowledge and ideas with other competitors while striving to develop

the best solution. Crowdsourcing communities are also different from general OCs where members share information about common interests (e.g., travel, photograph, sports) and life experiences. Unlike crowdsourcing, members of general communities typically do not compete with one another for rewards. Crowdsourcing communities differ from professional communities as well. While members of a professional online community (e.g., the database community) may work for competing companies (e.g., Oracle and IBM), they do not compete for rewards as members of a crowdsourcing community do. Since competition among participants is peculiar to crowdsourcing communities, the motivations and consequences of rendering help to other members could be quite different from what we know about other OCs. Of particular interest in crowdsourcing settings is the question of whether the sharing of knowledge and ideas with rivals would diminish the contributor's competitive advantage. It is therefore likely that the effects of helping behaviors in cooperative crowdsourcing communities may be more complicated than in other online forums that have been studied so far. Previous literature on crowdsourcing is conspicuously lacking in this regard.

Prior empirical studies on crowdsourcing contests have examined various factors that may influence outcomes, including output quality (Boudreau, Lacetera, & Lakhani, 2011), a solver's project completion rate (Yang, Chen, & Pavlou, 2009), a solver's probability of winning a contest (Mo, Zheng, & Geng, 2011; Yang, Chen, & Banker, 2010) and the number of participating solvers (Yang et al., 2009). It has also been demonstrated that contestants' characteristics such as skills (Archak, 2010) and effort (Mo et al., 2011), contest reward structure (Archak, 2010), and the total number of solvers (Archak, 2010; Boudreau et al., 2011) influence a solver's contest performance and the chance of winning. Füller et al. (2014) explored the heterogeneity of participants in a crowdsourcing context and its relationship to quality and type of contributions. However, to the best of our knowledge, the dynamics of knowledge and information sharing in crowdsourcing communities and its impacts have not been rigorously examined in empirical studies.

Prior research on OCs has primarily examined knowledge sharing behavior in cooperative environments. While many of the studies conducted so far have investigated the antecedents (e.g., motivations or intentions) of knowledge sharing, very few have empirically examined the consequences (e.g., benefits) of sharing (i.e., using knowledge sharing as an independent variable) in OCs. For instance, Huang and Zhang (2016) showed that contribution in a knowledge community results in job-hopping. Table 1 compares a sample of related studies on online knowledge collaboration published in premier IS journals.

Table 1. Online Knowledge Sharing Literature

References	Knowledge contribution		Context		
	IV	DV	Collaboration/ Non-competition	Competition	Cooperation
Bateman, Gray, & Butler (2011)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Chen & Huang (2010)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Chiu, Hsu, & Wang (2006)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Chang & Chuang (2011)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Faraj et al. (2015)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Füller et al. (2014)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Huang & Zhang (2016)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hutter et al. (2011)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Wasko & Faraj (2005)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Tsai & Bagozzi (2014)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Our study	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Notes: <input checked="" type="checkbox"/> represents “applied” or “considered.” <input type="checkbox"/> represents “not applied” or “not considered.”					

Although the focus of this study is on the dynamics of knowledge sharing in a cooperative crowdsourcing environment, we used text modeling to get a sense of how empirical studies have used knowledge sharing in other contexts, such as online forums, virtual communities, and open source software development.

Specifically, we used relevant search terms to extract abstracts from the Web of Science (WOS) database. It must be noted that the search terms¹ were motivated by the need to retrieve articles on knowledge sharing in specific contexts (online communities, competitions, and software development). A total of 245 abstracts were downloaded from WOS. While there are many ways to perform topic modeling (e.g., non-negative matrix factorization, latent semantic analysis/indexing, and latent Dirichlet allocation or LDA), we chose LDA because of its popularity (for example, see Debortoli et

al., 2016). The interested reader may refer to Debortoli et al. (2016) and Blei (2012) for details of LDA.

The downloaded abstracts were preprocessed before they were analyzed. Preprocessing included normalizing the text and removing stopwords, punctuation and digits. Ldatuning,² a package for tuning LDA model parameters in R, identifies the optimum number of topics using different algorithms. After considering the suggestions of several models and manually examining different numbers of topics, we found 20 topics that were reasonably interpretable. Topics were generated using the LDA implementation in a popular toolkit called MALLET (machine learning for language toolkit).³ A summary of the 20 topics that we extracted using LDA is shown in Table A (see Appendix A). It is apparent from the table that knowledge sharing has been investigated in a variety of contexts, including supply chain, open source software development, healthcare, social media, customer co-

¹ TOPIC: (“knowledge sharing” and “open source software”) or (“knowledge exchange” and “open source software”) or (“knowledge sharing” and “software development”) or (“knowledge exchange” and “software development”) or (“knowledge sharing” and “competition”) or (“knowledge exchange” and “competition”) or (“knowledge sharing” and “virtual communities”) or (“collaboration” and “virtual communities”) or (“knowledge sharing” and “online community”).

Refined by: WEB OF SCIENCE CATEGORIES: (MANAGEMENT OR COMPUTER SCIENCE INFORMATION SYSTEMS OR BUSINESS OR OPERATIONS RESEARCH MANAGEMENT SCIENCE) AND DOCUMENT TYPES: (ARTICLE OR REVIEW) Timespan: 1900-2016. Indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI.

² See <https://cran.r-project.org/web/packages/ldatuning/ldatuning.pdf>

³ See <http://mallet.cs.umass.edu/>

creation, and other virtual communities; however, a majority of these studies were conducted in collaborative or noncompetitive environments. There is a paucity of studies that have investigated competitive environments, and the examination of cooperative knowledge sharing is mainly limited to interorganizational settings. Furthermore, the results appear to support our claim that prior studies mainly focused on the motives or intention behind knowledge sharing rather than on the realized consequences or benefits.

Prior studies on collaborative, noncompetitive communities have demonstrated that helping others often leads to several benefits, including reciprocal benefits (Kankanhalli, Tan, & Wei, 2005), reputational and learning benefits (Chen, Xu, & Whinston, 2011; Khansa et al. 2015; Sproull, Conley, & Moon, 2005), improved professional status, increased number of professional contacts, enhanced self-image, access to expert advice, and an increased level of confidence in one's knowledge (Bateman et al., 2011). However, as mentioned earlier, competition and cooperation are simultaneously present in crowdsourcing contests. In some crowdsourcing competitions, solvers help each other through sharing information and knowledge (cooperation), while competing with one another for a monetary reward (competition). Therefore, the benefits that accrue to a contributor by virtue of helping others in such cooperative environments may be different from the advantages that one may get by rendering assistance in other communities. To further understand knowledge sharing behavior in cooperative environments, we examine cooperative knowledge sharing literature in other domains.

2.2 Cooperative Knowledge Sharing

Brandenburger & Nalebuff (1996) originally introduced the concept of "coopetition," which refers to being simultaneously involved in both competitive and cooperative activities. Cooperative knowledge sharing is extensively studied and applied at an organizational level in many industry sectors such as high-tech industries, healthcare, automotive, air transport, and food (Ritala, 2012). Cooperative knowledge sharing has spawned several breakthroughs and generated many benefits for organizations. An oft-cited success story is the coopetition between Samsung Electronics and Sony Corporation to develop LCD TV panels (Gnyawali & Park, 2011). This challenges the traditional view that a pure structure—competitive or cooperative—is superior to hybrid structures (Ghobadi & D'Ambra, 2011).

Research studies have identified many benefits and risks associated with knowledge sharing in cooperative environments. On the one hand, reciprocal knowledge sharing could result in synergistic effects and benefit all parties involved in knowledge sharing activities. On

the other, knowledge spillover and opportunistic behaviors of knowledge receivers could negatively impact knowledge contributors (Ilvonen & Vuori, 2013). Carayannis & Alexander (1999) viewed knowledge sharing as a positive-sum game.

Prior studies have indicated that coopetition is particularly important for knowledge-intensive, highly complex, and dynamic networks (Carayannis & Alexander, 1999). Crowdsourcing competitions networks are very dynamic in nature. The exit or entry barriers in such contests is low. Specifically, the platform that we selected is characterized by knowledge-intensive and complex tasks. In this study, we attempt to extend the literature on coopetition by understanding how it can be applied beyond inter- and intraorganizational networks to online crowdsourcing competitions.

2.3 Impact of Social Exchanges (Helping Behavior)

In a crowdsourcing community, a competitor can request help from other solvers by posting a message in an open forum. Participants often use these forums to share ideas with and render help to competitors who seek answers to specific contest-related questions. Such helping behavior is regarded as a form of generalized social exchange involving multiple participants (Fulk, Flanagan, Kalman, Monge, & Ryan, 1996), and its effects can be explained using multiple theoretical lenses, such as social capital theory and social exchange theory (Kankanhalli et al., 2005). Specifically, social exchange theory has been applied to investigate knowledge contribution and knowledge-seeking behaviors in online communities (Phang, Kankanhalli, & Sabherwal, 2009). Unlike in economic exchanges, obligations in social exchanges are not clearly specified (Blau, 1964). In social exchanges, people render favors with general expectations of gratitude and a possible return of the favor; however, such reciprocity is not guaranteed. This is what occurs in online crowdsourcing communities. In these communities, solvers voluntarily help peers by answering questions—through messages that they post in community forums—without specific expectations of future returns. According to social exchange theory, prosocial behavior of this nature can be rewarding in three ways: (1) By enhancing the likelihood of receiving help in the future (i.e., reciprocity), (2) by reducing personal distress of the contributor, and (3) by gaining social approval and enhancing self-worth (Aronson et al., 2006).

Consistent with social exchange theory, there is ample evidence of the benefits that accrue to contributors who exhibit prosocial behavior. For instance, empirical studies in management and organizational behavior have found that such helping behavior may increase employees' psychological well-being, protect them

from emotional exhaustion, and help them increase performance, persistence, and productivity (Grant, 2008; Grant & Sonnentag, 2010). Psychological research shows that when driven by intentions to help others, people come up with novel and useful ideas, thereby resulting in more creative performance (Forgeard & Mecklenburg, 2013; Grant & Berry, 2011). Also, helping behavior has been found to energize people by satisfying their fundamental psychological needs for competence, relatedness, and autonomy (Weinstein & Ryan, 2010). Based on these insights, we expect crowdsourcing solvers' helping behavior to enhance their creativity and work effectiveness, thereby enhancing their performance in contests.

In addition, according to social exchange theory, helping behavior leads to reciprocal benefits and increase the chances of receiving help from peers in the future (Aronson et al., 2006; Kankanhalli et al., 2005). Both direct and indirect reciprocity could result from these social exchanges. Direct reciprocity occurs when one's help is directly reciprocated by the recipient, while indirect reciprocity means one's giving is indirectly reciprocated by a third party. Prior studies have identified several reasons for indirect social exchanges in online communities, namely, distributed knowledge and uneven expertise, visibility, and social solidarity and norms (Faraj & Johnson, 2011). Furthermore, helping behavior contributes to the development of trusting and high-quality relationships (Ozer, 2011) and increases the social status of the helper (Hardy & Van Vugt, 2006). These, in turn, may lead to high-quality intellectual exchanges between contributors and their peers (Chen & Hung, 2010; Chiu et al., 2006). People who render help to others build both a good reputation and social capital, which further contributes to their success (Grant, 2013).

There are two features of social exchanges—appropriability and generalized reciprocity—that can be used to explain how help received improves the contributor's performance through enhancing the social capital of the contributor. Appropriability refers to the usefulness of social ties beyond the purpose for which they were originally created (Huang & Zhang, 2016). Consider, for example, a neighborhood safety patrol comprising members from the community. Although the social ties among these members were established for the sole purpose of keeping the neighborhood safe, these connections may be beneficial to members in different circumstances (e.g., getting their personal computers fixed or finding a good person to mend their fence). Generalized reciprocity refers to help being rendered without any expectation of an immediate return, anticipating perhaps that the favor will be returned at a later date (Nahapiet, 2008). Solvers can save considerable time and energy by soliciting help when they experience

technical difficulties. This would enable them to devote their valuable time to develop superior solutions to the problem. In addition, the varied background of the competitors may provide insights that would otherwise not be available to those seeking help.

While helping peers leads to performance and reciprocal benefits, social exchange theory suggests that providers of help incur certain actual and opportunity costs. Helping rivals by sharing knowledge and information in competitive environments may be more generate costs because of knowledge spillover. For instance, the study by Ghobadi and D'Ambra (2011) on knowledge sharing in organizational settings showed that the contributor might experience “negative reverse impact” (i.e., competitors can also use that knowledge), which diminishes the value of knowledge to the contributor. However, regardless of these costs, people may be inclined to share knowledge in high-profile competitive environments, such as Kaggle, because of their desire to express themselves and be identified as valuable members of the community (e.g., Cyr & Choo, 2010). Constant, Kiesler, and Sproull (1994, p. 400) note that, “sharing expertise may depend on people's own self-expressive needs.” Furthermore, they observe that, “expertise contributes to a person's self-identity, and that sharing expertise allows for personal benefits arising from self-expression and self-consistency” (p. 412).

As illustrated earlier, crowdsourcing contests exhibit characteristics of a cooperative environment. Thus, competitors in crowdsourcing contests could incur costs when they render help to others. For instance, not only do contributors have to expend time and effort to share relevant knowledge, but their sharing of know-how could also reduce their competitive advantage. Opportunity costs are incurred because of the loss of time and effort (Kankanhalli et al., 2005) that could have otherwise been used to improve their own solutions. It is also reasonable to expect that both the costs of helping others in contest-based communities and the attendant benefits are contingent on the level of competition (i.e., competitive intensity).

In light of the preceding discussions, we identify two main effects of a contributor's helping behavior: (1) *help given* influences the contributor's performance, and (2) *help given* influences the help received by the contributor. Motivated by the works of Wasko & Faraj (2005) and Chiu et al. (2006), we distinguished between the quantity and quality of help. *Help given* in this context refers to the quantity and quality of answers given by a solver when others raise questions related to the contest. *Help received* refers to quantity and quality of answers received from others when a solver poses a question. Wasko and Faraj's conceptualization of the knowledge contribution

comprises both helpfulness (analogous to quality in our case) and volume (i.e., quantity) of contributions based on users' message postings. In our study, quantity represents how often users engage in knowledge sharing activities, while quality represents the value of the knowledge shared, as measured by the number of votes that a message receives. *Performance* refers to the relative quality of a solver's solutions. Given these arguments, we propose that: (1) help given by competitors in a crowdsourcing community is

positively related to their contest performance (H1) as well as to the help received by them (H2), and (2) help received by competitors is positively related to their contest performance (H3). Furthermore, we hypothesize that the competitive intensity of a contest moderates the relationship between (1) help given and contest performance (H4), and (2) help given and help received (H5). Figure 1 summarizes the research model.

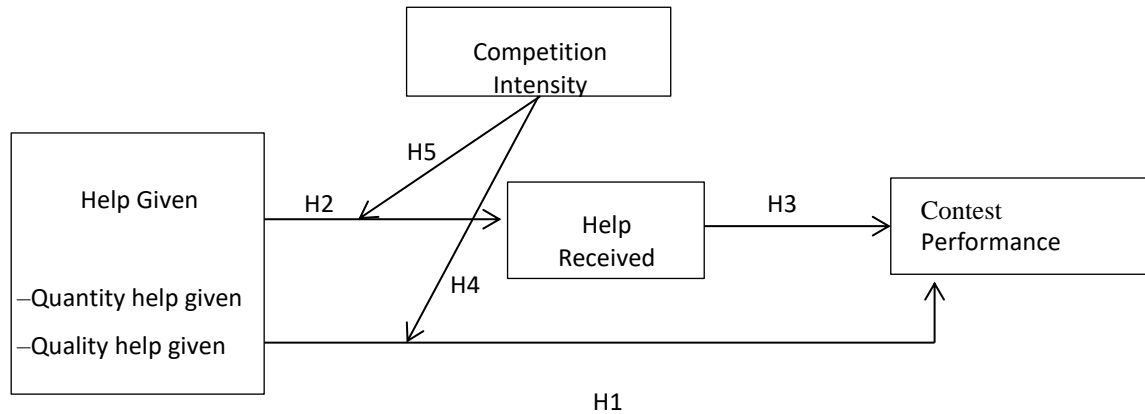


Figure 1. Research Model

2.4 Hypotheses Development

Crowdsourcing tasks are often knowledge intensive and intellectually demanding. The success of these projects depends on the skills and creativity of solvers. Even though helping others in crowdsourcing is not a required aspect of task performance, doing so may enhance an individual's cognitive processing and ability to generate creative ideas "by increasing persistence and vigor for the task at hand" (Forgeard & Mecklenburg, 2013, p. 262). Helping others encourages individuals to spend more time and energy on a task, which eventually enhances their performance by increasing creativity and effectiveness (Grant, 2008). Engaging in knowledge and information sharing through answering questions may also help solvers feel more confident about their own knowledge and reduce their stress levels. As a consequence, they might persist and perform better on their tasks (Bateman et al., 2011; Grant & Sonnentag, 2010; Xu, Jones, & Shao, 2009). Moreover, those engaged in prosocial help-giving behaviors may engage in "perspective-taking" (Forgeard & Mecklenburg, 2013) or the ability to objectively understand the needs and aspirations of those seeking assistance. This, in turn, enables them to

widen their perspectives and potentially come up with more creative solutions. Thus, we hypothesize:

H1: Help given by competitors in a crowdsourcing community is positively related to their contest performance.

H1a: The quantity of help given by competitors in a crowdsourcing community is positively related to their contest performance.

H1b: The quality of help given by competitors in a crowdsourcing community is positively related to their contest performance.

As per social exchange theory, contributors are likely to receive direct or indirect reciprocal benefits from others over time (Chen & Hung, 2010; Flynn, 2003). In our context, this suggests that help rendered by competitors would increase their likelihood of receiving help from their peers in the future (Aronson et al., 2006; Kankanhalli et al., 2005). Specifically, there are several situational factors that facilitate reciprocal behaviors in crowdsourcing competitions. First, unlike in organizational settings, there are no defined roles or expertise for the job. Any individual can participate in these competitions irrespective of their skills or expertise. Thus, there is likely to be a

huge variance in the knowledge and expertise among these participants. Second, the reciprocal benefit of prosocial behavior is particularly salient when helpful response is visible to the public. In the crowdsourcing context, solvers help each other by asking and answering questions and posting messages in discussion forums that are open to the public. The high visibility of these social exchanges suggests that help given is likely to result in an increase in help received. Third, help given could also enhance a solver's social status/technical reputation, thus enabling them to grow their social network (Flynn, 2003). All these situational factors enhance the likelihood that the other solvers in the community would directly or indirectly reciprocate the help that they received. Thus, we hypothesize:

H2: Help given is positively related to help received in a contest.

H2a: The quantity of help given is positively related to help received in a contest.

H2b: The quality of help given is positively related to help received in a contest.

Crowdsourcing contests, such as the ones on Kaggle, require a deep understanding of machine learning, predictive analytics, statistics, mathematics, and programming. In addition, they are knowledge intensive, often necessitating an integration of diverse pieces of information to evolve an innovative and effective solution to the problem at hand. It is reasonable to assume that participants who seek specific help are trying to fill crucial gaps in the knowledge that they need to develop an effective solution. According to Newell and Simon's (1972) theory of human problem solving, solvers navigate a potentially large search space as they evaluate solutions that reliably map on to the problem space being addressed. The availability of pertinent knowledge, such as "know-how," algorithms, and heuristics, facilitates the search process and enables participants to evolve an efficacious solution (for example, see Mangalaraj et al., 2014). Therefore, while the help rendered (i.e., posts) is available to all participants, including lurkers who don't engage in knowledge sharing/receiving, the ones who benefit the most are likely to be those who sought help to get their specific technical questions answered. Also, the time and effort they save can be directed toward improving their own solutions.

Terwiesch & Xu (2008) showed that one of the advantages of crowdsourcing is that it yields a very diverse set of solutions. Prior studies have shown that different points of view expressed by people from diverse backgrounds lead to new insights and ideas that foster creativity (Fischer, Scharff, & Ye, 2004). Thus, we hypothesize:

H3: Help received by competitors in a crowdsourcing community is positively related to their contest performance.

Prior studies have noted that "giving away knowledge eventually causes the possessor to lose his or her unique value relative to what others know" (Wasko & Faraj, 2005, p. 38). Since both competition and cooperation coexist in crowdsourcing communities, sharing can diminish the value of knowledge and blunt the competitive edge of competitors who render help (Ghobadi & D'Ambra, 2011). This represents the cost factor of social exchanges in contest-based communities. Such a phenomenon could be even more apparent in a highly competitive contest in which a large number of teams with similar levels of skills (i.e., low skill distribution) compete with one another. Sharing knowledge in a highly competitive environment is more likely to result in knowledge spillover. In addition, competitors incur higher opportunity costs when they spend time and effort in helping others in a contest where the intensity of competition is very high. Thus, we hypothesize:

H4: Competitive intensity negatively moderates the relationship between help given and contest performance such that the effect of help given on contest performance is less positive in highly competitive contests than in less competitive contests.

H4a: Competitive intensity negatively moderates the relationship between quantity of help given and contest performance such that the effect of quantity of help given on contest performance is less positive in highly competitive contests than in less competitive contests.

H4b: Competitive intensity negatively moderates the relationship between quality of help given and contest performance such that the effect of quality of help given on contest performance is less positive in highly competitive contests than in less competitive contests.

However, the help that solvers render in highly competitive contests is more likely to be reciprocated than in contests where the intensity of competition is low. In such circumstances, not only will the knowledge shared benefit a larger number of participants but it will also be more visible. In an intense competition where specialized knowledge can make a difference, solvers are more likely to understand and appreciate the risk that was taken by those who share knowledge. Hence, help given in highly competitive contests is more likely to be recognized by the community and will thus enhance the probability of receiving help. Furthermore, in highly competitive contests, solvers are more likely to have comparable skills (Dissanayake, Zhang, & Gu, 2015) and will thus be more likely to understand and

appreciate the knowledge and information shared. In contrast, in less competitive contests, solvers' skill distribution would be skewed, and it is likely that some solvers with low skill levels will have difficulty assimilating the knowledge shared because low skill levels act as a barrier (e.g., low absorptive capacity) to the reception of knowledge (Paulin & Suneson, 2012). Hence, help given in highly competitive contests is more likely to result in reciprocal benefits than in less competitive contests. Thus, we hypothesize:

H5: Competitive intensity positively moderates the relationship between help given and help received such that the effect of help given on help received is more positive in highly competitive contests than in less competitive contests.

H5a: Competitive intensity positively moderates the relationship between the quantity of help given and help received such that the effect of quantity of help given on help received is more positive in highly competitive contests than in less competitive contests.

H5b: Competitive intensity positively moderates the relationship between quality of help given and help received such that the effect of quality of help given on help received is more positive in highly competitive contests than in less competitive contests.

3 Data

We used data⁴ from Kaggle.com, a specialized crowdsourcing community that mainly deals with predictive modeling tasks. It has a member base of more than 100,000 data scientists from all over the world. Since its launch in 2010, Kaggle has worked with many companies, including Walmart, Allstate, Expedia, and Mercedes-Benz, to run analytics competitions to seek the best predictive models for a variety of problems, such as improving sales forecasting, predicting customer choices, optimizing search processes, and accelerating product testing (Kaggle.com; Dissanayake et al., 2015; Dissanayake et al., 2018).

Companies, government agencies, and researchers provide Kaggle with datasets, a description of the problem to be solved, and the reward they are willing to pay. Then, Kaggle sets up contests (Dissanayake et al., 2015). Each solver or team can submit multiple solutions throughout the contest. Kaggle evaluates all submissions and provides solvers instant feedback through a live scorecard, which gives solvers information on the predictive accuracy of their models and their relative positions (i.e., rank) in the contest. Kaggle's website features an online profile of each

solver, which shows a solver's personal information and overall performance score based on rankings in the contests in which they have participated.

In addition, each contest has a forum where solvers can initiate or participate in multiple discussion threads on various topics. When a solver needs help, they can initiate a thread by posting the question (initial post). Then, other participants voluntarily respond to the question by creating multiple posts (Figure 2). This is the main avenue for sharing knowledge in order to help others. Figure 3 shows a screenshot of a thread of topics from a forum. The helpfulness of a post may be assessed by the number of votes it receives from solvers. Thus, the number of votes that a post receives reflects its quality. Kaggle also reports the number of replies received by each thread topic. The total number of votes that a post has received is displayed below the post. Figure 4 shows a screenshot of an initial post on a discussion thread that solicits help from other participants.

3.1 Measurement

Help given refers to total forum posts made by a solver in response to help sought through a thread initiated by his or her peers in a given context (Figure 5). Help given was assessed in terms of the *quantity of help given* and the *quality of help given*. Following prior empirical studies, we used the number of posts as the *quantity of help given* and the votes received for those posts as the *quality of help given*. For example, Chiu et al. (2006) used average volume of posts as a measure of the quantity of knowledge shared. Sproull et al. (2005) used the identified "message" as the basic unit of contribution and measured "active participation" by the number of posts. Bateman et al., (2011) used the number of replies posted as a measurement of community participation, while Tsai & Bagozzi, (2014) used the number of messages posted to measure quantity contribution. Chen et al. (2014) used data from four online news communities to investigate the impact that feedback on users' posts has on their subsequent behaviors. Feedback was measured in terms of the number of likes or votes that a post received from other community members, and the proportion of up-votes was used as a measure of the quality of a post.

Help received is based on total forum replies received by a solver who posts a question and initiates a forum thread in a given contest (Figure 5). While considering replies received, we eliminated all subsequent posts made by the solver who posed the question (i.e., the thread initiator), including further clarifications and thank you notes to the respondents. Since the quantity and quality of help received are likely to be influenced

⁴ Kaggle's public data: <https://www.kaggle.com/kaggle/meta-kaggle>

by help given, we used principal component analysis (PCA) to derive help received based on the total number of replies as well as the total number of votes received for those replies. In addition, we separately investigated the impact using the total number of replies received as a measure of *quantity of help received* and the total number of votes received as a measure of *quality of help received*. These results are reported in the section on robustness tests.

Contest performance refers to the performance score of a solver in a given contest. Based on each solver's performance and on the characteristics of the contest, Kaggle computes the performance score. Specifically, the characteristics of the contest include its level of difficulty and the total number of participants. According to Kaggle, "the current formula for each competition splits the points among the team members, decays the points for lower finishes, and adjusts for the number of teams that entered the competition."⁵

Competitive intensity: Following Dissanayake et al. (2015), we used the Herfindahl index (HHI) to

measure the level of competition in a contest. It has been commonly used to measure market concentration in the economics literature. A higher HHI indicates a higher degree of market concentration and a lower level of competition. In our context, HHI measures the concentration of teams' intellectual capital⁶ within a contest, and competitive intensity was determined using Equation (1):

$$\text{Competitive Intensity}_j = 1 - \text{HHI}_j = 1 - \frac{\sum_{i=1}^n \text{TeamIC}_{ij}^2}{(\sum_{i=1}^n \text{TeamIC}_{ij})^2} \quad (1)$$

where i , j , and n denote team, contests, and the total number of teams participating in the contest, respectively. HHI measures the level of a team's skill in relation to the contest and indicates the extent of competition among participating teams. The competitive intensity (i.e., $(1 - \text{HHI})$) increases when more teams are involved and participants have comparable skills, whereas it goes down when the skill distribution of participants is more diverse.

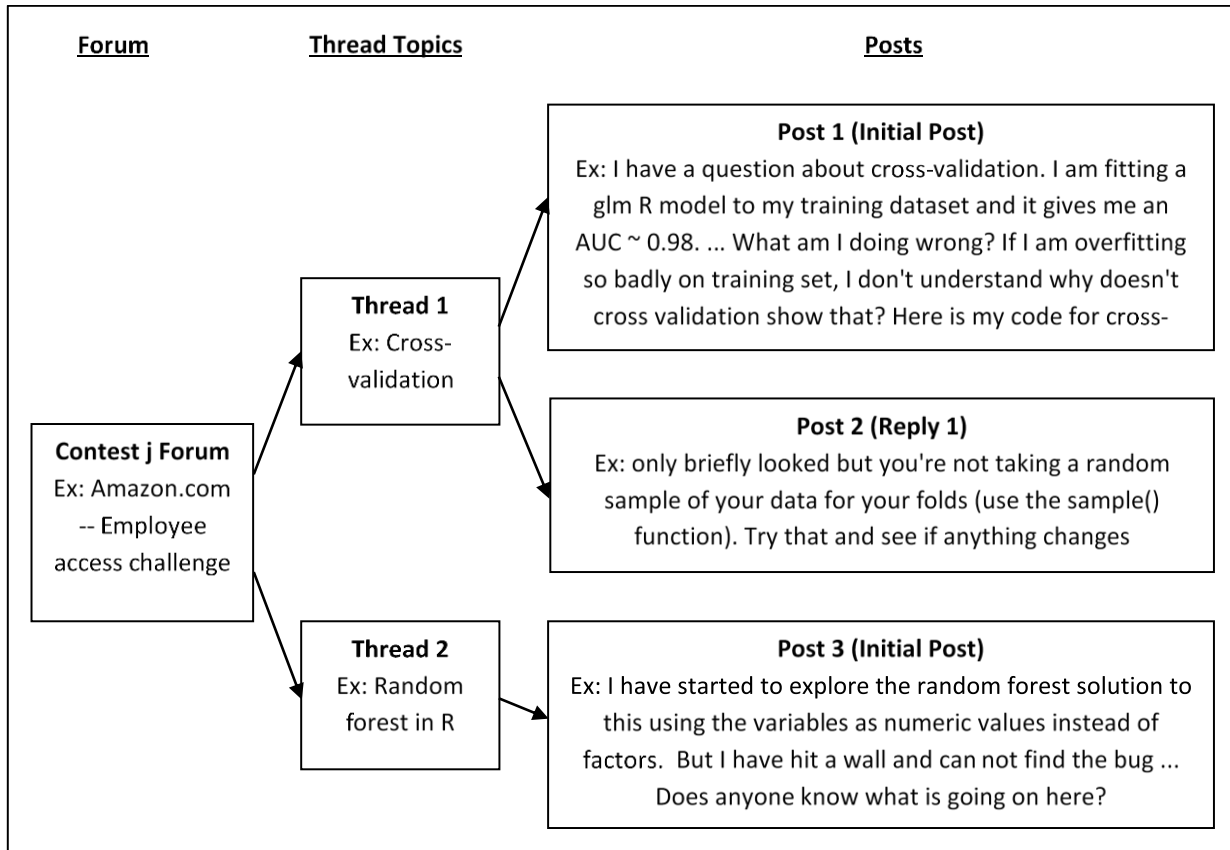


Figure 2. Forum Tree Diagram

⁵ Kaggle user ranking and tier system is available at <https://www.kaggle.com/wiki/UserRankingAndTierSystem>.

⁶ We used the formula given by Kaggle to calculate the intellectual capital of a team given a contest.

Votes	98 topics, 1,240 posts	Replies
2	.892 AUC Python code and question <i>by Tanay Tandon, 2 years ago</i>	1
1	R : Convert sparse.model.matrix to SVMlight format <i>by BarrenWuffet, 2 years ago</i>	2
0	cross validation <i>by Anna, 2 years ago</i>	6
1	Random Forest in R <i>by ferg96, 2 years ago</i>	2

Figure 3. Screenshot of Thread Topics in a Contest Forum (Source: Kaggle.com)

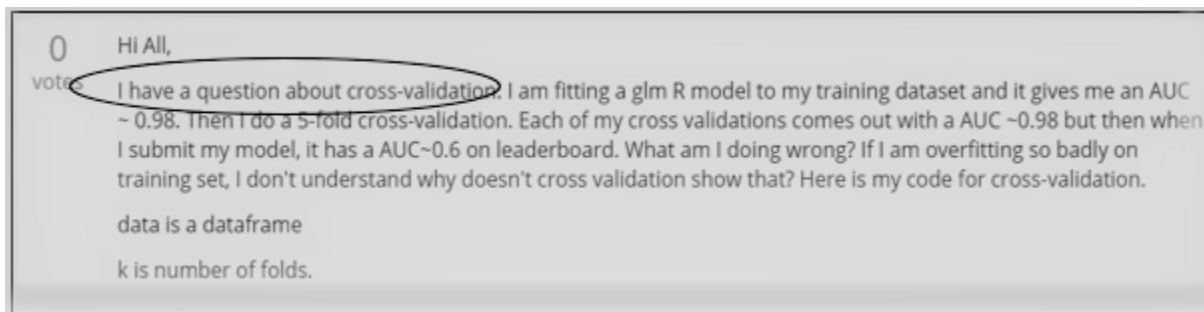


Figure 4. Screenshot of a Question Post on a Thread (source: kaggle.com)

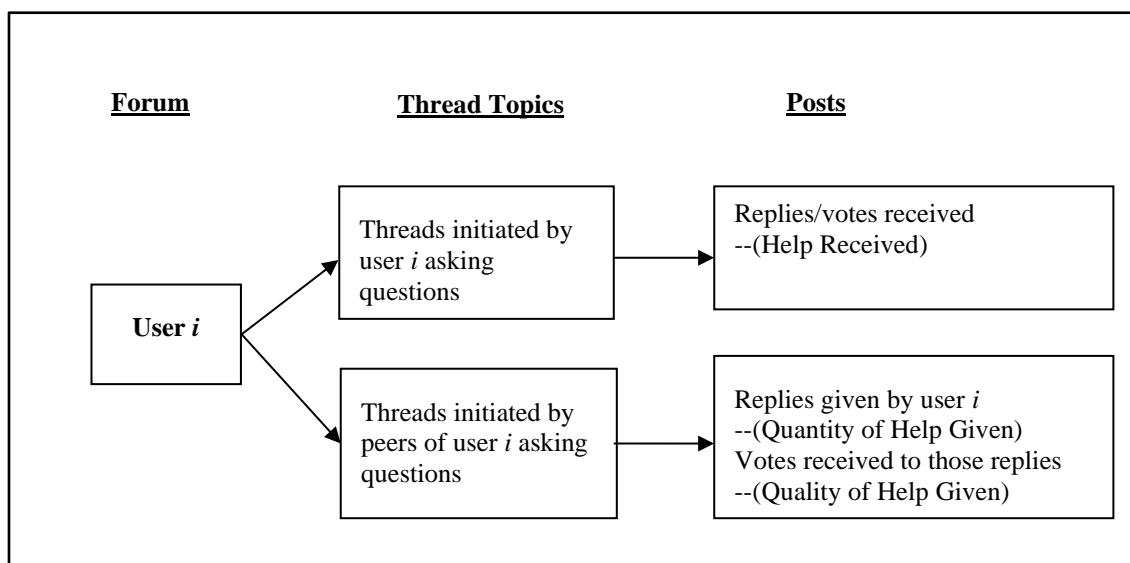


Figure 5. Help Given and Help Received

3.2 Control Variables

We used fixed effects to control for contest-specific effects. Moreover, we controlled for a solver's domain expertise or skill score, tenure, and the number of submissions of the solver in each contest. Solver skill score (i.e., skill) refers to a solver's profile score. This score reflects domain expertise and is derived based on one's cumulative performance in Kaggle competitions. The platform considers factors such as solvers' final ranking in each competition, the number of teams that participated, the number of members in the team, and the timing of the competition when calculating the score. The profile score was log-transformed to account for scaling effects. Prior literature in crowdsourcing has shown that solver skill positively influences their performance (Archak, 2010; Terwiesch & Xu, 2008). Hence, we used solver skill to control for the fact that their performance and recognition are affected by their domain expertise. Moreover, prior studies in online communities have identified tenure as a significant driver of knowledge contribution (Ma & Agarwal, 2007). It is also conceivable that high tenure individuals may receive more help by virtue of being well known in the community. In light of this, we added tenure as a control variable. Solver tenure in years was

dynamically calculated by taking the difference between the date a solver registered with the platform and the contest start date. Tenure was regarded as 0 if the solver registered after the contest start date. *The number of submissions* (i.e., *submission*) refers to the number of solutions a solver submitted in a given contest. Previous studies have indicated that the number of submissions affects performance (Mo et al., 2011).

We also entered a lagged dependent variable—*Performance_(k-1)*—in the regressions to control for persistency. This is the solver's performance score from his or her previous contest. For most solvers, performance in contests could be correlated due to factors such as work ethic, abilities, and other unobservable traits (Dissanayake et al., 2018). If a solver performed well in a previous contest, it is likely that he or she would perform well in the current time period as well. Thus, the lagged dependent variable can account for these individual-level effects. In addition, this also helps to control for the fact that high performing solvers may contribute more knowledge.⁷

We collected information on 18612 solvers and their related 130 contests. Tables 2 and 3 summarize the descriptive statistics and correlation matrix, respectively.

Table 2. Descriptive Statistics

Variable	Mean	Std. Dev.	Min.	Max.
Performance	166.17	435.03	1.17	10294.20
Help given (# posts)	0.44	2.41	0	123.00
Help given (# votes)	0.22	1.70	0	55
Help received (# posts)	0.55	4.35	0	231
Help received (# votes)	0.33	4.11	0	243
Submissions	11.78	20.15	1	965
Skill	6.44	2.24	0	12.32
# contests participated	1.70	2.28	1	57
Tenure (yr.)	0.55	0.83	0	4

Table 3. Correlation Matrix

		1	2	3	4	5	6	7	8
1	Contest performance	1.0000							
2	Quantity help given	0.1220	1.0000						
3	Quality help given	0.1603	0.5629	1.0000					
4	Quantity help received	0.0852	0.4648	0.2509	1.0000				
5	Quality help received	0.0585	0.2349	0.2186	0.7138	1.0000			
6	Submissions	0.1509	0.2033	0.1800	0.1236	0.1003	1.0000		
7	Skill	0.1394	0.1620	0.1666	0.1103	0.0920	0.2549	1.0000	
8	Tenure	0.0498	0.0778	0.1013	0.0591	0.0605	0.1064	0.4333	1.0000

⁷ We also tested the main models by controlling for the initial value of performance. The results were almost identical.

4 Regression Models and Estimation Results

We estimated the model coefficient using seemingly unrelated regression (SUR) analysis. The SUR corrects for any cross-equation error correlations that might be present. Furthermore, we controlled for contest-specific characteristics using a fixed-effects model. In the robustness test section, we also reported results with individual-fixed effects to control for solvers' heterogeneity.

4.1 Impact of Help Given on Performance and Help Received

First, we investigated how a solver's helping behavior influences his or her performance and the likelihood of his or her receiving help. Equations (2) and (3) describe the regression model:

$$\begin{aligned} Performance_{ik} = & \alpha_{10} + \alpha_{11} Help_Given_{ik} + \\ & \alpha_{12} Help_Received_{ik} + \\ & \alpha_{13} Skill_i + \alpha_{14} Submissions_{ik} + \\ & \alpha_{15} Performance_{i(k-1)} + \alpha_{16} Tenure_{ik} + \\ & Contest_{1j} + \varepsilon_{1ik} \end{aligned} \quad (2)$$

$$\begin{aligned} Help_Received_{ik} = & \alpha_{20} + \alpha_{21} Help_Given_{ik} + \\ & \alpha_{22} Skill_i + \alpha_{23} Performance_{i(k-1)} + \\ & \alpha_{24} Tenure_{ik} + Contest_{2j} + \varepsilon_{2ik} \end{aligned} \quad (3)$$

where $\alpha_{1m} (m = 0 \dots 6)$ in Equation (2) and $\alpha_{2m} (m = 0 \dots 4)$ in Equation (3) represents the coefficients of the variables, $Contest_{1,2j}$ is the coefficient for the fixed effects of the contest j , and j represents the k th contest that solver i has participated in.

Table 4a summarizes the effect of the *quantity of help given* on performance and help received. Model 2 controls for persistency of performance and eliminates observations with no prior competition details, thus making the sample size of Model 2 smaller than Model

1. Both Model 1 ($\alpha_{21} = 0.20, p < 0.01$) and Model 2 ($\alpha_{21} = 0.21, p < 0.01$) show that help given has a positive and significant impact on help received. Thus, H2a is supported, suggesting that help given enhances the probability of receiving help from others. The effect of help given on contest performance is significant in both Model 1 ($\alpha_{11} = 9.51, p < 0.01$) and Model 2 ($\alpha_{11} = 7.82, p < 0.01$). Thus, H1a is supported. Consistent with H3a, help received has a positive and significant impact on performance in both Model 1 ($\alpha_{12} = 8.00, p < 0.01$) and Model 2 ($\alpha_{12} = 7.03, p < 0.01$). The Sobel-Goodman mediation test confirmed that help received acts as a partial mediator. The proportion of the direct effect of quantity help given that is mediated through help received is 12%.

Table 4b shows the effect of the *quality of help given* on performance and the likelihood of receiving help from peers. Model 2 controls for persistency of performance. Both Model 1 ($\alpha_{21} = 0.18, p < 0.01$) and Model 2 ($\alpha_{21} = 0.18, p < 0.01$) shows that help given has a positive and significant impact on help received. Thus, H2b is supported. The effect of the quality of help given on contest performance is positive and significant in both Model 1 ($\alpha_{11} = 30.42, p < 0.001$) and Model 2 ($\alpha_{11} = 28.94, p < 0.01$). Thus, H1b is supported. Help received has a positive and significant impact on performance in both Model 1 ($\alpha_{12} = 5.23, p < 0.01$) and Model 2 ($\alpha_{12} = 3.03, p < 0.1$). Thus, H3b is supported. In addition, we tested whether help received mediates the relationship between the *quality of help given* and performance. The Sobel-Goodman mediation test confirmed that help received acts as a partial mediator. The proportion of the direct effect of the quality that is mediated through help received is 6%. This shows that both *quality* and *quantity of help given* result in performance and reciprocal benefits to the contributor in general. We also explored how help given enhances the contributors' performance through strengthening their help-seeking network. These results are discussed in the additional analysis section.

Table 4a. Seemingly Unrelated Regression Results for Help Quantity

DV	Help received		Contest performance	
	Model 1	Model 2	Model 1	Model 2
Help given (quantity)	0.1993 ***	0.2107 ***	9.5098 ***	7.8199 ***
Help received			8.0012 ***	7.0292 ***
Contest performance(k-1)		-0.0000		0.0704 ***
Submissions			2.3935 ***	2.5926 ***
Skill	0.0258 ***	0.0418 ***	37.7791 ***	41.4180 ***
Tenure	0.0260 ***	0.0291	-4.7215 *	12.3008 ***
Observations	31693	13082	31693	13082
R-squared	0.1460	0.1375	0.4326	0.4534
Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$				

Table 4b. Seemingly Unrelated Regression Results for Help Quality

DV	Help received		Contest performance	
	Model 1	Model 2	Model 1	Model 2
Help given (quality)	0.1838 ***	0.1829 ***	30.4175 ***	28.9381 ***
Help received			5.2325 ***	3.0296 *
Contest performance _(k-1)		0.0000		0.0662 ***
Submissions			2.2338 ***	2.3531 ***
Skill	0.0439 ***	0.0655 ***	36.3046 ***	38.4803 ***
Tenure	0.0286 ***	0.0349 *	-6.1257 **	11.4071 ***
Sample size	31693	13082	31693	13082
R-squared	0.0726	0.0721	0.4430	0.4686

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

4.2 Moderating Effects of Competitive Intensity

Prior literature has shown that help given leads to benefits (performance, reciprocal benefits) in collaborative communities (Xu et al., 2009). However, *competition* is unique to contest-based communities. Thus, we further investigated how *competitive intensity* affects these relationships. Equations (4) and (5) describe the regression model:

$$\begin{aligned} Performance_{ik} = & \alpha_{10} + \alpha_{11} Help_Given_{ik} + \\ & \alpha_{12} Help_Received_{ik} + \\ & \alpha_{13} Help_Given_{ik} * Competitive_Intensity_j + \\ & \alpha_{14} Competitive_Intensity_j + \alpha_{15} Skill_i + \\ & \alpha_{16} Submissions_{ik} + \alpha_{17} Performance_{i(k-1)} + \\ & \alpha_{17} Tenure_{ik} + Contest_{1j} + \varepsilon_{1ik} \end{aligned} \quad (4)$$

$$\begin{aligned} Help_Received_{ik} = & \alpha_{20} + \alpha_{21} Help_Given_{ik} + \\ & \alpha_{22} Help_Given_{ik} * Competitive_intensity_j + \\ & \alpha_{23} Competitive_Intensity_j + \alpha_{24} Skill_i + \\ & \alpha_{25} Performance_{i(k-1)} + \alpha_{26} Tenure_{ik} + \\ & Contest_{2j} + \varepsilon_{2ik} \end{aligned} \quad (5)$$

where $\alpha_{1m} (m = 0 \dots 7)$ in Equation (4) and $\alpha_{2m} (m = 0 \dots 5)$ in Equation (5) represents the coefficients of the variables. $Contest_{1,2j}$ is the coefficient for the fixed effects of the contest j , and j represents the k th contest that solver i has participated in.

Table 5a summarizes the results of the seemingly unrelated regression of Equation (4) and (5) where the *quantity of help given* is used as the main independent variable. After accounting for competitive intensity, the direct effect of the *quantity of help given* on performance is positive and significant in both Model 3 ($\alpha_{11} = 224.81, p < 0.01$) and Model 4 ($\alpha_{11} = 295.26, p < 0.01$). Moreover, the interaction effect of the *quantity of help given* and competitive intensity on performance is negative and significant in both Model 3 ($\alpha_{13} = -225.83, p < 0.01$) and Model 4 ($\alpha_{13} = -299.11, p < 0.01$). Thus, when competitive intensity is low (i.e., when competitive intensity is closer to 0), the overall effect

of help given on performance could be positive. However, this positive effect of *quantity of help given* on contest performance weakens and may even flip as the competitive intensity increases (i.e., when competitive intensity gets closer to 1). Thus, H4a is supported.

The interaction effect of the *quantity of help given* and competitive intensity on help received is positive and significant in both Model 3 ($\alpha_{22} = 0.62, p < 0.01$) and Model 4 ($\alpha_{22} = 1.31, p < 0.01$), while the direct effect of *quantity of help given* on help received is negative and significant in both Model 3 ($\alpha_{21} = -0.39, p < 0.01$) and Model 4 ($\alpha_{21} = -1.05, p < 0.01$). This suggests that the overall effect of help given on help received would be more positive in highly competitive contests, whereas it would be less positive in contests where the level of competition is low. Thus, H5a is supported. We mean-centered the variables to reduce any collinearity effects arising from the introduction of interaction terms.

Alternatively, we used the *quality of help given* as the main independent variable and reestimated the models (Table 5b). Consistent with the previous results with the *quantity of help given*, the direct effect of *quality of help given* on performance is positive and significant in both Model 3 ($\alpha_{11} = 1750, p < 0.01$) and Model 4 ($\alpha_{11} = 1674, p < 0.01$), while the interaction effect of *quality of help given* and competitive intensity on performance is negative and significant in both Model 3 ($\alpha_{13} = -1759, p < 0.01$) and Model 4 ($\alpha_{13} = -1683, p < 0.01$). Thus, H4b is supported. Moreover, the interaction effect of the *quality of help given* and competitive intensity on help received is not significant. Thus, H5b is not supported. Table 6 provides a summary of the results.

We plotted these interaction effects for three levels of competitive intensity for quantity (Figure 6a) and quality (Figure 6b) of help given. We fixed competitive intensity at the 10th percentile (low), 50th percentile (medium), and 90th percentile (high). The interaction plots show that both *quantity and quality of*

help given are related to better performance when competitive intensity is low, while they result in poorer performance when competitive intensity is high. Our results also show that the *quantity of help given* enhances the *quantity of help received* when competitive intensity is high.

Additionally, we used a bootstrap method to analyze the impact of help given on performance and help at different levels of competitive intensity (H4). Given Equation (4), we can write the following partial derivative to estimate the effect of help given on performance (Equation 6):

$$\frac{\partial \text{Performance}}{\partial \text{HelpGiven}} = \hat{\alpha}_{11} + \hat{\alpha}_{13} \text{CompetitiveIntensity} \quad (6)$$

Where $\hat{\alpha}_{11}$ shows the direct impact of help given on performance. We also expect help given to have an

indirect impact on performance ($\hat{\alpha}_{13}$), depending on the level of competitive intensity. To further evaluate these effects, we estimated the help given effect on performance at various percentiles of competitive intensity (i.e., 90th, 50th, and 10th), where lower percentiles represent contests that are less competitive. We computed standard errors using a bootstrapping method introduced by Efron (1979). Table 7 shows the effect of quantity and quality of help given on performance.

As the results indicate, competitive intensity negatively moderates the relationship between help given and performance. In other words, solvers get more performance benefits through sharing knowledge in less competitive contests rather than in high-competitive contests.

Table 5a. SUR Results for Help Quantity with Competitive Intensity Moderation

DV	Help received		Contest performance	
	Model 3	Model 4	Model 3	Model 4
Help given (quantity)	-0.3926 ***	-1.0540 ***	224.8094 ***	295.2586 ***
Help received			8.8262 ***	8.2645 ***
Competition* HG.	0.6200 ***	1.3143 ***	-225.8285 ***	-299.1137 ***
Contest performance _(k-1)		-0.0000		0.0702 ***
Submissions			2.4610 ***	2.6890 ***
Skill	0.0249 ***	0.0395 ***	37.8835 ***	41.5588 ***
Tenure	0.0244 **	0.0280	-4.0862	12.5865 ***
Observations	31693	13082	31693	13082
R-squared	0.1480	0.1424	0.4350	0.4568

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5b. SUR Results for Help Quality with Competitive Intensity Moderation

DV	Help received		Contest performance	
	Model 3	Model 4	Model 3	Model 4
Help given (quality)	0.0853	0.0529	1750.0470 ***	1673.9720 ***
Help received			5.2953 ***	3.1151 *
Competition* HP.	0.1008	0.1329	-1759.7940 ***	-1683.0620 ***
Contest performance _(k-1)		0.0000		0.0622 ***
Submissions			2.2911 ***	2.4103 ***
Skill	0.0440 ***	0.0655 ***	35.2127 ***	38.0116 ***
Tenure	0.0286 ***	0.0350 *	-6.6361 **	10.2198 **
Observations	31693	13082	31693	13082
R-squared	0.0726	0.0721	0.4603	0.4896

Notes:*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 6. Summary Results

Hypothesis	Result table number	Relationship direction	Support
H1a: Help given (quantity) → Performance (+)	4a	(+)	Supported
H1b: Help given (quality) → Performance (+)	4b	(+)	Supported
H2a: Help given (quantity) → Help received (+)	4a	(+)	Supported
H2b: Help given (quality) → Help received (+)	4b	(+)	Supported
H3a: Help received (quantity) → Performance (+)	4a	(+)	Supported
H3b: Help received (quality) → Performance (+)	4b	(+)	Supported
H4a: Competition moderation of H1a (-)	5a	(-)	Supported
H4b: Competition moderation of H1b (-)	5b	(-)	Supported
H5a: Competition moderation of H2a (+)	5a	(+)	Supported
H5b: Competition moderation of H2b (+)	5b	n.s.	Not Supported

Notes: (+) positive relationship; (-) negative relationship; n.s. nonsignificant relationship

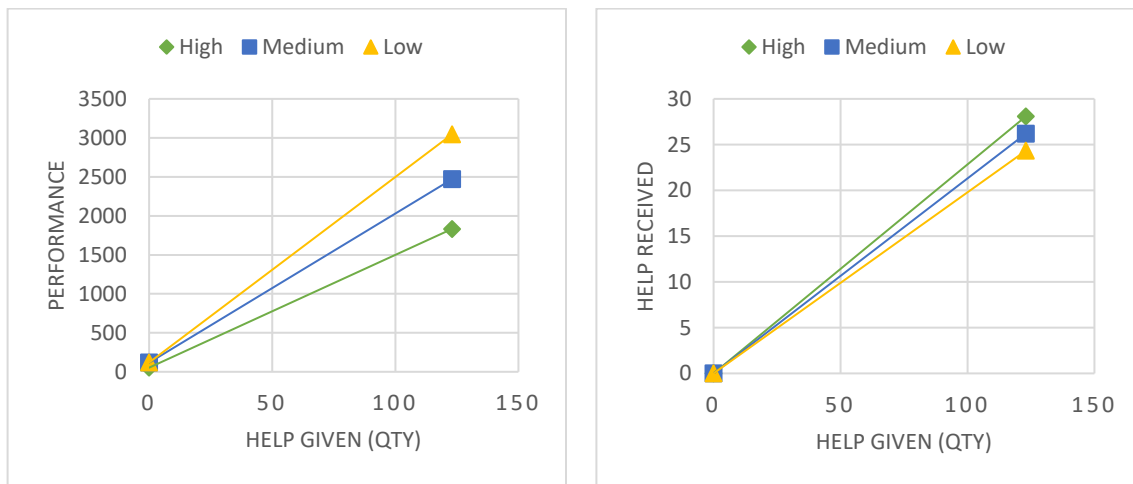
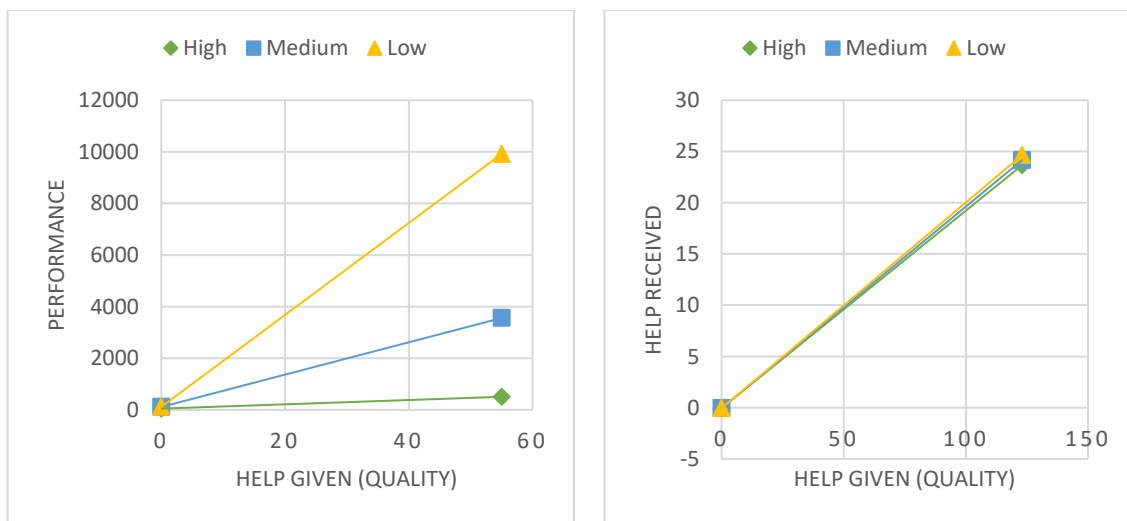

Figure 6a. Interaction Plots of Competitive Intensity (Quantity Help)

Figure 6b. Interaction Plots of Competitive Intensity (Quality Help)

Table 7. Competition Moderation of Help Given and Performance Results

Competitive intensity	Quantity of help given impact	Quality of help given impact
90th percentile	0.629	3.100
50th percentile	4.310 ***	31.784 ***
10th percentile	20.389 ***	157.082 ***
Observations	31693	31693

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

4.3 Additional Analysis:

Motivated by the works of Faraj et al. (2015), we conducted additional analyses to deepen our understanding of: (1) the differential effects of the nature of knowledge shared, and (2) the effect of structural social capital on the extent of help received. In addition, we also tested our model for potential reverse causality effect.

4.3.1 Effects of Knowledge Type

In order to have a clear understanding of the nature of knowledge being shared on these forums, we conducted a topic analysis of the messages posted by users. We combined all the posts for each user and then identified 200 topics, which were categorized into two broad groups: technical knowledge and context clarification knowledge. Technical knowledge includes discussions directly related to technical aspects of solution design. For instance, discussions about algorithms, programming codes, and statistical/predictive models and methods were evident in the messages. Some common keywords in the “technical knowledge” category were *neural networks*, *logistic regression*, *markov chain*, *algorithm*, *matlab*, *python*, *java*, *json*, *correlations*, and *matrix*. The “context clarification” category included general questions related to context, platform, and data. Some common keywords in this category were *team*, *leaderboard*, *rules*, *submission*, *variables*, and other keywords directly related to project-specific topics and variables (e.g., *hospital*, *days-in-hospital*, *claim*, *csv*, and *data*). We also identified the top five contributors to each topic. Out of the sample of 923 users, 49% contributed only technical knowledge, 48%

contributed only context clarification knowledge, and 3% contributed both technical and context clarification knowledge.

Tables 8a, 8b, 9a, and 9b show the effects of the quantity and quality of technical and context-clarification knowledge sharing on performance outcomes, respectively. Overall, our results show that solvers who contribute either type of knowledge (technical or context clarification) in any form (quantity or quality) are likely to have their sharing reciprocated. Furthermore, our results suggest that the quantity of contextual help given is more likely to impact performance through reciprocated help (i.e., help received). As for technical help rendered, both quantity and quality are likely to directly impact performance. Not surprisingly, solvers who share knowledge in a less competitive contest likely achieve more performance gains than those who share knowledge in a very highly competitive contest. Sharing technical knowledge (both quantity and quality) in a highly competitive contest can have a less positive or an adverse impact on solvers’ overall performance. A plausible explanation for this is that other competitors may be able to exploit the technical knowledge that has been shared, thus engendering a knowledge spillover effect. Therefore, solvers must be mindful when they share technical knowledge in extremely competitive contests.

The model controlled for contest-specific effects. It should be noted that solvers who contributed both technical and context-clarification knowledge were not included in the analysis. Thus, the results cannot be compared with the main analysis. The purpose of this analysis is mainly to provide some additional insight into the effects of the type of knowledge being shared.

Table 8a. SUR Results for Help Quantity (Technical Knowledge)

DV	Help received		Contest performance	
	Model 1	Model 2	Model 1	Model 2
Help given (quantity)	0.123 ***	3.928 ***	6.431 **	1556.958 ***
Help received			8.247	-0.317
Competition* HP.		-3.902 ***		-1589.073 ***
Contest performance _(k-1)	-0.000	-0.000	0.044 **	0.047 **
Submissions			3.028 ***	3.271 ***
Skill	0.060	0.056	33.828 ***	31.753 ***
Tenure	-0.052	-0.055	38.459 **	36.858 **
Observations	1019	1019	1019	1019
R-Squared	0.176	0.193	0.667	0.688

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 8b. SUR Results for Help Quantity (Context Knowledge)

DV	Help Received		Contest Performance	
	Model 1	Model 2	Model 1	Model 2
Help given (quantity)	0.103 ***	0.531 ***	2.364	-49.528
Help received			23.420 **	24.482 **
Competition* HP.		-0.458 **		55.496
Contest performance _(k-1)	0.000 *	0.000 **	-0.005	-0.006
Submissions			4.112 ***	4.039 ***
Skill	0.024	0.033	21.754 **	20.819 **
Tenure	-0.039	-0.031	13.826	12.615
Observations	592	592	592	592
R-Squared	0.405	0.410	0.710	0.710

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 9a. SUR for Help Quality (Technical Knowledge)

DV	Help Received		Contest Performance	
	Model 1	Model 2	Model 1	Model 2
Help given (quality)	0.065 ***	-0.650	24.718 ***	1974.320 ***
Help received			7.458	8.677
Competition* HP.		0.730		-1989.885 ***
Contest performance _(k-1)	-0.000	-0.000	0.039 *	0.037 *
Submissions			2.742 ***	2.761 ***
Skill	0.065	0.064	29.991 ***	31.390 ***
Tenure	-0.058	-0.056	35.601 *	30.326 *
Observations	1019	1019	1019	1019
R-Squared	0.100	0.101	0.681	0.703

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 9b. SUR for Help Quality (Context Knowledge)

DV	Help Received		Contest Performance	
	Model 1	Model 2	Model 1	Model 2
Help given (quality)	0.087 ***	8.156 **	41.046 ***	1657.077 ***
Help received			21.834 **	15.139
Competition* HP.		-8.295 ***		-1660.854 ***
Contest performance _(k-1)	0.000	0.000	-0.004	-0.003
Submissions			3.537 ***	3.611 ***
Skill	0.101 ***	0.101 ***	21.179 **	21.416 **
Tenure	-0.001	0.004	10.344	11.679
Observations	592	592	592	592
R-Squared	0.253	0.277	0.725	0.731

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 10a. OLS Regression Results (Social Capital in Help-Seeking Network)

DV	Betweenness centrality		Degree centrality	
Help given (quality)	24.7144 ***		0.7044 ***	
Help given (quantity)		27.1980 ***		0.7949 ***
Number of contests	23.5470 ***	6.9962 ***	0.4709 ***	0.7732 ***
R-squared	0.2132	0.4488	0.7453	0.4862

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 10b. Zero-Inflated Negative Binomial Regression (Indegree)

DV	Help received (quantity)	Help received (quality)
In-degree	0.0507 ***	0.0430 ***
Number of contests	0.0288 ***	0.0356 ***
Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$		

Table 10c. OLS Regression (Performance)

DV	Overall performance
Betweenness	1.83 ***
In-degree	128.50 ***
Out-degree	70.52 *
Number of contests	1427.14 ***
Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$	

4.3.2 Effects of Help Given on Help-Seeking Network

To further explore how help given influences the contributor's performance through help received, we analyzed the help-seeking network of users by taking into account the overall crowdsourcing platform (pool data). In this network, a directed edge between users A and B ($A \rightarrow B$) exists if B responds to A's request for help. In other words, the directed edge from A to B implies that A receives help from B. The results (see Table 10a) show that structural social capital accrues to those who share knowledge, eventually leading to superior overall performance. Following prior research, we used betweenness centrality to measure structural social capital (Faraj et al., 2015). Solvers who appear frequently on the geodesic (i.e., shortest path) between other pairs of solvers would enjoy the benefits of betweenness centrality (Freeman, 1979). Such solvers generally bridge otherwise isolated solvers and/or networks of solvers, thus being positionally privileged in terms of access to and control of information (Borgatti, Everett, & Johnson, 2013). In their articulation of various network measures of social capital, Borgatti, Jones, and Everett (1998) note the positive relation between betweenness centrality and social capital. Prior studies have also noted that betweenness centrality is highly correlated with structural holes (Everett & Borgatti, 2005) and that players (i.e., solvers) who bridge structural holes perform better (Zaheer & Bell, 2005). Consistent with theory, our results show that knowledge contribution is positively associated with the degree centrality of the contributor in the knowledge-seeking network.

We also tested the relationship between in-degree and help received (Table 10b). In our network, high in-degree (i.e., number of edges coming into a node) refers to someone who renders a lot of help. We found a positive and significant relationship between in-

degree and help received. In addition, our results show a significant positive relationship between centrality measures (in-degree & out-degree and betweenness) and performance (Table 10c). This further affirms that help given leads to performance gains through help received (i.e., enhanced structural social capital). Since we used directional networks, degree centrality was broken down into in-degree and out-degree in the performance analysis.

4.3.3 Correcting for Endogeneity

Our SUR estimates show a positive and significant relationship between the help given and performance scores, after controlling for other observed sources of performance. The primary challenge in estimating the performance impact of help given variables is that unobservable solver characteristics (such as innate quality or ability, ambition, intelligence, industriousness, and level of knowledge and skills) are likely to determine both the performance and help given. It is reasonable perhaps to expect high performers to be more willing to render help to others. The unobservable factors are known to the utility-maximizing solvers but not to others when they make their decisions regarding how much help to give. Those endowed with skills/innate ability and having positive expectations about their performance are more likely to give help. In other words, there are alternative sources of the relationship between performance and help provided to others; solvers differ not only in their performance and help provided to others but also in a variety of factors. Thus, the unobserved common determinants of performance and help given would confound the relationship. The measurement error can also cause a correlation between the error term and help given, particularly if there were errors in help given arising from misreporting or other reasons.

These sources of the endogeneity of help given imply that the parameter estimates obtained from the equations estimated by OLS and/or SUR are likely to be biased and inconsistent, as the disturbance term is correlated with help given. Specifically, given the existence of these sources of endogeneity, the performance impact of help given obtained by OLS or SUR will be biased upward if the correlation between these unobserved common factors and help given is positive because the coefficient of help given would not only pick up the performance effect of help given but also the hidden effect of the common unobservable factors subsumed in the error term that determines both performance and help given.

In light of the preceding discussions, we used an instrumental variable method to account for the endogeneity of help given. We control for the endogeneity of the help given variables by using an instrumental variable obtained from the average levels of the word counts variable for the other contests j ($j \neq k$) that user i participated in at time t (i.e., $\frac{1}{n} \sum_{j=1}^{n-1} \text{Wordcount}_j$). More specifically, suppose that Solver 1 participated in three contests, all with a deadline at time t (e.g., 2010). Furthermore, suppose that the word counts for these three teams are x_1, x_2 , and x_3 , respectively. Then, the level of the instrumental variable for Solver 1's first contest would be the mean value of x_2 and x_3 . Similarly, for Solver 1's second contest it would be the mean of x_1 and x_3 , and for the third contest it is the average of x_1 and x_2 . Word for word, instead of using the word count from the current observation itself, we used the mean of word counts computed using all observations for the solver except the one from the current observation at time t .

The validity of this instrumental variable hinges on the view that the average word counts will not affect the focal solver's performance directly but will affect it indirectly through its impact on the help given variables. Given this, we expect this instrument to be strongly related to help given but not correlated with the error terms of the outcome equations. Since these averages are for the other contests that solver i has participated in, they are relevant (i.e., they are correlated with the endogenous variable help given). The average word count variable is also arguably uncorrelated with the error term in the estimating equation for solver i 's k focal contest, as it is the average of the other contests in which a solver participated (i.e., they are unaffected by idiosyncratic factors of solver i 's focal contest). Thus, under the condition that average word count affects performance only through help given, as required for

a valid instrument, we employed an instrumental variable approach to examine the performance impact of help given.

To test the assumption that the average word count variable should only indirectly affect performance through help given, we ran performance on average word count, help given, and the exogenous variables. The results reveal that there is not a significant (statistically or economically) relationship between average word count and performance. The coefficient of average word count in these regressions is almost zero and statistically insignificant, indicating that average word count variable only indirectly affects performance.

We also tested the instrument relevance and exogeneity assumptions separately. The first stage regressions show that the instrument has a very strong impact on both quantity and quality of help given, indicating it is a strong or relevant instrument. Specifically, the t -stat for the coefficient on the instrumental variable in the quality of help given equation is 43.36 (3.3 or higher indicates that the instruments are relevant) and the t -stat for the coefficient on the instrumental variable in the quantity of help equation is 126.73 (3.3 or higher indicates that the instruments are relevant).

Since we only used one instrument, we were not able to perform an overidentification test (Sargan or Hansen J) because our model is exactly identified, resulting in a degree of freedom of zero. As discussed above, it is intuitive to argue that average word count would only affect the focal solver's performance indirectly through its impact on help given, but not directly. Thus, the instrument is both relevant and exogenous, indicating that it is valid. We summarize the results for quantity and quality of help given (with and without the interaction of competitive intensity) in Table 11a and Table 11b, respectively. The results generally support our main hypotheses.⁸

In order to perform the overidentification test to evaluate whether the instrument correlates with the error term and to check the robustness of the results, we also employed a recently advocated moment-based instrumental variable (IV) method (See Lewbel, 2012). The moment-based IV estimator exploits heteroscedasticity in the error term from the first-stage regressions (e.g., regressions of the help given variables to exogenous covariates), to control for the endogeneity even in the absence of exclusion restrictions. We employed this approach for the model without the interaction term. The results are illustrated in Table 11c.

⁸ We also estimated our models using three-stage least squares (3SLS). The results were almost identical.

Table 11a. 2SLS Results for Help Quantity

	(1)	(2)
DV: Performance	IV	IV_INT
Help given	8.3213***	430.9076***
Competition*HG		-449.8108***
Help received	5.3982***	11.0044***
Submission	2.0345***	2.2802***
Performance ₀	0.3448***	0.3425***
Tenure	-12.8972***	-11.1847***
Skill	21.5331***	22.6050***
Constant	1,018.6650***	1,017.0392***
Observations	31,693	31,693
R-squared	0.5122	0.5099
Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$		

Table 11b. 2SLS Results for Help Quality

	(1)	(2)
DV: Performance	IV	IV_INT
Help given	27.4911***	5,564.8836***
Competition*HG		-5,703.0788***
Help received	2.7701	13.3709***
Submission	1.8907***	2.4746***
Performance ₀	0.3400***	0.3230***
Tenure	-14.0686***	-12.9814***
Skill	20.3956***	20.7245***
Constant	1,031.8367***	1,056.9965***
Observations	31,693	31,693
R-squared	0.5203	0.4233
Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$		

Table 11c. 2SLS with Moment-based Instrumental Variable Method

	(1)	(2)
Variables	Moment-based IV quantity	Moment-based IV quality
Help given	10.6991***	31.9468***
Help received	3.9067	1.4981
Submission	1.9939***	1.8444***
Tenure	-13.0292***	-14.3333***
Skill	21.2801***	20.0677***
Performance ₀	0.3441***	0.3388***
Constant	1,237.0958***	1,240.1697***
Observations	31,693	31,693
R-squared	0.5119	0.5199
Cragg-Donald Wald F -stat	1.6e+04	4.5e+04
Sargan / Hansen J p -value	0.6338	0.8659
Notes: We used robust standard errors. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$		

To employ the Lewbel (2012) method, we specified two auxiliary estimating equations: one for the quality of help given and one for the quantity of help given. As pointed out by Lewbel (2012), the model is identified if the error terms in these first-stage regressions are heteroscedastic. Explicitly, there should be a correlation between some of the exogenous variables and the residuals in the first-stage regressions, but these variables should not be correlated with the covariance between the error terms of these two equations and the error term in the second-stage equation. According to Lewbel (2012), the residuals from the first-stage quality and quantity of help given equations multiplied by each of the mean-centered exogenous variables would be valid instruments for the quality and quantity of help given.

We used the Breusch and Pagan (1979) test to corroborate the existence of the heteroscedasticity in the first-stage regressions. We then multiplied the residuals from the auxiliary equations and the mean-centered submission, tenure, and skill variables as additional instruments. As illustrated in Table 11c, the results from the test of overidentifying restrictions (the Sargan or Hansen J test), employed to test for correlation between instruments and error term, justify the use of these additional instruments (i.e., the p -value is greater than 0.10, failing to reject the presence of instrument exogeneity).

In summary, we first used the average word count variable as an instrument for help given and estimated our model by employing the two-stage least squares method to causally identify the performance impact of help given. We believe that this instrument is intuitively both strong and exogenous. Furthermore, we also employed the moment-based IV method to conduct the test of overidentifying restrictions and checked the robustness of our results. All of our regressions also include competition-fixed effects that helped us rule out the likelihood that differences in competition-specific unobservable factors drove the relationship between performance and help given.

It is also likely that the solver performance is persistent, which can stem from solver heterogeneity or state dependence. In order to control the confluence of persistence in user performance, we included the initial performance of solvers (the average of the performance of each user in the first year they participated in a contest), which can control for the unobserved factors that remain relatively stable over time across solvers, such as attitude toward helping others. The inclusion of this variable is needed to control for the potential serial correlation. It is likely that the current levels of performance are a function of

lagged levels of performance. This variable, however, may be endogenous as well (although in this analysis we include the initial levels of the performance, not the lagged levels by one period, which is less susceptible to endogeneity). Thus, we also ran our models without including the initial levels of performance in our regressions. The coefficients of the main variables of interest still had the expected magnitude and significance.

4.4 Additional Robustness Tests

In addition to the above main results and additional analyses, we also conducted various supplementary tests using alternative measures and methods. All these checks consistently demonstrated that the results are robust.

First, we retested our main model including individual-level fixed effects. Individual-fixed effects account for the heterogeneity of solvers and help control for endogeneity issues that arise because of solver traits and other unobservable variables. In other words, we study the effect of a given solver (with the same skill and individual characteristics) participating in multiple projects with differing competitive intensity. Equations (7) and (8) describe the regression model with individual-fixed effects. Results are summarized in Tables 12a and 12b.

$$\begin{aligned} Performance_{ik} = & \alpha_{11} Help_Given_{ik} + \\ & \alpha_{12} Help_Received_{ik} + \\ & \alpha_{13} Help_Given_{ik} * CompetitionHHI_j + \\ & \alpha_{14} CompetitionHHI_j + \alpha_{15} Submission_{ik} + \\ & Solver_i + Contest_{1j} + \varepsilon_{1ik} \end{aligned} \quad (7)$$

$$\begin{aligned} Help_Received_{ik} = & \alpha_{21} Help_Given_{ik} + \\ & \alpha_{22} Help_Given_{ik} * CompetitionHHI_j + \\ & \alpha_{23} CompetitionHHI_j + Solver_i + Contest_{2j} + \\ & \varepsilon_{2ik} \end{aligned} \quad (8)$$

As expected, competitive intensity negatively moderates the effect of both quantity ($\alpha_{13} = -380.62, p < 0.01$) and quality ($\alpha_{13} = -1657.59, p < 0.01$) of help given on performance. Moreover, competitive intensity positively moderates the effect of quantity ($\alpha_{13} = 1.18, p < 0.01$) of help given on help received. Therefore, the results are robust when controlling for individual-specific effects.

Second, we separately investigated the effects of *quantity* and *quality of help given* on *quantity* and *quality of help received*. Table 13 summarizes these results. The results show that both *quantity* and *quality of help given* increased the *quantity and quality of help received* and improved contest performance.

Table 12a. SUR Results for Help Quantity with Individual-Fixed Effects

DV	Help received	Contest performance
Help given (quantity)	-0.9095 ***	375.1905 ***
Help received		6.3164 ***
Competition* HG.	1.1785 ***	-380.6183 ***
Submissions		3.7365 ***
Sample size	8559	8559
Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ We only considered individuals who have participated in at least five contests due to limitation of handling too many variables.		

Table 12b. SUR Results for Help Quality with Individual-Fixed Effects

DV	Help received	Contest performance
Help given (quality)	-0.2226	1645.4020 ***
Help received		4.1167 *
Competition* HG.	0.3772	-1657.5860 ***
Submissions		3.3628 ***
Sample size	8559	8559
Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ We only considered individuals who have participated in at least five contests due to limitation of handling too many variables.		

Table 13. Seemingly Unrelated Regression Results (Quantity and Quality Help Received)

DV	Help received				Contest performance			
	Model 1	Model 2	Model 3	Model 4	Model 1	Model 2	Model 3	Model 4
Help given (quantity)	0.799 ***	0.404 ***			9.232 ***	10.209 ***		
Help given (quality)			0.605 ***	0.497 ***			30.260 ***	30.818 ***
Help received (quantity)					2.332 ***		1.864 ***	
Help received (quality)						2.219 ***		1.093 **
Submissions					2.400***	2.393 ***	2.228 ***	2.245 ***
Skill	0.067 ***	0.086 ***	0.159 ***	0.105 ***	37.809 ***	37.795 ***	36.255 ***	36.386 ***
Tenure	0.060 *	0.095 **	0.079 **	0.091 ***	-4.648 *	-4.725 *	-6.128 **	-6.069 **
Sample size	31693	31693	31693	31693	31693	31693	31693	31693
Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$								

5 Discussion

With the rapid development of crowdsourcing technologies and practices, more data are continuously becoming available to researchers, affording them an opportunity to empirically investigate design and managerial issues in crowdsourcing. This study takes an important step in this direction by examining how helping behavior affects solver performance and reciprocal benefits in crowdsourcing communities.

The primary motivation for this study was to investigate how solvers benefit from sharing information and knowledge in contest forums. We found that solvers gained reciprocal benefits by

helping others, which, in turn, contributed to performance gains. However, the impact on performance of helping others may not be straightforward, for it may be contingent on the environment. While rendering help to peers in collaborative environments may positively influence performance, in competitive environments, our results suggested a weaker positive effect on performance. Indeed, in an extremely competitive contest, helping behaviors may result in knowledge spillover that could adversely affect performance. Our results also suggest a significant moderating effect of competitive intensity on the relationship between the quantity of help given and help received.

We also found that help received has a significant and positive relationship with performance. Furthermore, in our study, help received partially mediated the positive effect of help given on contest performance. Our analysis of the help-seeking network underlying a cooperative crowdsourcing environment shows that social capital accrues to those who render help, which, in turn, helps to improve their overall performance.

In order to provide a deeper understanding of the nature of knowledge shared, we performed a text analysis of users' posts on the discussion forum and identified two broad categories of knowledge: namely, technical and context clarification. Our results suggest that, compared to the quantity of context clarification knowledge shared, the quantity of technical knowledge shared has a greater impact on performance.

Furthermore, to check whether high-skill solvers are likely to withhold their knowledge in highly competitive contests, we plotted percentile shares of quantity and quality of help given against competitive intensity (see Appendix B). Our results do not reveal evidence of such behavior. Based on the plots, low-skill solvers are more inclined to share knowledge (quantity) in less competitive contests.

Our study utilized a unique dataset from Kaggle, a leading predictive analytics contest community. The data afforded us objective measures of help given, help received, and contest performance on every participant in a given contest over a period of time. This rich dataset allowed us to control for individuals' persistency and contest heterogeneity in estimating the effects of help given by the contributor.

5.1 Theoretical Implications

This study offers several theoretical contributions. First, the findings of this study extend the boundaries of the extant literature on knowledge sharing in online and collaborative communities to include the dynamics of helping behaviors in a cooperative environment. Specifically, our study contributes to the growing literature on prosocial behaviors in online communities. We examine such behaviors in the context of a crowdsourcing environment in which cooperation and competition simultaneously exist. Prior research has focused largely on collaborative communities (Huang & Zhang, 2016) and the motivation for members to participate in such communities (Wasko & Faraj, 2005). Furthermore, past studies, for the most part, implicitly assume that benefits accrue to those who share. By elucidating the dynamics of knowledge sharing (i.e., help given, help received, and performance) in such cooperative settings, our study makes a valuable contribution to the extant literature. Our study theorizes and clarifies the actual benefits of community contributions. We empirically demonstrate that helping behavior in

cooperative communities does have an important effect on contest performance as well as on the help that one receives.

Second, we demonstrate the enduring value of social exchange theory in explaining knowledge sharing behavior in a cooperative environment. Although social exchange theory explains the dual effects (costs and benefits) of social exchanges, it has rarely been applied in the same context to understand the overall impact of these dual forces. Prior studies have mainly focused on the benefits of social exchanges. For example, earlier studies suggest that prosocial behavior benefits contributors in noncompetitive environments (Grant & Sonnentag, 2010; Sproull et al., 2005; Xu et al., 2009). We have not only applied the theory in a new cooperative context to study the dual effect but have also used the same context to study how the effects vary according to the intensity of the competition. Specifically, we extend the boundaries of existing knowledge by demonstrating that rendering help in less competitive settings enhances performance, whereas the potential cost of time and effort in sharing knowledge in settings characterized by very high levels of competition could be unfavorable to performance.

Third, our additional analyses also shed light on the impact of structural social capital in cooperative settings. Previous studies have demonstrated the positive impact of structural social capital (operationalized as degree centrality in Wasko & Faraj, 2005) and as betweenness centrality in Faraj et al., 2015) on knowledge sharing (Wasko & Faraj, 2005) and leadership roles (Faraj et al., 2015) in online communities. Our study affirms the positive impact of both degree and betweenness centrality in a cooperative setting that is quite different from the one used in previous research.

Finally, our research also contributes to cooperative knowledge sharing literature in interorganizational settings (Gnyawali & Park, 2011; Ritala, 2012) by extending it to a completely different context—an online crowdsourcing community. The nature of cooperation is very different in online crowdsourcing communities, compared to interorganizational settings where knowledge sharing often happens between two rival organizations. The knowledge shared in such cases is not available to anyone other than the two parties involved in the process. Thus, the probability of knowledge spillover is minimal. In crowdsourcing contest setting, on the other hand, shared knowledge is available to all the competing solvers, thereby increasing the chances of knowledge spillover.

5.2 Practical Implications

Our findings have strong implications for the design of crowdsourcing platforms as well as for the participants

who engage in crowdsourcing competitions. Since the goal of crowdsourcing sites such as the one we studied (i.e., Kaggle) is to help organizations, researchers, and the general public find innovative and effective solutions to their challenging problems, it is imperative that the designers of such platforms foster a climate that encourages the exchange of ideas and knowledge. It is, therefore, important for them to understand both the motivations of participants to share knowledge and the effects or consequences of doing so. Our study provides insight into *help-giving* and *help-receiving* behaviors and their relationships with performance. Further, we clarify how the intensity of competition may affect these dynamics. Specifically, under conditions of low competitive intensity solvers will benefit from helping others. Understandably, some solvers may be hesitant to share knowledge in highly competitive environments because of the possibility of being outperformed by others who may benefit from the knowledge that is shared. Platform providers should use different techniques to motivate solvers to share knowledge in such environments. For example, platforms could introduce reward systems that take into account not only the frequency of knowledge contributions but also other factors such as quality of the contribution, knowledge type (e.g., context clarification vs. technical), and the level of competition. A solver takes much higher risks by sharing technical knowledge in an extremely competitive contest and is thus worthy of a higher reward than one who shares context clarification knowledge in a less competitive contest. On the positive side, helping others in highly competitive environments is more likely to be reciprocated.

In summary, our study shows that *help giving* has an impact on *help receiving* (i.e., reciprocity) and that the latter has a positive impact on performance. The upshot is that a platform that is conducive to the exchange of ideas and information can, in general, enhance the performance of participants, thereby leading to several benefits, including superior solutions to the problem at hand. Furthermore, designers of these platforms should find ways (for example, through gamification) or develop rules to motivate and engage participants so that they are more inclined to share their knowledge (e.g., Khansa et al., 2015) and form valuable ties with other members. Incorporating gaming elements (Simões, Redondo, & Vilas, 2013) such as rewards for sharing knowledge, leaderboards that display the top contributors of ideas, and formal recognition of those who helped the ultimate winner are examples of some of the things that designers could do to motivate even the most reticent of participants to share their knowledge.

Our results suggest that individuals can enhance their social and intellectual capital and increase their performance by actively engaging in information and

knowledge sharing in online crowdsourcing communities. The social capital that accrues to those who share knowledge in such environments may confer several benefits on them, including increased reciprocity, opportunities for new ideas, and the potential for cultivating ties that could be useful in other situations (i.e., “appropriability”).

Our findings also have broader implications for how knowledge is shared in other online and virtual communities. In organizational settings, project managers may also apply these insights to enhance the performance of individuals on their work teams. A key insight that our study offers is that knowledge sharing is beneficial to both collaborating and competing work teams. For instance, managers could facilitate technology-mediated knowledge and information sharing in inter- and intradepartmental work teams. Thus, our study has strong implications for both theory and practice.

6 Conclusion and Future Studies

In a hypercompetitive business environment where agility is paramount and the strategic advantage that organizations enjoy is fleeting, it is essential for firms to continually innovate to offer superior, differentiated products/services. At the heart of this imperative is the ability of companies to offer personalized and/or customized products/services based on keen insights derived from data. Data analytics, in general, and predictive modeling/analytics, in particular, are crucial in these endeavors. Given that there is an acute shortage of analytics talent and that many organizations lack the resources and/or capabilities to harness the enormous potential of data, companies often rely on crowdsourcing platforms to provide actionable insights. The contests on these platforms are intense, and knowledge and expertise can separate the winner(s) from the rest. It would seem to be a contradiction then that participants would be willing to expend time and effort to share ideas and knowledge with their rivals when doing so could hurt their chances of winning. Given the importance of understanding knowledge sharing under such cooperative conditions, it is surprising that there is little or no theoretically grounded, empirical research that elucidates this phenomenon. Our study expressly addresses this concern and represents a small but important step towards clarifying the dynamics of knowledge sharing in cooperative environments.

As outlined in the implications section, our research has strong implications for both theory and practice. In summary, our findings suggest that: (1) rendering help can increase the likelihood of receiving help; (2) the extent of help received can positively influence performance; (3) competitors who help others in a highly competitive contest, as opposed to one with low competitive intensity, are more likely to receive help,

but are less likely to see an improvement in their performance; (4) help received partially mediates the relationship between help given and performance; and (5) consistent with prior studies, structural social capital—measured in terms of degree and betweenness centrality—has a positive impact on knowledge sharing behaviors. Thus, our study makes a valuable contribution to the existing literature and provides insight to academics and practitioners alike.

As with many other empirical studies, this study has some limitations that future research could address. However, we believe these limitations are minor and do not detract from our contributions in any major way. First, our study is based on data from a data mining/predictive modeling crowdsourcing website. It would be useful to generalize our study to other types of crowdsourcing communities. Second, our data only

include information that is publicly available on the website. Passive participation or “lurking” could improve the intellectual performance of individual participants in the community. However, this is not visible to researchers. Thus, we encourage further investigation of helping behaviors using additional methods such as follow-up surveys. Third, we measured help received (by a solver) according to the number of forum replies to a question posed in a thread initiated by the solver. It is conceivable that some of the forum responses may be to related questions asked by solvers while responding to the original question, and may, therefore, inflate the help received count. Fourth, while our study provides some evidence of the long-term effects of knowledge contributions, it would be interesting to further investigate the consequences of the effects of help given that go beyond help received and performance.

References

- Archak, N. M. (2010). Glory and cheap talk: Analyzing strategic behavior of contestants in simultaneous crowdsourcing contests on TopCoder.com. *Proceedings of the 19th International Conference on World Wide Web*, 21-30.
- Ardichvili, A., Page, V., & Wentling, T. (2003). Motivation and barriers to participation in virtual knowledge-sharing communities of practice. *Journal of Knowledge Management*, 7(1), 64-77.
- Aronson, E., Wilson, T. D., & Akert, R. M. (2006). *Social Psychology*. http://wps.prenhall.com/hss_aronson_socpsych_6/64/16429/4205880.cw/index.html
- Bateman, P. J., Gray, P. H., & Butler, B. S. (2011). The impact of community commitment on participation in online communities. *Information Systems Research*, 22(4), 841-854.
- Blau, P. M. (1964). *Exchange and power in social life*. Wiley.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84.
- Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2013). *Analyzing social networks*. SAGE.
- Borgatti, S. P., Jones, C., & Everett, M. G. (1998). Network measures of social capital. *Connections*, 21(2), 27-36.
- Boudreau, K. J., Lacetera, N., & Lakhani, K. R. (2011). Incentives and problem uncertainty in innovation contests: An empirical analysis. *Management Science*, 57(5), 843-863.
- Brandenburger, A. M., & Nalebuff, B. J. (1996). *Co-opetition*. Doubleday.
- Breusch, T., & Pagan, A. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47, 1287-1294.
- Bullinger, A. C., Neyer, A. K., Rass, M., & Moeslein, K. M. (2010). Community based innovation contests: Where competition meets cooperation. *Creativity and Innovation Management*, 19(3), 290-303.
- Carayannis, E. G., & Alexander, J. (1999). Winning by co-opeting in strategic government-university-industry R&D partnerships: The power of complex, dynamic knowledge networks. *The Journal of Technology Transfer*, 24(2), 197-210.
- Chang, H. H., & Chuang, S. (2011). Social capital and individual motivations on knowledge sharing: Participant involvement as a moderator. *Information & Management*, 48(1), 9-18.
- Chen, C. J., & Hung, S. W. (2010). To give or to receive? Factors influencing members' knowledge sharing and community promotion in professional virtual communities. *Information & Management*, 47(4), 226-236.
- Chen, J., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2014). How community feedback shapes user behavior. *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*. <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8066/8104>
- Chen, J., Xu, H., & Whinston, A. B. (2011). Moderated online communities and quality of user-generated content. *Journal of Management Information Systems*, 28(2), 237-268.
- Chiu, C. M., Hsu, M. H., & Wang, E. T. (2006). Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision Support Systems*, 42(3), 1872-1888.
- Constant, D., Kiesler, S., & Sproull, L. (1994). What's mine is ours, or is it? A study of attitudes about information sharing. *Information Systems Research*, 5(4), 400-421.
- Crowdsourcing.org. (2015). *10 of the 11 best global brands use creative crowdsourcing*. <http://www.crowdsourcing.org/editorial/10-of-the-11-best-global-brands-use-creative-crowdsourcing-/16935>
- Cyr, S., & Choo, C. W. (2010). The individual and social dynamics of knowledge sharing: An exploratory study. *Journal of Documentation*, 66(6), 824-846.
- Debortoli, S., Junglas, I., Müller, O., & vom Brocke, J. (2016). Text mining for information systems researchers: An annotated topic modeling tutorial. *Communications of the Association for Information Systems*, 39(1), 100-135.
- Dissanayake, I., Zhang, J., & Gu, B. (2015). Task division for team success in crowdsourcing contests: Resource allocation and alignment effects. *Journal of Management Information Systems*, 32(2), 8-39.
- Dissanayake, I., Zhang, J., Yasar, M., & Nerur, S. P. (2018). Strategic effort allocation in online innovation tournaments. *Information & Management*, 55(3), 396-406.
- Efron, B. (1979). Bootstrap. Another look at the Jackknife. *Annals of Statistics*, 7(1), 1-26.

- Everett, M., & Borgatti, S. P. (2005). Ego network betweenness. *Social Networks*, 27(1), 31-38.
- Faraj, S., Kudaravalli, S., & Wasko, M. (2015). Leading collaboration in online communities. *MIS Quarterly*, 39(2), 393-412.
- Faraj, Samer, & Johnson, S. L. (2011). Network exchange patterns in online communities. *Organization Science*, 22(6), 1464-1480.
- Faraj, Samer, von Krogh, G., Monteiro, E., & Lakhani, K. R. (2016). Special Section Introduction: Online Community as Space for Knowledge Flows. *Information Systems Research*, 27(4), 668-684.
- Fischer, G., Scharff, E., & Ye, Y. (2004). Fostering social creativity by increasing social capital. In M. Huysman & V. Wulf (Eds.) *Social capital and information technology* (pp. 355-399). MIT Press.
- Flynn, F. J. (2003). How much should I give and how often? The effects of generosity and frequency of favor exchange on social status and productivity. *Academy of Management Journal*, 46(5), 539-553.
- Forgeard, M. J., & Mecklenburg, A. C. (2013). The two dimensions of motivation and a reciprocal model of the creative process. *Review of General Psychology*, 17(3), 255.
- Freeman, L. C. (1979). Centrality in social networks conceptual clarification. *Social Networks*, 1(3), 215-239.
- Fulk, J., Flanagan, A. J., Kalman, M. E., Monge, P. R., & Ryan, T. (1996). Connective and communal public goods in interactive communication systems. *Communication Theory*, 6(1), 60-87.
- Füller, J., Hutter, K., Hautz, J., & Matzler, K. (2014). User roles and contributions in innovation-contest communities. *Journal of Management Information Systems*, 31(1), 273-308.
- Ghobadi, S., & D'Ambra, J. (2011). Cooperative knowledge sharing: An analytical review of literature. *The Electronic Journal of Knowledge Management*, 9(4), 307-317.
- Gnyawali, D. R., & Park, B.-J. (Robert). (2011). Co-opetition between giants: Collaboration with competitors for technological innovation. *Research Policy*, 40(5), 650-663.
- Grant, A. (2013). *Give and take: Why helping others drives our success*. Penguin Books.
- Grant, A. M. (2008). Does intrinsic motivation fuel the prosocial fire? Motivational synergy in predicting persistence, performance, and productivity. *Journal of Applied Psychology*, 93(1), 48-58.
- Grant, A. M., & Berry, J. W. (2011). The necessity of others is the mother of invention: Intrinsic and prosocial motivations, perspective taking, and creativity. *Academy of Management Journal*, 54(1), 73-96.
- Grant, A. M., & Sonnentag, S. (2010). Doing good buffers against feeling bad: Prosocial impact compensates for negative task and self-evaluations. *Organizational Behavior and Human Decision Processes*, 111(1), 13-22.
- Hardy, C. L., & Van Vugt, M. (2006). Nice guys finish first: The competitive altruism hypothesis. *Personality and Social Psychology Bulletin*, 32(10), 1402-1413.
- Howard, A. (2010). *Crowdsourcing national challenges with the new challenge.gov*. http://readwrite.com/2010/08/31/crowdsourcing_national_challenges_with_the_new_challengegov
- Huang, P., & Zhang, Z. (2016). Participation in open knowledge communities and job-hopping: Evidence from enterprise software. *MIS Quarterly*, 40(3), 785-806.
- Hutter, K., Hautz, J., Füller, J., Mueller, J., & Matzler, K. (2011). Communitation: The tension between competition and collaboration in community-based design contests. *Creativity and Innovation Management*, 20(1), 3-21.
- Ilvonen, I., & Vuori, V. (2013). Risks and benefits of knowledge sharing in co-opetitive knowledge networks. *International Journal of Networking and Virtual Organisations*, 13(3), 209-223.
- Kankanhalli, A., Tan, B. C., & Wei, K. K. (2005). Contributing knowledge to electronic knowledge repositories: An empirical investigation. *MIS Quarterly*, 29(1), 113-143.
- Khansa, L., Ma, X., Liginlal, D., & Kim, S. S. (2015). Understanding members' active participation in online question-and-answer communities: A theory and empirical analysis. *Journal of Management Information Systems*, 32(2), 162-203.
- Kohler, T. (2015). Crowdsourcing-based business models: How to create and capture value. *California Management Review*, 57(4), 63-84.
- Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mis-measured and endogenous regressor models. *Journal of Business & Economic Statistics*, 30(1), 67-80

- Lukyanenko, R., Parsons, J., Wiersma, Y., Wachinger, G., Huber, B., & Meldt, R. (2017). Representing crowd knowledge: Guidelines for conceptual modeling of user-generated content. *Journal of the Association for Information Systems*, 18(4), 297-339.
- Mangalaraj, G., Nerur, S., Mahapatra, R., & Price, K. (2014). An experimental investigation of the role of design patterns and collaboration. *MIS Quarterly*, 38(1), 249-274.
- Meng Ma, & Agarwal, R. (2007). Through a glass darkly: Information technology design, identity verification, and knowledge contribution in online communities. *Information Systems Research*, 18(1), 42-67.
- Mo, J., Zheng, Z., & Geng, X. (2011). *Winning crowdsourcing contests: A micro-structural analysis of multi-relational networks*. Presented at the 5th China Summer Workshop on Information Management, Harbin, China.
- Nahapiet, J. (2008). The role of social capital in inter-organizational relationships. In S. Cropper, M. Ebers, C. Huxham, and P. Smith Ring (Eds.). *The Oxford handbook of inter-organizational relations* (pp. 580-606). Oxford University Press.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Prentice-Hall.
- Ozer, M. (2011). A moderated mediation model of the relationship between organizational citizenship behavior and job performance. *Journal of Applied Psychology*, 96(6), 1328-1336.
- Paulin, D., & Suneson, K. (2012). Knowledge transfer, knowledge sharing and knowledge barriers-Three blurry terms in KM. *The Electronic Journal of Knowledge Management*, 10(1), 81-91.
- Phang, C., Kankanhalli, A., & Sabherwal, R. (2009). Usability and sociability in online communities: A comparative study of knowledge seeking and contribution. *Journal of the Association for Information Systems*, 10(10), 721-747.
- Ritala, P. (2012). Coopetition strategy: When is it successful? Empirical evidence on innovation and market performance. *British Journal of Management*, 23(3), 307-324.
- Saxton, G. D., Oh, O., & Kishore, R. (2013). Rules of crowdsourcing: Models, issues, and systems of control. *Information Systems Management*, 30(1), 2-20.
- Schlagwein, D., & Bjorn-Andersen, N. (2014). Organizational learning with crowdsourcing: The revelatory case of LEGO. *Journal of the Association for Information Systems*, 15(11), 754-778.
- Seltzer, E. (2012). Citizen participation, open innovation, and crowdsourcing: Challenges and opportunities for planning. *Journal of Planning Literature*, 28(1), 3-18.
- Simões, J., Redondo, R. D., & Vilas, A. F. (2013). A social gamification framework for a K-6 learning platform. *Computers in Human Behavior*, 29(2), 345-353.
- Sproull, L., Conley, C., & Moon, J. Y. (2005). Prosocial behavior on the net. Y. Amichai-Hamburger (Ed.), *The social net: understanding human behavior in cyberspace* (pp. 139-161). Oxford University Press
- Sullivan. (2010). E. A group effort: More companies are turning to the wisdom of the crowd to find ways to innovate. *Marketing News*, 44(2), 22-28.
- Terwiesch, C., & Xu, Y. I. (2008). Innovation contests, open innovation, and multiagent problem solving. *Management Science*, 54(9), 1529-1543.
- Tsai, H. T., & Bagozzi, R. P. (2014). Contribution behavior in virtual communities: Cognitive, emotional, and social influences. *MIS Quarterly*, 38, 143-163.
- Tsai, W. (2002). Social structure of "coopetition" within a multiunit organization: Coordination, competition, and intraorganizational knowledge sharing. *Organization Science*, 13(2), 179-190.
- Wasko, M., & Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly*, 29(1), 35-57.
- Weinstein, N., & Ryan, R. M. (2010). When helping helps: Autonomous motivation for prosocial behavior and its influence on well-being for the helper and recipient. *Journal of Personality and Social Psychology*, 98(2), 222-244.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT Press.
- Xu, B., Jones, D. R., & Shao, B. (2009). Volunteers' involvement in online community based software development. *Information & Management*, 46(3), 151-158.
- Yan, Z., Wang, T., Chen, Y., & Zhang, H. (2016). Knowledge sharing in online health communities: A social exchange theory

- perspective. *Information & Management*, 53(5), 643-653.
- Yang, Y., Chen, P. Y., & Banker, R. (2010). Impact of past performance and strategic bidding on winner determination of open innovation contest. *Proceedings of the 21st Workshop on Information Systems and Economics*.
- Yang, Y., Chen, P. Y., & Pavlou, P. (2009). Open innovation: Strategic design of online contests. *Proceedings of the Workshop on Information Systems and Economics*.
- Zaheer, A., & Bell, G. G. (2005). Benefiting from network position: Firm capabilities, structural holes, and performance. *Strategic Management Journal*, 26(9), 809-825.

Appendix A

Table A1. Topic Analysis

Topics	Main theme	Collaboration/ non-competition	Competition	Co-opetition
knowledge management sharing organizational organizations technology factors framework analysis organization characteristics business collaborative role practice successful practices information understanding empirical	Organizational knowledge sharing (collaborative)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
network firms knowledge networks capabilities market firm business role structure exchange competition competitive resources ties industries units cooperation view transformation	Firm networks (competitive, cooperation)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
motivation intention intrinsic extrinsic altruism selfefficacy motivations factors rewards contributors recipients reward satisfaction injustice enjoyment justice contextual helping levels lurkers	Motives and intentions (altruism, enjoyment)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
outsourcing barriers district industrial success strategies relationships perspectives assimilation protection closely aspects vcops ipr realworld local contingency positions productivity smos	Outsourcing	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
practices objects boundary activities saas brokering requirements activity service release projects shared strategies boundaries software trust indepth dynamics coworkers customers	Customer service (cocreation)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
social users information media capital snss creativity facebook creative status sites bookmarking contributions attacks security capabilities data engagement survey library	Social Media	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
software oss developers project source open projects network firms community participants development work openness externalities formal strategic production model volunteer	Open source software development	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
systems public framework health web architectural infrastructure architecture development pair models data einnovation configuration highlevel services cqa files solo mining	Healthcare and software-development (pair, solo)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
development software process knowledge learning system processes companies systems engineering tools integration time information work exchange concepts networks explicit task	Software development (exchange, integration)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
information engine users travel inferior agents service share knowledgesharing superior answer internet people market competition database mobile answers quality engines	Knowledge sharing services (search engines)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
public behaviour organizations sector brand share consumers consumer cost specific capability costs behaviours employees involvement malaysian iwe moderating innovativeness practitioners	Public sector organizations and brand communities	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
cops group transfer groups efficiency organization experience status consortia participants agricultural benchmarking department cop agile private consulting productivity icts corporate	Communities of practice	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
community virtual communities online members social users participation interaction contribution reciprocity information behavior quality vcs attitudes professional learning internet understanding	Virtual communities	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
online analysis forum content core members structures specification map communication usage debate distributed providers time community innovative als farmers expert	Online communities (forum)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

knowledge sharing social model trust effect theory data members factors influence relationship perceived individuals theoretical knowledgesharing behaviors practical behavior positive	Motives and intentions (e.g. trust, perceived expertise, perceived relative advantages, perceived ease of use, perceived credibility)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
competition learning cooperation cooperative coopetition competitive model firms alliances advantage alliance effective simultaneous crossfunctional levels task open resource control interorganizational	Interorganizational/ cross- functional groups (cooperative, coopetition, competitive)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
web students identity activities dimensions selfefficacy intention seeking expectations knowledgewithholding education crossorganizational pressure mechanism extent sites sense ttest samples sociability	Motives and intentions (identity, self-efficacy)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
communication supply chain strategy collaboration type firm leadership tools digital focal features chains professional interviews interactive identity interactions supplier view	Supply chain (collaboration & communication)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
team teams coordination project activities communication systems development competition projects gsd dispersed group collaboration isd cultural hrm teamwork distributed global	Project team (coordination, collaboration, competition)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
innovation open innovations analysis clustering imp governance internal concept ideas returns incentives forms approaches closed remixing form source networking pci	Open innovation (private- collective innovation—PCI)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Appendix B

The percentile shares are based on the proportion of quantity and quality of help given that falls into each percentile of competitive intensity. We grouped observation into two groups based on solvers' skill scores, namely, high skill and low skill. Specifically, the top 25% (or above the 75 percentile) of skill scores were considered high-skill solvers, while the bottom 75% were regarded as low-skill solvers.

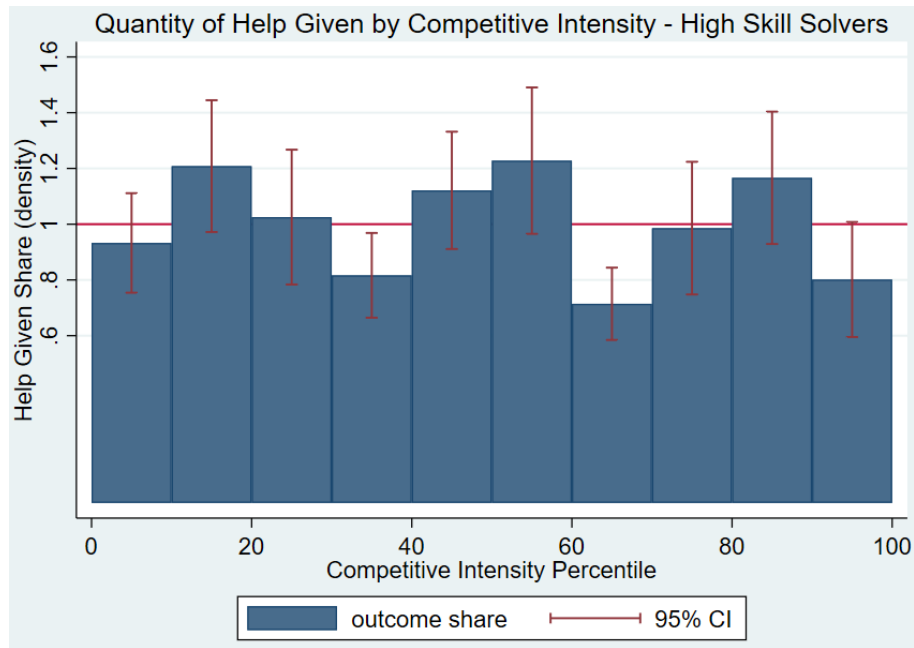


Figure B1. Quantity of Help Given by Competitive Intensity (High-Skill Solvers)

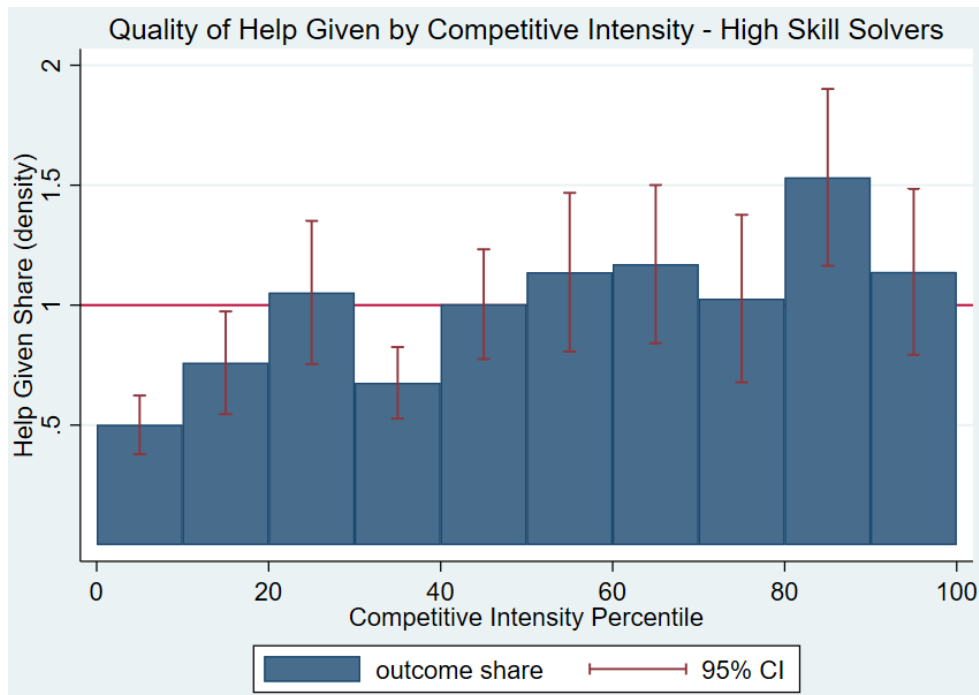


Figure B2. Quality of Help Given by Competitive Intensity (High-Skill Solvers)

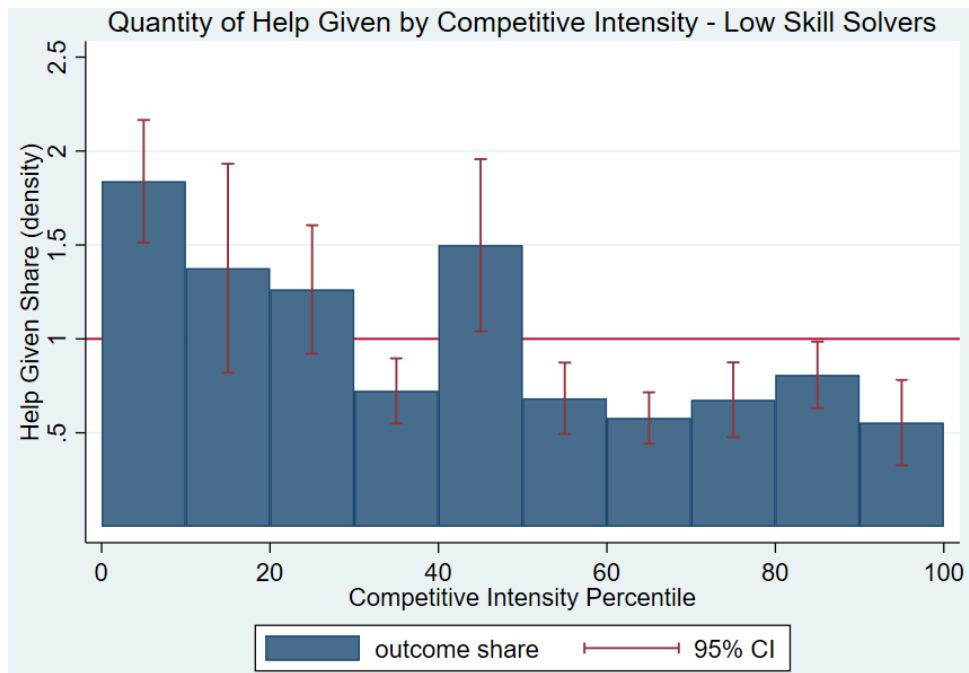


Figure B3. Quantity of Help Given by Competitive Intensity (Low-Skill Solvers)

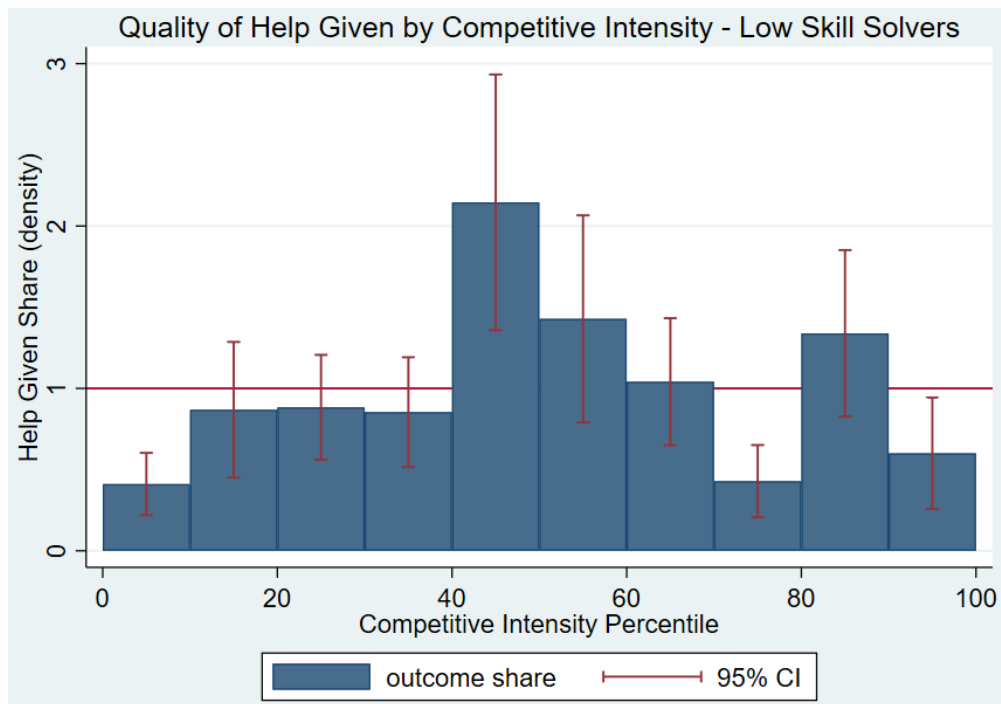


Figure B4. Quality of Help Given by Competitive Intensity (Low-Skill Solvers)

About the Authors

Indika Dissanayake is an assistant professor of information systems and supply chain management at the Bryan School of Business and Economics, University of North Carolina Greensboro. She received her PhD in information systems from the College of Business Administration at the University of Texas at Arlington. Her research interests include crowdsourcing, sharing economies, social media, and virtual communities. Her research has appeared in *Journal of Management Information Systems*, *Journal of the Association for Information Systems*, *Information & Management*, and *Information Systems Management*.

Sridhar Nerur is the Goolsby-Virginia and Paul Dorman Endowed Chair in Leadership and professor of information systems at the University of Texas at Arlington (UTA). He is the chair of the Graduate Studies Committee on Business Analytics at UTA. His research has been published in premier journals/magazines such as *MIS Quarterly*, *Strategic Management Journal*, *Communications of the ACM*, *European Journal of Information Systems*, *Information & Management*, *IEEE Software*, *Journal of the Association for Information Systems*, and *Journal of International Business Studies*. He has served on the editorial boards of the *European Journal of Information Systems* and the *Journal of the Association for Information Systems*. His research and teaching interests include social networks, machine learning/AI/deep learning, text analytics, neuroeconomics, and agile software development.

Jingguo Wang is a professor of information systems at the University of Texas at Arlington. He earned his PhD in management science and systems from the University at Buffalo, the State University of New York. His current research interests are in the areas of information systems, cybercrime, and information security. His work has been published in *MIS Quarterly*, *Information Systems Research*, *Journal of the Association for Information Systems*, *Journal of Management Information Systems*, among others. His research has been supported by the University of Texas at Arlington and the National Science Foundation.

Mahmut Yasar is a professor of economics at the College of Business Administration, the University of Texas at Arlington (UTA). Before coming to UTA, he taught at the Goizueta Business School Department of Finance and the Department of Economics at Emory University. His primary research interests center around the microeconomics of trade and investment, productivity, knowledge and technology transfer, innovation, and applied microeconometrics. He has also worked on issues in environmental economics and corporate finance. His research has appeared in journals such as *Information & Management*, *Journal of Business and Economic Statistics*, *Journal of International Economics*, and *World Development*. He served as an associate editor of the *International Economic Journal* and *Emerging Markets Finance and Trade*.

Jie Jennifer Zhang is a professor of information systems in the College of Business at the University of Texas at Arlington. She received her PhD in computer information systems from the University of Rochester. She employs analytical and empirical techniques to closely examine issues in advanced and business applications of information technologies, for example, trends and business impacts of social media, web advertising, consumer online behaviors, online channel selection and pricing strategies, software licensing policies, and business alliances. Her research has appeared in *MIS Quarterly*, *Journal of Economics and Management Strategies*, *Information Systems Research*, *Journal of Management Information Systems*, and *Decision Support Systems*, among other outlets. She has received grants and awards recognizing her research impacts.

Copyright © 2021 by the Association for Information Systems. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and full citation on the first page. Copyright for components of this work owned by others than the Association for Information Systems must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or fee. Request permission to publish from: AIS Administrative Office, P.O. Box 2712 Atlanta, GA, 30301-2712 Attn: Reprints, or via email from publications@aisnet.org.