

Accepted for publication in Journal of the American Society for Information Science and Technology (JASIST)

**You Scratch Someone's Back and We'll Scratch Yours:
Collective Reciprocity in Social Q&A Communities**

Philip Fei Wu, Surrey Business School, University of Surrey, Guildford, United Kingdom.

Email: f.wu@surrey.ac.uk

Nikolaos Korfiatis, Institute for Informatics and Mathematics, Goethe University Frankfurt,

Frankfurt am Main, Germany. Email: korfiatis@em.uni-frankfurt.de

Abstract

Taking a structuration perspective while integrating reciprocity research in economics, this study examines the dynamics of reciprocal interactions in social Q&A communities. We postulate that individual users of social Q&A constantly adjust their kindness into the direction of the observed benefit and effort of others. Collective reciprocity emerges from this pattern of conditional strategy of reciprocation and helps form a structure that guides the very interactions that give birth to the structure. Based on a large sample of data from Yahoo! Answers, our empirical analysis supports the collective reciprocity premise, showing that the more effort (relative to benefit) an asker contributes to the community, the more likely the community will return the favor. On the other hand, the more benefit (relative to effort) the asker takes from the community, the less likely the community will cooperate in terms of providing answers. We conclude that a structuration view of reciprocity sheds light on the duality of social norms in online communities.

Introduction

The increasing popularity of online knowledge-sharing websites has attracted much academic interest in recent years. Past research has looked at general knowledge-sharing sites such as Wikipedia (Nov, 2007) and Answerbag (Gazan, 2010), as well as specialized professional electronic networks (e.g., Chiu, Hsu, & Wang, 2006; Wasko, Teigland, & Faraj, 2009). One of the baffling questions to researchers is the motivation for voluntary contribution in these online spaces, for it seems irrational to contribute time and effort to help strangers online (Wasko & Faraj, 2005). Among many proposed motivational accounts, reciprocity is one that has been frequently mentioned in online community studies (Constant, Sproull, & Kiesler, 1996; Kankanhalli, Tan, & Wei, 2005; Kollock, 1999; Wasko & Faraj, 2005). Nevertheless, there is surprisingly little empirical evidence of reciprocity as an established norm of behavior at the community level in the online spaces. In other words, the existence of reciprocity (or lack thereof) has been researched mostly from the perspective of individual psychology (typically through a small sample of research participants), while the macro-level, collective patterns of reciprocation in online communities are largely unexamined. The aim of this research is to define and understand *collective reciprocity* by examining interaction patterns in social question and answer (SQA) communities – a prominent type of online knowledge-sharing community. Drawing on structuration theory in sociology and game theory in economics, we propose that members of SQA communities follow a “virtual order” (Giddens, 1984) of conditional cooperation that governs the reciprocal interactions in these online spaces. Empirical analyses based on a large sample of questions and answers in Yahoo! Answers largely support our

theoretical claims.

This paper is organized as follows: We first discuss the characteristics of SQA within the intellectual context of communities of practice (CoPs) research. Then we propose collective reciprocity as a conceptual cornerstone for understanding reciprocal interactions in SQA. Next, we present data collected from Yahoo! Answers and test hypotheses based on the collective reciprocity premise. Finally, we discuss how this work contributes to scholarly understanding of reciprocal behavior in large-scale social networks.

Are SQA Websites Communities of Practice?

With more and more people turn to the Web for information, opinion, and advice, SQA websites have been thriving in the past decade. Popular SQA sites such as Yahoo! Answers and WikiAnswers each attracts tens of millions of unique monthly visitors (comScore, 2012). There also exist many specialized SQA sites such as Avvo (for asking legal and health questions of lawyers and doctors) and MathOverflow (for mathematicians). These SQA websites may differ in design and user demographics, but they all provide an online space where users express their information need in the form of natural language in anticipating public response (Shah, Oh, & Oh, 2009). In addition, most SQA sites provide facilitating features that allow a user to create an online profile and to rate the quality of answers.

Are SQA sites online CoPs? Decades of academic debate in community studies has shown the difficulty of determining what is a community, and this task becomes even more daunting now as

Internet-based communication technologies afford ubiquitous and large-scale social interactions. Indeed, some online social networking services have millions of users who are geographically dispersed and may never meet face-to-face. It is therefore difficult for those people to develop a “sense of community” in its traditional sense (McMillan & Chavis, 1986; Reich, 2010). Wellman et al. (2003) go even further to claim that greater online access and interactions lead to networked individualism rather than community solidarities. A more recent stream of research, however, has been using the term “community” to describe the large online spaces such as MySpace (boyd, 2006), Youtube (Rotman & Preece, 2010), and Yahoo! Answers (Rosenbaum & Shachaf, 2010).

A similar discrepancy arises when we look closely at the concept of CoPs. The term was coined by Lave and Wenger (1991) to refer to sets of professionals within a well-bounded community who have similar responsibilities. Most studies in the knowledge management field use the term in its original meaning (Cox, 2005; Wenger, McDermott, & Snyder, 2002). However, Murillo (2008) applied Wenger’s core constructs to assessing whether Usenet groups are CoPs and concluded that “true communities of practice” can emerge spontaneously in the online social spaces without face-to-face interaction. Hara, Shachaf, and Stoerger (2009) also try to extend the CoPs framework from organizational settings to open online communities that are not constrained by any organizational context. Brown and Duguid (2001) draw a distinction between “network of practice” and “community of practice”: Networks of practice form from informal and emergent social relations in which individuals interact for practice-related reasons, while CoPs is a specialized subset of networks of practice in which co-located individuals build strong, communal ties. In a similar vein, Wasko and Faraj (2005) differentiate a CoP from an electronic

network of practice, arguing that the former requires strong interpersonal ties and direct reciprocity while the latter consists of loosely knit members with little direct interaction.

For the purpose of this study, we adopt Bruckman's (2006) proposition of conceptualizing community as a prototype category in which some social groups are prototypical examples of the category whereas others are less typical examples. Just like a robin is a better example of a bird than a penguin, an online health support group (e.g., Bob's Kneeboard studied by Maloney-Krichmar & Preece, 2005) may have a higher "degree of prototypicality" (Rosch, 1978) in the community category than a general purpose SQA group such as Yahoo! Answers. This nuanced view of the community category helps us to step back from heated debate about the connotations of the term "community" and focus on mechanisms of online social interactions. Once we accept the notion that there are various degrees of membership in a category rather than a clear-cut inclusion–exclusion boundary, we may begin to examine the characteristics of SQA sites in the theoretical lens of CoPs and identify their resemblance to CoPs.

Collective Reciprocity

In SQA communities, an asker seeks knowledge from "kind strangers" (Constant et al., 1996) who answer the question without anticipating the asker to return the favor. In fact, a large number of questions posted on SQA sites are answered by a small army of "answer persons" (Turner, Smith, Fisher, & Welser, 2005) whose main goal seems to only help others, whereas the majority of users are one-time askers who have never contributed any answers (Adamic, Zhang, Bakshy, & Ackerman, 2008). Prior literature suggests that the asymmetry of contribution is a

common phenomenon across various types of online communities (Nonnecke & Preece, 2000; Shachaf, 2009; Whittaker, 1998).

The voluntary contribution in online communities seems irrational from the economic perspective, as the contributors receive little immediate rewards. Some early Internet studies attempted to understand the motivations for voluntary contribution through the lens of gift economy (Barbrook, 1998; Raymond, 1998; Rheingold, 1993). For example, Rheingold (1993) suggests that virtual communities operate by the norms of gift economy, in which “people do things for one another out of a spirit of building something between them, rather than a spreadsheet-calculated quid pro quo” (p. 59). But Kollock (1999) argues that the acts of information sharing and knowledge contributing online are different from traditional gift exchange, as information is a special type of gift which can be given to more than one recipient without added cost. A gift of information is usually not from one individual to another but from one to many anonymous recipients in an open network. Thus, from the recipients’ point of view, they may tap into the gifts when in need, without bearing an obligation of repaying the gift givers. Furthermore, as an individual can join or leave an online community at any time, the obligation of reciprocity is almost impossible to enforce.

Some sociologists have also argued that information sharing in online spaces is likely to form networked individualism rather than communal reciprocity (Wellman, 2001; Reich, 2010).

Unlike symmetric social networks such as Facebook in which friendship is always directly reciprocal, the question–answer relationship in large-scale SQA communities is usually one-directional and non-reciprocal. The lack of direct reciprocity means that there is no assurance

that an asker gets a good answer, nor does an answerer know if her favor will ever be returned. As a result, Wasko and Faraj (2005) and other researchers (e.g., Chiu et al., 2006; Lampel & Bhalla, 2007) turn to social capital elements such as reputation and status for explaining the voluntary contributions in online social spaces.

Yet, there seems to be some kind of equilibrium in information giving and taking that keeps the online SQA communities healthy and sustainable. Indeed, if everyone consumes knowledge contributed by others without giving back, SQA sites such as Yahoo! Answers would not have thrived for many years. Kollock (1999) contends that a balanced reciprocity with each individual participant is not possible, but there should be a rough balance with a community as a whole over time. Similarly, Constant et al. (1996) and Kankanhalli et al. (2005) argue that while a contributor provides a gift of information without anticipating immediate reciprocation, a future reciprocation is somewhat expected. Wasko et al. (2009) hypothesize that the creation of knowledge in electronic networks is characterized by a pattern of generalized exchange, in which a giving is reciprocated by a third party rather than the recipient.

We conceptualize the generalized exchange observed in Wasko et al. (2009) as *collective reciprocity*. In line with recent economic literature, we view reciprocity as a behavioral response to perceived kind or unkind intentions in *bilateral* interactions (Falk & Fischbacher, 2006; Stanca, 2010). Collective reciprocity, therefore, is a collective patterning in responding to kind or unkind intentions in *multilateral* interactions in a social network. Just as in bilateral interactions, fairness is desirable in multilateral interactions where kind intentions are rewarded and unkind intentions are punished (Cox, Friedman, & Gjerstad, 2007; Falk, Fehr, & Fischbacher, 2008;

Fehr & Gächter, 2000).

While “collective reciprocity” and “generalized exchange” both refer to the social exchange in an n -person network, we prefer the former term for the following reasons. First, generalized exchange is a core concept in social exchange theory (Ekeh, 1974) that focuses on explaining social dilemma puzzles under the assumption of rational choice (e.g., Yamagishi & Kiyonari’s (2000) study of in-group favoritism in a Prisoner’s Dilemma game and Takahashi’s (2000) model of “pure-generalized exchange” in computer simulations). By contrast, we emphasize the emergent *social norm* of collective reciprocity on which generalized exchange can be based. Second, although we share Takahashi’s (2000) view that generalized exchange results from individual actors’ “fairness-based selective giving”, we are more interested in the collective nature of such giving at the community-level rather than the chain-like indirect reciprocity among actors (Bearman, 1997; Nowak & Sigmund, 2005). Third, in online SQA communities a question often receives more than one answer from multiple users, so the term “collective reciprocity” better captures the collectiveness of the reciprocation effort.

The notion of collective reciprocity also moves beyond some conventional conceptualizations of voluntary contribution in online communities. The gift economy framework connotes a returned favor in either present or future interactions (e.g., Lampel & Bhalla, 2007), but we argue that such an exchange of favor is not necessarily needed for multilateral interactions in online communities. Collective reciprocity is also different from altruism, although both may be classified as “other-oriented motives” (Peddibhotla & Subramani, 2007). In fact, collective reciprocity may be easily mistaken as altruism because community-level interaction patterns are

not immediately apparent to an individual observer. Altruism is a form of unconditional kindness that is independent of other people's intentions or acts, while collective reciprocity depends on and emerges from constant interactions with others. Altruism is certainly an important motive in online knowledge sharing, but we shall not automatically assume its existence when direct reciprocity is not observed.

The Duality of Collective Reciprocity

We propose collective reciprocity as a conceptual cornerstone for understanding multilateral interactions in online knowledge-sharing communities. Thanks to digital archives of messages and user profiles in these communities, collective reciprocity is observable and measurable. In a typical SQA community, for example, the questions and answers along with details of the question–answer interactions are recorded. A complete record of these interactions constitutes a collective history that reveals the community's trajectory over time. It is important to view the trajectory of such communities as a dynamic process in which normative rules are *being constructed in practice*. Collective reciprocity, therefore, is not an external rule imposed to a community; rather, it is being “written into being” (boyd, 2006) as community participants post questions and provide answers. Wenger (1999) describes this process of “congealing our experience into ‘thingness’” (p. 58) as the “process of reification.” However, reification is not merely a process of translating our experience into explicit expressions. There are official “guidelines” and “rules” created and posted by site owners or moderators in many online communities (e.g., http://answers.yahoo.com/info/community_guidelines), but the users are unlikely to read them. Rather, participation and reification are a duality that interplays and

defines each other. From this viewpoint, collective reciprocity is a reification of what individuals in a community do, but at the same time it defines what the individuals ought to do. This duality is fundamental to our understanding of collective reciprocity.

Rosenbaum and Shachaf (2010) see the duality of participation and reification as an essential contact point between the concept of CoPs and structuration theory. In Giddens' (1984) terms, social structure is "both medium and outcome" (p. 25) of social practices: Structure gives shape and form to social interactions, while at the same time these interactions reproduce and generate the structure. In a similar vein, Wenger states that an underlying structure should be understood as "an emerging property of interacting practices rather than their generative infrastructure" (p. 289), and he acknowledges that his position is close to Giddens' notion of structuration. For both Giddens and Wenger, normative rules of social structure are "virtual orders" (p. 17) that exist independently of any given individual, but the ability to follow the rules is grounded in individuals' reflections and actions. In fact, Giddens (1984) distinguishes between reciprocity between co-present actors and reciprocity between "collectivities across extended time-space" (p. 28). The latter one is an account of structural properties of social systems in which reciprocity is not established by normative code or co-present agents but through "reflectively monitoring" the flow of interaction with one another and then acting upon it. This structuration view is the foundation for our understanding of the collective reciprocity in online communities.

Hypotheses

We designed our empirical study to examine the prominence and characteristics of the "virtual

order” of collective reciprocity in SQA communities. The reification of collective reciprocity is a behavioral patterning in responding to kind or unkind intentions in multilateral interactions. In the context of SQA, both the kind intention and the reciprocation are reified through question–answer acts. Since SQA sites keep track of the acts and display the summary information as part of each user’s online profile, it becomes possible for the community members to reflectively monitor each other’s kindness and act upon it. Thus, we postulate that collective reciprocity is being dynamically constructed in question–answer interactions, where the kindness of a potential answerer is contingent upon the past effort and benefit of the asker. This is in line with Fehr and Schmidt’s (1999) observation of “inequity aversion,” a social norm formed on the basis that overcompensated users will be punished by not receiving the same level of reciprocation as the users who have over contributed. The ratio of past benefit (in terms of receiving answers from the community) to past effort (in terms of posting answers to the community) is a simple indicator of a user’s reciprocity status. We expect that an asker with a low benefit–effort ratio will be reciprocated better by the community than someone with a high benefit–effort ratio and vice versa.

Since the only form of reciprocation in a SQA community is answer-providing, it is reasonable to assume that the number of answers the asker receives is a proxy for the effect of collective reciprocity, which is contingent upon the user’s reciprocity status. Thus, we hypothesize:

H1a. There is a positive association between an asker’s past **benefit–effort ratio** and the number of answers she receives.

Almost all SQA sites organize the posted questions into categories to facilitate navigation and information seeking. When posting questions or answers, some users focus on one category only, while others may be active in different categories. An early study on mass interaction in Usenet found that newsgroup users often go beyond a particular newsgroup to engage in diverse conversations (Whittaker, Terveen, Hill, & Cherny, 1998). Arguello et al. (2006) also observed that Usenet users cross-post to different threads in order to increase her visibility. A few recent studies looked at cross-participation in online discussion groups as “boundary spanning” (Barcellini, Détienne, & Burkhardt, 2008) or “boundary reshaping” (Butler & Wang, 2012). However, to date little is known about the extent of and the motivation for cross-posting in SQA communities. From our economic-behaviorist perspective, it is reasonable to postulate that some kind SQA users might endeavor to answer questions across categories and try to be helpful to as many people as possible, whereas those who ask questions in many different categories are likely to have an intention of exploiting the commons. The former shows a high variation of effort, while the latter obtains a high variation of benefit. We expect that these effort/benefit variations have influences on the reciprocity status of an individual user:

H1b. There is a positive association between an asker’s **effort variation** and the number of answers she receives.

H1c. There is a negative association between an asker’s **benefit variation** and the number of answers she receives.

For most information seekers, it is often the quality of information that matters the most: A question may receive many answers, but the asker only needs one good answer to meet her

information need. Many SQA sites allow the asker to pick the “best answer” from the answer pool, indicating that the question has been resolved (Adamic et al., 2008; Shah & Pomerantz, 2010). Thus, whether or not a question is resolved can be an indicator of the quality of the answers received. Popular online SQA communities receive hundreds or even thousands of questions each day. For example, Yahoo! Answer’s official blog claims that the average rate of posting on the site is 10 messages (questions or answers) per second (Yanswersblog, 2010). Prior research also discovered that a large portion of questions in knowledge-sharing communities are answered by a small army of well-versed “answer persons” (Turner et al., 2005). In this context, a potential “answer person” has to be selective in answering questions, even if the question topic falls in her area of expertise. We have discussed that an answerer’s kindness is contingent upon the asker’s past effort and benefit, and we can further expect that the likelihood of getting a “best answer” is influenced by the asker’s past acts as well. The more the well-versed answerers engage in answering a question, the more likely the question gets a “best answer.” We hypothesize:

H2a. There is a positive association between an asker’s past **benefit–effort ratio** and the likelihood of getting a “best answer.”

H2b. There is a positive association between an asker’s **effort variation** and the likelihood of getting a “best answer.”

H2c. There is a negative association between an asker’s **benefit variation** and the likelihood of getting a “best answer.”

Methodology

Yahoo! Answers

We chose Yahoo! Answers as our primary data source for the empirical investigation. Yahoo! Answers typifies the question–answer mechanism in online SQA: A registered user posts a question in one of the pre-defined categories to solicit answers from other users; if multiple answers are provided, the user may select a particular answer as the “best answer” and the system marks the question as “resolved”; if a question does not get any answer four days after being posted, it will automatically expire and be deleted from the system. The basic function model of the Yahoo! Answers is illustrated in Figure 1.

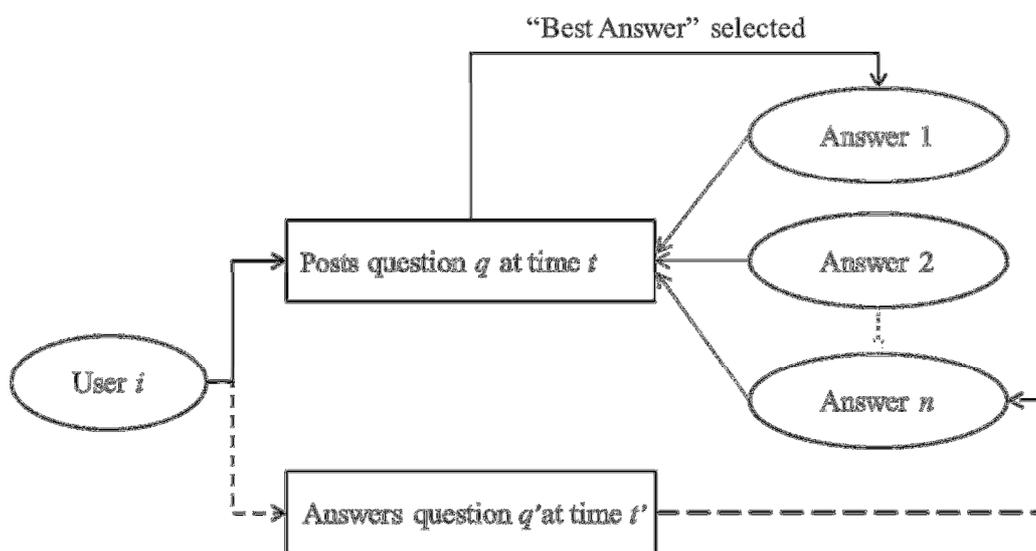


Figure 1. Yahoo! Answers' Function Model.

The user interface of Yahoo! Answers resembles a typical online forum in which messages are threaded in chronological order. However, once an answer has been chosen as the “best answer,”

it moves to the top of the thread and becomes more prominent than the rest of the answer candidates. Each user is identifiable through an alias and an avatar, which are hyperlinks to the user's profile page where a history of her activities is displayed. The profile pages are accessible to everyone including non-registered visitors.

Data Collection

We programmed a Web crawler to collect questions and answers across 23 different categories on Yahoo! Answers. The data collection started on November 2009 and lasted for one month.

For a user i , we define two actions: u_i^q when the user posts a question and u_i^a when the user posts an answer. Since each question/answer was associated with a unique user ID and a timestamp of when the question/answer was posted, we were able to construct a profile for each user in an offline process. In our panel data, we record the number of questions $N(u_i^q)_t$ and the number of answers $N(u_i^a)_t$ that the user has posted *at the time she asks a new question*. For each question q_x , we record the number of answers received ($NumAnswer(q_x)$) and the status of resolution ($Resolved(q_x)$). The status of resolution is a binary measure of whether the question has been resolved or not (i.e., whether a “best answer” is identified).

We used a reversed Herfindahl–Hirschman Index (HHI) (Hirschman, 1964) to measure users' distributed benefit $BenefitVar(u_i^q)$ and distributed effort $EffortVar(u_i^a)$. The original HHI is widely applied in market analysis to determine whether a market is an oligopoly. A high HHI generally indicates a concentration of market power, whereas a low HHI indicates high competition and market diversity. Butler (2001) adopted the HHI in his analysis of listserv messages and reversed the index ($1 - HHI$) to measure the daily topic variation in the sample

listservs. Inspired by Butler, we view each individual user as a firm and apply the reversed HHI in measuring the category variations of each user's postings. The higher the reversed index, the more the category variations in user's contributions. The advantage of using the reversed HHI as a variation measure is the ability to capture variation within a subset of question/answers while taking the total number of them into account. The formulas are as follows:

$$\mathit{BenefitVar}(u_i^q) = 1 - \mathit{HHI} = 1 - ([S_1^2 + S_2^2 + \dots + S_n^2]/\mathit{QuestionCount})$$

$$\mathit{EffortVar}(u_i^a) = 1 - \mathit{HHI} = 1 - ([S_1^2 + S_2^2 + \dots + S_n^2]/\mathit{AnswerCount})$$

S_n is the percentage of questions a user has posted in category $[1 \dots n]$. $\mathit{QuestionCount}$ and $\mathit{AnswerCount}$ are the total number of questions and the total number of answers the user has posted in the community, respectively.

To summarize, we collected data on the following variables:

- $\mathit{NumAnswer}(q_x)$: The number of answers received for question q_x .
- $\mathit{Resolved}(q_x)$: A binomial variable for recording whether question q_x has been resolved.
- $N(u_i^q)_t$: The number of questions the user i has posted before time t which is the system timestamp when the new question q_x is posted.
- $N(u_i^a)_t$: The number of answers the user i has posted before time t which is the system timestamp when the new question q_x is posted.
- $\mathit{BenefitVar}(u_i^q)$: The variation of user benefit in terms of posting questions.
- $\mathit{EffortVar}(u_i^a)$: The variation of user effort in terms of answering questions.

Data Analysis and Results

Descriptive Statistics

We collected 129,331 questions (N_q) that received at least one answer. These questions were posted by 92,862 unique users across 23 question categories. A total of 139,055 users contributed 682,725 answers (N_a). On average, each user contributed 1.39 questions and 4.91 answers. The most prolific user has provided as many as 356 answers to the community. The number of resolved questions ($Resolved(q_x)$) is only 11,448, or 8.85% of the total number of questions in the dataset. A summary of descriptive statistics of the variables in the dataset is presented in Table 1.

Table 1. Descriptive Statistics.

Variable	Mean	SD	Min	Max
$NumAnswer(q_x)$	4.61	5.83	1	338
$N(u_i^q)_t$	0.71	2.27	0	68
$N(u_i^a)_t$	2.37	7.48	0	356
$N(u_i^q)_t / N(u_i^a)_t$	0.29	0.84	0	33
$BenefitVar(u_i^q)$	0.43	0.28	0	1
$EffortVar(u_i^a)$	0.16	0.25	0	1

Note: $N_q = 129,331$, $N_a = 682,725$.

Hypothesis Testing

We use two regression models to estimate the hypothesized effects of collective reciprocation.

Model 1 is an ordinary least squares linear regression model for testing hypotheses H1a–c, and Model 2 is a logistic regression model for testing hypotheses H2a-c. Note that we include $NumAnswer(q_x)$ in Model 2 as a control variable because a larger pool of candidate answers is more likely to include a “best answer”.

Model 1:

$$NumAnswer(q_x) = \beta_0 + \beta_1 * (N(u_i^q)_t / N(u_i^a)_t) + \beta_2 * BenefitVar(u_i^q) + \beta_3 * EffortVar(u_i^a) + \varepsilon$$

Model 2:

$$P(Resolved(q_x)) =$$

$$1 / (1 + e^{-(\beta_0 + \beta_1 * (N(u_i^q)_t / N(u_i^a)_t) + \beta_2 * BenefitVar(u_i^q) + \beta_3 * EffortVar(u_i^a) + \beta_4 * NumAnswer(q_x))})$$

The regression analysis results are included in Table 2.

TABLE 2. Regression Results.

	Model 1		Model 2		
	Coefficient	<i>t</i>	Coefficient	<i>z</i>	Odds ratio
(Constant)	5.73***	110.07	-2.55***	-80.35	0.08
$N(u_i^q)_t / N(u_i^a)_t$	-0.202***	-9.52	-0.11***	-7.06	0.90
$BenefitVar(u_i^q)$	-1.67***	-20.13	-0.13**	-2.85	0.88
$EffortVar(u_i^a)$	2.21***	28.04	0.68***	16.66	1.96
$NumAnswer(q_x)$			0.03***	21.83	1.03
Goodness of fit	$R^2 = 0.148$		McFadden’s pseudo $R^2 = 0.369$		

** $p < 0.01$, *** $p < 0.001$.

Consistent with hypothesis H1a, Model 1 estimates a statistically significant negative relationship between an asker's past benefit–effort ratio ($N(u_i^q)_t / N(u_i^a)_t$) and the number of received answers. The coefficient value of -0.202 means that for every unit increase in benefit–effort ratio, we expect to see a decrease of 20.2% in number of answers. The negative coefficient for $BenefitVar(u_i^q)$ indicates that the more a user posts questions across categories, the less answers she receives for her own question, supporting H1b; on the other hand, the positive coefficient of $EffortVar(u_i^a)$ suggests that the more answers the asker has contributed in diverse topic areas, the more answers she is likely to receive, supporting H1c.

The estimations from Model 2 also provide strong support for hypotheses H2. The negative coefficients of β_1 and β_2 mean that increases of the benefit–effort ratio and benefit variations both decrease the likelihood of getting a “best answer.” This supports H2a and H2b. The positive coefficient of β_3 supports H2c, confirming a positive association between an asker's effort variation and the odds of receiving a “best answer.” All these relationships are statistically significant, controlling the total number of answers received for the question.

Discussion

In our research context, the quantity of answers received for a question indicates the volume of collective reciprocation, while the quality of answers suggests the effectiveness of the reciprocation. The highly significant associations between the askers' past acts and the quantity and quality of answers support our premise of collective reciprocity: The more effort (relative to benefit) an asker contributes to a SQA community, the more likely the community will return the

favor; on the other hand, the more benefit (relative to effort) the asker takes from the community, the less likely she will be reciprocated well in terms of getting answers for her own question. In the terms of game theory, collective reciprocity is the result of a behavioral mechanism that includes both conditional cooperation and punishment of non-cooperation (Gintis, 2000; Gintis, Bowles, Boyd, & Fehr, 2003).

While research in anthropology has confirmed that direct reciprocity is routinely harnessed in the support of cooperation in small-scale societies (Boehm, 1986; Henrich, Boyd, Bowles, Camerer, Fehr, & Gintis, 2004; Wiessner, 1982), indirect reciprocity is more common in complex and large-scale social networks in which individuals meet randomly and may never interact with the same person twice (Faraj & Johnson, 2011; Henrich et al., 2010; Nowak & Sigmund, 2005). In Yahoo! Answers and other large-scale SQA communities, indirect reciprocity means that each individual may take the role of asker or answerer in different times but never with the same partner. In this context, a person can follow either an unconditional strategy, such as always be an “answer person” (Turner et al., 2005) or always be a free rider, or a conditional strategy that discriminates among the askers according to their past interactions with others in the community. While answer persons and free riders are both commonly seen in online spaces, our data provide indication that the users of Yahoo! Answers tend to take a conditional strategy of reciprocation—one reciprocates provided others reciprocate as well.

It is worth noting that while indirect reciprocity and collective reciprocity are two related concepts, they have different theoretical roots and each emphasizes on different aspects of reciprocal interactions. Indirect reciprocity has been discussed in economic literature, and the

discussions center on reciprocity as an intrinsic motivation for individual acts. Therefore, the concept is often used to explain the fragility of conditional cooperation in social experiments (Fischbacher & Gächter, 2010). By contrast, we ground the concept of collective reciprocity on structuration theory and use it to depict both a community norm and the reification of the norm. The constant evaluation of individual participants' acts and the subsequent conditional reciprocation create the dynamics of collective reciprocity in a social network; at the same time, the observable dynamics give birth to an implicit community norm that reinforces the conditional cooperation.

Another related concept in social network research is collective efficacy, which was proposed by community sociologist Robert Sampson and his colleagues (Sampson, Raudenbush, & Earls, 1997) and later applied to electronic communities by John Carroll and his colleagues (Carroll & Reese, 2003; Carroll, Rosson, & Zhou, 2005; Kavanaugh et al., 2005). Collective efficacy is a concept originated from studies of local communities and denotes a neighborhood's capacity in dealing with challenges such as violent crimes and socioeconomic inequality. The theory of collective efficacy moves away from a focus on strong social ties and emphasizes on "shared expectations and mutual engagement" (Sampson, 2004, p. 108) by residents in the neighborhood. This theoretical angle is similar to ours in that we view collective reciprocity as a shared norm among loosely connected users in online communities, who interact with each other based on the principle of collective reciprocity rather than strong, direct reciprocity.

Of course, how a question performs in terms of attracting answers depends on many other factors. The relatively low R^2 values from the regressions suggest that the models have limited

predictive power. Topicality, timing of posting, visibility of the asker's online ID, and qualitative characteristics of the question (such as wording, grammar, and length) can all affect how a question is perceived and attended in online communities (Kim, 2010; Butler & Wang, 2012). For example, advice-seeking or conversational questions tend to generate more responses than factual questions (Harper, Weinberg, Logie, & Konstan, 2010). Another possible influence on the actual effects of collective reciprocity is the social network structure of the category sub-groups (Kang, Kim, Gloor, & Bock, 2011). Users active in certain category groups (for example, the "Pregnancy & Parenting" category in Yahoo! Answers) may have relatively dense network structures and are more engaged in question-answer interactions. Also, a user who posts frequently gains visibility and familiarity among community members, so a question from her is more likely to get a response than a question from a new ID. Future research may be carried out to analyze the network structure of meaningful sub-groups (for instance, based on a taxonomy of question types proposed by Harper et al., 2010) in SQA communities to examine the effect of network structure on the reciprocal behaviors.

Conclusion

Taking a structuration perspective while integrating reciprocity research in economics, the present study examines the dynamics of reciprocal interactions in online question answering. Communities as social entities in time and space are made up of individuals, and the parameters that characterize and shape collective behavior in the communities ultimately derive from individual behaviors. The empirical work presented in this paper shows that individuals in a SQA community constantly adjust their kindness into the direction of the observed benefit and

effort of others. When many individuals in the community follow this conditional strategy of reciprocation, a pattern of collective reciprocity emerges. Moreover, this emergent pattern of reciprocal interaction reinforces the normative order of conditional reciprocation.

The study contributes on enhancing our understanding of reciprocal behavior in online spaces and provides empirical evidences to support the collective reciprocity premise. The concept of collective reciprocity not only enriches our vocabulary in scholarly discussion of online communities but also provides insights for community designers. Our observation of conditional reciprocation echoes findings in Ma and Agarwal (2007) and Chen, Harper, Konstan, and Li (2010) in that member profiling and visibility of contributions are crucial elements in designing online CoPs.

Previous studies of online communities tend to view reciprocity as one of the individual motivations behind knowledge-sharing behaviors. We contend that reciprocity in online communities can be understood as a normative rule being constructed and reified through social interactions. The seemingly altruistic conduct of knowledge contribution by many “kind strangers” is not unconditional; rather, the habitual style is shaped by a network context in which social exchange is partially governed by the principle of “fairness-based selective giving”. Thus, our conception of online CoPs attempts to provide an explanatory framework that encompasses both the micro-level human behavior and the macro-level social structure. We show that different literatures in sociology and economics can be brought together in the analysis of structural forms of online communities.

Finally, we would like to conclude this paper by quoting Wenger (1999), who elegantly describes the duality of CoPs and the reciprocal relations in these communities: “Our experience and our world shape each other through a reciprocal relation that goes to the very essence of who we are. The world as we shape it, and our experience as the world shapes it, are like the mountain and the river. ... The river only carves and the mountain only guides, yet in their interaction, the carving becomes the guiding and the guiding becomes the carving” (p. 71).

References

- Adamic, L. A., Zhang, J., Bakshy, E., & Ackerman, M. S. (2008). Knowledge sharing and Yahoo Answers: Everyone knows something. *Proceedings of the 17th international conference on World Wide Web, WWW'08* (pp. 665–674). New York, NY: ACM.
- Arguello, J., Butler, B. S., Joyce, E., Kraut, R., Ling, K. S., Rosé, C., & Wang, X. (2006). Talk to me: Foundations for successful individual–group interactions in online communities. *Proceedings of the SIGCHI conference on Human Factors in computing systems, CHI'06* (pp. 959–968). New York, NY: ACM.
- Barbrook, R. (1998). The hi-tech gift economy. *First Monday*, 3, 12.
- Bearman, P. (1997). Generalized Exchange. *American Journal of Sociology*, 102(5), 1383–1415.
- Boehm, C. (1986). *Blood revenge: The enactment and management of conflict in Montenegro and other tribal societies*. Philadelphia, PA: University of Pennsylvania Press.
- boyd, D. (2006, February) Identity production in a networked culture: Why youth heart MySpace. Paper presented at conference of the American Association for the Advancement of Science, St Louis, MO.
- Brown, J. S., & Duguid, P. (2001). Knowledge and organization: A social-practice perspective. *Organization Science*, 12(2), 198–213.
- Bruckman, A. (2006). A new perspective on “community” and its implications for computer-mediated communication systems. *CHI'06 extended abstracts on Human factors in computing systems*, pp. 616–621. New York, NY: ACM.
- Butler, B. S. (2001) Membership size, communication activity, and sustainability: A resource-based model of online social structures. *Information Systems Research*, 12(4), 346–362.

- Butler, B. S., & Wang, X. (2012). The cross-purposes of cross-posting: Boundary reshaping behavior in online discussion communities. *Information Systems Research*, 23(3), 993–1010.
- Carroll, J. M., & Reese, D. D. (2003). Community collective efficacy: Structure and consequences of perceived capacities in the Blacksburg Electronic Village. Paper presented at the 36th Hawaii International Conference on System Sciences (HICSS), Big Island, Hawaii (10 pp.).
- Carroll, J. M., Rosson, M. B., & Zhou, J. (2005). Collective efficacy as a measure of community. *Proceedings of the SIGCHI conference on Human Factors in computing systems, CHI'05* (pp. 1–10). New York, NY: ACM.
- Chen, Y., Harper, F. M., Konstan, J., & Li, S. X. (2010). Social comparisons and contributions to online communities: A field experiment on MovieLens. *The American Economic Review*, 100(4), 1358–1398.
- Chiu, C.-M., Hsu, M.-H., & Wang, E. T. G. (2006). Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision Support Systems*, 42(3), 1872–1888.
- comScore. (2012). comScore MMX Ranks Top 50 U.S. Web Properties for August 2012. Retrieved September 24, 2012, from http://www.comscore.com/Press_Events/Press_Releases/2012/9/comScore_Media_Metri_x_Ranks_Top_50_U.S._Web_Properties_for_August_2012
- Constant, D., Sproull, L., & Kiesler, S. (1996). The kindness of strangers: The usefulness of electronic weak ties for technical advice. *Organization Science*, 7(2), 119–135.
- Cox, A. (2005). What are communities of practice? A comparative review of four seminal works.

- Journal of Information Science, 31(6), 527–540.
- Cox, J. C., Friedman, D., & Gjerstad, S. (2007). A tractable model of reciprocity and fairness. *Games and Economic Behavior*, 59(1), 17–45.
- Ekeh, P. P. (1974). *Social exchange theory: The two traditions*. Cambridge, MA: Harvard University Press.
- Falk, A., Fehr, E., & Fischbacher, U. (2008). Testing theories of fairness—intentions matter. *Games and Economic Behavior*, 62(1), 287–303.
- Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, 54(2), 293–315.
- Faraj, S., & Johnson, S. L. (2011). Network exchange patterns in online communities. *Organization Science*, 22(6), 1464–1480.
- Fehr, E., & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives*, 14(3), 159–182.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3), 817–868.
- Fischbacher, U., & Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review*, 100(1), 541–556.
- Barcellini, F., Détienne, F., & Burkhardt, J.-M. (2008). User and developer mediation in an Open Source Software community: Boundary spanning through cross participation in online discussions. *International Journal of Human-Computer Studies*, 66(7), 558–570.
- Gazan, R. (2010). Microcollaborations in a social Q&A community. *Information Processing and Management*, 46(6), 693–702.
- Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration*.

- Cambridge, UK: Polity Press.
- Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology*, 206(2), 169–179.
- Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior*, 24(3), 153–172.
- Hara, N., Shachaf, P., & Stoerger, S. (2009). Online communities of practice typology revisited. *Journal of Information Science*, 35(6), 740–757.
- Harper, F. M., Weinberg, J., Logie, J., & Konstan, J. A. (2010). Question types in social Q&A sites. *First Monday*, 15(7).
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., & Gintis, H. (2004). Foundations of human sociality: Economic experiments and ethnographic evidence from fifteen small-scale societies. Oxford, UK: Oxford University Press.
- Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J. C., et al. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *Science*, 327(5972), 1480–1484.
- Hirschman, A. O. (1964). The paternity of an index. *The American Economic Review*, 54(5), 761–762.
- Kang, M., Kim, B., Gloor, P., & Bock, G. (2011). Understanding the effect of social networks on user behaviors in community-driven knowledge services. *Journal of the American Society for Information Science and Technology*, 62(6), 1066–1074.
- Kankanhalli, A., Tan, B. C. Y., & Wei, K.-K. (2005). Contributing knowledge to electronic knowledge repositories: An empirical investigation. *MIS Quarterly*, 29(1), 113–143.
- Kavanaugh, A., Carroll, J. M., Rosson, M. B., Zin, T. T., & Reese, D. D. (2005). Community

- networks: Where offline communities meet online. *Journal of Computer-Mediated Communication*, 10(4).
- Kim, S. (2010). Questioners' credibility judgments of answers in a social question and answer site. *Information Research*, 15(1). Retrieved from <http://informationr.net/ir/15-2/paper432.html>
- Kollock, P. (1999). The economies of online cooperation: Gifts and public goods in cyberspace. In M. Smith & P. Kollock (Eds.) *Communities in Cyberspace* (pp. 220–239). London: Routledge.
- Lampel, J., & Bhalla, A. (2007). The role of status seeking in online communities: Giving the gift of experience. *Journal of Computer-Mediated Communication*, 12(2), 434–455.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
- Ma, M., & 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.
- Maloney-Krichmar, D., & Preece, J. (2005). A multilevel analysis of sociability, usability, and community dynamics in an online health community. *ACM Transactions on Computer–Human Interaction*, 12(2), 201–232.
- McMillan, D. W., & Chavis, D. M. (1986). Sense of community: A definition and theory. *Journal of Community Psychology*, 14(1), 6–23.
- Murillo, E. (2008). Searching Usenet for virtual communities of practice: Using mixed methods to identify the constructs of Wenger's theory. *Information Research*, 13(4).
- Nonnecke, B., & Preece, J. (2000). Lurker demographics: Counting the silent. *Proceedings of the*

- SIGCHI conference on Human Factors in computing systems, CHI'00 (pp. 73–80). New York, NY: ACM.
- Nov, O. (2007). What motivates Wikipedians? *Communications of the ACM*, 50(11), 60–64.
- Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, 437(7063), 1291–1298.
- Peddibhotla, N. B., & Subramani, M. R. (2007). Contributing to public document repositories: A critical mass theory perspective. *Organization Studies*, 28(3), 327–346.
- Raymond, E.S. (1998). The cathedral and the bazaar. *First Monday*, 3, 3.
- Reich, S. M. (2010). Adolescents' sense of community on MySpace and Facebook: A mixed-methods approach. *Journal of Community Psychology*, 38(6), 688–705.
- Rheingold, H. (1993). *The virtual community: Homesteading on the electronic frontier*. Reading, MA: Addison-Wesley.
- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Lawrence Erlbaum.
- Rosenbaum, H., & Shachaf, P. (2010). A structuration approach to online communities of practice: The case of Q&A communities. *Journal of the American Society for Information Science and Technology*, 61(9), 1933–1944.
- Rotman, D., & Preece, J. (2010). The WeTube in YouTube: Creating an online community through video sharing. *International Journal of Web Based Communities*, 6(3), 317–333.
- Sampson, R. (2004). Neighbourhood and community. *New Economy*, 11(2), 106–113.
- Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*, 277(5328), 918–924.
- Shachaf, P. (2009). The paradox of expertise: Is the Wikipedia Reference Desk as good as your

- library? *Journal of Documentation*, 54. Retrieved January 10, 2009, from <http://ella.slis.indiana.edu/~shachaf/paradox.pdf>
- Shah, C., Oh, S., & Oh, J. S. (2009). Research agenda for social Q&A. *Library & Information Science Research*, 31(4), 205–209.
- Shah, C., & Pomerantz, J. (2010). Evaluating and predicting answer quality in community QA. *Proceedings of the 33rd International ACM SIGIR conference on Research and development in information retrieval* (pp. 411–418). Geneva, Switzerland: ACM Press.
- Stanca, L. (2010). How to be kind? Outcomes versus intentions as determinants of fairness. *Economics Letters*, 106(1), 19–21.
- Takahashi, N. (2000). The emergence of generalized exchange. *American Journal of Sociology*, 105(4), 1105–1134.
- Turner, T. C., Smith, M. A., Fisher, D., & Welser, H. T. (2005). Picturing Usenet: Mapping computer-mediated collective action. *Journal of Computer-Mediated Communication*, 10(4).
- Wasko, M. 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.
- Wasko, M. M., Teigland, R., & Faraj, S. (2009). The provision of online public goods: Examining social structure in an electronic network of practice. *Decision Support Systems*, 47(3), 254–265.
- Wellman, B. (2001). Physical place and Cyberplace: The rise of personalized networking. *International Journal of Urban and Regional Research*, 25(2), 227–252.
- Wellman, B., Quan-Haase, A., Boase, J., Chen, W., Hampton, K., Díaz, I., & Miyata, K. (2003). The social affordances of the Internet for networked individualism. *Journal of Computer-*

- Mediated Communication, 8(3).
- Wenger, E. (1999). *Communities of practice: Learning, meaning, and identity*. Cambridge, UK: Cambridge University Press.
- Wenger, E., McDermott, R., & Snyder, W. M. (2002). *Cultivating communities of practice*. Boston, MA: Harvard Business Press.
- Whittaker, R. J. (1998). *Island biogeography: Ecology, evolution and conservation* (p. 285). Oxford, UK: Oxford University Press.
- Whittaker, S., Terveen, L., Hill, W., & Cherny, L. (1998). The dynamics of mass interaction. *Proceedings of the 1998 ACM conference on Computer supported cooperative work, CSCW'98* (pp. 257–264). New York, NY, USA: ACM.
- Wiessner, P. (1982). Risk, reciprocity and social influences on! Kung San economics. In E. Leacock & R. Lee (Eds.), *Politics and history in band societies* (pp. 61–84). Cambridge, UK: Cambridge University Press.
- Yamagishi, T., & Kiyonari, T. (2000). The group as the container of generalized reciprocity. *Social Psychology Quarterly*, 63(2), 116–132.
- Yanswersblog. (2010). 1 Billion Answers Served! Retrieved September 25, 2012, from <http://yanswersblog.com/index.php/archives/2010/05/03/1-billion-answers-served/>