

A Little Help from My Friends: How Receiving Assistance Affects Participation in Online Knowledge- Sharing Communities

Eunkwang Seo
Frank Nagle
Sonali K. Shah

Working Paper 21-026



A Little Help from My Friends: How Receiving Assistance Affects Participation in Online Knowledge-Sharing Communities

Eunkwang Seo
University of Illinois at Urbana-Champaign

Frank Nagle
Harvard Business School

Sonali K. Shah
University of Illinois at Urbana-Champaign

Working Paper 21-026

Copyright © 2020 by Eunkwang Seo, Frank Nagle, and Sonali K. Shah.

Working papers are in draft form. This working paper is distributed for purposes of comment and discussion only. It may not be reproduced without permission of the copyright holder. Copies of working papers are available from the author.

Funding for this research was provided in part by Harvard Business School.

A Little Help from My Friends: How Receiving Assistance Affects Participation in Online Knowledge- Sharing Communities

Eunkwang Seo

University of Illinois at Urbana-Champaign
eseo7@illinois.edu

Frank Nagle

Harvard Business School
fnagle@hbs.edu

Sonali K. Shah

University of Illinois at Urbana-Champaign
sonali@illinois.edu

Keywords: Online communities, knowledge development, innovation, reciprocity

Acknowledgements: We thank Aravinda Garimella, Hyeonsuh Lee, Hari Sundaram, Nataliya Langburd Wright and participants at the 2019 Strategic Management Society Conference, the 2019 Boston College Digital Innovation Workshop, the 2019 Harvard Business School Strategy Alumni Workshop, and the 2020 Academy of Management Conference for feedback on these ideas. All authors contributed equally.

A Little Help from My Friends:

How Receiving Assistance Affects Participation in Online Knowledge-Sharing Communities

ABSTRACT

Online knowledge-sharing communities are important arenas for the development and exchange of knowledge; both firm participation and sponsorship in communities are increasing, as is scholarly interest in understanding how interactions between members shape subsequent participation. We contribute to this literature by examining the relationship between two increasingly relevant, yet underexplored, sets of factors: we examine the effects of with whom an interaction occurs (a user or an employee) on two types of participatory outcomes (future asking versus future replying). Specifically, we examine whether receiving a reply to a question—and from whom the reply is received—affects a *new member's* likelihood of asking another question or replying to a question. We examine fifteen years of data from Statalist, the online community that formed around the Stata statistical software package. Our findings show that receiving a reply from another user is positively associated with future questioning, but negatively associated with future replying. However, receiving a reply from a firm employee is positively associated with future replying. To help address endogeneity concerns, we conducted qualitative interviews with Stata employees, used fixed-effects models, implemented an instrumental variable-based approach, utilized a two-stage Heckman selection model, and applied coarsened exact matching; the findings remain consistent. These findings suggest that social interactions play a key, albeit nuanced, role in shaping subsequent participation and the continued health of a community.

Keywords: Online communities, knowledge development, innovation

INTRODUCTION

Online knowledge-sharing communities are increasingly important forums for knowledge exchange (Lee and Cole 2003, Wasko and Faraj, 2005, Bagozzi and Dholakia 2006, Zhang et al. 2013, Aaltonen and Seiler 2016, Bauer et al. 2016). They allow individuals with common interests to voluntarily and freely share valuable information beyond spatial, temporal, and social boundaries using a variety of computer-mediated communication technologies (*i.e.*, listservs, bulletin board systems, and online chat rooms). Such communities have formed to support the creation of novel technologies and products (*e.g.*, a hyperloop transportation system¹, the Linux computer operating system², and a ventilator³), as well as to provide support and extend existing products produced by corporations (*e.g.*, Apple's Support Community⁴, and LEGO's IDEAS platform⁵). When communities form, they provide value not only for participants, but also for the firms around whose products those communities form (Gabriel and Goldman 2005, Jeppesen and Frederiksen 2006, Shah 2006, Franke et al. 2008, Di Gangi and Wasko 2009, Bayus 2013, Manchanda et al. 2015, Nagle 2018, Dahlander et al. 2019, Shah and Nagle 2020). Firms can, for example, derive strategic benefits by using online communities to strengthen their innovation process, increase brand recognition, and reduce support costs. Given the value that online communities can generate for firms, identifying actions that a firm can undertake to promote the health of a community is critical (Shaikh and Levina 2019).

While many studies conducted to date have examined why individuals *join* a community (Dholakia et al. 2004, Ridings and Gefen 2004, Butler et al. 2007), fewer studies examine the factors that lead individuals to *continue* participating (Joyce and Kraut 2006, Zhang et al 2013). However, sustaining continued participation is critical to the long-term viability of an online community (*e.g.*, Worthen 2008, Ren et al. 2012): most communities are composed of a small set of core members and a much larger

¹ <https://www.hyperloopptt.com/>

² <https://www.linuxfoundation.org/>

³ <https://opensourceventilator.ie/>

⁴ <https://discussions.apple.com/welcome>

⁵ <https://ideas.lego.com/>

number of peripheral members, many of whom make only a single post to the community (Kuk 2006, Dahlander and Frederiksen 2012, Rullani and Haefliger 2013). For example, a study of six different online support groups finds that only 44% of individuals who contribute once make a second contribution (Joyce and Kraut 2006). Scholars interested in identifying ways to encourage continued participation have begun examining the effects of social interactions between members on continued participation, finding mixed evidence for whether or not receiving a reply increases a member's likelihood of posting again (Joyce and Kraut 2006, Zhang et al. 2013). Studies also find that emotional tone of the reply is not significantly related to continued contribution (Joyce and Kraut 2006).⁶

We build on and extend research examining how interactions can contribute to continued participation in two ways. Existing studies tend to view online contributions in aggregate. However, asking questions and replying to questions are distinct acts and likely driven by different factors. Recognizing this heterogeneity, we examine whether or not receiving a reply to the *first* question a new member posts influences their likelihood of contributing again by asking a question or replying to another member's question. In addition, the importance of firm contributions to online communities has been recognized by academics and industry practitioners alike (Kim 2000, Gu and Jarvenpaa 2003, Porter and Donthu 2008), but has not been empirically examined: "Two popular strategies to increase the likelihood of contribution are (1) having a firm's technical support staff and other employees respond to posted questions from time to time... When a firm's technical staff responds to questions, the hope is that customers will reciprocate by replying to questions and hence increase the overall knowledge base" (Gu and Jarvenpaa 2003, p. 110). We therefore investigate the effect of firm participation in online communities. Specifically, we ask: *When a new member poses their first question to the community, how do differences in from whom a reply is received—a user or firm employee—influence the new member's*

⁶ In addition, existing research in the context of online health support groups finds that receiving informational support (*e.g.*, facts about a disease, treatments, side effects, insurance and financial matters, etc.) is negatively related to the *length of time* that a user continues to participate in a community, whereas receiving emotional support is positively related (Wang et al. 2012). The need for emotional support is likely to differ by context.

likelihood of continuing to ask questions? Of replying to questions? Examining these issues is critical as building and maintaining participation is central to the health and continued functioning of online communities.

We investigate these questions in the context of the Statalist community, an online forum dedicated to sharing knowledge related to the Stata statistical software package. We examine the behaviors of 8,157 new members who made their first post during the 33-month time period from March 2014 through November 2016. Based on logit regression analyses, we find that when a new member receives a reply from another user who is not employed by the firm around which the community is organized (referred to as a user in the remainder of the paper) to the first question that they post, their likelihood of asking another question in the future increases by 13% and the likelihood of replying to another user's question decreases by 83% compared to if they had not received a reply at all. We then examine how receiving a reply from a firm employee affects these behaviors: receiving a reply from an employee does not appear to affect the likelihood of asking another question, but raises the likelihood of replying to another user's question by 50% compared to if they had not received a reply at all. To help address endogeneity concerns related to how employees choose the questions to which they reply, we conduct qualitative interviews with Stata employees, run fixed-effect analyses, and implement an instrumental variable-based approach. We also analyze the data utilizing a two-stage Heckman selection model and coarsened exact matching to provide evidence suggesting a more causal interpretation of the results. In addition to our key findings, we also note a few additional observations that will be of interest to scholars. Consistent with past studies, we find that response tone does not impact future behavior. We also find that new members behave similarly whether they receive a reply from a user or super-user. And, we also find that the length of time until a reply is received is negatively correlated with the likelihood of asking another question, but has no relationship with the likelihood of replying to the questions of others.

These findings contribute to the literature in three ways. To the literature on social interactions in online communities, our findings show that disaggregating future contributions by type (asking versus replying behaviors) and considering the role of the interaction partner in the community are important as

these factors can influence the direction of the future contribution, even causing changes in directionality of effects. To the literature on firm participation in online communities, we show that small levels of effort by firms can contribute to the health and continued vibrancy of the community. To the literature on reciprocity, we show that, at least in online-knowledge sharing communities, actual levels of reciprocity are much lower than that suggested by studies examining behavioral intentions. These findings also have practical implications for firms, community members, and policy-makers interested in seeding and maintaining vibrant communities: at low levels, participation by employees of a firm around whose product a community is structured has a positive effect on user participation (and does not detract from user activity).

THEORY & HYPOTHESES DEVELOPMENT

Social interactions within an online knowledge-sharing community are an important factor shaping members' decision to continue participating in the community (von Hippel and von Krogh 2003, Gächter et al. 2010, Dewan et al. 2017, Giordani et al. 2018). In particular, receiving a reply to a posted question tends to lead to continued participation: amongst new members, receiving a reply increases the likelihood of continued participation (Joyce and Kraut 2006); and amongst existing members, receiving a reply increases the likelihood of users', but not modifiers', continued participation (Zhang et al. 2013).⁷ Existing empirical studies tend to measure participation as *any* post an individual makes to a community (*i.e.*, asking questions or replying to questions), thereby aggregating various participation behaviors. Although prior studies do not distinguish between them, the two behaviors are quite different and are likely to be driven by different mechanisms (reviewed below). In addition, existing studies tend to focus on the actions of users. In most communities, users make up the bulk of the membership and contribute the majority of posts to the community. However, as firms become increasingly involved in online

⁷ *Users* refer to “those individuals in the OSSD [Open Source Software Development] user community who plan to adopt or have adopted the software, typically with no intention as yet of modifying the software themselves”, while *modifiers* refer to “those individuals in the OSSD user community who have modified (or plan to modify) the software code” (Zhang et al 2013, pp.1117-1118).

knowledge sharing communities for strategic purposes, there is a need to understand how the actions of employees influences social interactions within the community.

Therefore, in this study, we seek to extend this stream of research on social interactions and continued participation by making two major contributions. We differentiate two types of continued participation in online knowledge-sharing communities: (1) asking questions and (2) replying to the questions of other users. By differentiating more sharply on these dimensions, we seek to provide a more nuanced understanding of these relationships. Also, we differentiate between *who* replies to a new member's first question: employees of the firm around whose product the community has formed, or users who are not employed or formally associated with the firm around whose product the community is built. We use the term community member to refer to all participants in a community. Below, we theorize the expected effects of these factors.

Effects of Receiving a Reply from a User on Asking Questions in the Future

We expect that receiving a reply from a user(s) will increase the likelihood that a new member will continue to ask questions in an online knowledge-sharing community. To explain why, we, like other scholars (Chen 2007, Cheung and Lee 2007, Jin et al. 2009, Jin et al. 2010), draw on expectation confirmation theory and conceptualize new members' future questioning as a "repurchase" behavior. Initially created to explain the psychological underpinnings of consumer satisfaction, expectation confirmation theory argues that satisfaction with a particular product or service is a function of expectations, perceived performance, and belief confirmation (Oliver 1980): specifically, consumers form an initial expectation about a specific product or service prior to using them; as they use the product or service, they assess the extent to which their initial expectation is confirmed or disconfirmed; and satisfied consumers are likely to repurchase or continue to use the product or service, whereas dissatisfied consumers do not.

In the context of online knowledge sharing communities, a new member might observe the question answering and asking activity going on and form an initial expectation that they will receive a reply from another user should they post a question. If they get no reply, their initial expectation is

disconfirmed and thus they will not post another question. If they get a reply, their expectation is confirmed and they might post another question. We therefore expect that receiving a reply from a user would increase the likelihood of a new member asking another question in the future.

Hypothesis 1. When new members post their first question and receive a reply from another user (vs. receiving no reply), they are more likely to ask another question in the future.

Effects of Receiving a Reply from a User on Replying to Questions in the Future

We now turn to understanding how receiving a reply from a user(s) will affect a new member's likelihood of replying to questions; here, it is unclear what pattern will emerge. On one hand, norms around reciprocity might lead a new member to reply to others' questions (Gouldner 1960), because it represents what the user feels s/he ought to do as a member of the community. Reciprocity is a driver of behavior in a wide variety of contexts (Nowak and Sigmund 1998, Greiner and Levati 2005, Seinen and Schram 2006, Khadjavi 2017): "reciprocity means that in response to friendly actions, people are frequently much nicer and much more cooperative than predicted by the self-interest model; conversely, in response to hostile actions they are frequently much more nasty and even brutal" (Fehr and Gächter 2000: p.159). When reciprocity is repaid to someone other than the giver—as is almost always the case in online knowledge sharing communities (Bock et al. 2005, Chiu et al. 2006, Cheung and Lee 2007, Jin et al. 2013)—it is known as indirect reciprocity (Gouldner 1960, Yamagishi and Cook 1993, Bearman 1997, Nowak and Sigmund 1998). Surveys show that some community members believe that norms of reciprocity—that others would help them and that they should help others—exist in their communities (Wasko and Faraj 2000, Jin et al. 2013). Scholars find evidence that individuals who strongly believe that a norm of reciprocity exists in their online community *intend* to share more knowledge (Bock et al. 2005, Cheung and Lee 2007, Jin et al. 2013). Therefore, if reciprocity were at play, we would expect that receiving a reply from a user would increase the likelihood that a new member replies to questions in the future.

Hypothesis 2a. When new members post their first question and receive a reply from another user (vs. receiving no reply), they are more likely to reply to the questions of others in the future.

However, empirical evidence in support of norms of reciprocity leading to an increase in actual contributions (measured as the number of postings, *i.e.*, asking and replying to questions) is mixed. Chiu et al. (2006) find evidence that individuals who believe strongly in norms of reciprocity in their community contribute more, however, Wasko and Faraj (2005) find no significant relationship between norms of reciprocity and subsequent contributions. And, Kankanhalli et al. (2005) find that reciprocity is positively related to contributions only when pro-sharing norms are weak.⁸

Given the mixed data on the effects of reciprocity in online knowledge sharing communities, it's worth considering other factors that might influence the behavior of a new member. Considering the counterfactual is useful: how might a new member behave if they posted a question and *did not* receive a reply? Wang et al. (2012) found that once people received or found the information they needed, they stop searching and actively monitoring the community webpage.⁹ On the other hand, those who did not yet receive a reply could continue checking the community for replies as well as explore archives for an answer. During this process, they might encounter questions that they can help answer and proceed to reply to those. Alternatively, or in addition, the new member might begin to wonder if their chances of receiving a reply might increase if they first helped others and behave accordingly. In both these scenarios, not receiving a reply (or having to wait for an extended period of time before receiving a reply) might lead a new member to *reply* to questions. If this is the case, then a user who receives a reply may be *less* likely to reply to others (particularly if they receive a reply quickly and stop monitoring the community). We therefore expect that receiving a reply from a user might decrease the likelihood of a new member replying to questions in the future.

Hypothesis 2b. When new members post their first question and receive a reply from another user (vs. receiving no reply), they are less likely to reply to the questions of others in the future.

⁸ All the aforementioned studies focus on how an individual's understanding of reciprocity in their community lead to actual or intended behavior. Our focus is slightly different, we look at how receiving help affects knowledge sharing behavior (with reciprocity potentially being the unobserved mechanism).

⁹ Of course, if and when they have another question, they can come back and post another question to the community, as covered in Hypothesis 1.

Effects of Receiving a Reply from a Firm Employee on Asking Questions in the Future

We expect that receiving a reply from an employee will increase the likelihood that a new member will continue to ask questions in an online knowledge sharing community. We believe this to be the case as receiving a reply, consistent with the arguments made in support of Hypothesis 1, confirms the expectation the new member has that they will receive a reply and may lead them to ask more questions in the future (Oliver 1980, Chen 2007, Cheung and Lee 2007, Jin et al. 2009, Jin et al. 2010). That said, the new member's expectation is likely that they will receive a reply from another *user* and receiving a reply from an employee may be especially appreciated as firm employees are knowledgeable about the product and thus likely to provide a clear and nuanced reply and are likely to be perceived as a highly credible source of information, both factors that have been found to lead to satisfaction and therefore continued participation in online communities (Chen 2007, Jin et al. 2009). Because the new member's expectations have been met—and perhaps even surpassed—we therefore expect that receiving a reply from an employee would increase the likelihood of a new member asking another question in the future.

Hypothesis 3. When new members post their first question and receive a reply from a firm employee (vs. receiving no reply), they are more likely to ask another question in the future.

Effects of Receiving a Reply from a Firm Employee on Replying to Questions in the Future

We expect that receiving a reply from an employee will increase the likelihood that a new member will reply to questions in an online knowledge sharing community.¹⁰ We expect that two theorized mechanisms may act individually or in combination to promote future replying: the reaction to receiving the “gift” of an employee reply or the sense of being observed. We discuss each possible mechanism below.

Consistent with classic anthropological and sociological theorizing around gift exchange (Mauss 2002), it is possible that receiving a reply from an employee might encourage new members to contribute

¹⁰ Our theorizing here is different from that examining the effects of receiving a reply from a *user* on answering behavior (H2a and H2b). This is because reciprocity is unlikely to be at play (in contrast to H2a) and, while the same logic applied in H2b may apply to individuals who do not get a reply, the opposite may not apply to those individuals who do get a reply.

more to the community. New members, expecting to receive a reply from another user (H1), might interpret a reply from an employee as a gift, because firm employees are likely to be perceived as both highly knowledgeable and credible sources of information; hence they have received something of greater value than what they expected to receive when they posted their question. In Mauss's conceptualization, when a gift is given, it creates a debt that needs to be repaid: "In theory, these are voluntary, in reality they are given and reciprocated obligatorily (Mauss 2002, p.3)." Thus, to the extent that the new member considers the reply a gift, s/he would feel obligated to return the gift or offer a gift in return.¹¹ We therefore suggest that receiving a reply from a firm employee may be perceived as a gift and therefore lead them to reciprocate by replying to questions.

In addition, and in line with the classic Hawthorne effect, it is possible that receiving a reply from an employee might encourage new members to contribute more to the community. Prior research has found that people become more productive when observed by people in authority (Blalock and Blalock 1982, Festinger and Katz 1953, McCarney et al. 2007): the psychological stimulus of being singled out and made to feel important can increase people's motivation to contribute more.¹² In this vein, receiving a reply from firm employees could increase continued participation of new members.

Together, these reasons lead us to suggest that receiving a reply from a firm employee might increase the likelihood that a new member replies to questions in the future.

Hypothesis 4. When new members post their first question and receive a reply from a firm employee (vs. receiving no reply), they are more likely to reply to the questions of others in the future.

¹¹ Realistically, returning the gift is an unlikely outcome as it would be awkward and knowledge, as a non-excludable good, cannot be returned once shared.

¹² Along similar lines, Gallus (2016) finds that new Wikipedia contributors who receive a symbolic award—another form of being singled out and made to feel important—are more likely to continue participating in the community.

METHODS

Empirical Setting

We examine online community interactions on Statalist, a forum dedicated to sharing knowledge related to Stata. Stata—created by StataCorp in 1985—is an integrated statistical software package for data analysis, data management, and graphics. Statalist is an ideal context for our analysis, because it allows us to examine community members’ behavioral responses to interactions with other members, as well as with firm employees. Moreover, employee replies appear relatively random: StataCorp employees occasionally contribute to Statalist, although they are not required to do so and do not seem to contribute in a systematic manner (we discuss this further below). This is made clear on Statalist’s splash page, which states: “Statalist is run and moderated by Stata users and maintained by StataCorp. If you have private comments about contents, policy, or the software, contact the forum administrators.” Figure 1 shows an example of a question and reply in Statalist Forum.

[Insert Figure 1 about here]

Statalist was established by users in August 1994 and was originally run as a single archived email list hosted on a server at Harvard University. In March 2014, Statalist was migrated to servers maintained by StataCorp and the format of Statalist changed from an archived email list to a web-based forum; however, Statalist continued to be moderated by Stata users. Our analysis focuses on activity occurring *after* the format change in order to ensure consistency within our sample, but we perform robustness checks using earlier data from the mailing list as well.

Data and Sample

Statalist has three sub-forums: General, Mata, and Sandbox.¹³ We focus on the General sub-forum, which contains the vast majority (98%) of posts. We began by collecting all posts made to Statalist during the fifteen-year period from June 2002 to May 2017.¹⁴ We then manually removed posts in which

¹³ The Mata subforum is devoted to discussing Stata’s matrix programming language Mata. The Sandbox subforum allows new members to make test posts and practice with the forum software before posting.

¹⁴ Although Statalist started in 1994, historical data is only available starting in June 2002.

a question was not asked, including announcements, conference advertisements, and program updates, resulting in 241,792 posts in which a question was posed by 25,985 users. We then set November 2016 as the end of our sampling period to allow for a 6-month window in which to gauge subsequent interactions, that is for replies to be posted and for the original posters to make a subsequent contribution to the community. Then, to focus on the subsequent behaviors of new members, we further restricted the dataset to focus on new members who made their first post to the community between March 2014 to November 2016, resulting in 8,632 posts/participants.¹⁵ Finally, we identified new members whose first contribution to the community involved posting a new question, rather than responding to someone else's question, resulting in 8,326 posts/users.¹⁶ In addition, we also excluded 169 new members (out of 8,326 total) who posted another question within 10 minutes of posting their first question. When examining the data, we found that such posts were made for one of two reasons: they were either made mistakenly with the same question being posted more than once or the new member had more than one question they needed help with and posted them nearly simultaneously. Both of these actions are unrelated to our core research interest (understanding the impact of getting a reply to your question) as the additional questions they posted were either a mistake or independent of whether or not they got a reply to their first question. As a result, our final sample includes the first question posted by 8,157 new members to Statalist from March 2014 to November 2016. We also capture replies to these questions and the subsequent behavior of the focal members over the ensuing 6 months.

Measures

Dependent Variables

We measure two behaviors that are important to the health and continued success of online knowledge-sharing communities. Our first dependent variable, *Future Questioning*, equals one if a new

¹⁵ We ensure that these individuals are first-time users by comparing usernames against the full list of individuals posting to Statalist since the start of available data in June 2002.

¹⁶ The vast majority of users'—96.46% (8,326 of 8,632)—first post involved posting a question. The 3.54% of users whose first post involved answering a question are excluded in our main analyses, as they behave differently from others and may have different underlying characteristics. We examine these outliers in our robustness checks.

member posted another question within 6 months of their original question and zero otherwise. Our second dependent variable, *Future Replying*, equals one if a new member posted a reply to someone else's question within 6 months of their original question and zero otherwise. We limit our timeframe to 6 months after the first question was posted to account for the possibility that any effect from getting a reply to that first question may wear off over time and/or other community interactions could play a role. Robustness checks using shorter (1 month and 3 months) and longer windows (1 year and all time) show consistent results (not shown due to space constraints; available from the authors upon request). That said, it appears that, on average, most subsequent behaviors occur within a month of the first posting: in our sample, 31% of new members posted another question within 6 months (30 days on average), while only 6% posted a reply to someone else's question within 6 months (36 days on average). As a robustness check, we examine the number of future questions and replies (rather than a simple binary) as a dependent variable using Poisson regression models and find consistent results.

Explanatory Variables

Received Reply is a binary explanatory variable indicating whether or not a new member's first question received a reply from either a user or a Stata employee. It is coded as zero if the first post received no reply and coded as one if the post received a reply (from either a user or an employee) within six months.¹⁷ Of the 8,157 new members, 78% (6,382 new members) received a reply. *Received User Reply* is a binary explanatory variable indicating whether or not a new member's first question received a reply from a user. It is coded one if the post received a reply from a user and zero otherwise. *Received Employee Reply* is a binary explanatory variable indicating whether or not a new member's first post received a reply from a StataCorp employee.¹⁸ It is coded one if the post received a reply from an

¹⁷ In the rare occasion where a user asks a question, but does not receive a reply to that question until *after* they have asked another question, they are included in the control group and not the treatment group. This is because, they performed the outcome measure of "future participation"—asking a question—before being "treated." Therefore, their behavior of asking an additional question (without having received a reply) contributes to the baseline likelihood of asking a second question.

¹⁸ As shown in Figure 1, users can easily recognize employee responses, because these are clearly marked as coming from a "StataCorp Employee."

employee and zero otherwise. The data show that only 2% of new members (173 out of 8,157 new members) received a reply from a Stata employee.

A single question can receive replies from multiple people, therefore, we consider *all* replies to a question up until the point the new member engages in future participation (by asking another question or replying to the question of another user), or six months, whichever comes first; over this time period, the majority of posts receive a reply. As a result, the *Received User Reply* and *Received Employee Reply* variables are *not* mutually exclusive. The data show that only 1% of new members (112 out of 8,157 new members) received replies from *both* a user and employee. To simplify this analysis, we also run a robustness check where we only consider the first reply to a question (rather than all replies), and find consistent results.¹⁹ Finally, *Elapsed Time* is a continuous variable that captures the time (in minutes) between the first question and the first reply.

Control Variables

We controlled for five characteristics of the first question—total word count, emotional tone, analytic score, proportion of words that are six letters or more, and the proportion of words in the post that are dictionary words—using the Linguistic Inquiry & Word Count (LIWC) text analysis program.²⁰ These linguistic attributes have been shown to be useful for capturing differences across written texts (Joyce and Kraut 2006). In addition, we controlled for topic-specific heterogeneities by using Latent Dirichlet Allocation (LDA) topic modeling: an analysis of 26,638 questions posted in Statalist from March 2014 to May 2017 resulted in the identification of 10 topics. Including the probability of each post fitting into a given topic in our regression models allows us to account for their potential impact on both

¹⁹ Excluding the observations that have both user and employee replies results in broadly consistent results, with the one exception of an insignificant correlation between receiving an employee reply and our instrumental variable (i.e., posting in a working hour) in the first stage. The insignificant correlation is likely due to reduced statistical power due to the low number of employee replies.

²⁰ The emotional tone is a measure of how positive or negative the sentiments in the text are, based on dictionaries of positive and negative words. The measure ranges from 0 (very negative) to 100 (very positive). The analytic score captures the degree to which the post demonstrates logical, formal, or hierarchical thinking and ranges from 0 (low analytical thinking) to 100 (high analytical thinking).

the likelihood of receiving a reply and of continued participation. Lastly, we controlled for temporal shocks by including year-fixed effects in our regression models.

Estimation

We estimate the probability of our binary outcomes using Logit models, specified as follows:

$$Prob(Y = 1|\mathbf{x}) = \frac{\exp(\mathbf{x}'\beta)}{1 + \exp(\mathbf{x}'\beta)}$$

where Y refers to our binary dependent variable (either by future questioning or future replying), β denotes the coefficient vector, and \mathbf{x} represents the set of explanatory and control variables including year-fixed effects. To account for heteroskedasticity, we use robust standard errors.

RESULTS

Tables 1 and 2 provide summary statistics and correlations between variables. With respect to receiving a reply, 78% of new members received a reply(ies) from others (either another user or an employee), 2% received a reply(ies) from an employee, and 22% did not receive a reply.²¹ The average time to receive a reply was about one day (approximately 28 hours or 1,701 minutes), although the median time was only about one hour (62 minutes).²² Comparing the average elapsed times of getting a reply from a user versus an employee, the difference is less than 1 minute. This tells us that we can rule out the possible explanation that employees choose to respond to unanswered questions (*i.e.*, questions that users have not chosen to or been able to answer). With respect to subsequent participation, only 34% of new members participate again, with 31% of new members posting another question within 6 months and 6% posting a reply to someone else's question within 6 months. For those who do contribute again, the average time to a second contribution was 28 days and the median time was 8 days. Figure 2 summarizes our hypotheses and the main results of our regression analyses.

[Insert Tables 1 and 2 and Figure 2 about here]

²¹ 77% received a reply from only users, 1% received a reply from only Stata employees, and 1% received replies from both users and Stata employees.

²² This variable is extremely right-skewed: 1st quartile = 20 mins, 2nd quartile = 62 mins, 3rd quartile = 256 mins.

Table 3 shows the results of a logit analysis predicting the likelihood that a new member's post will receive a reply from other users or Stata employees. The results suggest that post characteristics are related to the likelihood of receiving a reply, so we control for these characteristics in our primary regressions. Users and Stata employees respond similarly to most post characteristics, but notably differently to one characteristic: specifically, a higher percentage of *Analytic* words increases the likelihood of getting a reply from a Stata employee, while it decreases the likelihood of getting a user's reply. This result suggests that Stata employees might behave slightly differently from users when it comes to selecting posts to reply to. Because of the potential endogeneity concern resulting from this difference, we not only control for these characteristics, but also conduct a number of robustness checks to allow for a more causal interpretation of our results. These include conducting interviews with Stata employees to better understand when they reply to questions on Statalist and how they choose questions to which to reply, a fixed effects analysis, an instrumental variable based approach, a Heckman selection model, and coarsened exact matching (described below).

[Insert Table 3 about here]

In Table 4, we test our primary hypotheses. As a baseline, receiving a reply from *anyone* is positively related to future questioning (Columns 1, 2, 3). Column 4 disaggregates this finding to show the effects of receiving a reply from users and employees. Consistent with Hypothesis 1, receiving a reply from a user increases continued questioning behavior: according to the analysis of average marginal effects (Column 4), receiving a reply from a user increases the likelihood of asking another question from 31% (the sample average) to 35%, a 13% increase ($p\text{-value}=0.004$). The data do not, however, confirm Hypothesis 3: getting a reply from a Stata employee does not increase the likelihood of continued questioning in a statistically significant manner ($p\text{-value}=0.601$). Although this result does not add support for Hypothesis 3, the point estimate for the employee reply ($\beta_{\text{employee}}=0.092$) has an overlapping

confidence interval with the user effect ($\beta_{\text{user}}=0.184$) and could be of a similar size, but might suffer from a power problem due to the small sample size of replies coming from employees (2% of the sample).²³

In Column 5, we explore the impact of how long it takes to get a reply on future behavior. We limit our analysis to those posts that received a reply. We find that the elapsed time between the first post and the first reply is negatively associated with future questioning (Column 5). When the elapsed time increases by 1%, the likelihood of asking another question in the future is shown to decrease from 31% to 28%, a 10% decrease ($p\text{-value}<0.001$). This result suggests that there is significant heterogeneity even amongst those who received a reply within 6 months, depending on the elapsed time: in support of the idea that expectation-satisfaction drives on continued participation, we observe that new members who receive a reply sooner are more likely to post another question.

We now shift our focus to examining the effects of receiving a reply on future replying to someone else's question. Receiving a reply from *anyone* is negatively related to replying to others' questions (Column 6, 7, 8). Digging deeper, when we analyze *who* the reply comes from (Column 9), a stark difference in behavior as a result of getting replies from users versus employees become clear. Consistent with Hypothesis 2b but not 2a, receiving a reply from a user has a strong negative effect. An analysis of average marginal effects (Column 9) shows that getting a user reply decreases the likelihood of responding to the question from 6% (the sample average) to 1%, about an 83% decrease ($p\text{-value}<0.001$). In contrast, receiving a reply from an employee, consistent with Hypothesis 4, has a strong, positive, and statistically significant impact on the new member's future replying to someone else's questions in the future. Specifically, receiving a reply from a Stata employee increases the likelihood of providing a reply to another user in the future from 6% to 9%, a nearly 50% increase ($p\text{-value}=0.032$).²⁴ Unlike the negative effect of elapsed time on future questioning (Column 5), the time elapsed between

²³ β_{employee} is not statistically different from β_{user} ($p\text{-value}=0.617$).

²⁴ Given that only 2% of the sample receives a reply from a Stata user, the fact that this result shows up as strongly significant implies that the effect is indeed quite strong.

asking a question and receiving a reply does not have a significant impact on replying to the questions of others (Column 10).

[Insert Table 4 about here]

Table 5 shows the results of linear probability models for our binary dependent variables, showing consistent results with the Logit models presented in Table 4. In Table 6, we examine the number of posts (asking and replying to questions) that a new member made after the first question within 6 months as a continuous dependent variable rather than considering this activity in only a binary manner as in Tables 4 and 5. Three observations have extreme outcomes (greater than 200 posts). To mitigate the effect of these outliers, we excluded these observations in our analyses of future replying (Table 6, Columns 4, 5, and 6). As shown in Table 6, Poisson models with robust standard errors provide consistent results with our main findings.

[Insert Table 5 and 6 about here]

Addressing Potential Endogeneity Concerns: Fixed-Effects, Interviews & Instrumental Variable Analyses

Because our explanatory variables are not randomly assigned, our regression analysis might produce biased estimates. Specifically, an endogeneity concern arises if Stata employees purposefully or inadvertently reply to questions from users who are predisposed to continue contributing to the community by replying to questions (*e.g.*, perhaps these users are more knowledgeable, confident, and/or community-oriented users and this is somehow signaled in their first posts) (H3 and H4). However, if employees reply to questions in a manner that is *not* correlated with a particular user's likelihood of continued participation, then our results can be interpreted in a more causal manner. To help address this concern, we did multiple things: we conducted qualitative interviews with Stata employees; we analyzed users who received a reply from a Stata employee to a later question, but not their initial question using fixed-effects; and we implemented an instrumental variable based approach. In addition, we utilized a two-stage Heckman selection model, and applied coarsened exact matching.

To the extent that endogeneity is an issue in that Stata employees are somehow replying to users who are predisposed to contributing more, we investigate the behaviors of users who did not receive an employee reply to their initial question, but did receive an employee reply to a later question (N=2,152). If endogeneity is a substantial cause for concern, we would expect that these individuals would contribute more than other users in the interim time period. Looking at the within-individual comparison using individual-fixed effects, we find results that are consistent with our main findings, helping to limit endogeneity concerns.²⁵ Specifically, we find that, when a user who received only user replies to prior questions then receives a reply from an employee for the current question, the effect on future replying is positive and significant ($\beta_{\text{employee}}=0.078$; $p\text{-value}=0.044$).²⁶

In addition, to help address this endogeneity concern, we corresponded with several Stata employees who had replied to user questions on Statalist to better understand when and why they replied to community questions. 29 unique employees in our dataset had participated in the forum (out of an estimated 100-250 employees total), predominately during working hours (8am to 6pm local time). Through our interviews we learned that there are several reasons employees reply to community questions, however none of these reasons appear to be tied to the likelihood of the user's future participation. Some employees reported answering questions on Statalist as part of their onboarding activities:

Part of the training for statistical developers is to participate in answering tech-support questions. Our replies are reviewed. After our training is completed, we are encouraged to participate on Statalist. We usually ask a senior developer to review a reply before posting it to Statalist (Developer 1).

The employees we interviewed noted that their participation tends to focus on looking for bug reports (that only a Stata employee can address in the main code), questions that cannot be answered based on existing Stata-provided documentation, or new commands. They also clearly indicated that Stata is a user

²⁵ To estimate the coefficients, we use linear panel models because we can difference out the fixed effects without loss of any observations in the data.

²⁶ Consistent with our main results, receiving a reply from a user is negatively related to future replying ($\beta_{\text{employee}}=-0.025$; $p\text{-value}=0.008$). Full results are not shown due to space constraints and are available from the authors upon request.

forum and that participation is not required and is a voluntary act above and beyond their official responsibilities.

*I mainly look for posts that point out when Stata is not working properly. I used to try to answer more involved programming questions, but they tend to take too much time away from some of my other responsibilities.... **I tend to participate more on Statalist when I'm waiting for a task or when I am between projects**... All StataCorp developers are encouraged to participate on Statalist. Some developers are more systematic about when they browse the forum. Others are like me, peeking in from time to time (Developer 1).*

*We have a company policy that Statalist is supposed to be a user forum and not a StataCorp directed forum. While many of us look at the posts to see if there is anything interesting, **we do not respond to posts unless there is a bug report, a report of Stata doing something correctly, but in an odd fashion, or if there is a question about a computation which cannot be answered from our documentation**... We also respond if we have written unofficial commands (a.k.a. community-contributed software) about which someone has a question. Most of my replies have fallen into this last case. Otherwise **there isn't much of a decision process** that goes on (Developer 2).²⁷*

Based on these replies, we believe that the endogeneity concern (that Stata employees were systematically answering the questions of users more likely to participate in the future) is partially allayed. However, to further address this concern, we use an instrumental variable approach. Our instrument for receiving a reply from a Stata employee consisted of whether or not a question was posted during standard working hours at Stata Corp's headquarters in College Station, Texas (*i.e.*, a dummy variable for whether or not the post was made between 8am and 6pm Central Time on weekdays that were not a national holiday). We chose this IV because it is unlikely that a user would specifically wait until these hours to post a question on Statalist, but instead would post their question when it arose (noting that users are spread out across the world and therefore time zones). Further, our discussions with Stata employees indicated that they primarily answer questions while they are at work, rather than on their own

²⁷ To clarify, this statement implies there are certain types of questions this Stata employee is more likely to answer: namely questions where it appears that a user would have a difficult time implementing a solution if there was indeed a problem (and not a misunderstanding). This makes sense as a way for this employee to focus on questions that s/he, as a Stata employee, can uniquely answer. This helps decrease our worries of there being an endogeneity concern related to which questions Stata employees answer. Specifically, for this to increase rather than decrease the endogeneity concern, one would have to believe that such questions were more frequently asked by users who are more likely to reply to the questions of others. However, this is unlikely to be the case (e.g., unofficial commands are widely used by Stata users of all levels, which limits the likelihood that when a Stata employee answers a user's question, that user is more or less likely to answer the questions of others; likewise, if there was an underlying problem with the Stata software itself, the likelihood that someone who encounters this problem and posts a question about it is correlated with the likelihood of helping people in the future is low).

time (*e.g.*, nights, weekends, and holidays). Therefore, whether or not it is working hours for Stata employees can be seen as uncorrelated with the likelihood of a new member's future participation on Statalist. Although this instrument is not perfect, it allows us to exploit a plausibly exogenous variation in the likelihood that a user receives a reply from a Stata employee.

To implement this instrument, we use a Bivariate Probit model. In the first stage, a strong correlation is found between the instrument and the chance of getting a reply from an employee. As shown in Column 1 of Table 7, a participant posting their first question during Stata working hours increases the likelihood of getting a reply from a Stata employee from 2% to 3%, about a 50% increase ($p\text{-value}=0.002$). Using Bivariate Probit models, we estimated the impact of receiving a reply from an employee (Wooldridge, 2010). Table 7 shows the results of this instrumental variable approach. Consistent with previous models, the results show that receiving a reply from a Stata employee significantly impacts new members' future replying (Column 3). Specifically, getting a reply from a Stata employee increases the likelihood of posting a reply to other users in the future from 6% to 45%, about a 650% increase ($p\text{-value}=0.097$). Our analyses using instrumental variables suggests that not accounting for the possibility of endogeneity in our initial estimates (Table 4, Columns 4 and 9) may underestimate the impact of receiving a reply from Stata employees on new members' future replying. This attenuation towards zero in the OLS estimates may occur due to measurement error, which the IV can help address in a manner similar to Bloom, Schankerman, and Van Reenen (2013).

[Insert Table 7 about here]

Robustness Checks

Lastly, we performed a series of robustness checks in Table 8. First, we re-ran our main regression analysis with an expanded sample period, from June 2002 to November 2016, including the posts in both the old mailing-list format and the new forum format. As shown in Columns 1 and 7, we found consistent results with 18,866 observations. Second, we used alternative measures for our explanatory variables. Specifically, focusing on the first reply to our sample posts, we measured *Received Employee Reply* and *Received User Reply* depending on whether the first reply was made by a Stata

employee or a user. The first reply has been shown to have an outsized effect in online discussion boards (Muchnik, Aral, and Taylor 2013). Among 6,382 posts that received a reply, 87 posts received their first reply from a Stata employee and 6,295 received their first reply from a user. With these measures, we still found qualitatively consistent findings, which are reported in Columns 2 and 8. Third, we conducted analyses only using new members who had posted a reply to other's question before they posted their first question, which we dropped in previous models as these are a small number of users and can be considered outliers. As Columns 3 and 9 suggest, these users do indeed behave differently from other new members on multiple dimensions, although there are power issues in our analyses, likely due to the small number of observations. In Columns 4 and 10, we conducted a survival analysis, which differentiates censored and uncensored observations to estimate the conditional probability of an event at a certain time. The survival analysis allows us to estimate *how quickly* continued participation occurs after a user receives a reply. The Cox-Hazard models are consistent with our main findings. The results suggest that getting a reply from other users is associated with new members asking questions sooner, and getting a reply from an employee is associated with new members responding to others' questions sooner. Fifth, instead of Bivariate Probit models, we used two-stage Heckman selection models to conduct the instrumental variable analysis. The positive impact of Stata employees' replies on new members' continued participation remains consistent in these models (Columns 5 and 11). Lastly, we further investigated the impact of Stata employees' replies by using coarsened exact matching (Iacus, King, and Porro 2012). Among 6,382 posts that received a reply, we matched 173 posts replied to by Stata employees to 5,214 posts replied by users, based on the control variables that were shown to affect the likelihood of receiving a reply from Stata employees differently than the likelihood of receiving a reply from users.²⁸ As shown in Columns 6 and 12, we find consistent results and still find that the positive impact of Stata employees' replies on new members' future replying holds with the matched sample.

²⁸ As shown in Table 3, *Topic 2*, *Topic 4*, *Topic 6*, *Topic 8*, and *Analytic* have different (significant) impacts on the likelihoods. For the topic variables, we use the cutpoint of 0.5 to define the coarsening because the value of those variables represents the probability of each topic.

[Insert Table 8 about here]

We also considered additional aspects of timing and found results consistent with our primary analysis²⁹: we expanded and contracted the six-month window after a question was asked to one month, three months, and twelve months allowing shorter and longer times for replies and future behaviors. We also looked for any difference in the results over the years and although we found that the impact of getting a reply from a community user on the likelihood of asking another question in the future got slightly stronger over the years, the results were broadly consistent with our primary analysis.

We also explored whether or not results from different types of users are consistent, finding that they are. To do this, we took advantage of the fact that participation in online communities is highly uneven (Faraj and Johnson 2011, Dahlander and Frederiksen 2012, Rullani and Haefliger 2013, Jabr and Rahman 2018), with many users contributing just once or a few times and a handful of “super-users” contributing often. In the web-forum version of Stata-list, super-users can be readily identified by novice users, because the number of posts they have made is prominently displayed with their reply. We ran an additional analysis to see if receiving a reply from a super-user has an effect more similar to regular users or employees. We identified the top 15 users in our sample; in aggregate, they responded to 2,926 total questions (35.87%). Results show that receiving a reply from a super-user had qualitatively the same impact as receiving a reply from any user when considering both future questioning and future replying behavior.^{30 31}

Finally, we also examined if the tone of the reply (from community users or employees) had any impact on future behavior. One might anticipate that replies that were more positive would lead to higher levels of future participation than negative replies, but we found that the tone of the reply did not have

²⁹ Results available from the authors upon request (not shown due to space constraints).

³⁰ Results available from the authors upon request (not shown due to space constraints).

³¹ We structure our arguments in Hypothesis 4 based on perceived, rather than actual, response quality and source credibility and these results suggest that it may well be *perception* that is driving these effects. Specifically, if one assumes that super users provide responses of equal quality and are equally credible sources of information as employees, one would expect new members to react equivalently to super-users and employees; instead, we see new members reacting equivalently to users and super-users.

any significant impact on continued participation.³² This finding is consistent with prior research that has also shown the tone of replies does not influence continued participation in a different setting (Joyce and Kraut 2006).

DISCUSSION

Because online knowledge-sharing communities are dependent upon users' voluntary contributions to thrive (Jin et al. 2013, Rullani and Haefliger 2013), it is important to understand what drives new community members to remain involved in and contribute to communities. Building upon extant work that recognizes the importance of social interactions as a key driver of continued participation (Joyce and Kraut 2006, Zhang et al. 2013), this study digs deeper into the effects of receiving a reply on new members continued participation.

Examining continued participation by asking another question, we find that receiving a reply from another user versus not receiving a reply results in a 13% *increase* in participation, whereas receiving a reply from an employee results in a small, but insignificant increase. Examining continued participation by replying to questions, we find that receiving a reply from another user versus not receiving a reply results in an 83% *decrease*, while, conversely, receiving a reply from a Stata employee results in a 50% *increase*. Comparing across these values, rather than to the baseline of not receiving a reply at all, receiving a reply from a Stata employee versus another user leads to an 890% increase in the likelihood of responding to another users' question.

The Effects of Receiving Assistance on Future Contributions

This study contributes to the literature on online communities by examining how receiving assistance impacts a new member's propensity to engage in two social interactions important to the health of the community: asking subsequent questions and replying to questions. The existing literature explores the importance of receiving help on future postings in aggregate (Joyce and Kraut 2006, Zhang et al.

³² Results available from the authors upon request (not shown due to space constraints).

2013). However, asking questions and replying to questions are distinct behaviors. We show that receiving assistance has different effects on these behaviors.

Our findings show that receiving a reply from a user increases the likelihood of asking questions in the future and receiving a reply from an employee is directionally similar, but insignificant (possibly because of the sample size). This suggests that receiving a reply—regardless of from whom—creates satisfaction with the community as a help-seeking tool and leads to continued use of the community as such (contingent on the user having another question).

Our findings also show that receiving a reply from a user decreases the likelihood of replying to questions, while receiving a reply from an employee does the opposite, with both effects being statistically significant. These findings are important as they shed light on behavioral dynamics in online knowledge sharing communities: it appears that new participants take receiving help from other users and super-users in stride (or possibly even for granted), while reacting differently to help received from employees. This suggests that getting new members to contribute their knowledge to the community is a key challenge. Having employees respond to questions promotes continued participation, however, these individuals tend to be few and far between, representing just 2% of replies in our study and hence their overall impact is limited by their numbers. Overall, this paper represents an initial step in research seeking to identify interaction-based levers for increasing the number of members who respond to the questions of others. Future research might examine whether replies from elected leaders in communities—individuals who are also likely seen as high quality and credible sources of information—also make a positive difference.

Receiving Assistance from *Employees* Boosts Contribution

This study also contributes to the small, but growing literature on firms' engagement in online communities (Dahlander et al. 2008, Colombo et al. 2013, Altman et al. 2015, Nagle 2018, Shah and Nagle 2020). Prior research suggests that firms benefit from working within or sponsoring online communities (Gabriel and Goldman 2005, Jeppesen and Frederiksen 2006, Franke et al. 2008, Manchanda et al. 2015, Nagle 2018, Shah and Nagle 2020), but little work has examined the effects of

firms on communities. We begin to examine this question. Our results suggest that by having employees respond to users questions, firms can increase users' likelihood of responding to other users' questions (i.e., boost the likelihood that a user will contribute knowledge to the community). Our findings suggest that such participation does not harm user interactions in online knowledge sharing communities, and may provide a small lever for firms interested in seeding new communities or invigorating existing ones. However, our theorizing suggests that this benefit might hold only when receiving replies from employees is somewhat unexpected—whether interpreted as a gift or being observed; future experimental studies might alter the frequency of employee participation in the community to further investigate this effect. Interestingly, however, the overall impact of this effect is not nearly as large as the volume of questions that are answered by super-users, suggesting that users are at the heart of a healthy community and various individual motives need to be satisfied to maintain a healthy community. To the extent that, in order to function effectively, a community needs to attract and retain members who of their own accord answer questions (Wasko and Faraj 2005, Shah 2006, Butler et al. 2007, Hwang et al. 2015), future research might seek to examine how such individuals choose which communities to participate in as an important next step.

Reciprocity in Online Knowledge-Sharing Communities

We contribute a deeper understanding of how reciprocity functions in online knowledge-sharing contexts. Past studies have found mixed results regarding the effects of reciprocity on actual contributions, and strong results regarding the effects of a belief in the norm of reciprocity on *intended* contributions. Examining actual contributions, we find that receiving a reply from another user *decreases* the likelihood of a new member replying to someone else's question. This suggests that contextual issues stemming from the decision to not monitor or dig into the listserve/forum to search for an answer if one receives a reply (and particularly a fast one; the average time to reply in our sample is approximately 28 hours) may outweigh espoused norms in guiding replying behavior. We leave it to future research to examine whether this is due to new members' eagerness to apply their knowledge and move on with their projects, a choice to not abide by community norms, an inability to find questions to which they feel

comfortable responding, or a combination. Regardless of the mechanism, we, like Wang et al. (2012) who study an online community health page, see little evidence of replying behavior being motivated by reciprocity. Future research might examine other social-interaction based-levers for cultivating replying behavior (*e.g.*, the quality of FAQs, navigability of the site, overall community culture on dimensions such as psychological safety, friendliness/aggression, etc.).

Limitations & Generalizability

As with most studies, our study has limitations that leave open avenues for future research. For example, we are unable to comment on the percentage of employee replies to questions that would be optimal in encouraging continued contribution in a community (the percentage in our study is 2%). Further, whether the mechanism for the differing effect on future replying depends on whether a reply is from a user (strongly negative effect) or an employee (strongly positive effect) is the result of gift exchange or the feeling of being watched cannot be fully teased apart given our data; future research might survey users immediately after they receive a reply to understand the mechanisms at play. In addition, our study is, in effect, a case study of a single community. Future research is needed to see whether our results and insights are generalizable to other knowledge-sharing communities, particularly those not affiliated with a corporate product or those focused on the development (rather than improvement) of an artifact. That said, Statalist has structural similarities to other online knowledge-sharing communities—for example, it is question and answer based, it is focused on a particular product or service offered by an organization, and most questions are answered by users—that help improve the generalizability and external validity of our findings to other communities. Finally, although the use of observational data does not allow us to claim a fully causal relationship, we run multiple analyses in an effort to provide evidence that such a relationship is plausible.

CONCLUSION

Knowledge is increasingly being built and disseminated through online knowledge-sharing communities. We contribute to understanding how interactions between community members of various types shape subsequent contributions and interactions, showing both that receiving a reply has different

effects on a new member's subsequent asking and replying behaviors, and that the role of the community member who provides help—employee or user—also matters. These findings are relevant to community members, firms, and policy-makers interested in creating vibrant, impactful communities.

REFERENCES

- Aaltonen A, Seiler S (2016) Cumulative growth in user-generated content production: Evidence from Wikipedia. *Management Science* 62(7):2054–2069.
- Altman E, Nagle F, Tushman M (2015) Innovating without information constraints: Organizations, communities, and innovation when information costs approach zero. Hitt MA, Shalley C, Zhou J, eds. *Oxford Handbook of Creativity, Innovation, and Entrepreneurship* (Oxford University Press, Oxford, UK).
- Bagozzi RP, Dholakia UM (2006) Open source software user communities: A study of participation in Linux user groups. *Management Science* 52(7):1099–1115.
- Bauer J, Franke N, Tuertscher P (2016) Intellectual property norms in online communities: How user-organized intellectual property regulation supports innovation. *Information Systems Research* 27(4):724–750.
- Bayus BL (2013) Crowdsourcing new product ideas over time: An analysis of the Dell IdeaStorm community. *Management Science* 59(1):226–244.
- Bearman P (1997) Generalized exchange. *American Journal of Sociology* 102(5):1383–1415.
- Blalock AB, Blalock HM (1982) *Introduction to Social Research*. (Prentice-Hall, Englewood Cliffs, NJ).
- Bloom N, Schankerman M, Van Reenan J (2013). Identifying technology spillovers and product market rivalry. *Econometrica* 81(4):1347–1393.
- Bock GW, Zmud RW, Kim YG, Lee JN (2005) Behavioral intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological forces, and organizational climate. *MIS Quarterly* 29(1):87–111.
- Butler BS, Sproull L, Kiesler S, Kraut RE (2007) Community effort in online groups: Who does the work and why? Weisband S, eds. *Leadership at a Distance: Research in Technologically-Supported Work* (Lawrence Erlbaum, Mahwah, NJ), 171–194.
- Chen IY (2007). The factors influencing members' continuance intentions in professional virtual communities—a longitudinal study. *Journal of Information Science* 33(4):451–467.
- Cheung CM, Lee MK (2007) What drives members to continue sharing knowledge in a virtual professional community? The role of knowledge self-efficacy and satisfaction. Zhang Z, Siekmann J, eds. *Knowledge Science, Engineering and Management. KSEM 2007. Lecture Notes in Computer Science*. (Springer, Berlin, DE), 472–484.
- Chiu CM, Hsu MH, Wang ET (2006) Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision Support Systems* 42(3):1872–1888.
- Colombo MG, Piva E, Rossi-Lamastra C (2013) Authorising employees to collaborative with communities during working hours: When is it valuable for firms? *Long Range Planning* 46(3):236–257.
- Dahlander L, Frederiksen L (2012) The core and cosmopolitans: A relational view of innovation in user communities. *Organization Science* 23(4):988–1007.
- Dahlander L, Frederiksen L, Rullani F (2008) Online communities and open innovation: Governance and symbolic value creation. *Industry and Innovation* 15(2):115–123.
- Dahlander L, Jeppesen LB, Piezunka H (2019) How organizations manage crowds: Define, broadcast, attract, and select. Sydow J, Berends H, eds. *Managing Inter-organizational Collaborations: Process Views (Research in the Sociology of Organizations Volume 64)* (Emerald Publishing Limited, Bingley, UK), 239–270.
- Dewan S, Ho Y-J, Ramaprasad J (2017) Popularity or proximity: Characterizing the nature of social influence in an online music community. *Information Systems Research* 28(1):117–136.
- Dholakia UM, Bagozzi RP, Pearo LK (2004) A social influence model of consumer participation in network-and small-group-based virtual communities. *International Journal of Research in Marketing* 21(3):241–263.

- Di Gangi PM, Wasko M (2009) Steal my idea! Organizational adoption of user innovations from a user innovation community: A case study of Dell IdeaStorm. *Decision Support Systems* 48(1):303–312.
- Faraj S, Johnson SL (2011) Network exchange patterns in online communities. *Organization Science* 22(6):1464–1480.
- Fehr E, Gächter S (2000) Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives* 14(3):159–181.
- Festinger LE, Katz DE (1953) *Research Methods in the Behavioral Sciences* (Dryden Press, New York).
- Franke N, Keinz P, Schreier M (2008) Complementing mass customization toolkits with user communities: How peer input improves customer self-design. *Journal of Product Innovation Management* 25(6):546–559.
- Gabriel RP, Goldman R (2005) *Innovation Happens Elsewhere: Open Source as Business Strategy* (Elsevier/Morgan Kaufman, San Francisco, CA).
- Gallus J (2017) Fostering public good contributions with symbolic awards: A large-scale natural field experiment at Wikipedia. *Management Science* 63(12):3999–4015.
- Gächter S, von Krogh G, Haefliger S (2010) Initiating private-collective innovation: The fragility of knowledge sharing. *Research Policy* 39(7):893–906.
- Giordani PE, Rullani F, Zirulia L (2018) Endogenous growth of open collaborative innovation communities: A supply-side perspective. *Industrial and Corporate Change* 27(4):745–762.
- Gouldner AW (1960) The norm of reciprocity: A preliminary statement, *American Sociological Review* 25(2):161–178.
- Greiner B, Levati MV (2005) Indirect reciprocity in cyclical networks: An experimental study. *Journal of Economic Psychology* 26(5):711–731.
- Gu B, Jarvenpaa S (2003). Online discussion boards for technical support: the effect of token recognition on customer contributions. *ICIS 2003 Proceedings* 10:110–120.
- Hwang EH, Singh PV, Argote L (2015) Knowledge sharing in online communities: Learning to cross geographic and hierarchical boundaries. *Organization Science* 26(6):1593–1611.
- Iacus SM, King G, Porro G (2012) Causal inference without balance checking: Coarsened exact matching. *Political Analysis* 20(1):1–24.
- Jeppesen LB, Frederiksen L (2006) Why do users contribute to firm-hosted user communities? The case of computer-controlled music instruments. *Organization Science* 17(1):45–63.
- Jin XL, Cheung CM, Lee MK, Chen HP (2009) How to keep members using the information in a computer-supported social network. *Computers in Human Behavior* 25(5):1172–1181.
- Jin XL, Lee MK, Cheung CM (2010) Predicting continuance in online communities: model development and empirical test. *Behaviour & Information Technology* 29(4):383–394.
- Jin XL, Zhou Z, Lee MK, Cheung CM (2013) Why users keep answering questions in online question answering communities: A theoretical and empirical investigation. *International Journal of Information Management* 33(1):93–104.
- Jabr W, Rahman MS (2018) Can 'top reviews' save online review systems? Evidence from empirical analyses and a quasi-natural field experiment on Amazon. Working Paper. <https://ssrn.com/abstract=3200803> or <http://dx.doi.org/10.2139/ssrn.3200803>.
- Joyce E, Kraut RE (2006) Predicting continued participation in newsgroups. *Journal of Computer-Mediated Communication* 11(3):723–747.
- Kankanhalli A, Tan BC, Wei KK (2005) Contributing knowledge to electronic knowledge repositories: an empirical investigation. *MIS Quarterly* 29(1):113–143.
- Khadjavi M (2017) Indirect reciprocity and charitable giving: Evidence from a field experiment. *Management Science* 63(11):3708–3717.
- Kim AJ (2000) *Community Building on the Web* (Peachpit Press, Berkeley, CA).
- Kuk G (2006) Strategic interaction and knowledge sharing in the KDE developer mailing list. *Management Science* 52(7):1031–1042.

- Lee GK, Cole RE (2003) From a firm-based to a community-based model of knowledge creation: The case of the Linux kernel development. *Organization Science* 14(6):633–649.
- Manchanda P, Packard G, Pattabhiramaiah A (2015) Social dollars: The economic impact of customer participation in a firm-sponsored online customer community. *Marketing Science* 34(3):367–387.
- Mauss M (2002) *The Gift: Forms and Functions of Exchange in Archaic Societies*. (Routledge, London, UK).
- McCarney R, Warner J, Iliffe S, Van Haselen R, Griffin M, Fisher P (2007). The Hawthorne Effect: A randomised, controlled trial. *BMC Medical Research Methodology* 7:30.
- Muchnik L, Aral S, Taylor SJ (2013) Social influence bias: A randomized experiment. *Science* 341(6146):647–651.
- Nagle F (2018) Learning by contributing: Gaining competitive advantage through contribution to crowdsourced public goods. *Organization Science* 29(4):569–587.
- Nowak MA, Sigmund K (1998) Evolution of indirect reciprocity by image scoring. *Nature* 393(6685):573–577.
- Oliver RL (1980) A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research* 17(4):460–469.
- Porter CE, Donthu N (2008) Cultivating trust and harvesting value in virtual communities. *Management Science* 54(1):113–128.
- Ren Y, Harper FM, Drenner S, Terveen L, Kiesler S, Riedl J, Kraut RE (2012) Building member attachment in online communities: Applying theories of group identity and interpersonal bonds. *MIS Quarterly* 36(3):841–864.
- Ridings C, Gefen D (2004) Virtual community attraction: Why people hang out online. *Journal of Computer-Mediated Communication* 10(1): JCMC10110, <https://doi.org/10.1111/j.1083-6101.2004.tb00229.x>.
- Rullani F, Haefliger S (2013) The periphery on stage: The intra-organizational dynamics in online communities of creation. *Research Policy* 42(4):941–953.
- Seinen I, Schram A (2006) Social status and group norms: Indirect reciprocity in a repeated helping experiment. *European Economic Review* 50(3):581–602.
- Shah SK (2006) Motivation, governance, and the viability of hybrid forms in open source software development. *Management Science* 52(7):1000–1014.
- Shah SK, Nagle F (2020) Why do user communities matter for strategy? *Strategic Management Review* 1(2):305–353.
- Shaikh M, Levina N (2019) Selecting an open innovation community as an alliance partner: Looking for healthy communities and ecosystems. *Research Policy* 48(8):103766.
- von Hippel E, von Krogh G (2003) Open source software and the “private-collective” innovation model: Issues for organization science. *Organization Science* 14(2):209–223.
- Wang YC, Kraut R, Levine JM (2012) To stay or leave? The relationship of emotional and informational support to commitment in online health support groups. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* :833–842.
- Wasko MM, Faraj S (2000) “It is what one does”: Why people participate and help others in electronic communities of practice. *The Journal of Strategic Information Systems* 9(2-3):155–173.
- Wasko MM, Faraj S (2005) Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly* 29(1):35–57.
- Wooldridge JM. (2010). *Econometric Analysis of Cross Section and Panel Data (2nd Edition)* (The MIT Press, Cambridge, MA).
- Worthen B (2008) Why most online communities fail. *Wall Street Journal*. (July 16), <http://blogs.wsj.com/biztech/2008/07/16/why-most-online-communities-fail/>.
- Yamagishi T, Cook KS (1993) Generalized exchange and social dilemmas. *Social Psychology Quarterly* 56(4):235–248.

Zhang C, Hahn J, De P (2013) Continued participation in online innovation communities: Does community response matter equally for everyone? *Information Systems Research* 24(4):1112–1130.

Table 1. Summary Statistics

Variable	Obs.	Mean	SD	Min	Max
<i>Future Questioning (Binary)</i>	8,157	0.310	0.462	0	1
<i>Future Questioning (Count)</i>	8,157	0.860	2.415	0	74
<i>Future Replying (Binary)</i>	8,157	0.065	0.246	0	1
<i>Future Replying (Count)</i>	8,157	0.316	5.377	0	306
<i>Received Reply</i>	8,157	0.782	0.413	0	1
<i>Received User Reply</i>	8,157	0.775	0.417	0	1
<i>Received Employee Reply</i>	8,157	0.021	0.144	0	1
<i>Word Counts</i>	8,157	176.840	145.920	14	4,198
<i>Analytic</i>	8,157	78.545	15.563	3.55	99
<i>Tone</i>	8,157	53.728	25.893	1	99
<i>Six Letter Words (%)</i>	8,157	19.431	5.911	0	42.61
<i>Dictionary Words (%)</i>	8,157	65.754	11.066	4.19	89.51
<i>Elapsed Time (min)</i>	6,382	1,701	13,969	0	248,080

Note: *Elapsed Time* is calculated based on 6,382 observed posts that received a reply (from either a user or employee).

Table 2. Correlations

	1	2	3	4	5	6	7	8	9	10	11	12
1 <i>Future Questioning (Binary)</i>												
2 <i>Future Questioning (Count)</i>	0.532											
3 <i>Future Replying (Binary)</i>	0.202	0.245										
4 <i>Future Replying (Count)</i>	0.065	0.176	0.224									
5 <i>Received Reply</i>	0.061	0.050	-0.077	-0.006								
6 <i>Received User Reply</i>	0.060	0.049	-0.087	-0.015	0.979							
7 <i>Received Employee Reply</i>	-0.007	-0.013	0.037	0.058	0.078	-0.045						
8 <i>Word Count</i>	0.012	-0.004	0.017	-0.001	-0.062	0.011	-0.064					
9 <i>Analytic</i>	-0.026	-0.023	0.001	-0.004	-0.083	0.003	-0.082	0.155				
10 <i>Tone</i>	0.009	0.016	-0.003	0.003	0.025	-0.035	0.031	-0.185	0.005			
11 <i>Six Letters</i>	-0.069	-0.066	-0.008	-0.017	-0.136	0.013	-0.135	0.040	0.114	-0.057		
12 <i>Dictionary Words</i>	-0.008	0.010	-0.010	-0.001	0.068	-0.017	0.070	-0.231	-0.245	0.128	0.195	
13 <i>Elapsed Time (logged)</i>	-0.121	-0.110	-0.028	-0.025		0.046	-0.065	0.113	0.051	-0.059	0.168	-0.032

Note: Correlations with *Elapsed Time* are calculated for the 6,382 observations that received a reply, and thus the correlation between *Elapsed Time* and *Received Reply* is undefined.

Table 3. Predicting the Likelihood of Getting a Reply (Logit Models)

DV:	(1)	(2)	(3)
	<i>Received Reply</i>	<i>Received User Reply</i>	<i>Received Employee Reply</i>
<i>Year 2015</i>	-0.104 (0.072)	-0.093 (0.071)	-0.048 (0.176)
<i>Year 2016</i>	-0.087 (0.072)	-0.038 (0.071)	-0.941*** (0.227)
<i>Topic 1</i>	-0.068 (0.247)	-0.013 (0.244)	-1.514** (0.650)
<i>Topic 2</i>	2.082*** (0.260)	2.206*** (0.258)	-4.204*** (0.809)
<i>Topic 3</i>	-1.478*** (0.247)	-1.397*** (0.244)	-2.184*** (0.743)
<i>Topic 4</i>	0.906*** (0.254)	0.922*** (0.251)	-1.049* (0.602)
<i>Topic 5</i>	-0.309 (0.243)	-0.286 (0.240)	-0.758 (0.564)
<i>Topic 6</i>	0.793*** (0.248)	0.861*** (0.245)	-3.436*** (1.003)
<i>Topic 7</i>	0.290 (0.269)	0.397 (0.266)	-3.770*** (0.837)
<i>Topic 8</i>	0.906*** (0.272)	0.918*** (0.268)	-1.764** (0.715)
<i>Topic 9</i>	0.890*** (0.255)	0.729*** (0.249)	0.637 (0.495)
<i>Word Count</i>	-0.001*** (0.000)	-0.001*** (0.000)	0.001 (0.000)
<i>Analytic</i>	-0.013*** (0.002)	-0.013*** (0.002)	0.010* (0.005)
<i>Tone</i>	-0.000 (0.001)	0.000 (0.001)	-0.009*** (0.003)
<i>Six Letters</i>	-0.026*** (0.005)	-0.025*** (0.005)	0.018 (0.014)
<i>Dictionary Words</i>	0.002 (0.003)	0.002 (0.003)	0.008 (0.008)
<i>Constant</i>	2.543*** (0.332)	2.445*** (0.328)	-3.519*** (0.841)
Observations	8,157	8,157	8,157

Note: ***p<.01, **p<.05, *p<.1. All standard errors are heteroskedastic robust.

Table 4. Impact of Receiving a Reply on Future Participation (Logit Models)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
DV:	<i>Future Questioning</i>					<i>Future Replying</i>				
<i>Received Reply</i>	0.332*** (0.061)	0.289*** (0.062)	0.195*** (0.064)			-0.665*** (0.096)	-0.686*** (0.098)	-0.757*** (0.104)		
<i>Received User Reply</i>				0.184*** (0.063)					-0.815*** (0.102)	
<i>Received Employee Reply</i>				0.092 (0.176)					0.515** (0.239)	
<i>Elapsed Time (logged)</i>					-0.126*** (0.016)					-0.055 (0.034)
<i>Word Count</i>		0.000* (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)		0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
<i>Analytic</i>		-0.003 (0.002)	-0.005*** (0.002)	-0.005*** (0.002)	-0.003* (0.002)		-0.002 (0.003)	-0.000 (0.003)	-0.000 (0.003)	0.000 (0.004)
<i>Tone</i>		0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)		-0.000 (0.002)	-0.000 (0.002)	0.000 (0.002)	0.001 (0.002)
<i>Six Letters</i>		-0.022*** (0.004)	-0.009* (0.005)	-0.009* (0.005)	-0.009 (0.005)		-0.013 (0.008)	-0.000 (0.009)	-0.000 (0.009)	0.001 (0.011)
<i>Dictionary Words</i>		0.000 (0.002)	-0.004 (0.003)	-0.004 (0.003)	-0.004 (0.003)		0.001 (0.004)	-0.000 (0.005)	-0.000 (0.005)	0.001 (0.006)
<i>Constant</i>	-0.994*** (0.068)	-0.412* (0.235)	-0.323 (0.293)	-0.315 (0.292)	0.502 (0.331)	-2.028*** (0.104)	-1.713*** (0.437)	-2.616*** (0.561)	-2.611*** (0.560)	-2.990*** (0.564)
Topic Effects	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Year-Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,157	8,157	8,157	8,157	6,382	8,157	8,157	8,157	8,157	6,382

Note: ***p<.01, **p<.05, *p<.1. All standard errors are heteroskedastic robust. Columns 5 and 10 are a sub-sample analysis with the observations of new members who received a community reply (either a user or employee).

Table 5. Impact of Receiving a Reply on Future Participation (Linear Probability Models)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
DV:	<i>Future Questioning</i>					<i>Future Replying</i>				
<i>Received Reply</i>	0.068*** (0.012)	0.059*** (0.012)	0.038*** (0.012)			-0.046** (0.008)	-0.048*** (0.008)	-0.051*** (0.008)		
<i>Received User Reply</i>				0.036*** (0.012)					-0.056*** (0.008)	
<i>Received Employee Reply</i>				0.019 (0.036)					0.047* (0.025)	
<i>Elapsed Time (logged)</i>					-0.025*** (0.003)					-0.003* (0.002)
<i>Word Count</i>		0.000* (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)		0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
<i>Analytic</i>		-0.001 (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001* (0.000)		-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)
<i>Tone</i>		0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)		-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
<i>Six Letters</i>		-0.005*** (0.001)	-0.002* (0.001)	-0.002* (0.001)	-0.002 (0.001)		-0.001 (0.000)	-0.000 (0.001)	-0.000 (0.001)	0.000 (0.001)
<i>Dictionary Words</i>		-0.000 (0.001)	-0.001* (0.001)	-0.001* (0.001)	-0.001 (0.001)		0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)
<i>Constant</i>	0.272*** (0.014)	0.397*** (0.051)	0.415*** (0.061)	0.417*** (0.061)	0.585*** (0.061)	0.111*** (0.009)	0.130*** (0.028)	0.079** (0.031)	0.081*** (0.031)	0.049 (0.032)
Topic Effects	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Year-Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,157	8,157	8,157	8,157	6,382	8,157	8,157	8,157	8,157	6,382

Note: ***p<.01, **p<.05, *p<.1. All standard errors are heteroskedastic robust. Columns 5 and 10 are a sub-sample analysis with the observations of new members who received a community reply (either a user or employee).

Table 6. Impact of Receiving a Reply on Future Participation (Poisson Models for Count DVs)

	(1)	(2)	(3)	(4)	(5)	(6)
DV:	<i>Future Questioning</i>			<i>Future Replying</i>		
<i>Received Reply</i>	0.180* (0.103)			-0.909*** (0.268)		
<i>Received User Reply</i>		0.165* (0.099)			-1.195** (0.241)	
<i>Received Employee Reply</i>		-0.088 (0.173)			0.994** (0.486)	
<i>Elapsed Time (logged)</i>			-0.138*** (0.018)			-0.100 (0.083)
<i>Constant</i>	-0.055 (0.325)	-0.037 (0.324)	0.908*** (0.346)	-3.664 (1.318)	-3.599 (1.320)	-4.836*** (1.371)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Topic Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year-Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8,157	8,157	6,382	8,154	8,154	6,379

Note: ***p<.01, **p<.05, *p<.1. All standard errors are heteroskedastic robust. Three extreme outliers are excluded for the analyses of Column 4, 5, and 6.

Table 7. Impact of Receiving a Reply on Future Participation (IV Regression)

	(1)	(2)	(3)
DV:	<i>Received Employee Reply</i>	<i>Future Questioning</i>	<i>Future Replying</i>
<i>Posted in Working Hours</i>	0.217*** (0.069)		
<i>Received Employee Reply</i>		0.692 (1.571)	1.433** (0.601)
<i>Constant</i>	-1.991*** (0.366)	-0.122 (0.179)	-1.847*** (0.258)
Controls	Yes	Yes	Yes
Topic Effects	Yes	Yes	Yes
Year-Fixed Effects	Yes	Yes	Yes
Wald Chi2	Yes	Yes	Yes
Observations	8,157	8,157	8,157

Note: ***p<.01, **p<.05, *p<.1. All standard errors are heteroskedastic robust.

Table 8. Impact of Receiving a Reply on Future Participation (Robustness Checks)

	(1)	(2)	(3)	(4)	(5)	(6)
DV:	<i>Future Questioning</i>					
<i>Received User Reply</i>	0.203*** (0.053)	0.523*** (0.070)	-0.536** (0.272)	0.122** (0.055)		
<i>Received Employee Reply</i>	0.001 (0.155)	0.516** (0.246)	1.852** (0.766)	0.021 (0.139)	0.124 (0.213)	0.011 (0.036)
<i>Constant</i>	-0.240 (0.211)	-0.641** (0.293)	0.611 (1.509)		0.446*** (0.061)	0.494*** (0.082)
Model	Logit	Logit	Logit	Cox-Hazard	Heckman	OLS (CEM)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Topic Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year-Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,866	8,157	301	8,157	8,157	5,387

Note: ***p<.01, **p<.05, *p<.1. All standard errors are heteroskedastic robust.

	(7)	(8)	(9)	(10)	(11)	(12)
DV:	<i>Future Replying</i>					
<i>Received User Reply</i>	-0.744*** (0.082)	-0.144 (0.116)	-1.222*** (0.285)	-0.804*** (0.098)		
<i>Received Employee Reply</i>	0.539*** (0.202)	0.937*** (0.299)	0.379 (0.684)	0.463** (0.220)	0.346*** (0.107)	0.069** (0.025)
<i>Constant</i>	-1.526*** (0.306)	-3.148*** (0.563)	-0.240 (1.645)		0.016 (0.033)	0.017 (0.037)
Model	Logit	Logit	Logit	Cox-Hazard	Heckman	OLS (CEM)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Topic Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year-Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18,886	8,157	301	8,157	8,157	5,387

Note: ***p<.01, **p<.05, *p<.1. All standard errors are heteroskedastic robust.

Figure 1. An Example of a Question and Reply in the Statalist Forum

The screenshot displays the Statalist forum interface. At the top, the 'STATALIST The Stata Forum' logo is visible. Below the logo, there are navigation tabs for 'Forums' and 'FAQ', and a search bar. A yellow banner at the top contains an attention message: 'ATTENTION: Users of Google Chrome version 83 may experience a bug when posting to the forum. Please see this page for more details.' Below this, there is a '+ SUBSCRIBE' button. The main content area is titled 'POSTS' and 'LATEST ACTIVITY'. A 'Post Reply' button is visible. The forum post is titled 'Overall p-value in independent categorical variable in the imputation model' and is dated '27 Jul 2015, 09:41'. The post is by a user with a profile picture and a join date of 'Jan 2015'. The post content includes a greeting, a question about finding an overall p-value in a multiple imputation model, and a list of Stata commands: `xi: mi estimate, or: logistic sah1uia0 ageatdiagnosis aneurysm_size aneurysm_loc multipleA HTN antiHTN heartdisease PVD DM preCH preStroke hyperchol statinpre antiplatelet HRT OCP postmeno L.premrs Smoker3 I.Drinker fshahuaia druguse` and `testparm _IDrinker_1 _IDrinker_2`. The post also includes a request for help and a thank you. The reply is dated '29 Jul 2015, 11:24' and is by a user with the Stata logo and 'StataCorp Employee' title. The reply explains that the `-test-` and `-testparm-` commands are not supported by `-mi estimate-` and provides the appropriate commands: `-mi test-` and `-mi testtransform-`. It also provides a code block:

```
webuse mhouses1993s30,clear
mi estimate: regress price tax sqft i.nfeatures
mi test 1.nfeatures 2.nfeatures 3.features
```

 The final reply is dated '31 Jul 2015, 08:00' and is a thank you message from the original poster.

Figure 2. Summary of Hypotheses & Findings

		Effect on New Member's Subsequent Contributions to the Community	
		<i>Asks Additional Questions</i>	<i>Replies to Others' Questions</i>
Who Responds to a New Member's Initial Question? (baseline = no reply)	<i>User</i>	H1: Positive ✓	H2a: Positive ✗ H2b: Negative ✓
	<i>Employee</i>	H3: Positive (not significant)	H4: Positive ✓