

Q & A: Indirect reciprocity in an internet forum *

Takahiro Miura †

Abstract

This paper analyzes upstream indirect reciprocity in the field. I use data from Stack Overflow, one of the most significant internet fora on programming. I compare users whose questions receive answers with users whose questions do not receive answers. I find indirect reciprocity. Receiving a high-quality answer encourages the answering of questions of other users. The positive effects remain within a week. However, users who do not receive any satisfactory answers decline answering the questions of other users. The negative effects are statistically significant only on the day of the question post.

JEL Classification: H41

Keywords: Indirect reciprocity, Public goods provision

*I thank Takeshi Murooka, Kazuhito Ogawa, Fumio Ohtake, and Masaru Sasaki for their useful comments.

†RISS Postdoctoral fellow, RISS, Kansai University, 3-3-55, Yamane-cho, Suita, Osaka, Japan, E-mail: takamiura12@gmail.com, OCRIID:0000-0002-5926-0945

1 Introduction

Q & A websites, such as Stack Overflow and Yahoo Answers, are now common and play a critical role on the internet. Answers are public goods—meaning everyone can access them without payment.

Because many users are in the forum and the knowledge of a questioner and the answerer may be different, direct reciprocity—namely, bilateral cooperation—is unlikely playing a crucial role in those websites. While indirect reciprocity—namely, multilateral cooperation—can be a potential candidate to sustain the provision of the public goods (Boyd and Richerson, 1989; Leimar and Hammerstein, 2001). Many laboratory experiments support the existence of indirect reciprocity (Milinski et al., 2001; Seinen and Schram, 2006; Ule et al., 2009); however, the evidence lacks in the field.

There are two types of indirect reciprocity: upstream and downstream indirect reciprocities. In upstream indirect reciprocity, if Anna helps (harms) Bob, then he helps (harms) a third person; Charlie helps (punishes) Anna because Anna helped (harmed) Bob in downstream indirect reciprocity.

In this paper, I study upstream indirect reciprocity. In terms of Q & A websites, upstream indirect reciprocity encourage answering others' questions because the answerer's problem is solved by others. To capture upstream indirect reciprocity in the field, I use the data of Stack Overflow, one of the most significant internet fora on programming. Users can ask and answer questions in the forum, and all user answers and asking behavior are trackable. From the data, I compare the answer behavior between users whose first questions receive answers and those who do not. Examining the difference in the answer behavior between them, I distinguish indirect reciprocity.

Moreover, users may learn to program over time, which would reduce the cost of posting answers. In addition, throughout the experience of the internet forum, users learn how to write a clear question that can be answered. That may generate spurious correlations between answer behavior and the propensity of response to a question. To account for this endogenous problem, I run a fixed effect estimation by taking the within-differences of individual-monthly data.

The results show indirect reciprocity. For the user—receiving high-quality answers from other users to his/her first question—increases the probability of the user answering a question and the number of answers to other users' questions. The effect is the largest after one day from the question's post; and it decreases gradually one day later. From the result of the fixed effect estimation, I find a similar result for the user's first question. Users increase their answer behavior when their questions receive high-quality answers. Users who do not receive any satisfactory answers withdraw their answer behavior; However, the negative effects do not remain a long time.

2 Data

Stack Overflow, founded in 2008, is one of the most significant internet fora for programming. As of 2019, the number of registered users is about 10 million, and about 900 thousand users register each year. Stack Overflow employs Q&A style, and a question should be about a specific programming problem, algorithm, or issue—and not about a recommendation, be too specific, or cover too broad an issue. Through 2019, more than 16 million questions and more than 26 million answers have been documented on Stack Overflow. As of 2019, the rate of response within a week to a question was about 80%. Most questions—93% of the questions answered within a week—are answered within a day. The mean of the number of answers per question is 1.3.

Any user can answer a question. Asking a question, however, is restricted to registered users only. Moreover, registered users can comment on answers and questions, and they can vote on comments, answers, and questions. The purpose of the comments is to clarify relevant but minor information or constructively criticize.

Users can upvote and downvote on others' answers. The voting score—the number of upvoting minus the

number of downvoting—orders the answers to a question. The questioner can vote to accept the most useful answer, and each question can have only one accepted answer. Moreover, the accepted answer is displayed at the top of the solution regardless of its voting score. If the questioner accepts his/her answer, they cannot get an additional reputation point, and the order of the answer is not changed.

As well as answers, users can upvote and downvote on others' questions and also vote a favorite.

In the analysis, I used posting data on Stack Overflow in 2017. To account for the initial value problem, I focus on users who registered between January 2017 and November 2017.

3 Estimation Strategy

In Stack Overflow, registered users can ask a question and answer a question. Suppose that the users have indirect reciprocity. If another answers his/her question, then he/she will reciprocate in the community by solving another user's problem. In contrast, if others do not solve his/her problem, then he/she will not respond to others' questions compared with the former case. Even though the response rate to a question in Stack Overflow is high, some questions do not receive any answers. In this paper, I use the variation in response to a question to measure indirect reciprocity.

Most questions receive an answer within a day. Therefore, I define the question as to be answered if the question receives an answer from others within a day. Note that the questioners can answer their question, but I did not regard the question as to be answered in this paper.

To identify indirect reciprocity, I use two estimation strategies. First, I focus on the user's first question, and compare answer behavior among the users whose first questions are answered and users who do not receive answers. I track answer behavior seven days from the posting of the first question. The estimation equation is,

$$Answer\ to\ others_{it} = \sum_{d=0}^6 \beta_d Be\ Answered_i \times \mathbf{1}(t = d) + \sum_{d=1}^6 \gamma_d \mathbf{1}(t = d) + \mathbf{Q}_i + \delta_t + \epsilon_{it}, \quad (1)$$

where i is the individual, t is the date, and d is the elapsed days from the posting of the individual i 's first question (0 is the published day). *Answer to others* is the answer behavior to others' questions. *Be Answered* is the dummy for whether the individual i 's first question is answered. \mathbf{Q}_i is the characteristics of the first questions such as the question's vote scores and favorite counts. δ_t includes year \times week dummies and day of the week dummies. γ_d captures the difference in answer behavior between d days after posting the first question and the question post day. β_d captures the difference in answer behavior d days after posting the first question between users whose first question is answered and those whose questions are not. $\beta_d > 0$ holds if there is indirect reciprocity; users increase their answer behavior if they receive an answer to their first question.

Second, I use all users' data from the registration date. It is plausible that answer behavior correlates with the response rate of the question. The endogenous problem may arise; for example, answering others' questions is easier for a mature programmer; however, his/her question may not be easy to answer. Moreover, it is plausible that an individual's programming skill changes over time.

To account for the endogenous problem, I use the fixed effect estimation. I control the individual's monthly fixed effect assuming an individual's programming skill is maintained within a month. The variation of the response to the question within a month captures indirect reciprocity. The estimation equation is as follows:

$$Answer\ to\ others_{im} = \sum_{d=0}^6 \beta_d Be\ Answered_{im} \times \mathbf{1}(t = d) + \sum_{d=0}^6 \gamma_d \mathbf{1}(t = d) + \sum_{d=0}^6 \mathbf{Q}_{im,t-d} + \delta_t + \alpha_{im} + \epsilon_{im}, \quad (2)$$

where i is the individual, t is the date, d is the elapsed days from posting the question, and m is the month. α_{im} is the individual's monthly fixed effect. $\mathbf{Q}_{im,t-d}$ is the characteristics of questions that posted on d days before. γ_d captures the difference in answer behavior between d days after the posted question and before and seven days after posting the question. β_d is the difference in answer behavior between whether or not the question is answered d days after the posted question. $\beta_d > 0$ holds if there is indirect reciprocity because an individual is more likely to answer others' questions if his/her question is answered compared with his/her question not being answered.

4 Result

I collected 306287 first questions and 359704 questions in total. Note that the first week's data is dropped in total questions. I dropped uses who did not ask a question in the sample period. Moreover, in the observations for the user's first question, I dropped users who answered a question before asking a question. Therefore, in terms of reputation point, users have the same profile at the user's first question. In the observations for all questions, I dropped the user's monthly data if the user has no further activity—namely, no answers and no question posts. The response rates to user's first questions were 66%; the response rate is 80% overall. Users answer others' questions 15566 times within seven days after their first question post and 146095 times in total.

Figure 1 shows the main result. The dependent variables are the probability of answering other users' questions and the number of answers to other users' questions. The OLS and fixed effect estimations are used to estimate the extent of indirect reciprocity from the users' first questions and all questions, respectively. β_d in Equation 1 corresponds to the line in panel A in Figure 1, and β_d in Equation 2 corresponds to it in panel B in the figure. Moreover, the shaded area in the figure shows a 95% confidence interval for the coefficients.

The result shows the opposite direction of indirect reciprocity. The patterns of the effect on the probability of answering other users' questions are similar to that on the number of answers to other users' questions. Focusing on the first question, users whose first question is answered within a day significantly decrease their answering of other users' questions.

The negative effect is the largest on the day of the posting of the first question; qualitatively, the probability of answering other users' questions decreases by about 0.2 percentage points, and the number of answers decreases by about 0.003. After one day from the question's post, the effects are not statistically significant.

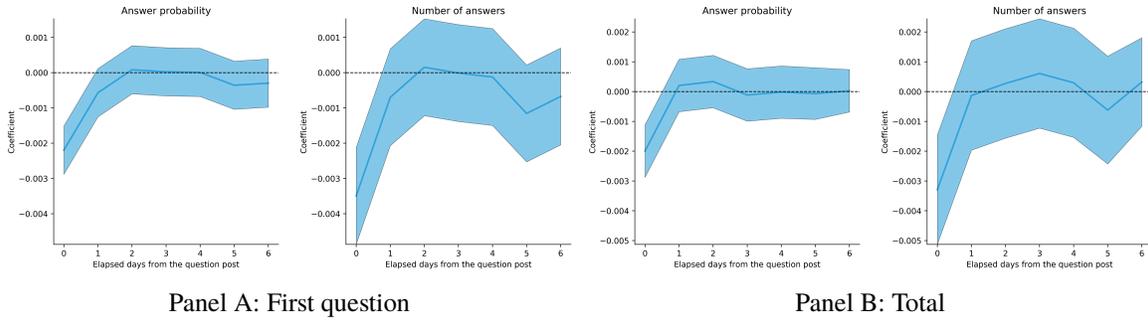
Focusing on the result for all questions, the pattern is similar to the result for the first question. Users, when their questions are answered within a day, significantly decrease their answers to other users' questions on the same day of the posting of the question. On the same day of posting the question, the effect is statistically significant. The effect is the largest after one day from posting the question; Qualitatively, the fixed effect estimation shows similar magnitudes as the result of the first question. The effects turn slightly positive after a day but are not statistically significant. The effects are not statistically significant after a day.

Overall, I cannot find indirect reciprocity that leads to answering other users' questions when the user's question has been answered. However, the quality of answers is heterogeneous among answers; some answers are not constructive, and some answers are satisfactory. It is plausible that users do not reciprocate if the quality of the answers is low.

To analyze the effect of the quality of answers on indirect reciprocity, I use subjective quality. Note that the subjective quality of the answer is its acceptance by the questioner.

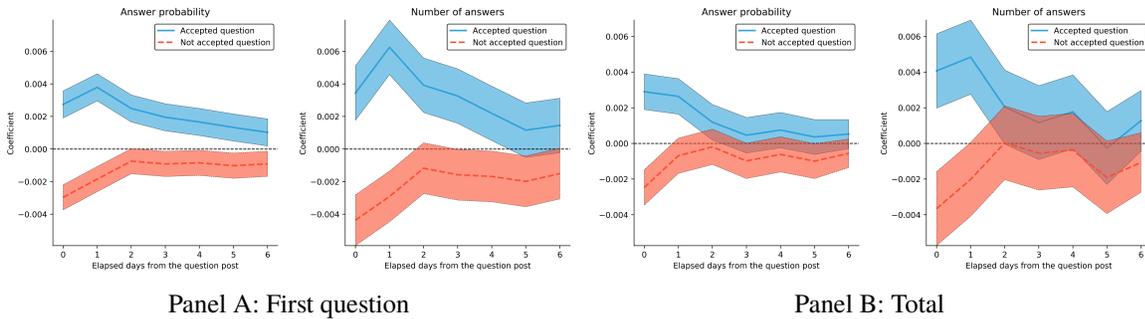
Figure 2 shows the effect of the subjective answer quality on indirect reciprocity. I separate answered questions into two types: accepted and not accepted. Accepted questions have an accepted answer, and not accepted questions do not. I control for the answer score and the number of answers in the analysis.

In the results for the first questions, users reciprocate more when they receive a high subjective quality answer. An accepted answer increases by about 0.3 percentage-points the probability of answering other



Shaded area shows the 99% confidence interval.

Figure 1: Main result



Shaded area shows the 99% confidence interval.

Figure 2: Subjective answer quality

users' questions and by about 0.004 answering other users' questions on the same day of the posted question. If users accept an answer to their questions, the positive effect on the probability of answering other users' questions remains for a week, and on the number of answers to other users' questions remains until four days after. In contrast to the accepted answer, if users do not accept any answers to their questions, they abstain from answering others; no accepted answers reduce by about 0.3 percentage points the probability of answering other users' questions and by about 0.004 answers to other users' questions on the same day of the posted question. The negative effect remains for a week.

For the results for all questions, the positive effect of receiving a high-quality answer does not remain for a week. On the same day of the question's post, receiving a high subjective quality answer increases the answer probability; qualitatively, accepting an answer increases by about 0.3 percentage-points answer probability and increases by about 0.004 answers to other users' questions. However, the decreasing trend of the effect is steeper than that for the first question; the effects on the probability of answering other users' questions are statistically insignificant after three days, and the effects on the number of answers to other users' questions are statistically insignificant after two days.

Receiving a low subjective quality of answer decreases answer probability; qualitatively, no accepted answers reduce by about 0.2 percentage-points answer probability and reduce by about 0.004 answers to other users' questions. The effect is statistically significant only on the day of the posting of the question.

In summary, if users receive an answer from other users, they reduce their behavior of answering other users' questions on average. However, the subjective and objective qualities of the answers affect answer behavior. The higher the quality of the answers the users receive, the more they reciprocate. Moreover,

receiving low subjective quality answers reduces answer behavior on the day of the question post.

5 Conclusion

At any time, new questions and answers are posted on Stack Overflow. Answers are public goods—meaning everyone can access them without payment. Although theoretical and experimental research indicates that indirect reciprocity can sustain such public goods provision, evidence in the field is lacking. This paper fills in the gap between empirical and other research.

The results in this paper provide evidence of indirect reciprocity in the field. Even if the population of the community is quite large (more than a million), receiving answers to questions increases questioners' answer behavior towards others.

This paper has a limitation as follows. Stack Overflow has a “reputation” score that evaluates the extent of an individual's cooperation. Reputation is important to maintain cooperation in indirect reciprocity and the effect depends on the reputation scoring rules (Ohtsuki and Iwasa, 2004 2006). However, I could not isolate the effect of the current reputation point rule in Stack Overflow. Future research is needed on what reputation method would be useful in the field.

References

- Robert Boyd and Peter J Richerson. The evolution of indirect reciprocity. *Social Networks*, 11(3):213–236, 1989.
- Olof Leimar and Peter Hammerstein. Evolution of cooperation through indirect reciprocity. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1468):745–753, 2001.
- Manfred Milinski, Dirk Semmann, Theo CM Bakker, and Hans-Jürgen Krambeck. Cooperation through indirect reciprocity: image scoring or standing strategy? *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1484):2495–2501, 2001.
- Hisashi Ohtsuki and Yoh Iwasa. How should we define goodness?—reputation dynamics in indirect reciprocity. *Journal of theoretical biology*, 231(1):107–120, 2004.
- Hisashi Ohtsuki and Yoh Iwasa. The leading eight: social norms that can maintain cooperation by indirect reciprocity. *Journal of Theoretical Biology*, 239(4):435–444, 2006.
- Ingrid Seinen and Arthur Schram. Social status and group norms: Indirect reciprocity in a repeated helping experiment. *European Economic Review*, 50(3):581–602, 2006.
- Aljaž Ule, Arthur Schram, Arno Riedl, and Timothy N Cason. Indirect punishment and generosity toward strangers. *Science*, 326(5960):1701–1704, 2009.