ソシオネットワーク戦略ディスカッションペーパーシリーズ ISSN 2434-9445
第 111 号　2024 年 3 月
RISS Discussion Paper Series
No.111　March, 2024

# Q&A: Upstream Indirect Reciprocity in a Large Online Community

Takahiro Miura

文部科学大臣認定　共同利用・共同研究拠点

関西大学ソシオネットワーク戦略研究機構

Research Institute for Socionetwork Strategies,
Kansai University
Joint Usage / Research Center, MEXT, Japan
Suita, Osaka, 564-8680, Japan
URL: https://www.kansai-u.ac.jp/riss/index.html
e-mail: riss@ml.kandai.jp
tel. 06-6368-1228
fax. 06-6330-3304

# Q&A: Upstream Indirect Reciprocity in a Large Online Community*

Takahiro Miura†

**Abstract**

This study explores the presence of upstream indirect reciprocity—the tendency that individuals who receive help from someone are more likely to help a third party—in a general field environment using Stack Overflow, a leading online Q&A forum. I investigate this by examining whether questioners who receive answers from others in the forum are more likely to engage in answering other users' questions. Initially, findings indicate that merely receiving answers does not lead to an increase in the tendency to help others. However, I find that the satisfaction of questioners with the answers they receive is crucial in fostering upstream indirect reciprocity. Questioners satisfied with the answers they receive are more likely to provide answers themselves, maintaining the quality of their contributions, and these effects are observed to persist even after a year. Conversely, questioners dissatisfied with received answers tend to be less helpful to others. This study highlights the importance of understanding upstream indirect reciprocity as a factor for promoting cooperative behavior in environments in which the role of direct reciprocity is often limited, such as online communities. *JEL Codes*: D83, D91, H41.

*An earlier version of this article was the fourth chapter of my dissertation at Osaka University. I would like to thank Shigeki Kano, Takeshi Murooka, Kazuhito Ogawa, Fumio Ohtake, Hideo Owan, Masaru Sasaki, and Shoko Yamane, as well as the participants of the 14th Seminar of Applied Econometrics, the 13th Conference of Behavioral Economics, the Wakayama University lunchtime seminar, and the RISS workshop for their valuable comments and insights. I would like to thank Editage for English language editing. This work was supported in part by JSPS KAKENHI Grant Number JP20K20757.

†Wakayama University and Research Institute for Socionetwork Strategies (RISS). E-mail: takam@wakayama-u.ac.jp. Mailing address: Department of Economics, 930, Sakaedani, Wakayama-city, Wakayama, Japan. ORCID: 0000-0002-5926-0945.

# I. Introduction

Online communities, as exemplified by Q&A forums and Open Source Software (OSS) projects, are vital in today's internet society. Online Q&A forums within the Stack Exchange network, including Stack Overflow, facilitate the exchange of knowledge and assist in problem-solving. The content in these platforms is not supported by monetary rewards but by users who voluntarily provide their expertise, creating an environment that fosters collective learning and support. This focus on volunteer-driven collaboration is also evident in OSS projects. In large-scale projects, such as the Python library pandas, which comprise numerous contributing participants, the scope for direct reciprocal exchanges is limited, suggesting that voluntary contributions are crucial to the project's development.[1] Both Q&A forums and OSS projects emphasize the significance of volunteer-led initiatives in online communities, highlighting the importance of collaboration beyond direct reciprocal relationships.

Reflecting on the voluntary nature of participation in online communities, this study delves into the role of upstream indirect reciprocity in situations in which the scope for direct reciprocity is limited. Upstream indirect reciprocity—characterized by a chain of altruistic actions—occurs when an act of kindness toward one individual inspires subsequent kindness toward others. In simpler terms, if individual A helps B, B may feel inclined to help a third person, C.[2] This phenomenon is supported by numerous laboratory experiments, demonstrating its prevalence in controlled settings (Dufwenberg et al., 2001: Bolton, Katok and Ockenfels, 2005: Bartlett and DeSteno, 2006: Engelmann and Fischbacher, 2009: Stanca, Bruni and Corazzini, 2009: Stanca, 2009: Herne, Lappalainen and Kestilä-Kekkonen, 2013: Bahr and Requate, 2014: Gray, Ward and Norton, 2014: Strang et al., 2016: Sun et al., 2020).[3] Despite its validation in experimental settings, real-world evidence of upstream indirect reciprocity, especially in online communities where direct reciprocity has less influence, remains scarce. This research aims to

[1] As of January 2024, pandas had 3,106 contributors. For more information, see https://github.com/pandas-dev/pandas.

[2] Various terms are used for upstream indirect reciprocity, including pay(ing)-it-forward (Gray, Ward and Norton, 2014: Horita et al., 2016), generalized (indirect) reciprocity (Rankin and Taborsky, 2009: Stanca, 2009), upstream reciprocity (Iwagami and Masuda, 2010), and pure indirect reciprocity (Greiner and Vittoria Levati, 2005).

[3] However, some studies have not identified upstream indirect reciprocity in laboratory experiments. For instance, Ben-Ner et al. (2004) conducted a two-stage dictator game and found no correlation between generosity in the first stage and the second stage. Similarly, Horita et al. (2016) observed upstream indirect reciprocity only at the beginning of a helping game. Moreover, Schnedler (2022) did not observe negative upstream indirect reciprocity across different tasks; individuals who were assigned dull tasks by someone else did not exhibit increased unkindness in subsequent dictator games involving a third party.

contribute to this gap by analyzing data from a large online community to explore the potential of upstream indirect reciprocity in fostering cooperative behaviors.

While empirical evidence of upstream indirect reciprocity in real-world settings is scarce, a field experiment by Mujcic and Leibbrandt (2017) in a supermarket parking area is an exception. In this experiment, the decision to yield to cars attempting to exit the parking lot was made randomly, and the effect of this behavior on the subsequent actions of others was observed. The results indicated that individuals who were yielded to were 119% more likely to reciprocate the kindness by yielding to others. This finding suggests the potential for upstream indirect reciprocity in real-world scenarios, although evidence in more general settings is yet to be discovered.[4]

To investigate upstream indirect reciprocity, I utilize data from Stack Overflow, the largest online Q&A forum for programming. This platform is ideal for studying social behaviors such as upstream indirect reciprocity, owing to its extensive user base and comprehensive data on user interactions. In this study, receiving answers from others is considered receiving help, and in turn, providing answers to other users' questions is viewed as helping. The existence of upstream indirect reciprocity in this forum would mean users who receive answers are more likely to subsequently help others by providing their responses.

This analysis centers on the impact of receiving answers to users' first questions on Stack Overflow. By focusing on users who have not posted answers before and have been active within a certain timeframe after registration, the study minimizes confounding factors for a more accurate estimation. To measure the impact comprehensively, I developed various outcome variables, including the likelihood of answering other users' questions, the number of answers posted, and the probability of providing high-quality responses as judged by the community and the questioners.

The initial findings indicated that simply receiving answers does not necessarily lead to upstream indirect reciprocity. The overall effect of receiving answers

[4]To my knowledge, only a few empirical studies on upstream indirect reciprocity and mixed results are reported. Baker and Bulkley (2014) analyzed a classroom Q&A board, finding a positive correlation: students who received responses to their questions in the past week to one month were more likely to answer other students' questions. Conversely, Van Apeldoorn and Schram (2016) reported no or a slightly negative correlation in an online community. Focusing on Stack Overflow, Yan and Jian (2017) found a negative correlation between receiving higher-scored answers to a first question and the likelihood of subsequently answering other users' questions. While my study uses a different dataset period but has some similar analytical strategies, such as examining the first question posted by users on the same platform, Stack Overflow, the main findings of my study contrast with theirs. I employ more rigorous sample restrictions and outcome variable definitions to mitigate confounder effects while utilizing various robustness checks to validate the results.

to a user's first question did not enhance their subsequent answering behavior. In fact, users who received answers were found to have a slightly decreased likelihood of answering other users' questions and a reduced number of answers posted.

However, upstream indirect reciprocity was evident where questioners were satisfied with the answers received. Users satisfied with responses to their initial questions showed a higher propensity to assist others, as evidenced by an increase in the number and quality of their answers. Conversely, unsatisfied users were more likely to refrain from answering than those who did not receive any answers. These patterns remained consistent across various robustness checks, including different subsample analyses and alternative outcome definitions, underscoring the role of questioners' satisfaction in encouraging helpful behavior.

Further analysis revealed distinct patterns in the short- and long-term effects of receiving answers. The daily impact transition patterns of receiving satisfying answers from others did not show backsliding or a shift from positive to negative effects. The positive impact of receiving satisfactory responses was observed even a year after posting a question, indicating sustained motivation to assist others in the community. This enduring effect suggests that upstream indirect reciprocity possibly impacts the collaborative environment within online communities over prolonged periods. Additionally, receiving answers, irrespective of satisfaction, positively influences users' subsequent engagement in the community.

This study's primary contribution is in providing the first empirical evidence of upstream indirect reciprocity in a general field environment, following various laboratory experiments, by analyzing data from Stack Overflow, a large Q&A forum. This study also complements the field experiment by Mujcic and Leibbrandt (2017), which demonstrated upstream indirect reciprocity in a specific situation. Unlike the previous study that highlighted considerably short-term effects, this study indicates that upstream indirect reciprocity can persist over time.

Moreover, this study potentially adds a new aspect to the reciprocity literature in behavioral economics. Much research in this field, as evidenced by numerous laboratory and field experiments, has focused on direct reciprocity and reciprocal behaviors in small-scale groups.[5] The proposed models have also primarily centered on explaining these behaviors (Rabin, 1993: Levine, 1998: Fehr and Schmidt, 1999: Bolton and Ockenfels, 2000: Charness and Rabin, 2002: Dufwenberg and

[5]For laboratory experiments, see meta-analyses of trust game by Johnson and Mislin (2011), meta-analyses of trust and gift exchange games by van den Akker et al. (2020), and meta-analyses of ultimatum game by Oosterbeek, Sloof and Van De Kuilen (2004) and Cooper and Dutcher (2011). For field experiments, see studies by Gneezy and List (2006), Falk (2007), Alpizar, Carlsson and Johansson-Stenman (2008), and DellaVigna et al. (2022). Earlier studies of reciprocity research in economics include those by Fehr and Gächter (2000) and Sobel (2005).

Kirchsteiger, 2004: Falk and Fischbacher, 2006: Cox, Friedman and Gjerstad, 2007: Cox, Friedman and Sadiraj, 2008). However, the applicability of these models to broader contexts, such as upstream indirect reciprocity in larger groups, remains uncertain. This study highlights the potential importance of upstream indirect reciprocity in scenarios where direct reciprocity is less prevalent, such as in online communities, peer reviews, and blood donations, wherein interactions are not directly reciprocated.

Furthermore, the main findings of this study provide insights into the mechanism of upstream indirect reciprocity. Most existing studies in this field have primarily focused on confirming its presence or evaluating the extent of its effects, rather than investigating the underlying mechanisms. There have been some attempts to apply insights from prosocial behavior and direct reciprocity research to upstream indirect reciprocity, such as examining the role of beneficiaries' intentions (Herne, Lappalainen and Kestilä-Kekkonen, 2013: Sun et al., 2020).[6] However, there remains a significant gap in our understanding.[7] This study contributes to bridging the gap. First, the observed upstream indirect reciprocity is not merely a result of social pressure, but the impact extends beyond the frequency of responses, affecting the quality of answers provided. This contrasts with the theoretical predictions of DellaVigna, List and Malmendier (2012) about minimal contributions under social pressure. Second, the satisfaction of recipients possibly fosters upstream indirect reciprocity, suggesting a potential alignment with the psychological literature, which emphasizes the role of gratitude in upstream indirect reciprocity. For instance, McCullough, Kimeldorf and Cohen (2008) advocates for the significance of gratitude in upstream indirect reciprocity, and Bartlett and DeSteno (2006) reports the positive effect of receiving assistance on gratitude and subsequent prosocial actions toward third parties.

This study also contributes to the increasing use of large-scale online community data in economics and related fields. It aligns with research utilizing data from various platforms such as mathematics Q&A forums (Bohren, Imas and Rosenberg, 2019), MovieLens (Chen et al., 2010), Reddit (Burtch et al., 2022), Yelp (Botelho and Gertsberg, 2021), Stack Overflow (Xu, Nian and Cabral, 2020: Smirnova, Reitzig and Sorenson, 2022), and Wikipedia (Gallus, 2017: Greenstein, Gu and Zhu, 2021: Linek and Traxler, 2021), situating the study within a broader

[6]For studies examining the impact of intentions in direct reciprocity, see McCabe, Rigdon and Smith (2003): Cox (2004): Cox and Deck (2005): Stanca, Bruni and Corazzini (2009).

[7]For instance, there is a lack of research on phenomena such as the moral wiggle room effect in the context of upstream indirect reciprocity despite its importance, as highlighted in studies on prosocial behavior (Broberg, Ellingsen and Johannesson, 2007: Dana, Weber and Kuang, 2007: Lazear, Malmendier and Weber, 2012) and reported in direct reciprocity studies (Malmendier, Te Velde and Weber, 2014: Regner, 2018).

context of digital community analysis.

Finally, this study's findings contribute to the literature on the evolution of cooperation in large groups. Focus on the role of indirect reciprocity in maintaining cooperative behaviors within large groups has been increasing. Nevertheless, this focus has been majorly on downstream indirect reciprocity in which past assistance increases the likelihood of receiving future help,[8] This predominant focus may stem from the perception that downstream indirect reciprocity is more crucial than upstream indirect reciprocity. Nowak and Roch (2007) posited that upstream indirect reciprocity is often considered merely a "hitchhiker" on other forms of reciprocity and cannot sustain group cooperation independently.[9] I demonstrate the existence of upstream indirect reciprocity in a broad and general large-scale setting, such as online Q&A forums in which direct reciprocity plays a minimal role. Although not primarily focused on the dynamics of cooperation, this research suggests that upstream indirect reciprocity could indeed be a significant driver in sustaining cooperation within groups.

The remainder of this paper is organized as follows. Section II. outlines Stack Overflow and discusses the identification strategy used to test upstream indirect reciprocity. Next, Section III. presents the summary statistics. Section IV. shows the results of analyzing upstream indirect reciprocity in Stack Overflow and examines the robustness of the main findings, followed by the analysis of the short- and long-term effects. Finally, Section V. provides the discussion and conclusion of this study. The appendix includes image examples of a Stack Overflow and supplementary results.

# II. Data and Methods

To test upstream indirect reciprocity in the field, I focus on a large online Q&A forum. Users in this forum can post their questions, which can be considered asking for help from others. Some of these users receive answers—considered being helped by someone else. Posting answers in the forum can be interpreted as

[8] See for example Nowak and Sigmund (1998): Leimar and Hammerstein (2001): Milinski et al. (2001): Nowak and Sigmund (2005): Nowak and Roch (2007): Hilbe et al. (2018): Clark, Fudenberg and Wolitzky (2020). Laboratory experiments evidencing downstream indirect reciprocity include those by Bolton, Katok and Ockenfels (2005): Seinen and Schram (2006): Engelmann and Fischbacher (2009): Stanca, Bruni and Corazzini (2009): Stanca (2009): Horita et al. (2016). Field experiments are reported by Van Apeldoorn and Schram (2016): Khadjavi (2017), with observational studies conducted by Wu and Korfiatis (2013): Baker and Bulkley (2014).

[9] Some studies have argued that upstream indirect reciprocity can sustain group cooperation under certain conditions (Pfeiffer et al., 2005: Rankin and Taborsky, 2009: Iwagami and Masuda, 2010: van Doorn and Taborsky, 2012).

an act of helping others. If upstream indirect reciprocity exists in the forum, users who received answers from others to their questions are more likely to respond to other users' questions. Furthermore, in a large online Q&A forum, users can easily find other people's questions, showing numerous opportunities to help others. Therefore, I consider that a large online Q&A forum is a suitable platform for testing upstream indirect reciprocity.

I use data from Stack Overflow to examine upstream indirect reciprocity in the field. This section briefly explains the structure of Stack Overflow and describes the available information to test upstream indirect reciprocity in the forum. Subsequently, I discuss the empirical strategy of upstream indirect reciprocity.

## II.A. Stack Overflow

Stack Overflow, established in 2008, is a premier online Q&A platform focused on programming. By 2022, it had grown to over 17 million registered users, with approximately 2.2 million questions and 3.3 million answers posted on the site. While any user can respond to queries, posting a question is a privilege reserved for registered members.[10] The scope of queries on Stack Overflow is narrowly defined: questions must pertain to specific, solvable issues. This policy excludes broader, opinion-based queries, such as "What is your favorite programming language?" Each post and registered user are assigned a unique ID and timestamp, enabling the tracking of all registered users' activities on Stack Overflow.

Stack Overflow hosts an extensive array of questions and answers, where the quality of posts ranges from highly informative to less useful, including incorrect answers or uninteresting questions. To ensure the quality of content, which is accessible to all users, Stack Overflow incorporates a rating system. This system allows contributing users to assess the quality of other users' posts through voting: upvotes for high-quality content and downvotes for poor-quality posts. The voting results impact the user's "reputation" points, reflecting their level of contribution to the forum. Higher reputation points confer privileges such as commenting, editing, and voting.[11] An upvote on a user's post increases their reputation points by ten, while a downvote decreases it by two. Furthermore, downvoting a question results in the voter losing one reputation point. Votes are anonymous and recorded with daily timestamps. The cumulative total of upvotes and downvotes, termed as a post's "score," provides a summary of its overall reception.

[10] Additionally, completing the Stack Overflow tutorial is a prerequisite for posting questions.

[11] Users need a minimum of 15 reputation points to upvote and 125 points to downvote. More information about these privileges is available at https://stackoverflow.com/help/privileges.

In addition to the voting system, Stack Overflow allows questioners to express satisfaction with the answers to their questions. They have the option to mark the most helpful answer as the "accepted answer." This choice is not mandatory; if a questioner is not satisfied with any of the responses, they can withhold their selection until a satisfactory answer is received. The user whose answer is selected as the accepted answer earns fifteen reputation points, while the questioner gains two reputation points. Notably, questioners can select their answers as accepted, but this does not affect their reputation points. Thus, Stack Overflow provides a comprehensive assessment of post quality, incorporating the objective quality, as measured by votes from other members, and the subjective quality, as indicated by the questioner's satisfaction.[12]

Answers on a question page in Stack Overflow are displayed in order of their score; nevertheless, the accepted answer is always displayed at the top regardless of its score. Online Appendix Figure A.I presents an example of a question page on Stack Overflow. The current score of a post is displayed at the top left of the post, between the upper and lower triangle icons.[13] At the lower right corner of the post, basic information about the submitter is shown: name, icon, reputation points, and the number of badges.[14] Users can earn badges by engaging in specific activities or accumulating high reputation points.

To analyze user behavior, I utilized the Stack Overflow data dumps. These data dumps are released quarterly on the Internet Archive and include information such as undeleted posts, user data, and anonymized voting data up to that point.[15]. Additionally, a duplicate of the data dump is available in Google Cloud's BigQuery public project.[16] I downloaded the data from BigQuery on January 10, 2023. The latest post in the dataset was created on September 25, 2022.

## II.B. Empirical Strategy

To investigate upstream indirect reciprocity on Stack Overflow, I examine the effect of receiving answers from others on the subsequent answering behavior of

[12]Questioners also have the option to express their evaluation of received answers by upvoting or downvoting. However, as I will discuss later, the sample used in my analysis is designed to minimize the probability of tainting due to such voting.

[13]Users can upvote and downvote a post by clicking the upper and lower triangles, respectively. The number below the star icon to the left of a question indicates how many other users have added the question as a favorite. However, as most questions have zero favorites, this information is not utilized in this study.

[14]Detailed user information, such as the number of answers and questions posted and the type of badges, can be viewed on the user's page, which is accessed by clicking on the user's name or icon.

[15]For more details, see https://archive.org/.

[16]For more details, see https://cloud.google.com/bigquery.

questioners. The effect is estimated by comparing the answering behavior of questioners who received answers from others to that of those who did not. My identification strategy utilizes the variation in responses to posted questions from others while relying on the exogeneity of the availability of these responses.

To bolster the reliability of the exogeneity assumption, I apply three sample restrictions: focusing on users' first questions, excluding users who posted their first answer before their first question, and restricting the analysis to users who asked a question within one year of registration. These restrictions ensure that at the time of posting their first questions, the sample users' profiles are almost identical to other members' profiles, particularly having only one reputation point and no prior activity on Stack Overflow.[17] This minimizes the likelihood that past contributions influence the chances of receiving responses. Moreover, as the sample users had no previous experience and were new registrants, their expectations about receiving responses were likely similar. This helps reduce the potential for self-selection bias in which some questioners may be hesitant to ask questions they believe are unlikely to receive responses. Therefore, while questioners may expect to receive answers, they have no prior knowledge of whether their questions will be responded to. For this study, I compile a sample of users who registered between 2015 and 2019, resulting in 1,639,304 users (questions).

To examine the effect of receiving answers from other users on the sample users' subsequent answering behavior, I focus on their answering behavior over the seven days following their first question submission. Four types of outcome variables are created to measure this effect. The first, $PostedAnswer$, is a dummy variable indicating whether a sample user answered other users' questions. The second, $AnswerCount$, measures the number of answers provided by a sample user to others, counting the number of other users' questions that they answered. The third, $HasPosScore$, is a dummy variable indicating whether any of a sample user's answers to other users' questions received a positive score, calculated based on the net number of upvotes minus downvotes received within one week of the answer's posting date. The fourth, $IsAccepted$, indicates whether any of a sample user's answers were chosen as the accepted answer.

As some questions on Stack Overflow were answered more than a week after their submission, I arbitrarily categorized questions based on whether they received answers within 24 hours of posting. This threshold is arbitrary, but it captures most responses, as most answers are provided within one day. Figure I shows a histogram of response times for questions answered within seven days of posting. The blue bars represent all questions posted between 2015 and 2020, while the red

[17] Registered users receive one reputation point upon completing Stack Overflow's tutorial.

bars depict the response times for the sample user's first question. The histogram reveals that approximately 93% of questions receive responses within a day, and the likelihood of receiving a response significantly decreases after one day. The patterns of response times are similar for the overall and sample questions.

[Insert Figure I Here]

Although questioners on Stack Overflow cannot predict responses to their questions in advance, the characteristics of the questions may be related to the response rate and the abilities of the questioners. For instance, amateur programmers might lack the skills to write comprehensible questions and may not possess sufficient knowledge to answer other users' questions. Conversely, experienced programmers may easily answer other users' questions, but their questions could be complex and less likely to receive responses. To account for these confounding factors, I create four control variables. The first two variables capture the evaluations of other members—the voting results. I collect the votes cast within one week of the question being posted. Two binary variables are defined: $UpVoted$ is 1 if the question is upvoted and 0 otherwise; $DownVoted$ is 1 if the question is downvoted and 0 otherwise.[18] The third variable is the number of words in the question's body. To help answerers understand their questions, questioners must provide detailed explanations. Despite being a rough measure, the word count is used as a proxy for question quality in online community analysis (Yan and Jian, 2017). I define a binary variable, $ManyWords$, which takes 1 if the question's word count is above the mean of that of the sample questions and 0 otherwise. The fourth variable is the use of a code block feature. When writing questions and answers, users can include explanatory text and programming and error codes. By using the code block feature, they can enhance the readability of their content.[19] I define another binary variable, $CodeBlocked$, which takes 1 if the question includes a code block and 0 otherwise.[20]

The following variables are added as additional control variables. Stack Overflow hosts various question topics in which the difficulty of questions and user skills

[18] While the questioners' ratings might influence the voting results, this possibility is minimized by the sample restriction rules. Sample users had only one reputation point when they posted their first questions, preventing them from voting. Additionally, the short period for collecting vote data likely precludes the questioners from gaining the necessary points to vote later.

[19] Online Appendix Figure A.I provides an example of how content appears when the code block feature is used.

[20] Smirnova, Reitzig and Sorenson (2022) incorporated the use of the code block feature as a control variable in the analysis of online communities.

may vary. To indicate the type of question they are asking, questioners can add up to three tags to each of their questions. I document the tags used in the sample questions and select the 10 most frequent tags.[21] I incorporate these popular tags as dummy variables into the control variables. Moreover, year-week dummies, day-of-the-week dummies, and hour dummies of question posts are included.[22]

I run the following regression for various outcome variables:

$$Y_i = \alpha + \beta Received_i + \gamma \mathbf{Q}_i + \theta \mathbf{Z}_i + \epsilon_i, \tag{1}$$

where $Y_i$ is the outcome variable representing the answering behavior of user $i$ after posting their first question. $Received_i$ is a binary variable that takes 1 if user $i$ received answers from others to their first question within 24 hours of posting and 0 otherwise. The vector $\mathbf{Q}_i$ represents the characteristics of the questions submitted by user $i$, including information on question quality and popular tag dummies. Time-fixed effects are included in the vector $\mathbf{Z}_i$. $\epsilon_i$ is a random shock.

In this study, the parameter of interest, $\beta$, is examined in the context of Stack Overflow, particularly regarding upstream indirect reciprocity. If upstream indirect reciprocity is present, questioners who receive answers to their initial questions from others may feel compelled to reciprocate by assisting others. This could manifest as an increased likelihood to post answers, potentially reflecting a social obligation to participate rather than a genuine desire to provide in-depth help. This hypothesis draws on the theoretical predictions by DellaVigna, List and Malmendier (2012) regarding donation behavior under social pressure. They suggest that people, when feeling socially pressured to donate, often contribute the minimal amount necessary to fulfill this obligation. This implies that the act of donating or, in the context of Stack Overflow, the act of answering questions is more a response to social expectations than an expression of altruistic intent. Consequently, while the probability of engaging in helping behavior (such as answering questions) may increase, the quality or depth of assistance (e.g., providing thorough answers) may not show a corresponding improvement. Therefore, I hypothesize that $\beta > 0$ for indicators like $PostedAnswer$ and $AnswerCount$, reflecting an increased frequency of responses. However, for metrics representing the quality of these responses, such as $HasPosScore$ and $IsAccepted$, the increase in $\beta$ may not be proportional, possibly remaining as $\beta \geq 0$. This aligns with the theory that under social pressure, the quantity of responses might increase but not necessarily their quality.

[21] The order by frequency is as follows: Python, JavaScript, Java, PHP, Android, HTML, C#, C++, jQuery, and CSS.

[22] The time of the question submission is recorded in UTC.

As indicated in studies including that by Stanca (2009), people tend to reciprocate more when they receive more help. Owing to the wide range of skills among users who answer questions on Stack Overflow, the quality of answers also varies greatly. The questioners' assessment of the answers received can be gauged by whether an answer is marked as accepted. If a questioner is satisfied with an answer, they may choose it as the accepted answer. As users at the time of posting their first question do not have voting privileges, selecting an accepted answer is their only way to express their evaluation of the response. As will be shown in the next section in Table I, there is considerable variation in the quality of answers: nearly half of the questioners expressed satisfaction with the answers they received—approximately half of the questioners were not satisfied with the received answers. I estimate the heterogeneous effect of questioners' satisfaction on upstream indirect reciprocity using the following equation:

$$Y_i = \alpha + \beta^S SatisfiedAns_i + \beta^U UnsatisfiedAns_i + \gamma \mathbf{Q}i + \theta \mathbf{Z}i + \epsilon_i. \quad (2)$$

Questions that received answers from others are divided based on the questioners' satisfaction with these answers. A binary variable $SatisfiedAns_i$ is defined as 1 if the questioner $i$ received answers to their first question within 24 hours and chose one of them as the accepted answer and 0 otherwise. Conversely, $UnsatisfiedAns_i$ takes 1 if the questioner $i$ received answers within this time frame but did not select any as the accepted answer. $\beta^S$ and $\beta^U$ capture the effects of receiving satisfying and unsatisfying answers, respectively, on the subsequent answering behavior. I hypothesize that $\beta^S > \beta^U$ and $\beta^S > 0$, assuming that questioners who are satisfied with the received answers will be more inclined to answer other users' questions. The value of $\beta^U$ is uncertain. It could be zero if the questioner does not feel helped due to dissatisfaction with the received answers. Nevertheless, it could be positive if the questioner perceives some level of assistance from the answers, even if they are not fully satisfactory. Therefore, I hypothesize $\beta^U \geq 0$.

# III. Summary Statistics

Here, I present summary statistics and graphs to aid in understanding the role of the sample users on Stack Overflow. Table I shows the summary statistics of the sample users' first question (Sample) and those of the posted questions between 2015 and 2020 (All). The response rate to the first questions posted by the sample users is high (approximately 65%), which is similar to typical Stack Overflow questions, but the quality of the received answers is lower than those

questions. For questioners who received answers from others, the acceptance rate is approximately 46% for sample users' first questions and approximately 57% for typical questions. The low acceptance rate is unlikely due to sample users being unaware of the acceptance option, as the probability of receiving an answer with a positive score is also lower for sample users' first question than for typical questions, approximately 40% and 57%, respectively. The low quality of received answers to sample questions may be attributed to the low quality of questions: those questions had a higher likelihood of receiving downvotes than typical questions, resulting in lower average scores. Additionally, the probability of using code block features in sample questions is lower than that of typical questions. Other characteristics of the sample questions, including the probability of receiving an upvote, the number of words in the question, and the tags used, are similar to those of the typical questions.

[Insert Table I Here]

Compared with users who have newly registered to the forum, sample users have a higher motivation to participate in the community, and their participation is mostly to ask questions. I define an "active user" as one who has posted an answer or a question within one year of registration. Panel A in Figure II shows the number of registered, active, and sample users from 2015 to 2019. While the number of annual registrations, represented by the solid blue line, increases, the number of active and sample users, represented by the yellow dashed and red dotted lines, respectively, remains constant or even decreases slightly. There is a large discrepancy between the number of registrations and that of active and sample users derived from the registrations. The share of active users among newly registered users is approximately 22% on average, indicating that most registered users do not contribute to the forum. The sample users constitute a large share of active users on Stack Overflow: only approximately 36.2% of active users are non-sample users (hereafter referred to as "non-sample active users").

[Insert Figure II Here]

Panel B shows the number of questions posted from 2015 to 2020 by all users and, specifically, active and sample users within one year of their registration.

Panel C displays the corresponding number of answers submitted by these user groups. Notably, I record sample users who registered between 2015 and 2019; the number of posts by sample users in 2015 and 2020 is smaller than in other sample years. As shown in Panel B, numerous questions are submitted by new users. From 2016 to 2019, a total of 7,972,936 questions are posted, with 39.7% (3,166,158) of these questions being posted by users within one year of registration. By construction, non-sample active users have a strong motivation to answer other users' questions because they answer these questions before they ask their own. Of the questions posted by users within one year of registration between 2016 and 2019, approximately 95.3% of them were posted by sample users. Panel C illustrates that the proportion of answers posted by new users is smaller compared to their questions, suggesting a more modest contribution in terms of answers. Over the same span, 10,925,486 answers were posted, of which 21.9% (2,402,792) were by users within their first year of registration. Contrary to the number of posted questions, the number of posted answers by active users and sample users from 2016 to 2019 varies. Although approximately two-thirds of the active users are sample users during this period, the sample users posted only approximately 32.5% of the answers compared to those posted by active users. Therefore, it can be inferred that sample and non-sample active users have different motivations: the motivation for sample users is inclined toward solving their own problems, while that of non-sample active users is inclined toward helping others solve their problems.

# IV. Results

In this section, I first show that the effect of receiving answers from others on the questioners' answering behavior appears to be contrary to the hypothesis of upstream indirect reciprocity. Specifically, as shown in Table II, the analysis shows that questioners who received an answer from others are less likely to post an answer to other users' questions compared to those who did not receive any answers.

However, when determining whether the questioners were satisfied with the answers they received, I find that the effect of receiving satisfactory answers aligns with the hypothesis of upstream indirect reciprocity. For instance, Table III illustrates that when questioners expressed satisfaction with the answers received, their subsequent answering behavior increased. Conversely, those who did not express satisfaction were less likely to post an answer than those who received no answers. These findings underscore the important role of the receivers' satisfaction

in upstream indirect reciprocity. The remainder of this section will demonstrate the robustness of these findings and provide additional insights through short-term and long-term impact analyses.

## IV.A. Effects of Answers on Questioners' Behavior

Table II demonstrates that receiving answers from others to the sample questioners' first posted question has a small but negative effect on their subsequent answering behavior. For instance, as shown in Column (1), questioners who received answers to their first question were 0.17 percentage points less likely to respond to other users' questions in the seven days following their initial submission ($p < 0.001$). This represents a 5.7% decrease compared to the 2.9% likelihood of responding for those who did not receive any answers. The frequency of posting answers also decreased, as illustrated in Column (2) ($p < 0.001$). The adverse effects on the probability of posting high-quality answers were much smaller, as seen in Columns (3) and (4) ($p = 0.079$ and $p = 0.002$, respectively).

[Insert Table II Here]

[Insert Table III Here]

Table III illustrates a nuanced view of how receiving answers from others can heterogeneously impact questioners' subsequent behavior. Supposedly, when questioners feel satisfied with the answers they received, their subsequent answering activity tends to increase. In Column (1), the probability of answering other users' questions within the seven days following the first question submission increases by approximately 0.84 percentage points for questioners satisfied with the answers received ($p < 0.001$). This increase is approximately 28% compared with that of questioners who did not receive any answers. Column (2) reports that the number of answers increases by approximately 0.018 for satisfied questioners ($p < 0.001$).

Moving to the qualitative aspects of the answering behavior, Columns (3) and (4) demonstrate that satisfaction increases the quantity of responses and the probability of providing high-quality answers (both $p < 0.001$). The probability of posting any answer with a positive score increases by 0.38 percentage points,

whereas that of posting any answer chosen by the questioner as the accepted answer increases by 0.29 percentage points, corresponding to approximately 38% and 36% increases, respectively, compared to probabilities of questioners who did not receive any answers.

The results suggest that receiving answers from others increases the likelihood of answering and the probability of posting a quality answer—the observed upstream indirect reciprocity may not be solely driven by social pressure. If questioners felt socially pressured to help others after receiving answers, they would likely exert minimal effort in answering other users' questions. However, the positive effects seen in Columns (3) and (4) suggest a different scenario.

Nevertheless, if the questioner was not satisfied with the answers received from others, a nonnegligible negative effect on the questioner's subsequent response behavior is observed; the magnitude of the effects shown in Columns (1)–(4) is larger or equal to the magnitude of effect of receiving satisfactory answers from others ($p < 0.001$ in all columns). Column (1) shows that if questioners received only unsatisfactory answers from others, the probability of answering behavior within seven days from their first question post decreases by approximately 1 percentage point. This decrease is approximately 34% compared with that of questioners who did not receive any answers. Column (2) reports that the number of answers decreases by approximately 0.0234 for unsatisfied questioners.

The negative effects of receiving an unsatisfactory answer from others are also evident in Columns (3) and (4), suggesting that the observed decrease in answering behavior is not merely a result of ceasing to post low-quality answers. Specifically, for those who received unsatisfactory answers, the probability of posting any answer with a positive score decreases by approximately 38% compared with that of questioners who did not receive any answers, as shown in Column (3). Similarly, the probability of posting any answer chosen by the questioner as the accepted answer decreases by approximately 42% for unsatisfied questioners, as indicated in Column (4). These significant reductions highlight the impact of unsatisfactory answers on discouraging quality contributions from the questioners.

In summary, receiving answers from others seems to be a double-edged sword. If questioners receive satisfactory answers from others, they are more likely to put effort into the community by posting answers to other users' questions. Nonetheless, if they receive answers but are unsatisfied, the questioners' subsequent answering behavior decreases compared to those who did not receive any answers. This suggests that questioner satisfaction is important for the occurrence of upstream indirect reciprocity. The robustness and mechanisms of these results are

discussed later.[23]

## IV.B. Robustness Checks

Subsection IV.A. demonstrates that a questioner's subsequent answering behavior depends on receiving responses from others and their overall satisfaction with those answers. Meanwhile, this subsection provides several supportive evidence for the main findings.

First, the observed pattern is not limited to specific programming languages, particular years, or users with high/low skills. Second, controlling for past answering behavior by analyzing the restricted sample users' second questions does not alter the observed pattern. Third, attrition bias may not be critical. Fourth, the results remain consistent with both alternative outcome definitions and varying timeframes for received answers and when effects are divided based on other community members' evaluations instead of the questioner's own. Finally, direct reciprocity plays a minimal role in the increased contribution behavior resulting from receiving answers from others.

### Subsample Analysis

Is the observed pattern shown in Table III driven by specific clusters within the community? The difficulty level and the characteristics of users may vary depending on the type of question. Therefore, I restrict the analysis to samples that contain the 10 most commonly used tags in the first questions posted by the users. The sample sizes among these 10 tags vary significantly, ranging from 55,194 to 173,089. Online Appendix Table A.II presents the results of this subsample analysis, including further divisions by each popular tag. Although the effects varied to some extent depending on the type of question, a consistent pattern of upstream indirect reciprocity is observed across all popular question tags.

Next, I divide the samples by the registration year to check whether the observed pattern is a transitory phenomenon or not. Online Appendix Table A.III shows the result. The qualitative results are similar to those in Table III, but the magnitude of the effects of receiving satisfactory/unsatisfactory answers from others on the subsequent questioners' answering behavior is decreasing. For instance, in Column (1), the effects on the probability of answering other users' questions are approximately a 1 and 1.2 percentage points increase and decrease in 2015,

[23] Online Appendix Table A.I also reports the results of control variables other than time-fixed effects. Questioners whose questions contain many words, who use the code block feature, and who are upvoted by other users tend to answer more questions for others. Additionally, there is some variation in answering behavior depending on the tags used by the questioner.

respectively; however, these reduce to 0.6 and 0.8 percentage points increase and decrease in 2019, respectively. This may be partly driven by the decline in the overall number of answers; indeed, the probability of answering others' questions for the samples who did not receive any answers from others is also decreasing yearly. Nevertheless, compared with those, the percentage ratios of the effects of receiving answers from others to the means have not decreased.

Does the extent of the response to receiving answers from others vary depending on the characteristics of the question content? For instance, amateur programmers may lack the ability to write comprehensible questions and the ability to answer other users' questions. The extent of their response to receiving answers from others may be small. I divide the sample by variables related to the comprehensibility of the questions—number of words and use of the code block feature—and rerun the analysis. Online Appendix Table A.IV shows the result of the subsample analysis divided by the number of words: I divide the sample according to whether the number of words in the question is higher than the sample average (Panel A) or not (Panel B). Online Appendix Table A.V shows the results according to whether they use the code block feature in their first question (Panel A) or not (Panel B). The observed pattern in these tables is similar to that in Table III, but the magnitude of effects varies. Interestingly, although the samples that use many words in their questions or use the code block feature are more likely to engage in answering other users' questions even if they did not receive any answers from others, the relative magnitude of the percentage ratios of the effects of receiving satisfactory answers from others to the mean outcomes for samples who did not receive any answers from others decreased. In addition, Online Appendix Table A.VI shows the result for samples whose question scores are nonnegative (Panel A) and positive (Panel B). Because most sample question scores are nonnegative, the result is similar to that in Table III. When I rerun the analysis using only samples for which questions' scores are positive, the percentage ratio of the effects of receiving satisfactory answers from others also decreases.

### Answer Exogeneity

While I endeavor to control for confounding effects by restricting samples and using control variables, the possibility of endogeneity due to unobserved ability still remains. If the unobserved users' ability correlates with the probability of receiving answers from others to their questions and the quality of the answers received and their own behavior in responding to other users' questions, the magnitude of the effects estimated in Table III regarding the impact of receiving satisfactory or

unsatisfactory answers may be overestimated. To indirectly assess the exogeneity assumption, I focus on the behavior of users who did not receive answers to their first questions and analyze their second question posting behavior. This approach serves as an indirect check on whether unobserved user ability affects the observed results. I use the users' answering behavior after their first question as a proxy for their ability. If this unobserved ability indeed causes endogeneity, then controlling for this proxy variable would attenuate the estimated effect. Therefore, I compare the magnitude of the effects with and without the inclusion of post-first-question answering behavior in the control variables, aiming to discern the influence of unobserved ability on the patterns observed in Table III.

The estimation equation is as follows:

$$\begin{aligned} Y_i = \alpha + \beta^S SatisfiedAns_i + \beta^U UnsatisfiedAns_i \\ + \eta Y_i^1 + \gamma \mathbf{Q}_i + \gamma^1 \mathbf{Q}_i^1 + \theta \mathbf{Z}_i + \epsilon_i, \end{aligned} \tag{3}$$

where $Y_i$ and $Y_i^1$ are the outcome variables representing questioner $i$'s answering behavior after their second and first question posts, respectively. $SatisfiedAns_i$ is a binary variable that equals 1 if the questioner $i$ received answers from others to their second question within 24 hours of posting and selected one of these as the accepted answer; otherwise, it equals 0. Conversely, $UnsatisfiedAns_i$ is a binary variable that equals 1 if the questioner $i$ received answers from others to their second question within the same timeframe but did not select any as the accepted answer; otherwise, it equals 0. The vectors $\mathbf{Q}_i$ and $\mathbf{Q}_i^1$ represent the characteristics of the second and first questions submitted by the user $i$, respectively. $\mathbf{Z}_i$ represents the time-fixed effects for the second question.

To facilitate the interpretation of the results, I impose the following sample restrictions. First, I only target sample users who had not received an answer to their first question; thus, I estimate the effect of receiving answers for the first time. Second, I limit the sample to those who posted their second question within one week to three months after posting their first question to avoid overlapping the period for measuring answering behavior after the first and second question posts. The number of observations in this analysis is 73,230.

Table IV provides supporting evidence for the main findings. Columns (1), (4), (7), and (10) show the effects of receiving answers on the subsequent answering behavior after the second question post, without controls for first question information. Although the relative magnitude of the effects of receiving satisfactory/unsatisfactory answers is less clear, a similar pattern to that in Table III is observed. For instance, in Column (1), questioners who received satisfactory or

unsatisfactory answers to their second question were more/less likely to answer others' questions over the seven days following the second question post by approximately 1.1 and 1.2 percentage points, respectively. Columns (2), (5), (8), and (11) present the results after adding the answering behavior following the first question post to the control variables. There is a strong positive correlation between the answering behaviors after the first and second question posts; nonetheless, adding this control only slightly attenuates the coefficients of interest, $\beta^S$ and $\beta^U$. For instance, the coefficient of $SatisfiedAns$ marginally decreased from 0.0119 in Column (1) to 0.0107 in Column (2), and the coefficient of $UnsatisfiedAns$ changed from -0.012 to -0.0101. Columns (3), (6), (9), and (12) further control for the characteristics of the first question; however, the coefficients remain relatively stable.

In summary, the analysis reveals a robust correlation between the answering behaviors after the first and second question posts, highlighting significant differences in user ability or motivation. Nevertheless, this correlation does not entirely account for the variations in answering behavior observed between users who received answers and those who did not. This suggests that the likelihood of receiving an answer from others is not readily predictable based on a user's capabilities or motivation.

[Insert Table IV Here]

### Attrition Bias

As the Stack Overflow data excludes deleted questions, the main analysis results may be subject to attrition bias. On Stack Overflow, questions can be deleted in four ways: (1) auto-deletion, (2) deletion by the questioner, (3) deletion upon account withdrawal, and (4) deletion by moderators.

First, Stack Overflow automatically deletes certain questions that remain unanswered for a specific period. Questions with negative scores are deleted 30 days after being posted if they have not received an answer. Moreover, after 365 days, the criteria for auto-deletion become more relaxed, and questions with a score of 0 and fewer views or comments below a certain threshold are also deleted.[24] These conditions are checked on a weekly basis. Notably, questions considered duplicates of previously asked questions are not deleted. All questions used in

[24] https://stackoverflow.com/help/roomba

this study were posted over a year ago, calculated from the last day of the data collection period as the starting point, and most have a score of 0.

Consequently, the remaining unanswered questions in the sample are those that have garnered a certain number of views or comments, indicating that questioners who did not receive an answer are those who asked questions of considerable interest to the community. However, this type of sample selection does not apply to questioners who did receive answers, potentially leading to asymmetric sample selection and bias in the estimates. To ensure the robustness of the main findings, the analysis is replicated with questions posted within the last year and whose scores had not fallen below zero since 30 days after posting.[25] Online Appendix Table A.VII presents the results. While similar implications to those in Table III are observed, the effects of receiving satisfactory answers appear to be even more pronounced in this more recent set of data.

Second, users can delete their own questions if there are no answers or the answers did not receive an upvote.[26] This asymmetric sample selection can also lead to attrition bias. If only users with low motivation retain deletable questions, the estimated effects may reflect differences in user motivation. The robustness of the main results of this type of sample selection is partially demonstrated in Table IV. The sample predominantly comprises users who did not receive answers to their first question but chose not to delete it. Furthermore, the coefficients remained relatively stable even when controlling for users' capabilities and motivation.

Third, users may withdraw from the community by deleting their accounts. It is plausible that questioners who did not receive any answers from others or only received unsatisfactory answers may withdraw from the community. If they leave immediately after submitting their question, the outcome variables for these questioners take zero, which is the minimum possible value for the outcome variables. Therefore, the more the dropouts, the higher the observed average of the outcome variables relative to their true value. This type of sample selection may lead to an underestimation of the positive effect of receiving satisfactory answers from others. However, the direction of bias toward the effect of receiving unsatisfactory answers is unclear.[27]

To examine the potential bias resulting from the withdrawal of questioners who did not receive answers, I conducted the following exercise. I assume that

[25]The most recent submission date for the data I am using is September 25, 2022. I collect questions submitted between March 1 and September 1, 2022.

[26]https://stackoverflow.com/help/deleted-questions

[27]As seen in Online Appendix Table V, those who received unsatisfactory answers tend to continue with community engagement slightly more. Thus, those who did not receive any answers might be more likely to leave, which implies that the estimated effect of receiving unsatisfactory answers is negatively biased.

questioners who did not receive answers leave the community with a probability of $p$, while those who received answers had a withdrawal rate of 0. Given the withdrawal rate $p$, suppose the observed sample size of questioners without answers is $N_C$. Thus, $\frac{1-p}{p}N_C$ would be missing from the sample. I added pseudo-data for the missing samples and reanalyzed the data.[28] For simplicity, I excluded control variables and ran the regression for equation 2. I set the outcome variables for the missing samples to zero, implying that the pseudo-data have zeros for all variables except the constant term. Online Appendix Figure A.III presents the change in the coefficients for $SatisfiedAns$ and $UnsatisfiedAns$ as $p$ varies from 0 to 0.9 in increments of 0.1. Note that the values at $p = 0$ show the result of estimating equation 2 without adding pseudo-data and excluding control variables.[29] The results indicate that the effect of receiving unsatisfactory answers becomes positive when the withdrawal rate is approximately 40% to 50%. While the actual withdrawal rates in Stack Overflow or other online Q&A forums are unknown, considering that membership continuity is free, this threshold rate seems quite high. Therefore, while withdrawal may introduce bias, its impact is likely limited. However, caution is necessary, as the actual size of the effect might differ from the estimated one.

Finally, certain users have the ability to delete questions.[30] These users typically remove spam, offensive content, or questions that are evidently off-topic. Such questioners are not considered in this study as they represent a different cohort from the general user base.

### Alternative Measures

The results in Table III might not accurately represent the causal relationship between receiving answers from others and subsequent answering behavior by the questioners. The definition of the outcome used in Table III includes the potential for reverse causality, where a questioner may post answers to other users' questions immediately after posting their first question and then receive answers within a day. To mitigate this issue, I reanalyze the data using an alternative definition of outcomes, which excludes any answering behavior by the questioners within one day of posting their question. Online Appendix Table A.VIII presents these results, and a similar pattern is observed even when focusing only on the answering behavior after the first day of posting.

[28] This approach is similar to the Fail-safe N analysis (Rosenthal, 1979) used in meta-analysis.

[29] The estimated effects closely align to those presented in Table III.

[30] Moderators can delete any question, and users with over 10,000 reputation points can cast delete votes on closed questions. For more information on moderators, see: https://stackoverflow.com/help/site-moderators.

In addition to varying the definition of the outcome, I conduct a reanalysis using a stricter criterion for when a questioner is considered to have received an answer. Specifically, I define two new binary variables: $SatisfiedAnsHour$ and $UnsatisfiedAnsHour$. $SatisfiedAnsHour$ takes 1 if a questioner received answers within an hour of their question posting and selected one as the accepted answer. However, $UnsatisfiedAnsHour$ is 1 if a questioner received answers in the same period but did not select any as the accepted answer. This adjustment reduced the proportion of questioners considered to have received an answer from 66% to 33%. The results of this analysis are presented in Online Appendix Table A.IX. Despite the decrease in the proportion of questioners receiving answers under this stricter definition, the observed patterns remain consistent with those seen in Table III.

Instead of using the questioners' subjective evaluations of the answers they received to assess the heterogeneous effects of receiving answers from others, I use evaluations from other community members to analyze the heterogeneity. I define two new binary variables: $PositiveAnswer$ and $NonPositiveAnswer$. $PositiveAnswer$ is a dummy variable indicating that the respondent received answers from others to their first question within 24 hours of posting, and at least one of these has a positive score. Conversely, $NonPositiveAnswer$ is a dummy variable indicating that the respondent received answers within the same timeframe, but none of these have a positive score. Online Appendix Table A.X shows the results. Although the magnitude of the coefficients is smaller than those in Table III, a similar pattern is observed. If the questioner received any answers from others that are highly evaluated by other members, they are more likely to answer other users' questions compared to those who did not receive any answers. Nevertheless, if they received answers, but none were highly evaluated, they are less likely to respond than those who received no answers at all. As shown in Table I, among samples who received answers from others, about 40% received at least one answer with a positive score, which is not far from the approximately 46% who selected one of the received answers as the accepted answer. The attenuated magnitude of coefficients in the results indicates that the other users' objective evaluation is coarse information and the questioners' subjective evaluation may be more influential in the occurrence of upstream indirect reciprocity.

### Direct Reciprocity

While the positive effect of receiving answers from others on the subsequent questioner's answering behavior is interpreted as evidence of upstream indirect reci-

procity, the effect may also be explained by direct reciprocity if the questioner posts answers to the questions of those who answered theirs. To assess the extent of direct reciprocity in the increased answering behavior, I check the answerer IDs of the first question of sample users and the questioner IDs answered by these same users over the seven days following their first question submission. I define an index of direct reciprocity for each user, which takes the value 1 if the user answered a question posed by someone who had previously answered one of their questions and 0 otherwise. The average of the direct reciprocity index across users provides insight into the extent of direct reciprocity practiced; it indicates the proportion of users who engage in direct reciprocity by responding to the questions of those who previously answered theirs. If the observed answering behavior is entirely due to direct reciprocity, the mean of the direct reciprocity index is 1; however, if no direct reciprocity was involved, the mean is 0.

The auxiliary analysis suggests that direct reciprocity does not play a significant role in the observed answering behavior. I limit the samples to those who received answers from others to their first question ($Received = 1$) and posted an answer to other users' questions within the seven days following the first question submission ($PostedAnswer = 1$). With 28,621 users who met this condition, only 0.9% (282 users) answered the question of a user who previously posted an answer to their question.

## IV.C. Temporal Effects: Daily and Long-term Impact

This subsection presents the results of two additional analyses: the short- and long-run effects of receiving answers from others on questioners' subsequent behavior following their first question post. In the short-term analysis, the effect of receiving answers is examined daily to track how the impact transitions over time. The long-term analysis extends this examination up to one year post initial question, assessing the sustained impact on answering behavior and the overall engagement with the community, including answering and question-asking activities.

### Daily Variation in Effects

In the short term, the effect of receiving satisfactory or unsatisfactory answers on the questioners' behavior in the week following their first question posting may vary. For instance, Allcott and Rogers (2014) report that the impact of energy-saving nudges induced by social norms diminishes over time. Similarly, Schmitz (2019) demonstrate intertemporal substitution in prosocial behavior. In

this context, the motivation of sample users to answer other users' questions might initially increase after receiving any form of an answer, including unsatisfactory ones, but this motivation might decrease over time. Consequently, even for those who received unsatisfactory answers, there might be a temporary surge in the inclination to answer other users' questions, possibly as an initial reaction to the act of receiving an answer.

I decompose the effects of receiving answers from others to their first questions into daily effects on the subsequent questioners' answering behavior over the seven days following the first question submission. The estimation equation 2 is rewritten as follows:

$$\begin{aligned} Y_{it} =& \alpha + \sum_{d=0}^{6} \beta_d^S SatisfiedAns_i \times 1(t=d) \\ &+ \sum_{d=0}^{6} \beta_d^U UnsatisfiedAns_i \times 1(t=d) \\ &+ \sum_{d=1}^{6} \delta_d 1(t=d) + \gamma \mathbf{Q}_i + \theta \mathbf{Z}_i + \epsilon_i, \end{aligned} \tag{4}$$

where $t$ represents the elapsed days from the first question post, ranging from 0 to 6, and $Y_{it}$ is the outcome variable representing the questioner $i$'s answering behavior after $t$ days from the first question post. $1(t=d)$ is an indicator function for $d$ days elapsed after the first question post. $\beta_d^S$ and $\beta_d^U$ capture the effects of receiving satisfactory and unsatisfactory answers, respectively, from others on the subsequent questioners' answering behavior $d$ days after the first question post, compared with that of questioners who did not receive answers from others.$\delta_d$ captures the change in the questioners' answering behavior $d$ days after the first question post. In all regression estimations, I use individual-level clustered robust standard errors to account for the serial correlation of $\epsilon_{it}$ within individuals.[31]

Figure III plots the transition patterns of $\beta_d^S$ (represented by the blue solid line) and $\beta_d^U$ (represented by the red dashed line). The details of the coefficients of regression are reported in Online Appendix Table A.XI. The transition patterns for both $\beta_d^S$ and $\beta_d^U$ across four different outcomes show similar trends: (1) no evidence of backsliding is observed, with effects not shifting from positive to negative; (2) the daily effects of receiving satisfactory answers are consistently greater than those of receiving unsatisfactory answers; (3) even after receiving satisfactory answers, the probability of answering within 24 hours of the question post often decrease; (4) those who received unsatisfactory answers experienced a particularly

[31] To avoid computational issues, 20% of the sample users were randomly selected for analysis.

significant decrease in their answering behavior within 24 hours of the question post ($p < 0.001$).

As shown in Online Appendix Table A.XI, questioners tend to answer others' questions around the same time they post their questions, with a significant drop in answering probability on subsequent days. This transition pattern is distinct from the effects of receiving answers—the increase in answering behavior among those who received satisfactory answers is likely due to an increase in the extensive margin, not the intensive margin. A plausible explanation for the decrease in immediate answering behavior upon receiving answers is that questioners might spend time solving their own problems based on the answers they received. Therefore, the significant decrease in answering behavior among those who received unsatisfactory answers possibly indicates confusion and a resulting lack of capacity to effectively answer other users' questions.

[Insert Figure III Here]

### Long-term Impact Analysis

To what extent does the experience of the response to the first user's question affect their behavior in the long run? Here, I analyze the long-run effect of receiving satisfactory/unsatisfactory answers from others to questioners' subsequent answering behavior, extending the tracking time to 365 days after their first question post. In addition, I examine the effect of receiving answers from others on questioners' engagement to the community. This is to explore whether the negative effect of receiving unsatisfactory answers from others on subsequent answering behavior means the questioner abandons the contribution to the community. For tracking any activity for users, I create a binary outcome variable, *Engaged*, that takes 1 if the questioner posted an answer or a question over a year following their first question post and 0 otherwise.

Even a year after the first question post, differences in questioners' answering behavior persist between those who received satisfactory/unsatisfactory answers from others and those who did not receive any answers. Table V presents the results. For example, in Column (1), receiving satisfactory answers from others leads to a 2.9 percentage point increase in the probability of posting an answer to other users' questions within a year after the first question, representing a 29% increase compared with the 10% probability of questioners who did not receive

any answers. Receiving unsatisfactory answers decreases this probability by approximately 3 percentage points, amounting to a 30% decrease. These effects are significantly larger than those shown in Table III, which focused on the seven days following the first question post. The most notable increase is seen in the number of answers posted, with the effect of receiving satisfactory or unsatisfactory answers being approximately 12 times or 7 times larger, respectively. These findings indicate that receiving answers and the quality of those answers have a long-term impact on questioners' answering behavior.

Moreover, as shown in Column (5), questioners who received answers from others tend to be more engaged in the community. Receiving satisfactory answers increases the likelihood of subsequent answering and asking behavior. The probability of posting either an answer or a question within a year from their first question increased by approximately 19.3 percentage points, more than half compared with the 35% for those who did not receive any answers. Interestingly, while questioners who received unsatisfactory answers show a decrease in subsequent answering behavior, as shown in Columns (1)–(4), their overall activity modestly increases when not considering withdrawal effects. This suggests that the negative impact of receiving unsatisfactory answers on subsequent answering behavior is not due to abandoning the community.

Furthermore, as shown in Column (2), the effect of receiving satisfactory answers from others on the number of answers posted is larger than that of receiving unsatisfactory answers. In Columns (3) and (4), the probability of posting a high-quality answer is also higher for those who received satisfactory answers. The total effects are shown in Online Appendix Table A.XII. The results indicate that the overall effect of receiving answers from others does not adversely affect the questioners' subsequent behavior, except for the probability of answering.

[Insert Table V Here]

# V. Discussion and Conclusion

This study aims to shed light on the importance of upstream indirect reciprocity—in which people who received help from someone are more likely to help others—in the field. To provide evidence of this phenomenon, I analyzed data from Stack Overflow, one of the most popular online Q&A forums. This platform was chosen

due to its vast user base and rich data on user interactions, making it an ideal setting to study social behaviors such as upstream indirect reciprocity. In the context of upstream indirect reciprocity, receiving answers from others is considered as receiving help and answering others' questions as helping others. In this regard, I estimated the effect of receiving answers on questioners' subsequent answering behavior. If upstream indirect reciprocity exists, a positive effect on answering behavior was expected. The analysis focused on the effects of answers received to the first question, allowing for a more precise examination by excluding confounding effects. Furthermore, to capture the nuances of this effect, I assessed the heterogeneity of the impact using questioners' satisfaction with the answers received.

In this forum, upstream indirect reciprocity was observed only when the questioner expressed satisfaction with the answers received. For instance, the effect of receiving satisfactory answers on the probability of answering other users' questions within seven days of the first question increased by approximately 28%, compared with the 2.9% probability for those who did not receive any answers. This positive effect extended beyond answering probability to the number of answers given and the likelihood of posting high-quality answers. However, the overall effect of receiving answers tended to be negative due to the adverse effects of unsatisfactory answers on subsequent answering behavior.

I provided supporting evidence for the main findings. The observed patterns proved robust across various subsample analyses, including adding proxies for questioners' ability or motivation to the control variables and accounting for attrition bias. Additionally, the robustness was maintained when using alternative definitions of the outcome, varying timeframes for received answers, and dividing effects based on other community members' evaluations. Furthermore, the increased answering behavior was unlikely due to direct reciprocity—questioners answered questions posted by individuals who had previously answered their questions.

The short-term and long-term effects provided additional insights into the main findings. The daily impact transition patterns of receiving answers from others did not show backsliding or a shift from positive to negative effects. The positive effects of receiving satisfactory answers were attributed to the increase in answering behavior after a day of question posting, as questioners who received answers reduced their answering behavior within 24 hours after posting. This contrasts with the findings of Mujcic and Leibbrandt (2017), who reported upstream indirect reciprocity immediately after receiving help. The effects of receiving satisfactory/unsatisfactory answers persisted for a year following the first

question post. The accumulated positive effects of receiving satisfactory answers were more significant than those of unsatisfactory answers; therefore, in the long run, the negative effect of receiving any answers, except for the probability of answering, dissipated. Additionally, receiving answers from others increased the community engagement rate.

Although my study does not directly explore the mechanism of upstream indirect reciprocity, the findings provide some implications regarding the underlying processes. First, the observed upstream indirect reciprocity is unlikely to be solely due to social pressure. The results indicated that as the probability of answering increased, so did the likelihood of questioners who received satisfactory answers providing high-quality responses. This finding contrasts with the theoretical predictions of DellaVigna, List and Malmendier (2012), who propose that prosocial behavior, such as donations, driven by social pressure might result in minimal contributions just sufficient to escape the pressure. In this context, questioners might post low-quality answers as a parallel to making minimal donations under social pressure.[32] Moreover, the observed reduction in effect when the definition of high-quality answers was based on other users' high evaluations, rather than questioners' satisfaction, implies that questioners' response to receiving satisfactory answers is not primarily driven by others' expectations. This suggests that the reaction of questioners to receiving answers is not simply a response to the community's evaluation but possibly influenced by their own perceptions of answer quality.

Second, the role of recipient satisfaction in fostering upstream indirect reciprocity is highlighted in this study. Specifically, it was found that only those questioners who received satisfactory answers were more likely to increase their subsequent answering behavior. This finding supports the notion that satisfaction may be a key factor in motivating such reciprocal behavior and may align with psychological literature emphasizing the role of gratitude in prosocial behavior and upstream indirect reciprocity. For example, McCullough, Kimeldorf and Cohen (2008) discusses the significance of gratitude in this context, while Bartlett and DeSteno (2006) demonstrates how receiving assistance influences gratitude and leads to further prosocial actions toward third parties.[33] Moreover, Watkins et al. (2006) suggests that gratitude is different from emotions like indebtedness

[32]This result could be explained by social pressure if questioners who received satisfactory answers felt an obligation to provide satisfying answers to others, thereby responding to a sense of duty rather than external pressure.

[33]However, Ma, Tunney and Ferguson (2017)'s meta-analysis found a weaker correlation for upstream indirect reciprocity in relation to gratitude compared to direct reciprocity, although the limited number of studies on upstream indirect reciprocity necessitates careful interpretation of these findings.

and obligation, which are often associated with social pressure to reciprocate. This additional insight indicates that the observed upstream indirect reciprocity may not be entirely attributed to social pressure. Conversely, questioners who received unsatisfactory answers showed decreased subsequent answering behavior. This outcome is unexpected because, despite the answers not fully addressing the questioners' problems, the answerers made some effort, which might have been perceived as kindness. The negative effect of receiving unsatisfactory answers may partly arise from questioners struggling more with their problems after such answers; however, this remains an open question. Therefore, studying the dynamics of emotional and psychological responses to receiving answers could provide deeper insights into the mechanisms behind observed behaviors, and future research could address this aspect.

Furthermore, it appears unlikely that the observed upstream indirect reciprocity is driven by information about social norms, as some studies like Frey and Meier (2004): Shang and Croson (2009): Allcott and Rogers (2014) suggested. Users on the platform can readily observe the amount and quality of answers on the question listing page, as shown in Online Appendix Figure A.II. Consequently, the presence or quality of responses to their own questions does not contribute additional normative information.

This research offers important implications for online community owners. It demonstrates that receiving help from others is a double-edged sword. When the help is satisfactory, recipients are more likely to assist others within the community; however, unsatisfactory help can lead to adverse effects. Therefore, community owners should focus on designs that enhance recipient satisfaction. This could include support for more effective communication of needs and assistance, as well as mechanisms for automatically filtering out suspiciously low-quality posts to prevent dissatisfaction.

This study has highlighted the significance of upstream indirect reciprocity. Future research could explore two distinct directions. First, there is a need for a theoretical model specifically dedicated to upstream indirect reciprocity. While existing models have primarily focused on direct reciprocity, effective in explaining interactions in small groups, they do not adequately capture behaviors within larger groups such as those involving upstream indirect reciprocity. Although Cox, Friedman and Gjerstad (2007)'s emotional state-dependent model may explain some aspects of this study's findings, particularly the increased prosociality following satisfactory responses, it falls short in outlining the full range of conditions that might elevate prosocial behavior. Second, a deeper exploration into the mechanism of upstream indirect reciprocity is crucial. For instance, con-

trolled experiments are needed to determine if motivations identified in direct reciprocity and other forms of prosocial behavior are also applicable in the context of upstream indirect reciprocity. Undertaking such research could clarify the distinctions between these forms of behavior and facilitate the development of more comprehensive theoretical models.

# References


**Allcott, Hunt, and Todd Rogers.** 2014. "The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation." *Am. Econ. Rev.*, 104(10): 3003–3037.

**Alpizar, Francisco, Fredrik Carlsson, and Olof Johansson-Stenman.** 2008. "Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica." *J. Public Econ.*, 92(5): 1047–1060.

**Bahr, Gunter, and Till Requate.** 2014. "Reciprocity and giving in a consecutive three-person dictator game with social interaction." *Ger. Econ. Rev.*, 15(3): 374–392.

**Baker, Wayne E, and Nathaniel Bulkley.** 2014. "Paying it forward vs. rewarding reputation: Mechanisms of generalized reciprocity." *Organization Science*, 25(5): 1493–1510.

**Bartlett, Monica Y, and David DeSteno.** 2006. "Gratitude and prosocial behavior: helping when it costs you." *Psychol. Sci.*, 17(4): 319–325.

**Ben-Ner, Avner, Louis Putterman, Fanmin Kong, and Dan Magan.** 2004. "Reciprocity in a two-part dictator game." *J. Econ. Behav. Organ.*, 53(3): 333–352.

**Bohren, J Aislinn, Alex Imas, and Michael Rosenberg.** 2019. "The dynamics of discrimination: Theory and evidence." *Am. Econ. Rev.*, 109(10): 3395–3436.

**Bolton, Gary E, and Axel Ockenfels.** 2000. "ERC: A Theory of Equity, Reciprocity, and Competition." *Am. Econ. Rev.*, 90(1): 166–193.

**Bolton, Gary E, Elena Katok, and Axel Ockenfels.** 2005. "Cooperation among strangers with limited information about reputation." *J. Public Econ.*, 89(8 SPEC. ISS.): 1457–1468.

**Botelho, Tristan L, and Marina Gertsberg.** 2021. "The Disciplining Effect of Status: Evaluator Status Awards and Observed Gender Bias in Evaluations." *Manage. Sci.*

**Broberg, Tomas, Tore Ellingsen, and Magnus Johannesson.** 2007. "Is generosity involuntary?" *Econ. Lett.*, 94(1): 32–37.

**Burtch, Gordon, Qinglai He, Yili Hong, and Dokyun Lee.** 2022. "How Do Peer Awards Motivate Creative Content? Experimental Evidence from Reddit." *Manage. Sci.*, 68(5): 3488–3506.

**Charness, Gary, and Matthew Rabin.** 2002. "Understanding Social Preferences with Simple Tests*." *Q. J. Econ.*, 117(3): 817–869.

**Chen, Yan, F Maxwell Harper, Joseph Konstan, and Sherry Xin Li.** 2010. "Social comparisons and contributions to online communities: A field experiment on MovieLens." *Am. Econ. Rev.*, 100(4): 1358–1398.

**Clark, Daniel, Drew Fudenberg, and Alexander Wolitzky.** 2020. "Indirect reciprocity with simple records." *Proc. Natl. Acad. Sci. U. S. A.*, 117(21): 11344–11349.

**Cooper, David J, and E Glenn Dutcher.** 2011. "The dynamics of responder behavior in ultimatum games: a meta-study." *Exp. Econ.*, 14(4): 519–546.

**Cox, James C.** 2004. "How to identify trust and reciprocity." *Games Econ. Behav.*, 46(2): 260–281.

**Cox, James C, and Cary A Deck.** 2005. "On the nature of reciprocal motives." *Econ. Inq.*, 43(3): 623–635.

**Cox, James C, Daniel Friedman, and Steven Gjerstad.** 2007. "A tractable model of reciprocity and fairness." *Games Econ. Behav.*, 59(1): 17–45.

**Cox, James C, Daniel Friedman, and Vjollca Sadiraj.** 2008. "Revealed altruism." *Econometrica*, 76(1): 31–69.

**Dana, Jason, Roberto A Weber, and Jason Xi Kuang.** 2007. "Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness." *Econom. Theory*, 33(1): 67–80.

**DellaVigna, Stefano, John A List, and Ulrike Malmendier.** 2012. "Testing for altruism and social pressure in charitable giving." *Q. J. Econ.*, 127(1): 1–56.

**DellaVigna, Stefano, John A List, Ulrike Malmendier, and Gautam Rao.** 2022. "Estimating Social Preferences and Gift Exchange at Work." *Am. Econ. Rev.*, 112(3): 1038–1074.

**Dufwenberg, Martin, and Georg Kirchsteiger.** 2004. "A theory of sequential reciprocity." *Games Econ. Behav.*, 47(2): 268–298.

**Dufwenberg, Martin, Uri Gneezy, Werner Guth, and Eric van Damme.** 2001. "Direct vs Indirect Reciprocity: An Experiment." *Homo Oeconomicus*, 18(AUGUST 2000): 19–30.

**Engelmann, Dirk, and Urs Fischbacher.** 2009. "Indirect reciprocity and strategic reputation building in an experimental helping game." *Games Econ. Behav.*, 67(2): 399–407.

**Falk, Armin.** 2007. "Gift exchange in the field." *Econometrica*, 75(5): 1501–1511.

**Falk, Armin, and Urs Fischbacher.** 2006. "A theory of reciprocity." *Games Econ. Behav.*, 54(2): 293–315.

**Fehr, Ernst, and Klaus M Schmidt.** 1999. "A Theory of Fairness, Competition, and Cooperation." *Q. J. Econ.*, 114(3): 817–868.

**Fehr, Ernst, and Simon Gächter.** 2000. "Fairness and Retaliation: The Economics of Reciprocity." *J. Econ. Perspect.*, 14(3): 159–181.

**Frey, Bruno S, and Stephan Meier.** 2004. "Social Comparisons and Pro-social Behavior: Testing "Conditional Cooperation" in a Field Experiment." *Am. Econ. Rev.*, 94(5): 1717–1722.

**Gallus, Jana.** 2017. "Fostering Public Good Contributions with Symbolic Awards: A Large-Scale Natural Field Experiment at Wikipedia." *Manage. Sci.*, 63(12): 3999–4015.

**Gneezy, Uri, and John A List.** 2006. "Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments." *Econometrica*, 74(5): 1365–1384.

**Gray, Kurt, Adrian F Ward, and Michael I Norton.** 2014. "Paying it forward: Generalized reciprocity and the limits of generosity." *J. Exp. Psychol. Gen.*, 143(1): 247–254.

**Greenstein, Shane, Grace Gu, and Feng Zhu.** 2021. "Ideology and Composition Among an Online Crowd: Evidence from Wikipedians." *Manage. Sci.*, 67(5): 3067–3086.

**Greiner, Ben, and M Vittoria Levati.** 2005. "Indirect reciprocity in cyclical networks: An experimental study." *J. Econ. Psychol.*, 26(5): 711–731.

**Herne, Kaisa, Olli Lappalainen, and Elina Kestilä-Kekkonen.** 2013. "Experimental comparison of direct, general, and indirect reciprocity." *J. Socio Econ.*, 45: 38–46.

**Hilbe, Christian, Laura Schmid, Josef Tkadlec, Krishnendu Chatterjee, and Martin A Nowak.** 2018. "Indirect reciprocity with private, noisy, and incomplete information." *Proc. Natl. Acad. Sci. U. S. A.*, 115(48): 12241–12246.

**Horita, Yutaka, Masanori Takezawa, Takuji Kinjo, Yo Nakawake, and Naoki Masuda.** 2016. "Transient nature of cooperation by pay-it-forward reciprocity." *Sci. Rep.*, 6(July 2015): 1–11.

**Iwagami, Akio, and Naoki Masuda.** 2010. "Upstream reciprocity in heterogeneous networks." *J. Theor. Biol.*, 265(3): 297–305.

**Johnson, Noel D, and Alexandra A Mislin.** 2011. "Trust games: A meta-analysis." *J. Econ. Psychol.*, 32(5): 865–889.

**Khadjavi, Menusch.** 2017. "Indirect reciprocity and charitable giving-evidence from a field experiment." *Manage. Sci.*, 63(11): 3708–3717.

**Lazear, Edward P, Ulrike Malmendier, and Roberto A Weber.** 2012. "Sorting in experiments with application to social preferences." *Am. Econ. J. Appl. Econ.*, 4(1): 136–163.

**Leimar, O, and P Hammerstein.** 2001. "Evolution of cooperation through indirect reciprocity." *Proceedings of the Royal Society B: Biological Sciences*, 268(1468): 745–753.

**Levine, David K.** 1998. "Modeling Altruism and Spitefulness in Experiments." *Rev. Econ. Dyn.*, 1(3): 593–622.

**Linek, Maximilian, and Christian Traxler.** 2021. "Framing and social information nudges at Wikipedia." *J. Econ. Behav. Organ.*, 188: 1269–1279.

**Ma, Lawrence K, Richard J Tunney, and Eamonn Ferguson.** 2017. "Does gratitude enhance prosociality?: A meta-analytic review." *Psychol. Bull.*, 143(6): 601–635.

**Malmendier, Ulrike, Vera L Te Velde, and Roberto A Weber.** 2014. "Rethinking reciprocity." *Annu. Rev. Econom.*, 6: 849–874.

**McCabe, Kevin A, Mary L Rigdon, and Vernon L Smith.** 2003. "Positive reciprocity and intentions in trust games." *J. Econ. Behav. Organ.*, 52(2): 267–275.

**McCullough, Michael E, Marcia B Kimeldorf, and Adam D Cohen.** 2008. "An Adaptation for Altruism: The Social Causes, Social Effects, and Social Evolution of Gratitude." *Curr. Dir. Psychol. Sci.*, 17(4): 281–285.

**Milinski, M, D Semmann, T C Bakker, and H J Krambeck.** 2001. "Cooperation through indirect reciprocity: image scoring or standing strategy?" *Proc. Biol. Sci.*, 268(1484): 2495–2501.

**Mujcic, Redzo, and Andreas Leibbrandt.** 2017. "Indirect Reciprocity and Prosocial Behaviour: Evidence from a Natural Field Experiment." *Econ J*, 128(611): 1683–1699.

**Nowak, M A, and K Sigmund.** 1998. "Evolution of indirect reciprocity by image scoring." *Nature*, 393(6685): 573–577.

**Nowak, Martin A, and Karl Sigmund.** 2005. "Evolution of indirect reciprocity." *Nature*, 437(7063): 1291–1298.

**Nowak, Martin A, and Sebastien Roch.** 2007. "Upstream reciprocity and the evolution of gratitude." *Proceedings of the Royal Society B: Biological Sciences*, 274(1610): 605–609.

**Oosterbeek, H, R Sloof, and G Van De Kuilen.** 2004. "Cultural differences in ultimatum game experiments: Evidence from a meta-analysis." *Exp. Econ.*

**Pfeiffer, Thomas, Claudia Rutte, Timothy Killingback, Michael Taborsky, and Sebastian Bonhoeffer.** 2005. "Evolution of cooperation by generalized reciprocity." *Proceedings of the Royal Society B: Biological Sciences*, 272(1568): 1115–1120.

**Rabin, Matthew.** 1993. "Incorporating Fairness into Game Theory and Economics." *Am. Econ. Rev.*, 83(5): 1281–1302.

**Rankin, Daniel J, and Michael Taborsky.** 2009. "Assortment and the evolution of generalized reciprocity." *Evolution*, 63(7): 1913–1922.

**Regner, Tobias.** 2018. "Reciprocity under moral wiggle room: Is it a preference or a constraint?" *Exp. Econ.*, 21(4): 779–792.

**Rosenthal, Robert.** 1979. "The file drawer problem and tolerance for null results." *Psychol. Bull.*, 86(3): 638–641.

**Schmitz, Jan.** 2019. "Temporal dynamics of pro-social behavior: an experimental analysis." *Exp. Econ.*, 22(1): 1–23.

**Schnedler, Wendelin.** 2022. "The broken chain: Evidence against emotionally driven upstream indirect reciprocity." *Games Econ. Behav.*, 136: 542–558.

**Seinen, Ingrid, and Arthur Schram.** 2006. "Social status and group norms: Indirect reciprocity in a repeated helping experiment." *Eur. Econ. Rev.*, 50(3): 581–602.

**Shang, Jen, and Rachel Croson.** 2009. "A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods." *Econ. J.*, 119(540): 1422–1439.

**Smirnova, Inna, Markus Reitzig, and Olav Sorenson.** 2022. "Building Status in an Online Community." *Organization Science*.

**Sobel, Joel.** 2005. "Interdependent preferences and reciprocity." *J. Econ. Lit.*, 43(2): 392–436.

**Stanca, Luca.** 2009. "Measuring indirect reciprocity: Whose back do we scratch?" *J. Econ. Psychol.*, 30(2): 190–202.

**Stanca, Luca, Luigino Bruni, and Luca Corazzini.** 2009. "Testing theories of reciprocity: Do motivations matter?" *J. Econ. Behav. Organ.*, 71(2): 233–245.

**Strang, Sabrina, Xenia Grote, Katarina Kuss, Soyoung Q Park, and Bernd Weber.** 2016. "Generalized negative reciprocity in the dictator game - How to interrupt the chain of unfairness." *Sci. Rep.*, 6(February): 10–14.

**Sun, Zhongqiang, Chuyuan Ye, Zhihui He, and Wenjun Yu.** 2020. "Behavioral Intention Promotes Generalized Reciprocity: Evidence From the Dictator Game." *Front. Psychol.*, 11(April): 1–8.

**Van Apeldoorn, Jacobien, and Arthur Schram.** 2016. "Indirect reciprocity; A field experiment." *PLoS One*, 11(4): 1–11.

**van den Akker, Olmo R, Marcel A L M van Assen, Mark van Vugt, and Jelte M Wicherts.** 2020. "Sex differences in trust and trustworthiness: A meta-analysis of the trust game and the gift-exchange game." *J. Econ. Psychol.*, 81: 102329.

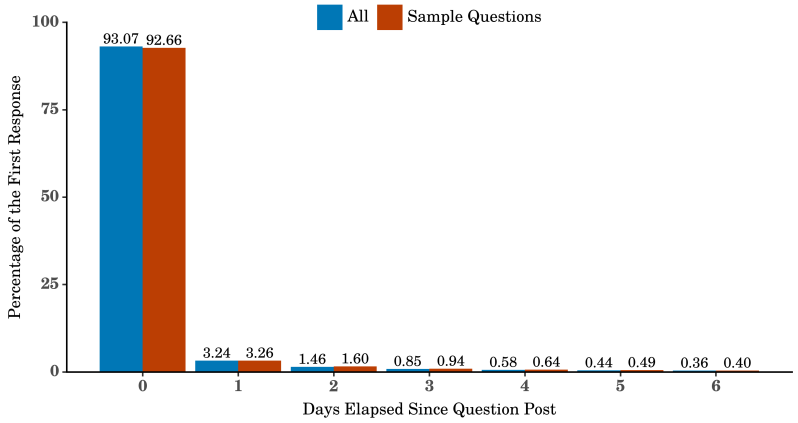
**van Doorn, Gerrit Sander, and Michael Taborsky.** 2012. "The evolution of generalized reciprocity on social interaction networks." *Evolution*, 66(3): 651–664.

**Watkins, Philip, Jason Scheer, Melinda Ovnicek, and Russell Kolts.** 2006. "The debt of gratitude: Dissociating gratitude and indebtedness." *Cognition and Emotion*, 20(2): 217–241.

**Wu, Philip Fei, and Nikolaos Korfiatis.** 2013. "You scratch someone's back and we'll scratch yours: Collective reciprocity in social Q&A communities." *J. Am. Soc. Inf. Sci. Technol.*, 64(10): 2069–2077.

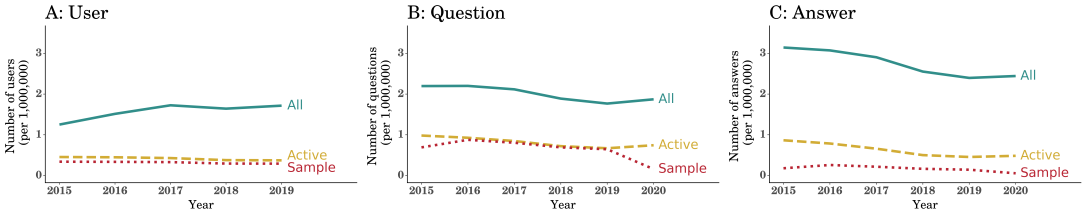
**Xu, Lei, Tingting Nian, and Luís Cabral.** 2020. "What makes geeks tick? A study of stack overflow careers." *Manage. Sci.*, 66(2): 587–604.

**Yan, Bei, and Lian Jian.** 2017. "Beyond reciprocity: The bystander effect of knowledge response in online knowledge communities." *Comput. Human Behav.*, 76: 9–18.



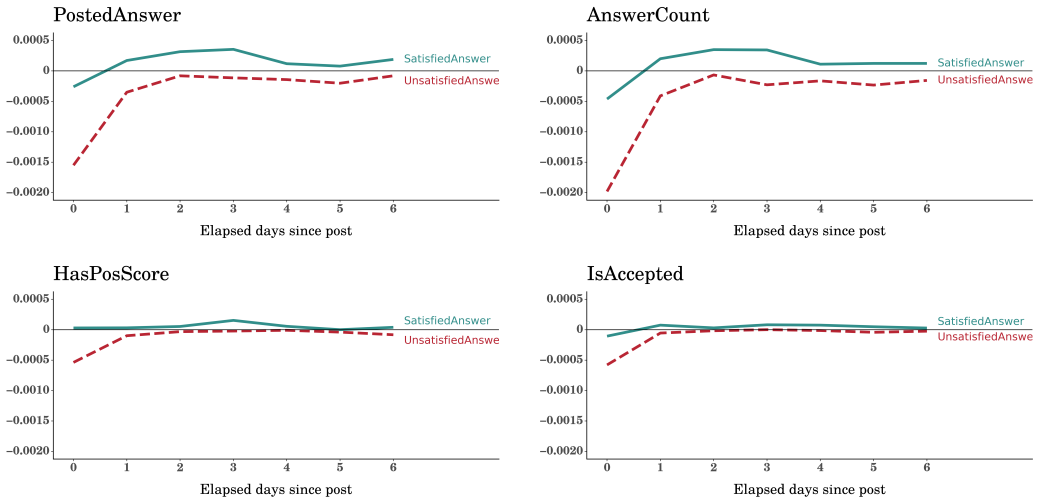

Notes: This figure displays histograms of the number of days taken to receive the first response from others for all questions (represented by left red bars) posted from 2015 to 2020 and the first questions posted by the sample (represented by right blue bars). Only questions that received a response within seven days of posting are included in the calculation. The y-axis indicates the percentage of questions that received their first response on day $d$ out of all questions that received a response within seven days.

Figure I: Response Time Distribution for All Questions vs. First Questions (2015-2020)




Notes: In Panel A, "All" indicates the annual number of new user registrations; "Active" represents the annual number of new registrations by active users; and "Sample' represents the annual number of new registrations by sample users. In Panels B and C, "All" represents the annual number of questions and answers posted, respectively; "Active" and "Sample" in these panels represent the count of posts made within one year of registration by active and sample users, respectively. The y-axis indicates counts per 100,000 for all panels.

Figure II: User Participation and Activity on Stack Overflow (2015-2020)




Notes: Each panel illustrates the transition of daily impact on the subsequent answering behavior of questioners from day 0 to day 6 following their initial question post. The values presented in the figure correspond to the coefficients detailed in Online Appendix Table A.XI. $SatisfiedAns$ (represented by a solid blue line) is a dummy variable indicating that the respondent received answers from others to their first question within 24 hours of posting and selected one of these as the accepted answer. Conversely, $UnsatisfiedAns$ (represented by a dotted red line) is a dummy variable indicating that the respondent received answers within the same timeframe but did not select any as the accepted answer.

Figure III: Transition of Daily Impact on Subsequent Answering Behavior

Table I: Summary Statistics of the First Question Posted by Sample Users and All Questions (2015-2020)

| | Sample | | All | |
|---|---|---|---|---|
| | *Mean* | *SD* | *Mean* | *SD* |
| Received | 0.664 | 0.472 | 0.655 | 0.475 |
| Acceptance Rate | 0.463 | 0.499 | 0.568 | 0.495 |
| Positive Score Answers | 0.404 | 0.491 | 0.573 | 0.495 |
| Number of Upvotes | 0.545 | 1.259 | 0.533 | 1.442 |
| Number of Downvotes | 0.425 | 0.993 | 0.242 | 0.717 |
| Question Score | 0.121 | 1.574 | 0.291 | 1.556 |
| Number of Words | 73.795 | 86.434 | 75.485 | 76.161 |
| CodeBlocked | 0.790 | 0.407 | 0.825 | 0.380 |
| JavaScript | 0.101 | 0.302 | 0.115 | 0.318 |
| Python | 0.112 | 0.316 | 0.100 | 0.300 |
| Java | 0.100 | 0.300 | 0.081 | 0.273 |
| C# | 0.057 | 0.232 | 0.061 | 0.239 |
| PHP | 0.066 | 0.249 | 0.060 | 0.237 |
| Android | 0.062 | 0.240 | 0.060 | 0.238 |
| HTML | 0.060 | 0.237 | 0.052 | 0.222 |
| jQuery | 0.035 | 0.184 | 0.038 | 0.191 |
| C++ | 0.037 | 0.189 | 0.031 | 0.173 |
| CSS | 0.035 | 0.183 | 0.034 | 0.182 |
| N | 1,592,162 | | 12,041,307 | |

Notes: "Sample" refers to the summary statistics of the first question posted by sample users between 2015 and 2020, while "All" represents the summary statistics of all questions posted on Stack Overflow during the same period. "Received" is a dummy variable indicating whether a question received any answers. "Acceptance Rate" is a dummy variable indicating whether an answer to a question was accepted, and "Positive Score Answers" is a dummy variable denoting whether any answer to a question has a score greater than zero, both calculated based on questions that received an answer ($Received = 1$). The counts of upvotes and downvotes are aggregated based on votes cast within one week from the question's posting date. Subsequently, the question score is calculated from these aggregated counts—as the difference between upvotes and downvotes. Other descriptive statistics are calculated based on all question data, regardless of whether they received an answer.

Table II: Total Impact of Receiving Answers on Subsequent Questioner Behavior

| | PostedAnswer | AnswerCount | HasPosScore | IsAccepted |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *Received* | −0.0017 | −0.0040 | −0.0003 | −0.0005 |
| | (0.0003) | (0.0011) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [0.079] | [0.002] |
| | < −5.7 > | < −6.5 > | < −3.0 > | < −6.1 > |
| Control mean | 0.0294 | 0.0623 | 0.0100 | 0.0079 |
| Response rate | | 0.66 | | |
| $N$ | | 1,592,162 | | |

Notes: OLS regressions of equation (1). *Received* is a dummy variable indicating whether the sample received answers from others to their first question within 24 hours of posting the question. The voting variable used in the analysis is based on votes cast within one week of the submission date of the post. In addition to the variables presented in the table, year-week fixed effects and question characteristics are included as controls. The list of question characteristics is provided in Online Appendix Table A.I. The coefficients reflect the impact on the subsequent answering behavior of questioners within seven days following their initial question post. Robust standard errors are presented in parentheses. $p$-values are indicated in square brackets; " $< 0.001$" denotes a $p$-value less than 0.001. "Control mean" denotes the average outcome for individuals who did not receive an answer within 24 hours, and the percentage ratios of the coefficients to control mean are indicated in angle brackets. "Response rate" represents the probability of receiving answers to their first question from others within 24 hours of posting the question.

Table III: Heterogeneous Effects of Answer Satisfaction on Subsequent Questioner Behavior

| | PostedAnswer | AnswerCount | HasPosScore | IsAccepted |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| $SatisfiedAns$ | 0.0084 | 0.0187 | 0.0038 | 0.0029 |
| | (0.0004) | (0.0014) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 28.5 > | < 30.1 > | < 38.4 > | < 36.2 > |
| $UnsatisfiedAns$ | −0.0103 | −0.0234 | −0.0038 | −0.0033 |
| | (0.0003) | (0.0011) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −34.9 > | < −37.6 > | < −38.4 > | < −42.1 > |
| Control mean | 0.0294 | 0.0623 | 0.0100 | 0.0079 |
| Response rate | | 0.66 | | |
| $N$ | | 1,592,162 | | |

Notes: OLS regressions of equation (2). $SatisfiedAns$ is a dummy variable indicating that the respondent received answers from others to their first question within 24 hours of posting and selected one of these as the accepted answer. Conversely, $UnsatisfiedAns$ is a dummy variable indicating that the respondent received answers within the same timeframe but did not select any as the accepted answer. The voting variable used in the analysis is based on votes cast within one week of the submission date of the post. In addition to the variables presented in the table, year-week fixed effects and question characteristics are included as controls. The list of question characteristics is provided in Online Appendix Table A.I. The coefficients reflect the impact on the subsequent answering behavior of questioners within seven days following their initial question post. Robust standard errors are presented in parentheses. $p$-values are indicated in square brackets; " $< 0.001$" denotes a $p$-value less than 0.001. "Control mean" denotes the average outcome for individuals who did not receive an answer within 24 hours, and the percentage ratios of the coefficients to control mean are indicated in angle brackets. "Response rate" represents the probability of receiving answers to their first question from others within 24 hours of posting the question.

Table IV: Impact of Responses on Second Questions for Non-responded First Posts

| | PostedAnswer | | | AnswerCount | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $SatisfiedAns$ | 0.0119 | 0.0107 | 0.0105 | 0.0313 | 0.0291 | 0.0294 |
| | (0.0021) | (0.0021) | (0.0021) | (0.0083) | (0.0081) | (0.0081) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 25.0 > | < 22.6 > | < 22.2 > | < 30.6 > | < 28.4 > | < 28.7 > |
| $UnsatisfiedAns$ | −0.0120 | −0.0101 | −0.0101 | −0.0230 | −0.0187 | −0.0182 |
| | (0.0018) | (0.0017) | (0.0017) | (0.0077) | (0.0075) | (0.0075) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [0.003] | [0.013] | [0.015] |
| | < −25.3 > | < −21.4 > | < −21.4 > | < −22.4 > | < −18.2 > | < −17.8 > |
| $Y^1$ | | 0.1981 | 0.1977 | | 0.1907 | 0.1908 |
| | | (0.0077) | (0.0077) | | (0.0231) | (0.0231) |
| | | [< 0.001] | [< 0.001] | | [< 0.001] | [< 0.001] |
| | | < 418.4 > | < 417.5 > | | < 186.0 > | < 186.1 > |
| Control | | | | | | |
| $\mathbf{Q}^0$ | N | N | Y | N | N | Y |
| Control mean | 0.0474 | 0.0474 | 0.0474 | 0.1025 | 0.1025 | 0.1025 |
| Response rate | | | 0.60 | | | |
| $N$ | | | 73,230 | | | |

Table IV: Impact of Responses on Second Questions for Non-responded First Posts (continued)

| | HasPosScore | | | IsAccepted | | |
|---|---|---|---|---|---|---|
| | (7) | (8) | (9) | (10) | (11) | (12) |
| $SatisfiedAns$ | 0.0040 | 0.0034 | 0.0033 | 0.0043 | 0.0040 | 0.0039 |
| | (0.0013) | (0.0013) | (0.0013) | (0.0012) | (0.0012) | (0.0012) |
| | [0.002] | [0.007] | [0.008] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 24.2 > | < 20.4 > | < 20.0 > | < 31.2 > | < 29.2 > | < 28.7 > |
| $UnsatisfiedAns$ | −0.0053 | −0.0048 | −0.0047 | −0.0039 | −0.0035 | −0.0034 |
| | (0.0010) | (0.0010) | (0.0010) | (0.0009) | (0.0009) | (0.0009) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −31.6 > | < −28.7 > | < −28.2 > | < −28.8 > | < −25.5 > | < −24.6 > |
| $Y^0$ | | 0.1552 | 0.1547 | | 0.1474 | 0.1466 |
| | | (0.0110) | (0.0110) | | (0.0122) | (0.0122) |
| | | [< 0.001] | [< 0.001] | | [< 0.001] | [< 0.001] |
| | | < 932.2 > | < 929.1 > | | < 1073.7 > | < 1068.1 > |
| Control | | | | | | |
| $\mathbf{Q}^0$ | N | N | Y | N | N | Y |
| Control mean | 0.0166 | 0.0166 | 0.0166 | 0.0137 | 0.0137 | 0.0137 |
| Response rate | | | 0.60 | | | |
| $N$ | | | 73,230 | | | |

Notes: OLS regressions of equation (3). $SatisfiedAns$ is a dummy variable indicating that the respondent received answers from others to their second question within 24 hours of posting and selected one of these as the accepted answer. Conversely, $UnsatisfiedAns$ is a dummy variable indicating that the respondent received answers within the same timeframe but did not select any as the accepted answer. $Y^1$ denotes the answering behavior of these individuals within seven days of posting their initial question. The voting variable used in the analysis is based on votes cast within one week of the submission date of the post. In addition to the variables presented in the table, year-week fixed effects and question characteristics are included as controls. The list of question characteristics is provided in Online Appendix Table A.I. The coefficients reflect the impact on the subsequent answering behavior of questioners within seven days following their second question post. The sample of questioners in Table III is limited to individuals who did not receive any answers to their initial question within 24 hours and who posted a second question 7 days to 3 months after their first post. In Columns (3), (6), (9), and (12), the models incorporate the characteristics of the first posted questions as control variables. Robust standard errors are presented in parentheses. $p$-values are indicated in square brackets; " < 0.001" denotes a $p$-value less than 0.001. "Control mean" denotes the average outcome for individuals who did not receive an answer within 24 hours, and the percentage ratios of the coefficients to control mean are indicated in angle brackets. "Response rate" represents the probability of receiving answers to their second question from others within 24 hours of posting the question.

Table V: Long-Term Impact Analysis: Posting Behavior within 365 Days after Initial Question

| | PostedAnswer | AnswerCount | HasPosScore | IsAccepted | Engaged |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| $SatisfiedAns$ | 0.0293 | 0.2392 | 0.0163 | 0.0143 | 0.1993 |
| | (0.0006) | (0.0196) | (0.0004) | (0.0004) | (0.0010) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 29.2 > | < 49.0 > | < 41.4 > | < 44.9 > | < 55.4 > |
| $UnsatisfiedAns$ | −0.0309 | −0.1834 | −0.0142 | −0.0123 | 0.0244 |
| | (0.0005) | (0.0102) | (0.0003) | (0.0003) | (0.0009) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −30.8 > | < −37.6 > | < −36.0 > | < −38.5 > | < 6.8 > |
| Control mean | 0.1003 | 0.4884 | 0.0393 | 0.0318 | 0.3596 |
| Response rate | | | 0.66 | | |
| $N$ | | | 1,592,162 | | |

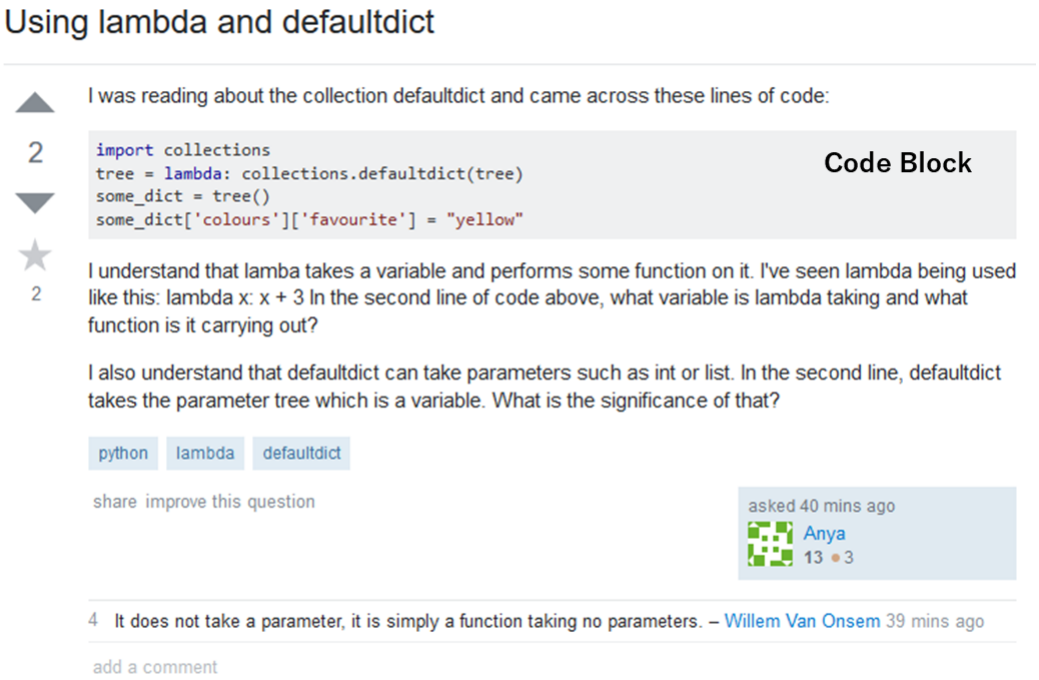
Notes: OLS regressions of equation (2). $SatisfiedAns$ is a dummy variable indicating that the respondent received answers from others to their first question within 24 hours of posting and selected one of these as the accepted answer. Conversely, $UnsatisfiedAns$ is a dummy variable indicating that the respondent received answers within the same timeframe but did not select any as the accepted answer. The voting variable used in the analysis is based on votes cast within one week of the submission date of the post. In addition to the variables presented in the table, year-week fixed effects and question characteristics are included as controls. The list of question characteristics is provided in Online Appendix Table A.I. The coefficients reflect the effect on the subsequent answering and asking behavior of questioners within 365 days following their initial question post. Robust standard errors are presented in parentheses. $p$-values are indicated in square brackets; " $< 0.001$" denotes a $p$-value less than 0.001. "Control mean" denotes the average outcome for individuals who did not receive an answer within 24 hours, and the percentage ratios of the coefficients to control mean are indicated in angle brackets. "Response rate" represents the probability of receiving answers to their first question from others within 24 hours of posting the question.

# Online Appendix

Q&A: Upstream Indirect Reciprocity in a Large Online Community

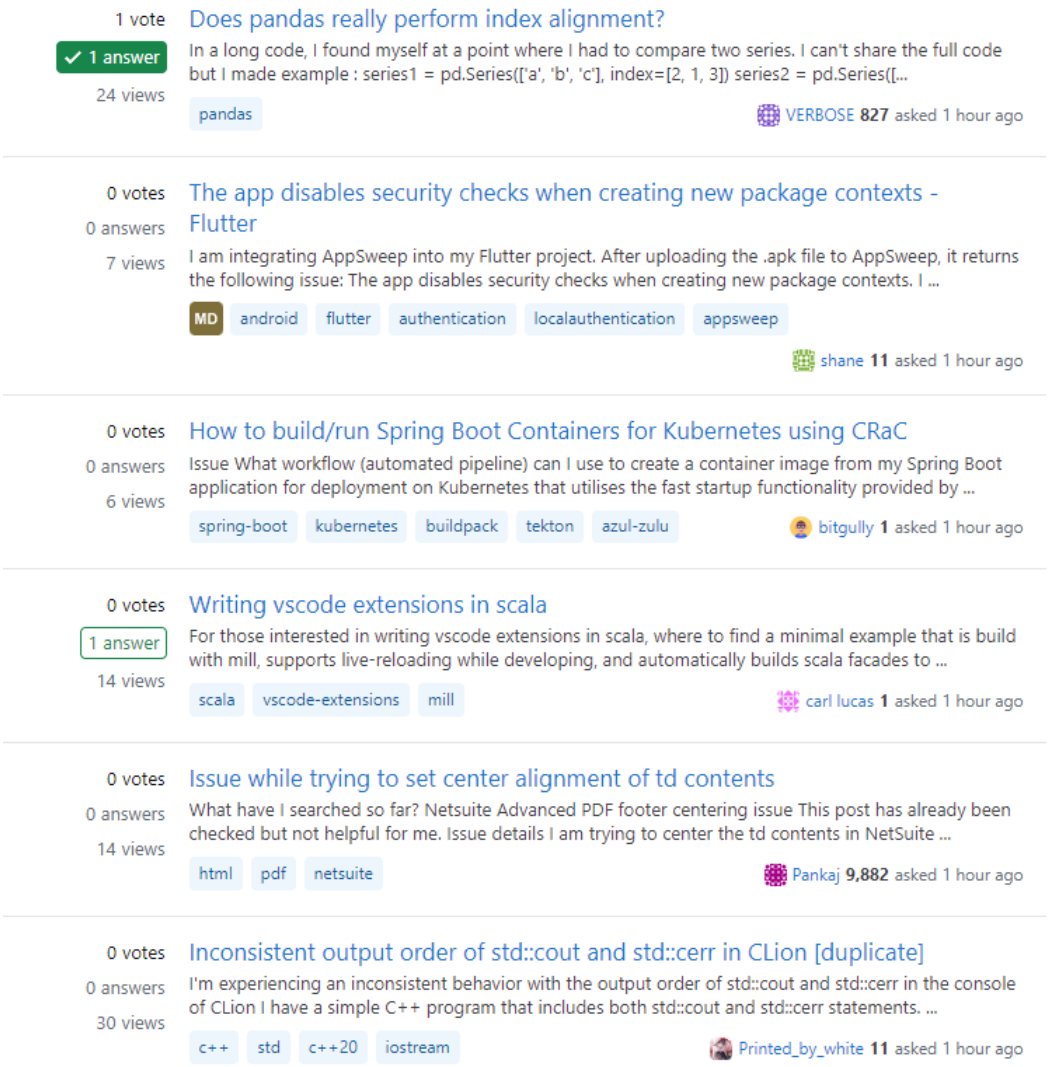
Takahiro Miura

# A Supplementary Tables and Figures



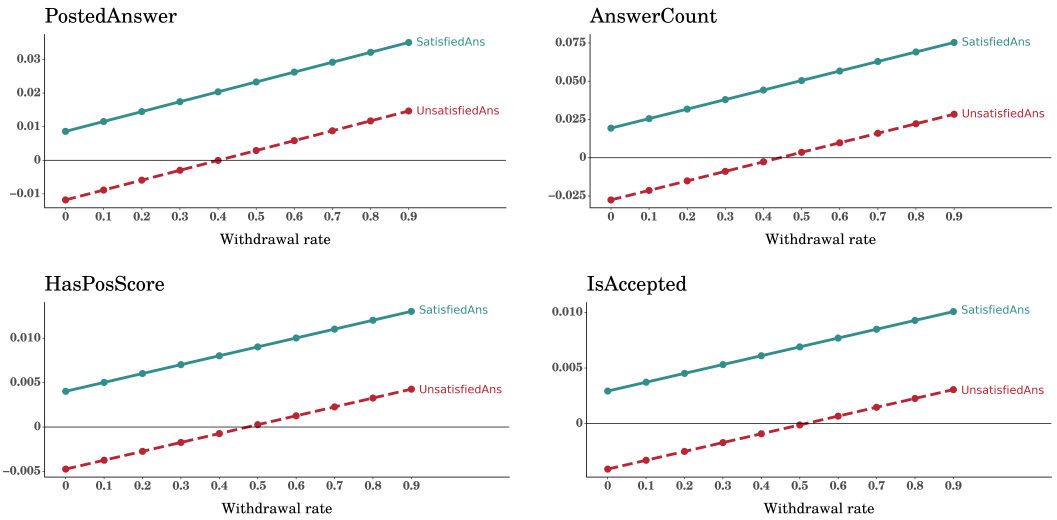

Notes: This figure represents an example of a Stack Overflow question page. Code blocks are highlighted with a gray background (the text "Code Block" was inserted by the author). The number enclosed between the left arrow icons indicates the current score of the question (the number of Upvotes minus Downvotes). The section surrounded by a blue background in the bottom right shows the user information of the questioner.

Figure A.I: Question Image




Notes: This figure exemplifies a question list page on Stack Overflow. Questions are separated by lines, with the number of answers indicated on the left. The answers count surrounded by green highlights indicates the presence of answers to the question. Additionally, questions marked with a check signify that one of the answers has been accepted as the accepted answer. Note that this image represents the page as of January 23, 2024, and the presentation may differ from earlier dates.

Figure A.II: Question List




Notes: Each panel illustrates how the estimated results change when the withdrawal rate of questioners who did not receive answers is $p$. The values presented in the figure correspond to the coefficients obtained from the OLS regression excluding control variables from equation 2. The values for $p = 0$ show the result using the original sample analysis, while the values for $p > 0$ depict the results of adding pseudo-data for the missing samples.

Figure A.III: Impact of Withdrawal Rates on Estimated Values

Table A.I: Full Result

| | PostedAnswer (1) | AnswerCount (2) | HasPosScore (3) | IsAccepted (4) |
|---|---|---|---|---|
| $SatisfiedAns$ | 0.0084 | 0.0187 | 0.0038 | 0.0029 |
| | (0.0004) | (0.0014) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 28.5 > | < 30.1 > | < 38.4 > | < 36.2 > |
| $UnsatisfiedAns$ | −0.0103 | −0.0234 | −0.0038 | −0.0033 |
| | (0.0003) | (0.0011) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −34.9 > | < −37.6 > | < −38.4 > | < −42.1 > |
| ManyWords | 0.0085 | 0.0256 | 0.0044 | 0.0039 |
| | (0.0003) | (0.0011) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 29.0 > | < 41.1 > | < 44.3 > | < 49.3 > |
| CodeBlocked | 0.0009 | 0.0031 | 0.0003 | 0.0006 |
| | (0.0003) | (0.0011) | (0.0002) | (0.0002) |
| | [0.006] | [0.006] | [0.063] | [< 0.001] |
| | < 3.0 > | < 4.9 > | < 3.5 > | < 7.6 > |
| UpVoted | 0.0099 | 0.0249 | 0.0060 | 0.0040 |
| | (0.0003) | (0.0011) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 33.6 > | < 40.0 > | < 59.6 > | < 50.8 > |
| DownVoted | −0.0044 | −0.0106 | −0.0024 | −0.0019 |
| | (0.0003) | (0.0011) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −15.0 > | < −17.0 > | < −24.4 > | < −23.9 > |
| JavaScript | 0.0055 | 0.0158 | 0.0019 | 0.0014 |
| | (0.0005) | (0.0019) | (0.0003) | (0.0003) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 18.6 > | < 25.4 > | < 19.0 > | < 18.0 > |
| Python | −0.0003 | −0.0013 | −0.0005 | −0.0004 |
| | (0.0004) | (0.0015) | (0.0002) | (0.0002) |
| | [0.473] | [0.390] | [0.039] | [0.071] |
| | < −1.0 > | < −2.0 > | < −5.0 > | < −4.9 > |
| Java | 0.0001 | −0.0023 | −0.0006 | −0.0006 |
| | (0.0005) | (0.0016) | (0.0003) | (0.0002) |
| | [0.837] | [0.138] | [0.034] | [0.005] |
| | < 0.3 > | < −3.7 > | < −5.6 > | < −8.1 > |
| C# | 0.0039 | 0.0077 | 0.0012 | 0.0011 |
| | (0.0006) | (0.0021) | (0.0004) | (0.0003) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 13.2 > | < 12.3 > | < 12.1 > | < 13.9 > |

Table A.I: Full Result (continued)

| | PostedAnswer (1) | AnswerCount (2) | HasPosScore (3) | IsAccepted (4) |
|---|---|---|---|---|
| PHP | 0.0068 | 0.0169 | 0.0018 | 0.0013 |
| | (0.0006) | (0.0021) | (0.0003) | (0.0003) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 23.1 > | < 27.2 > | < 17.6 > | < 16.7 > |
| Android | 0.0117 | 0.0254 | 0.0028 | 0.0023 |
| | (0.0006) | (0.0023) | (0.0004) | (0.0003) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 39.8 > | < 40.8 > | < 27.9 > | < 29.1 > |
| HTML | −0.0023 | −0.0089 | −0.0015 | −0.0012 |
| | (0.0007) | (0.0022) | (0.0004) | (0.0003) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −8.0 > | < −14.3 > | < −15.1 > | < −14.9 > |
| jQuery | 0.0009 | −0.0002 | −0.0001 | −0.0001 |
| | (0.0008) | (0.0028) | (0.0005) | (0.0004) |
| | [0.288] | [0.939] | [0.790] | [0.747] |
| | < 3.0 > | < −0.3 > | < −1.3 > | < −1.7 > |
| C++ | 0.0023 | 0.0072 | 0.0029 | 0.0017 |
| | (0.0007) | (0.0026) | (0.0005) | (0.0004) |
| | [0.001] | [0.006] | [< 0.001] | [< 0.001] |
| | < 7.9 > | < 11.6 > | < 29.1 > | < 20.8 > |
| CSS | 0.0024 | 0.0123 | 0.0019 | 0.0012 |
| | (0.0009) | (0.0033) | (0.0005) | (0.0005) |
| | [0.006] | [< 0.001] | [< 0.001] | [0.008] |
| | < 8.1 > | < 19.8 > | < 18.6 > | < 15.2 > |
| Control mean | 0.0294 | 0.0623 | 0.0100 | 0.0079 |
| Response rate | | 0.66 | | |
| $N$ | | 1,592,162 | | |

Notes: OLS regressions of equation (2). $SatisfiedAns$ is a dummy variable indicating that the respondent received answers from others to their first question within 24 hours of posting and selected one of these as the accepted answer. Conversely, $UnsatisfiedAns$ is a dummy variable indicating that the respondent received answers within the same timeframe but did not select any as the accepted answer. The voting variable used in the analysis is based on votes cast within one week of the submission date of the post. In addition to the variables presented in the table, year-week fixed effects are included as controls. The coefficients indicate the impact on and correlation with the answering behavior of questioners within seven days following their initial question post. Robust standard errors are presented in parentheses. $p$-values are indicated in square brackets; " < 0.001" denotes a $p$-value less than 0.001. "Control mean" denotes the average outcome for individuals who did not receive an answer within 24 hours, and the percentage ratios of the coefficients to control mean are indicated in angle brackets. "Response rate" represents the probability of receiving answers to their first question from others within 24 hours of posting the question.

Table A.II: Subsample Analysis of Heterogeneous Effects by Question Tag

| | PostedAnswer | AnswerCount | HasPosScore | IsAccepted |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **Panel A: Popular Tag** | | | | |
| $SatisfiedAns$ | 0.0080 | 0.0162 | 0.0034 | 0.0023 |
| | (0.0005) | (0.0019) | (0.0003) | (0.0003) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 24.5 > | < 22.7 > | < 30.4 > | < 26.1 > |
| $UnsatisfiedAns$ | −0.0110 | −0.0264 | −0.0042 | −0.0036 |
| | (0.0004) | (0.0016) | (0.0003) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −33.4 > | < −37.0 > | < −37.3 > | < −40.4 > |
| Control mean | 0.0328 | 0.0714 | 0.0113 | 0.0090 |
| Response rate | | 0.71 | | |
| N | | 858,078 | | |
| **Panel B: Python** | | | | |
| $SatisfiedAns$ | 0.0046 | 0.0092 | 0.0025 | 0.0012 |
| | (0.0010) | (0.0040) | (0.0006) | (0.0006) |
| | [< 0.001] | [0.022] | [< 0.001] | [0.039] |
| | < 16.1 > | < 15.3 > | < 25.7 > | < 14.2 > |
| $UnsatisfiedAns$ | −0.0123 | −0.0277 | −0.0043 | −0.0042 |
| | (0.0009) | (0.0033) | (0.0005) | (0.0005) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −42.5 > | < −45.9 > | < −44.5 > | < −51.7 > |
| Control mean | 0.0288 | 0.0604 | 0.0097 | 0.0081 |
| Response rate | | 0.72 | | |
| N | | 179,089 | | |
| **Panel C: JavaScript** | | | | |
| $SatisfiedAns$ | 0.0088 | 0.0174 | 0.0030 | 0.0018 |
| | (0.0012) | (0.0046) | (0.0007) | (0.0007) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [0.007] |
| | < 25.9 > | < 22.2 > | < 24.3 > | < 17.8 > |
| $UnsatisfiedAns$ | −0.0113 | −0.0308 | −0.0054 | −0.0046 |
| | (0.0011) | (0.0040) | (0.0006) | (0.0006) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −33.3 > | < −39.4 > | < −43.3 > | < −46.2 > |
| Control mean | 0.0341 | 0.0783 | 0.0124 | 0.0100 |
| Response rate | | 0.72 | | |
| N | | 161,547 | | |

Table A.II: Subsample Analysis of Heterogeneous Effects by Question Tag (continued)

| | PostedAnswer (1) | AnswerCount (2) | HasPosScore (3) | IsAccepted (4) |
|---|---|---|---|---|
| **Panel D: Java** | | | | |
| $SatisfiedAns$ | 0.0077 | 0.0173 | 0.0036 | 0.0027 |
| | (0.0012) | (0.0040) | (0.0007) | (0.0006) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 24.6 > | < 27.7 > | < 36.0 > | < 36.2 > |
| $UnsatisfiedAns$ | −0.0110 | −0.0215 | −0.0032 | −0.0023 |
| | (0.0010) | (0.0035) | (0.0006) | (0.0005) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −35.3 > | < −34.3 > | < −32.8 > | < −30.5 > |
| Control mean | 0.0313 | 0.0625 | 0.0099 | 0.0075 |
| Response rate | | 0.70 | | |
| N | | 159,425 | | |
| **Panel E: PHP** | | | | |
| $SatisfiedAns$ | 0.0132 | 0.0263 | 0.0051 | 0.0029 |
| | (0.0016) | (0.0057) | (0.0009) | (0.0008) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 39.2 > | < 33.8 > | < 44.2 > | < 30.5 > |
| $UnsatisfiedAns$ | −0.0082 | −0.0282 | −0.0045 | −0.0034 |
| | (0.0013) | (0.0046) | (0.0007) | (0.0007) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −24.2 > | < −36.2 > | < −38.8 > | < −36.5 > |
| Control mean | 0.0338 | 0.0779 | 0.0116 | 0.0094 |
| Response rate | | 0.70 | | |
| N | | 105,876 | | |
| **Panel F: Android** | | | | |
| $SatisfiedAns$ | 0.0123 | 0.0244 | 0.0050 | 0.0046 |
| | (0.0018) | (0.0064) | (0.0010) | (0.0009) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 31.0 > | < 28.6 > | < 41.2 > | < 48.5 > |
| $UnsatisfiedAns$ | −0.0109 | −0.0241 | −0.0034 | −0.0028 |
| | (0.0014) | (0.0054) | (0.0008) | (0.0007) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −27.6 > | < −28.3 > | < −28.2 > | < −29.7 > |
| Control mean | 0.0396 | 0.0852 | 0.0121 | 0.0095 |
| Response rate | | 0.63 | | |
| N | | 98,067 | | |

Table A.II: Subsample Analysis of Heterogeneous Effects by Question Tag (continued)

| | PostedAnswer (1) | AnswerCount (2) | HasPosScore (3) | IsAccepted (4) |
|---|---|---|---|---|
| **Panel G: HTML** | | | | |
| $SatisfiedAns$ | 0.0081 | 0.0173 | 0.0023 | 0.0020 |
| | (0.0016) | (0.0053) | (0.0010) | (0.0008) |
| | [< 0.001] | [0.001] | [0.015] | [0.014] |
| | < 26.9 > | < 27.1 > | < 22.1 > | < 25.3 > |
| $UnsatisfiedAns$ | −0.0106 | −0.0209 | −0.0039 | −0.0036 |
| | (0.0014) | (0.0053) | (0.0008) | (0.0007) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −35.2 > | < −32.8 > | < −37.4 > | < −45.0 > |
| Control mean | 0.0301 | 0.0638 | 0.0105 | 0.0081 |
| Response rate | | 0.79 | | |
| N | | 95,257 | | |
| **Panel H: C#** | | | | |
| $SatisfiedAns$ | 0.0062 | 0.0114 | 0.0039 | 0.0033 |
| | (0.0016) | (0.0057) | (0.0010) | (0.0009) |
| | [< 0.001] | [0.043] | [< 0.001] | [< 0.001] |
| | < 18.3 > | < 15.6 > | < 34.3 > | < 37.2 > |
| $UnsatisfiedAns$ | −0.0102 | −0.0244 | −0.0032 | −0.0024 |
| | (0.0014) | (0.0049) | (0.0008) | (0.0007) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −30.2 > | < −33.3 > | < −27.9 > | < −27.4 > |
| Control mean | 0.0338 | 0.0731 | 0.0114 | 0.0090 |
| Response rate | | 0.66 | | |
| N | | 90,571 | | |
| **Panel I: C++** | | | | |
| $SatisfiedAns$ | 0.0064 | 0.0160 | 0.0026 | 0.0012 |
| | (0.0021) | (0.0081) | (0.0014) | (0.0012) |
| | [0.002] | [0.048] | [0.067] | [0.318] |
| | < 18.5 > | < 20.4 > | < 16.2 > | < 10.1 > |
| $UnsatisfiedAns$ | −0.0140 | −0.0329 | −0.0067 | −0.0069 |
| | (0.0018) | (0.0066) | (0.0012) | (0.0010) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −40.0 > | < −42.0 > | < −42.3 > | < −56.0 > |
| Control mean | 0.0349 | 0.0783 | 0.0158 | 0.0123 |
| Response rate | | 0.74 | | |
| N | | 58,933 | | |

Table A.II: Subsample Analysis of Heterogeneous Effects by Question Tag (continued)

| | PostedAnswer (1) | AnswerCount (2) | HasPosScore (3) | IsAccepted (4) |
|---|---|---|---|---|
| **Panel J: jQuery** | | | | |
| $SatisfiedAns$ | 0.0092 | 0.0171 | 0.0043 | 0.0027 |
| | (0.0021) | (0.0071) | (0.0012) | (0.0011) |
| | [< 0.001] | [0.017] | [< 0.001] | [0.013] |
| | < 27.6 > | < 23.7 > | < 40.5 > | < 31.2 > |
| $UnsatisfiedAns$ | −0.0095 | −0.0201 | −0.0029 | −0.0029 |
| | (0.0019) | (0.0070) | (0.0011) | (0.0010) |
| | [< 0.001] | [0.004] | [0.006] | [0.003] |
| | < −28.5 > | < −27.8 > | < −27.5 > | < −33.6 > |
| Control mean | 0.0333 | 0.0722 | 0.0106 | 0.0086 |
| Response rate | | 0.74 | | |
| N | | 55,652 | | |
| **Panel K: CSS** | | | | |
| $SatisfiedAns$ | 0.0067 | 0.0070 | 0.0019 | 0.0020 |
| | (0.0022) | (0.0090) | (0.0013) | (0.0011) |
| | [0.002] | [0.436] | [0.143] | [0.085] |
| | < 21.2 > | < 9.2 > | < 16.2 > | < 22.4 > |
| $UnsatisfiedAns$ | −0.0113 | −0.0268 | −0.0046 | −0.0034 |
| | (0.0020) | (0.0089) | (0.0012) | (0.0010) |
| | [< 0.001] | [0.003] | [< 0.001] | [< 0.001] |
| | < −36.1 > | < −35.4 > | < −38.4 > | < −38.7 > |
| Control mean | 0.0314 | 0.0758 | 0.0120 | 0.0088 |
| Response rate | | 0.81 | | |
| N | | 55,194 | | |

Notes: OLS regressions of equation (2). $SatisfiedAns$ is a dummy variable indicating that the respondent received answers from others to their first question within 24 hours of posting and selected one of these as the accepted answer. Conversely, $UnsatisfiedAns$ is a dummy variable indicating that the respondent received answers within the same timeframe but did not select any as the accepted answer. The voting variable used in the analysis is based on votes cast within one week of the submission date of the post. In addition to the variables presented in the table, year-week fixed effects and question characteristics are included as controls. The list of question characteristics is provided in Online Appendix Table A.I. The coefficients reflect the impact on the subsequent answering behavior of questioners within seven days following their initial question post. The sample of questioners in Table III is limited to individuals whose first question's tag is among the 10 most popular question tags among sampled questioners. Panel A presents the results from an analysis of this entire subsample and Panel B that of the subsets of questioners who tagged their first question with Python, JavaScript, Java, PHP, Android, HTML, C#, C++, jQuery, or CSS. Robust standard errors are presented in parentheses. $p$-values are indicated in square brackets; " $< 0.001$" denotes a $p$-value less than 0.001. "Control mean" denotes the average outcome for individuals who did not receive an answer within 24 hours, and the percentage ratios of the coefficients to control mean are indicated in angle brackets. "Response rate" represents the probability of receiving answers to their first question from others within 24 hours of posting the question.

Table A.III: Subsample Analysis of Heterogeneous Effects by Registration Year

| | PostedAnswer | AnswerCount | HasPosScore | IsAccepted |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **Panel A: Registration Year = 2015** | | | | |
| $SatisfiedAns$ | 0.0108 | 0.0262 | 0.0046 | 0.0035 |
| | (0.0009) | (0.0034) | (0.0006) | (0.0005) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 28.5 > | < 31.5 > | < 33.2 > | < 32.5 > |
| $UnsatisfiedAns$ | −0.0124 | −0.0289 | −0.0056 | −0.0045 |
| | (0.0007) | (0.0029) | (0.0004) | (0.0004) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −32.9 > | < −34.7 > | < −40.3 > | < −41.8 > |
| Control mean | 0.0378 | 0.0833 | 0.0140 | 0.0107 |
| Response rate | | 0.69 | | |
| N | | 340,192 | | |
| **Panel B: Registration Year = 2016** | | | | |
| $SatisfiedAns$ | 0.0097 | 0.0201 | 0.0045 | 0.0035 |
| | (0.0009) | (0.0032) | (0.0005) | (0.0005) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 28.9 > | < 27.4 > | < 39.3 > | < 37.4 > |
| $UnsatisfiedAns$ | −0.0106 | −0.0255 | −0.0033 | −0.0033 |
| | (0.0007) | (0.0025) | (0.0004) | (0.0004) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −31.6 > | < −34.7 > | < −29.0 > | < −35.6 > |
| Control mean | 0.0336 | 0.0734 | 0.0115 | 0.0093 |
| Response rate | | 0.67 | | |
| N | | 336,187 | | |
| **Panel C: Registration Year = 2017** | | | | |
| $SatisfiedAns$ | 0.0080 | 0.0176 | 0.0034 | 0.0029 |
| | (0.0008) | (0.0029) | (0.0005) | (0.0004) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 28.4 > | < 30.2 > | < 38.2 > | < 39.7 > |
| $UnsatisfiedAns$ | −0.0103 | −0.0222 | −0.0037 | −0.0030 |
| | (0.0006) | (0.0024) | (0.0004) | (0.0003) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −36.7 > | < −38.3 > | < −41.2 > | < −42.3 > |
| Control mean | 0.0281 | 0.0581 | 0.0090 | 0.0072 |
| Response rate | | 0.66 | | |
| N | | 329,394 | | |

Table A.III: Subsample Analysis of Heterogeneous Effects by Registration Year (continued)

| | PostedAnswer (1) | AnswerCount (2) | HasPosScore (3) | IsAccepted (4) |
|---|---|---|---|---|
| **Panel D: Registration Year = 2018** | | | | |
| $SatisfiedAns$ | 0.0063 | 0.0160 | 0.0033 | 0.0027 |
| | (0.0008) | (0.0028) | (0.0005) | (0.0004) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 25.2 > | < 32.0 > | < 41.7 > | < 43.1 > |
| $UnsatisfiedAns$ | −0.0087 | −0.0191 | −0.0031 | −0.0027 |
| | (0.0006) | (0.0020) | (0.0004) | (0.0003) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −35.2 > | < −38.1 > | < −38.8 > | < −43.3 > |
| Control mean | 0.0248 | 0.0500 | 0.0080 | 0.0063 |
| Response rate | | 0.65 | | |
| N | | 294,227 | | |
| **Panel E: Registration Year = 2019** | | | | |
| $SatisfiedAns$ | 0.0062 | 0.0115 | 0.0029 | 0.0015 |
| | (0.0008) | (0.0027) | (0.0005) | (0.0004) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 27.7 > | < 25.2 > | < 38.7 > | < 24.7 > |
| $UnsatisfiedAns$ | −0.0082 | −0.0193 | −0.0031 | −0.0029 |
| | (0.0006) | (0.0020) | (0.0003) | (0.0003) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −36.8 > | < −42.1 > | < −40.9 > | < −47.7 > |
| Control mean | 0.0223 | 0.0458 | 0.0075 | 0.0061 |
| Response rate | | 0.64 | | |
| N | | 292,162 | | |

Notes: OLS regressions of equation (2). $SatisfiedAns$ is a dummy variable indicating that the respondent received answers from others to their first question within 24 hours of posting and selected one of these as the accepted answer. Conversely, $UnsatisfiedAns$ is a dummy variable indicating that the respondent received answers within the same timeframe but did not select any as the accepted answer. The voting variable used in the analysis is based on votes cast within one week of the submission date of the post. In addition to the variables presented in the table, year-week fixed effects and question characteristics are included as controls. The list of question characteristics is provided in Online Appendix Table A.I. The coefficients reflect the impact on the subsequent answering behavior of questioners within seven days following their initial question post. Each panel presents the results of a subsample analysis for the sample of questioners in Table III, segmented by their registration year from 2015 (Panel A) to 2019 (Panel E). Robust standard errors are presented in parentheses. $p$-values are indicated in square brackets; " $< 0.001$" denotes a $p$-value less than 0.001. "Control mean" denotes the average outcome for individuals who did not receive an answer within 24 hours, and the percentage ratios of the coefficients to control mean are indicated in angle brackets. "Response rate" represents the probability of receiving answers to their first question from others within 24 hours of posting the question.

Table A.IV: Subsample Analysis of Heterogeneous Effects by Question Length

| | PostedAnswer | AnswerCount | HasPosScore | IsAccepted |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **Panel A: Number of words > Mean(Number of words)** | | | | |
| $SatisfiedAns$ | 0.0050 | 0.0120 | 0.0027 | 0.0017 |
| | (0.0006) | (0.0026) | (0.0004) | (0.0004) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 13.8 > | < 14.7 > | < 20.0 > | < 15.0 > |
| $UnsatisfiedAns$ | −0.0135 | −0.0314 | −0.0054 | −0.0051 |
| | (0.0005) | (0.0022) | (0.0003) | (0.0003) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −37.1 > | < −38.3 > | < −39.9 > | < −45.2 > |
| Control mean | 0.0364 | 0.0818 | 0.0136 | 0.0112 |
| Response rate | | 0.60 | | |
| N | | 572,232 | | |
| **Panel A: Number of words ≤ Mean(Number of words)** | | | | |
| $SatisfiedAns$ | 0.0107 | 0.0237 | 0.0047 | 0.0038 |
| | (0.0004) | (0.0015) | (0.0003) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 44.1 > | < 49.5 > | < 63.3 > | < 68.2 > |
| $UnsatisfiedAns$ | −0.0082 | −0.0183 | −0.0029 | −0.0023 |
| | (0.0004) | (0.0011) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −33.6 > | < −38.4 > | < −38.7 > | < −41.2 > |
| Control mean | 0.0243 | 0.0478 | 0.0074 | 0.0055 |
| Response rate | | 0.70 | | |
| N | | 1,019,930 | | |

Notes: OLS regressions of equation (2). $SatisfiedAns$ is a dummy variable indicating that the respondent received answers from others to their first question within 24 hours of posting and selected one of these as the accepted answer. Conversely, $UnsatisfiedAns$ is a dummy variable indicating that the respondent received answers within the same timeframe but did not select any as the accepted answer. The voting variable used in the analysis is based on votes cast within one week of the submission date of the post. In addition to the variables presented in the table, year-week fixed effects and question characteristics are included as controls. The list of question characteristics is provided in Online Appendix Table A.I. The coefficients reflect the impact on the subsequent answering behavior of questioners within seven days following their initial question post. Each panel presents the results of a subsample analysis for the sample of questioners in Table III, segmented by the number of words in their first question. Panel A analyzes samples with a word count above the average of all sampled questions, while Panel B analyzes those with a word count below or equal to the average. Robust standard errors are presented in parentheses. $p$-values are indicated in square brackets; " < 0.001" denotes a $p$-value less than 0.001. "Control mean" denotes the average outcome for individuals who did not receive an answer within 24 hours, and the percentage ratios of the coefficients to control mean are indicated in angle brackets. "Response rate" represents the probability of receiving answers to their first question from others within 24 hours of posting the question.

Table A.V: Subsample Analysis of Heterogeneous Effects by Code Block Usage

| | PostedAnswer (1) | AnswerCount (2) | HasPosScore (3) | IsAccepted (4) |
|---|---|---|---|---|
| **Panel A: Using Code Block Feature** | | | | |
| $SatisfiedAns$ | 0.0071 | 0.0158 | 0.0034 | 0.0023 |
| | (0.0004) | (0.0015) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 22.9 > | < 23.9 > | < 31.6 > | < 27.4 > |
| $UnsatisfiedAns$ | −0.0109 | −0.0257 | −0.0042 | −0.0036 |
| | (0.0004) | (0.0013) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −35.4 > | < −38.9 > | < −39.3 > | < −42.5 > |
| Control mean | 0.0308 | 0.0660 | 0.0107 | 0.0086 |
| Response rate | | 0.68 | | |
| N | | 1,257,907 | | |
| **Panel B: Not Using Code Block Feature** | | | | |
| $SatisfiedAns$ | 0.0143 | 0.0311 | 0.0058 | 0.0052 |
| | (0.0009) | (0.0030) | (0.0005) | (0.0005) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 56.5 > | < 60.4 > | < 72.0 > | < 84.7 > |
| $UnsatisfiedAns$ | −0.0084 | −0.0165 | −0.0028 | −0.0025 |
| | (0.0006) | (0.0020) | (0.0003) | (0.0003) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −33.0 > | < −32.1 > | < −34.6 > | < −40.0 > |
| Control mean | 0.0253 | 0.0515 | 0.0081 | 0.0061 |
| Response rate | | 0.59 | | |
| N | | 334,255 | | |

Notes: OLS regressions of equation (2). $SatisfiedAns$ is a dummy variable indicating that the respondent received answers from others to their first question within 24 hours of posting and selected one of these as the accepted answer. Conversely, $UnsatisfiedAns$ is a dummy variable indicating that the respondent received answers within the same timeframe but did not select any as the accepted answer. The voting variable used in the analysis is based on votes cast within one week of the submission date of the post. In addition to the variables presented in the table, year-week fixed effects and question characteristics are included as controls. The list of question characteristics is provided in Online Appendix Table A.I. The coefficients reflect the impact on the subsequent answering behavior of questioners within seven days following their initial question post. Each panel presents the results of a subsample analysis for the sample of questioners in Table III, segmented based on whether the code block feature was used in the respondents' first question. Panel A analyzes samples that used the code block feature, whereas Panel B analyzes those that did not use the code block feature. Robust standard errors are presented in parentheses. $p$-values are indicated in square brackets; " $< 0.001$" denotes a $p$-value less than 0.001. "Control mean" denotes the average outcome for individuals who did not receive an answer within 24 hours, and the percentage ratios of the coefficients to control mean are indicated in angle brackets. "Response rate" represents the probability of receiving answers to their first question from others within 24 hours of posting the question.

Table A.VI: Subsample Analysis of Heterogeneous Effects by Question Score

| | PostedAnswer | AnswerCount | HasPosScore | IsAccepted |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **Panel A: Question Score $\geq$ 0** | | | | |
| $SatisfiedAns$ | 0.0093 | 0.0207 | 0.0043 | 0.0032 |
| | (0.0004) | (0.0015) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 31.7 > | < 33.1 > | < 42.0 > | < 40.2 > |
| $UnsatisfiedAns$ | −0.0100 | −0.0233 | −0.0040 | −0.0035 |
| | (0.0003) | (0.0012) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −33.9 > | < −37.3 > | < −39.3 > | < −43.1 > |
| Control mean | 0.0295 | 0.0626 | 0.0102 | 0.0080 |
| Response rate | | 0.62 | | |
| N | | 1,344,464 | | |
| **Panel B: Question Score > 0** | | | | |
| $SatisfiedAns$ | 0.0099 | 0.0225 | 0.0048 | 0.0030 |
| | (0.0007) | (0.0028) | (0.0005) | (0.0004) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 25.9 > | < 26.4 > | < 30.7 > | < 24.8 > |
| $UnsatisfiedAns$ | −0.0123 | −0.0288 | −0.0058 | −0.0052 |
| | (0.0007) | (0.0026) | (0.0004) | (0.0004) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −32.1 > | < −33.7 > | < −37.7 > | < −43.0 > |
| Control mean | 0.0383 | 0.0854 | 0.0155 | 0.0121 |
| Response rate | | 0.66 | | |
| N | | 459,894 | | |

Notes: OLS regressions of equation (2). $SatisfiedAns$ is a dummy variable indicating that the respondent received answers from others to their first question within 24 hours of posting and selected one of these as the accepted answer. Conversely, $UnsatisfiedAns$ is a dummy variable indicating that the respondent received answers within the same timeframe but did not select any as the accepted answer. The voting variable used in the analysis is based on votes cast within one week of the submission date of the post. In addition to the variables presented in the table, year-week fixed effects and question characteristics are included as controls. The list of question characteristics is provided in Online Appendix Table A.I. The coefficients reflect the impact on the subsequent answering behavior of questioners within seven days following their initial question post. The sample of questioners in Table III is limited to individuals whose first question's score is nonnegative. Panel A presents the results from an analysis of this entire subsample, while Panel B further restricts the sample to those posting their first question with a positive score. Robust standard errors are presented in parentheses. $p$-values are indicated in square brackets; " $< 0.001$" denotes a $p$-value less than 0.001. "Control mean" denotes the average outcome for individuals who did not receive an answer within 24 hours, and the percentage ratios of the coefficients to control mean are indicated in angle brackets. "Response rate" represents the probability of receiving answers to their first question from others within 24 hours of posting the question.

Table A.VII: Analysis of Recent First Questions: Newly Collected Sample from 2022/03/01 to 2022/09/01

| | PostedAnswer | AnswerCount | HasPosScore | IsAccepted |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| $SatisfiedAns$ | 0.0086 | 0.0216 | 0.0042 | 0.0034 |
| | (0.0011) | (0.0033) | (0.0006) | (0.0005) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 52.5 > | < 74.7 > | < 86.3 > | < 94.8 > |
| $UnsatisfiedAns$ | −0.0045 | −0.0093 | −0.0015 | −0.0011 |
| | (0.0007) | (0.0021) | (0.0004) | (0.0003) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [0.001] |
| | < −27.5 > | < −32.2 > | < −30.1 > | < −30.3 > |
| Control mean | 0.0163 | 0.0290 | 0.0049 | 0.0035 |
| Response rate | | 0.46 | | |
| N | | 168,101 | | |

Notes: OLS regressions of equation (2). $SatisfiedAns$ is a dummy variable indicating that the respondent received answers from others to their first question within 24 hours of posting and selected one of these as the accepted answer. Conversely, $UnsatisfiedAns$ is a dummy variable indicating that the respondent received answers within the same timeframe but did not select any as the accepted answer. The voting variable used in the analysis is based on votes cast within one week of the submission date of the post. In addition to the variables presented in the table, year-week fixed effects and question characteristics are included as controls. The list of question characteristics is provided in Online Appendix Table A.I. The coefficients reflect the impact on the subsequent answering behavior of questioners within seven days following their initial question post. The sample in this table has been recollected from questioners who posted their first question between 2022/03/01 and 2022/09/01. Similar sample restriction rules as mentioned in Section **??** are applied, including those who posted their first question within a year of registration and before posting their first answer. Furthermore, the sample is limited to those whose question scores did not fall below 0 from the time of the question post to the last date in the data (2022/09/25). Robust standard errors are presented in parentheses. $p$-values are indicated in square brackets; " $< 0.001$" denotes a $p$-value less than 0.001. "Control mean" denotes the average outcome for individuals who did not receive an answer within 24 hours, and the percentage ratios of the coefficients to control mean are indicated in angle brackets. "Response rate" represents the probability of receiving answers to their first question from others within 24 hours of posting the question.

Table A.VIII: Analysis with an Alternative Definition of Outcome: 1 to 7 Days Post Initial Question

| | PostedAnswer | AnswerCount | HasPosScore | IsAccepted |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| $SatisfiedAns$ | 0.0064 | 0.0140 | 0.0025 | 0.0021 |
| | (0.0003) | (0.0011) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 35.9 > | < 36.8 > | < 40.4 > | < 42.5 > |
| $UnsatisfiedAns$ | −0.0060 | −0.0139 | −0.0024 | −0.0019 |
| | (0.0002) | (0.0009) | (0.0001) | (0.0001) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −34.0 > | < −36.4 > | < −38.3 > | < −38.3 > |
| Control mean | 0.0177 | 0.0381 | 0.0062 | 0.0049 |
| Response rate | | 0.66 | | |
| N | | 1,592,162 | | |

Notes: OLS regressions of equation (2). $SatisfiedAns$ is a dummy variable indicating that the respondent received answers from others to their first question within 24 hours of posting and selected one of these as the accepted answer. Conversely, $UnsatisfiedAns$ is a dummy variable indicating that the respondent received answers within the same timeframe but did not select any as the accepted answer. The voting variable used in the analysis is based on votes cast within one week of the submission date of the post. In addition to the variables presented in the table, year-week fixed effects and question characteristics are included as controls. The list of question characteristics is provided in Online Appendix Table A.I. The coefficients reflect the impact on the subsequent answering behavior of questioners following their initial question post. Robust standard errors are presented in parentheses. $p$-values are indicated in square brackets; " $< 0.001$" denotes a $p$-value less than 0.001. "Control mean" denotes the average outcome for individuals who did not receive an answer within 24 hours, and the percentage ratios of the coefficients to control mean are indicated in angle brackets. "Response rate" represents the probability of receiving answers to their first question from others within 24 hours of posting the question.

Table A.IX: Analysis with Narrowed Timeframe for Received Answer Variables: Within One Hour

| | PostedAnswer | AnswerCount | HasPosScore | IsAccepted |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| $SatisfiedAnsHour$ | 0.0067 | 0.0153 | 0.0032 | 0.0018 |
| | (0.0004) | (0.0015) | (0.0003) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 23.4 > | < 25.5 > | < 32.7 > | < 22.7 > |
| $UnsatisfiedAnsHour$ | −0.0094 | −0.0202 | −0.0036 | −0.0031 |
| | (0.0003) | (0.0011) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −32.9 > | < −33.7 > | < −37.0 > | < −40.5 > |
| Control mean | 0.0286 | 0.0599 | 0.0098 | 0.0078 |
| Response rate | | 0.33 | | |
| N | | 1,592,162 | | |

Notes: OLS regressions of equation (2) where $SatisfiedAns$ and $UnsatisfiedAns$ are replaced with $SatisfiedAnsHour$ and $UnsatisfiedAnsHour$ respectively. $SatisfiedAnsHour$ is a dummy variable indicating that the respondent received answers from others to their first question within an hour of posting and selected one of these as the accepted answer. Conversely, $UnsatisfiedAnsHour$ is a dummy variable indicating that the respondent received answers within the same timeframe but did not select any as the accepted answer. The voting variable used in the analysis is based on votes cast within one week of the submission date of the post. In addition to the variables presented in the table, year-week fixed effects and question characteristics are included as controls. The list of question characteristics is provided in Online Appendix Table A.I. The coefficients reflect the impact on the subsequent answering behavior of questioners within seven days following their initial question post. Robust standard errors are presented in parentheses. $p$-values are indicated in square brackets; " $< 0.001$" denotes a $p$-value less than 0.001. "Control mean" denotes the average outcome for individuals who did not receive an answer within an hour, and the percentage ratios of the coefficients to control mean are indicated in angle brackets. "Response rate" represents the probability of receiving answers to their first question from others within an hour of posting.

Table A.X: Heterogeneous Effects of Answer's Evaluation by Other Members on Subsequent Questioner Behavior

| | PostedAnswer | AnswerCount | HasPosScore | IsAccepted |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *PositiveAnswer* | 0.0020 | 0.0101 | 0.0028 | 0.0018 |
| | (0.0004) | (0.0015) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < 6.9 > | < 16.2 > | < 28.2 > | < 22.3 > |
| *NonPositiveAnswer* | −0.0040 | −0.0127 | −0.0022 | −0.0019 |
| | (0.0003) | (0.0011) | (0.0002) | (0.0002) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −13.5 > | < −20.5 > | < −22.3 > | < −23.6 > |
| Control mean | 0.0294 | 0.0623 | 0.0100 | 0.0079 |
| Response rate | | 0.66 | | |
| N | | 1,592,162 | | |

Notes: OLS regressions of equation (2) where $SatisfiedAns$ and $UnsatisfiedAns$ are replaced with $PositiveAnswer$ and $NonPositiveAnswer$ respectively. $PositiveAnswer$ is a dummy variable indicating that the respondent received answers from others to their first question within 24 hours of posting, and at least one of these has a positive score. Conversely, $NonPositiveAnswer$ is a dummy variable indicating that the respondent received answers within the same timeframe, but none of these have a positive score. The voting variable used in the analysis is based on votes cast within one week of the submission date of the post. In addition to the variables presented in the table, year-week fixed effects and question characteristics are included as controls. The list of question characteristics is provided in Online Appendix Table A.I. The coefficients reflect the impact on the subsequent answering behavior of questioners within seven days following their initial question post. Robust standard errors are presented in parentheses. $p$-values are indicated in square brackets; " $< 0.001$" denotes a $p$-value less than 0.001. "Control mean" denotes the average outcome for individuals who did not receive an answer within 24 hours, and the percentage ratios of the coefficients to control mean are indicated in angle brackets. "Response rate" represents the probability of receiving answers to their first question from others within 24 hours of posting the question.

Table A.XI: Daily Impact Analysis: Answering Behavior within Seven Days Post Initial Question

| | PostedAnswer (1) | AnswerCount (2) | HasPosScore (3) | IsAccepted (4) |
|---|---|---|---|---|
| $SatisfiedAns \times 1(t=0)$ | −0.00026 | −0.00046 | 0.00003 | −0.00011 |
| | (0.00026) | (0.00032) | (0.00014) | (0.00013) |
| | [0.321] | [0.150] | [0.835] | [0.400] |
| | < −28.6 > | < −43.2 > | < 11.6 > | < −52.8 > |
| $SatisfiedAns \times 1(t=1)$ | 0.00017 | 0.00020 | 0.00003 | 0.00008 |
| | (0.00014) | (0.00017) | (0.00007) | (0.00006) |
| | [0.216] | [0.238] | [0.663] | [0.223] |
| | < 18.8 > | < 18.8 > | < 12.4 > | < 37.7 > |
| $SatisfiedAns \times 1(t=2)$ | 0.00032 | 0.00035 | 0.00005 | 0.00003 |
| | (0.00011) | (0.00013) | (0.00006) | (0.00005) |
| | [0.004] | [0.008] | [0.366] | [0.538] |
| | < 34.8 > | < 32.8 > | < 21.1 > | < 15.2 > |
| $SatisfiedAns \times 1(t=3)$ | 0.00035 | 0.00034 | 0.00015 | 0.00008 |
| | (0.00011) | (0.00015) | (0.00006) | (0.00005) |
| | [0.001] | [0.024] | [0.018] | [0.134] |
| | < 39.0 > | < 32.3 > | < 60.6 > | < 40.7 > |
| $SatisfiedAns \times 1(t=4)$ | 0.00012 | 0.00011 | 0.00006 | 0.00008 |
| | (0.00008) | (0.00010) | (0.00004) | (0.00004) |
| | [0.160] | [0.284] | [0.198] | [0.060] |
| | < 13.1 > | < 10.4 > | < 22.1 > | < 37.7 > |
| $SatisfiedAns \times 1(t=5)$ | 0.00008 | 0.00012 | 0.00000 | 0.00005 |
| | (0.00010) | (0.00013) | (0.00005) | (0.00005) |
| | [0.412] | [0.342] | [0.986] | [0.300] |
| | < 8.7 > | < 11.5 > | < 0.4 > | < 24.6 > |
| $SatisfiedAns \times 1(t=6)$ | 0.00019 | 0.00012 | 0.00004 | 0.00003 |
| | (0.00009) | (0.00011) | (0.00005) | (0.00004) |
| | [0.032] | [0.254] | [0.443] | [0.519] |
| | < 20.8 > | < 11.6 > | < 15.8 > | < 14.0 > |
| $UnsatisfiedAns \times 1(t=0)$ | −0.00155 | −0.00198 | −0.00054 | −0.00058 |
| | (0.00023) | (0.00028) | (0.00012) | (0.00010) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −170.7 > | < −185.4 > | < −211.9 > | < −288.9 > |
| $UnsatisfiedAns \times 1(t=1)$ | −0.00035 | −0.00041 | −0.00010 | −0.00005 |
| | (0.00011) | (0.00014) | (0.00006) | (0.00005) |
| | [0.002] | [0.003] | [0.087] | [0.241] |
| | < −38.6 > | < −38.4 > | < −38.8 > | < −27.3 > |
| $UnsatisfiedAns \times 1(t=2)$ | −0.00008 | −0.00007 | −0.00003 | −0.00002 |
| | (0.00009) | (0.00010) | (0.00005) | (0.00004) |
| | [0.346] | [0.528] | [0.504] | [0.677] |
| | < −8.8 > | < −6.2 > | < −12.4 > | < −8.4 > |
| $UnsatisfiedAns \times 1(t=3)$ | −0.00011 | −0.00023 | −0.00002 | 0.00000 |
| | (0.00008) | (0.00012) | (0.00004) | (0.00004) |
| | [0.155] | [0.050] | [0.626] | [0.982] |
| | < −12.6 > | < −21.4 > | < −8.5 > | < 0.5 > |
| $UnsatisfiedAns \times 1(t=4)$ | −0.00014 | −0.00016 | −0.00001 | −0.00001 |
| | (0.00006) | (0.00008) | (0.00003) | (0.00002) |
| | [0.022] | [0.041] | [0.756] | [0.527] |
| | < −15.7 > | < −15.3 > | < −3.7 > | < −6.6 > |

Table A.XI: Daily Impact Analysis: Answering Behavior within Seven Days Post Initial Question (continued)

| | PostedAnswer (1) | AnswerCount (2) | HasPosScore (3) | IsAccepted (4) |
|---|---|---|---|---|
| $UnsatisfiedAns \times 1(t=5)$ | −0.00020 | −0.00023 | −0.00004 | −0.00004 |
| | (0.00007) | (0.00010) | (0.00004) | (0.00003) |
| | [0.006] | [0.019] | [0.353] | [0.191] |
| | < −22.3 > | < −21.8 > | < −15.5 > | < −21.1 > |
| $UnsatisfiedAns \times 1(t=6)$ | −0.00008 | −0.00016 | −0.00008 | −0.00002 |
| | (0.00007) | (0.00009) | (0.00003) | (0.00003) |
| | [0.220] | [0.072] | [0.012] | [0.453] |
| | < −8.8 > | < −14.6 > | < −32.5 > | < −12.0 > |
| $1(t=1)$ | −0.00277 | −0.00326 | −0.00077 | −0.00070 |
| | (0.00020) | (0.00025) | (0.00011) | (0.00010) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −305.3 > | < −305.0 > | < −302.1 > | < −350.0 > |
| $1(t=2)$ | −0.00319 | −0.00374 | −0.00086 | −0.00075 |
| | (0.00020) | (0.00024) | (0.00010) | (0.00009) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −351.5 > | < −350.4 > | < −338.9 > | < −373.3 > |
| $1(t=3)$ | −0.00322 | −0.00369 | −0.00088 | −0.00075 |
| | (0.00019) | (0.00025) | (0.00010) | (0.00009) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −354.6 > | < −345.2 > | < −346.3 > | < −373.3 > |
| $1(t=4)$ | −0.00335 | −0.00391 | −0.00094 | −0.00081 |
| | (0.00019) | (0.00024) | (0.00010) | (0.00009) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −369.0 > | < −366.2 > | < −372.1 > | < −406.0 > |
| $1(t=5)$ | −0.00323 | −0.00377 | −0.00088 | −0.00077 |
| | (0.00019) | (0.00024) | (0.00010) | (0.00009) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −355.7 > | < −353.1 > | < −346.3 > | < −382.7 > |
| $1(t=6)$ | −0.00336 | −0.00388 | −0.00090 | −0.00078 |
| | (0.00019) | (0.00024) | (0.00010) | (0.00009) |
| | [< 0.001] | [< 0.001] | [< 0.001] | [< 0.001] |
| | < −370.0 > | < −363.5 > | < −353.7 > | < −387.3 > |
| Control mean | 0.00091 | 0.00107 | 0.00025 | 0.00020 |
| Response rate | | 0.66 | | |
| N | | 2,229,024 | | |

Notes: OLS regressions of equation (4). $SatisfiedAns$ is a dummy variable indicating that the respondent received answers from others to their first question within 24 hours of posting and selected one of these as the accepted answer. Conversely, $UnsatisfiedAns$ is a dummy variable indicating that the respondent received answers within the same timeframe but did not select any as the accepted answer. $1(t=d)$ is a dummy variable indicating $d$ days had elapsed since the first question post. The voting variable used in the analysis is based on votes cast within one week of the submission date of the post. In addition to the variables presented in the table, year-week fixed effects and question characteristics are included as controls. The list of question characteristics is provided in Online Appendix Table A.I. The coefficients reflect the daily impact on the subsequent answering behavior of questioners within seven days following their initial question post. Standard errors are clustered at the individual level and presented in parentheses. $p$-values are indicated in square brackets; " < 0.001" denotes a $p$-value less than 0.001. "Control mean" denotes the average outcome for individuals who did not receive an answer within 24 hours, and the percentage ratios of the coefficients to control mean are indicated in angle brackets. "Response rate" represents the probability of receiving answers to their first question from others within 24 hours of posting the question.

Table A.XII: Total Impact Analysis: Long-Term Posting Behavior within 365 Days after Initial Question

| | PostedAnswer | AnswerCount | HasPosScore | IsAccepted | Engaged |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| *Received* | −0.0032 | 0.0111 | −0.0001 | −0.0000 | 0.1049 |
| | (0.0005) | (0.0126) | (0.0003) | (0.0003) | (0.0008) |
| | [< 0.001] | [0.377] | [0.661] | [0.918] | [< 0.001] |
| | < −3.2 > | < 2.3 > | < −0.4 > | < −0.1 > | < 29.2 > |
| Control mean | 0.1003 | 0.4884 | 0.0393 | 0.0318 | 0.3596 |
| Response rate | | | 0.66 | | |
| *N* | | | 1,592,162 | | |

Notes: OLS regressions of equation (1). *Received* is a dummy variable indicating whether the sample received answers from others to their first question within 24 hours of posting the question. The voting variable used in the analysis is based on votes cast within one week of the submission date of the post. In addition to the variables presented in the table, year-week fixed effects and question characteristics are included as controls. The list of question characteristics is provided in Online Appendix Table A.I. The coefficients reflect the impact on the subsequent answering and asking behavior of questioners within 365 days following their initial question post. Robust standard errors are presented in parentheses. $p$-values are indicated in square brackets; " $< 0.001$" denotes a $p$-value less than 0.001. "Control mean" denotes the average outcome for individuals who did not receive an answer within 24 hours, and the percentage ratios of the coefficients to control mean are indicated in angle brackets. "Response rate" represents the probability of receiving answers to their first question from others within 24 hours of posting the question.