

The Effect of Bots on Human Interaction in Online Communities

Hani Safadi John Lalor, and Nicholas Berente
Accepted in MIS Quarterly

Abstract: We investigate how bots influence human-to-human interaction in online communities. In doing so, we distinguish between reflexive and supervisory bots delegated by community participants and moderators, respectively. We hypothesize that reflexive bot activity will reduce direct reciprocity and increase generalized reciprocity, and that supervisory bot activity will reduce preferential attachment among human participants. An analysis of almost 70 million posts on the discussion communities on Reddit, a popular platform for online discussions, supports the hypotheses.

Keywords: online communities, bots, agentic artifacts, network exchange, reflexive bots, supervisory bots, direct reciprocity, generalized reciprocity, preferential attachment, panel vector autoregression.

The Effect of Bots on Human Interaction in Online Communities

1 INTRODUCTION

There is broad scholarly interest in understanding the factors that affect the health and functioning of online communities (Faraj et al., 2011; e.g., Johnson et al., 2015; Kane et al., 2014; Kraut & Resnick, 2011; Oh et al., 2016). Yet, existing research has typically examined online communities where participation was driven by human users. This is not necessarily the case today, as agentic information systems (IS) such as “bots” are becoming ubiquitous and often interact with humans as fellow participants in online communities (Ferrara et al., 2016; Salge et al., 2022).

Bots (software robots) are software programs that autonomously perform tasks on behalf of community participants. Participants in online communities deploy bots for a variety of reasons – from frivolous entertainment bots to bots that help moderate participation in communities (Halfaker & Taraborelli, 2015). Bots post content on social media (Salge et al., 2022; Salge & Karahanna, 2018), organize software development (Hukal et al., 2019), and assist with editorial work (Halfaker and Riedl 2012; Geiger 2014). The proliferation of bots is impacting social interaction in online communities, and it is becoming important to understand how. Thus, in this research note, we ask: *how do bots influence human social interaction in online communities?*

To answer this question, we distinguish among types of bots and dimensions of human social interaction. Although there are various agentic IS artifacts (Baird & Maruping, 2021), bots in online communities primarily take two formsⁱ – reflexive or supervisory. Reflexive bots are those triggered by certain stimuli to automatically perform an action (e.g., Salge & Karahanna, 2018). Supervisory bots are endowed with decision rights to monitor a domain and influence or control activity in that domain (e.g., Hukal et al., 2019). Further, to conceptualize dimensions of human social interaction, we conceive of online communities in terms of network exchanges comprised of direct reciprocity, generalized reciprocity, and preferential attachment (Faraj & Johnson, 2011). We hypothesize that

reflexive bots diffuse human attention broadly, resulting in less direct reciprocity among community participants but more generalized reciprocity across the community overall. Further, supervisory bots substitute for negotiated decision-making with community moderators, resulting in reduced preferential attachment to human members of the community.

We test our hypotheses with panel data from Reddit, an online platform of topical communities (“subreddits”). Reddit’s subreddits are online communities focused on open discussion and the free sharing of ideas (Faraj et al., 2016; Sproull & Arriaga, 2012). Such discussion communities are computer-mediated environments that are minimally structured and social in nature – they are not primarily task-focused. Since bots are increasingly prevalent in discussion communities (Ferrara et al., 2016), this context is well-suited to explore how bots shape human social interaction.

Using panel vector autoregression, we find support for our hypotheses. These findings contribute to the nascent and growing scholarship on the impact of bots in online communities (Geiger, 2014; Hukal et al., 2019; Neff & Nagy, 2016; Salge et al., 2022; Salge & Karahanna, 2018). The remainder of the research note is organized as follows. First, we distinguish between two forms of bots and elicit how they differ from human participants in online communities. Then, we draw on network exchange theory to construct our hypotheses. We describe our data collection and results, then conclude by discussing the implications of this research.

2 BOTS IN ONLINE COMMUNITIES

Bots are agentic IS artifacts that act autonomously but in ways that are delegated by others. There are a variety of bots, including conversational agents (chatbots) that help to market products (Thomaz et al., 2020), decision-making bots that manage organizational activity (Kellogg et al., 2020), and online bots that disseminate information (Salge et al., 2022). Bots represent a new domain of study that departs from other information technologies by virtue of their autonomy – their activity is not necessarily “bracketed” by human interaction in the way that previous forms of

technology are (Berente et al., 2021). Indeed, bots and other autonomous algorithms act as independent agents in new ways as they manage, supervise, or otherwise influence human activity (Berente et al., 2021; Kellogg et al., 2020; Möhlmann et al., 2021).

Online communities are spaces for social interaction (Faraj et al., 2016). Recent decades have witnessed a proliferation of such communities (Ellison & boyd, 2013; Johnson et al., 2015), which can involve extensive interaction among human participants (Kraut & Resnick, 2011; Sproull & Arriaga, 2012). Participants often develop deep social ties (Preece & Maloney-Krichmar, 2003; Ren et al., 2012), self-organize around common causes (Nan & Lu, 2015; Vaast et al., 2017), and collaborate on innovation (Baldwin & von Hippel, 2011; Benkler, 2006).

Prior studies of online communities typically focus on humans, but now bots are playing a larger role. Understanding the implications of bots on human activity is thus an important frontier for research. Bots are instances of agentic IS artifacts “that have the ability to perceive and act, such as take on specific rights for task execution and responsibilities for preferred outcomes” (Baird & Maruping, 2021, p. 317). Bots satisfy the main properties of agentic IS artifacts: situatedness, autonomy, and flexibility (Jennings et al., 1998). Bots are *situated* in that they are located in specific digital environments. Bots are *autonomous* because they act without human supervision. Bots are *flexible* because their actions vary in response to environmental stimuli. These three properties are continuous rather than categorical. Recent advancements in machine learning make creating more situated, autonomous, and flexible bots possible. The capacity to learn and adapt allows bots the potential to be proactive and act opportunistically. The expanding nature of what bots and, more generally, agentic IS can do calls for a theoretical grounding of these capabilities.

Drawing on current generations of bots can help to establish a theoretical foundation for studying all sorts of bots in the future. Baird and Maruping (2021) offer a taxonomy of agentic IS based on increasing decision-making latitude. The four categories are reflexive, supervisory,

anticipatory, and prescriptive. Within this framework, we isolate more primitive (reflexive and supervisory) bots with an established track record (Table 1), leaving aside more advanced bots (anticipatory and prescriptive) for future work.

Table 1: Two Types of Bots in Online Communities

| Bot Type | Delegated By | Definition | References |
|------------------|--------------|--|---|
| Reflexive bots | Participants | Bots that respond to human activity by posting as community participants | (Baird & Maruping, 2021; Geiger, 2016; Salge et al., 2022; Salge & Karahanna, 2018) |
| Supervisory bots | Moderators | Bots that make decisions to manage and coordinate human activity | (Baird & Maruping, 2021; Geiger, 2014; Halfaker et al., 2013; Hukal et al., 2019) |

Reflexive bots act based on triggering events, performing specific, often repetitive tasks, but do not make decisions for the community. Reflexive bots are typically comprised of conditional statements (i.e., “if x then y”). They monitor an environment and act autonomously in response to triggers. As delegates of community participants (Geiger, 2016), they can act as participants in their own right (Seering et al., 2020) to respond to triggers to execute predefined activities. They can be used to amplify a participant’s influence in the community, such as when social media users employ bots to relay messages (Salge et al., 2022; Salge & Karahanna, 2018). Consider the Reddit bot WikiTextBot in Figure 1 (left) as an example. When a user posts a Wikipedia link in a comment, WikiTextBot automatically pulls the summary of the Wikipedia page and posts it as a reply.

Supervisory bots, on the other hand, automate the moderation of communities – they are decision-making delegates. Typically, supervisory bots act on behalf of platform moderators and administrators who seek to manage activity – “housekeeping” for the community (Hukal et al., 2019). Like reflexive bots, supervisory bots involve conditional statements and act in response to triggers. However, supervisory bots have broader decision-making authority to evaluate interactions and ensure compliance with norms and rules (Geiger, 2014). Tasks can include scanning contributions, applying quality standards (Halfaker & Taraborelli, 2015), flagging and removing low-quality content, and organizing content (Halfaker & Riedl, 2012). For example, Reddit’s AutoModerator bot in Figure 1 (right) can delete posts that do not conform to editorial standards.

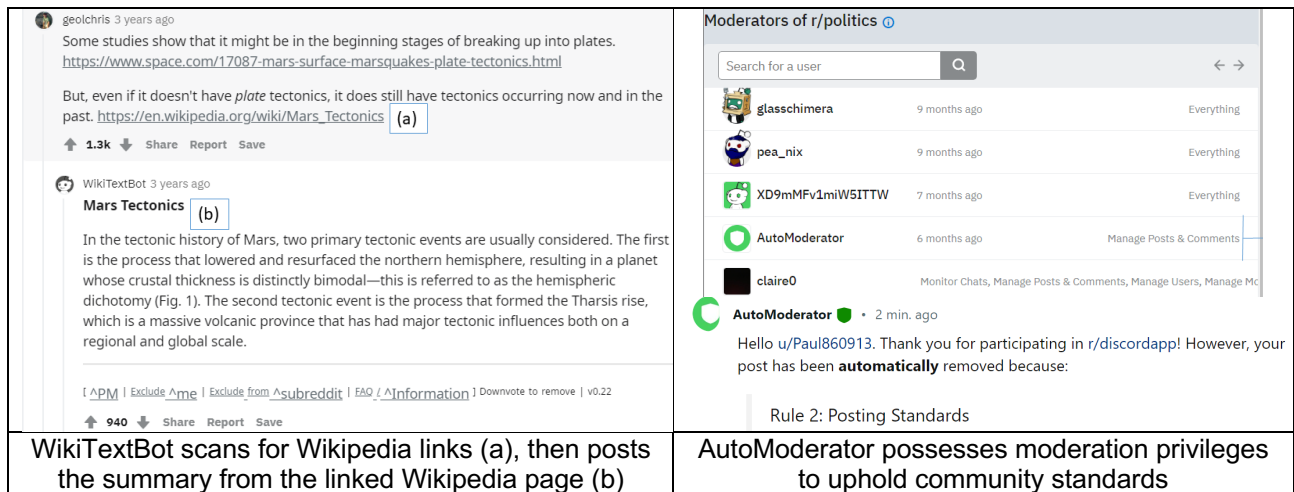


Figure 1: Examples of a reflexive bot (left) and a supervisory bot(right) in Reddit

2.1 Bots and Human Social Interaction in Online Communities

Humans exhibit distinct patterns of interaction with each other in online communities (e.g., Faraj & Johnson, 2011), and when bots participate in these environments, they influence these patterns of interaction. It is important to distinguish human-to-bot interaction from human-to-human interaction (i.e. “social” interaction). Existing research has begun to focus on human-to-bot interaction and points out that bots can influence humans in a variety of different ways. For example, human-to-bot interaction can elicit both negative and positive human reactions, based on how they are perceived (Clément & Guitton, 2015). Bots are responsible for a significant portion of posts on social media platforms, and as a result, bots can be influential (Ferrara et al., 2016; Salge & Karahanna, 2018), and their actions can increase human engagement with the content (Delkhosh et al., 2023). Bots can spread misinformation, and humans can be deceived as a result of interacting with bots (Lazer et al., 2018). Overall, prior work shows that bots can influence humans in various ways and highlights that further research is needed to fully understand the dynamics of this relationship (Salge et al., 2022). Through their interactions with humans, bots can contribute to entertaining, informing, or managing humans in a community, and their resulting behaviors will undoubtedly impact the communities. But what is the result of bots in an online community on human-to-human interactions? In other words, how does human interaction with bots spill over to social interaction with other humans?

To begin exploring how human interaction with bots can influence patterns of human social interaction, it is important to distinguish how bots behave in communities from how humans behave. Bots are delegates of humans, but they do not behave in communities precisely as their human delegators do. Three capabilities that distinguish between bots and humans are awareness, computational, and interfacing capabilities (Baird & Maruping, 2021; Jennings et al., 1998).

Awareness refers to the ability to constantly attend to available stimuli and cues across an environment. Bots can be always awake and pay attention to every post in an online community, humans cannot because their attention is limited and episodic (Simon, 1956). *Computational* capabilities of bots allow them to quickly and accurately process information across a greater breadth – the attention of bots has a wide reach (Salge & Karahanna, 2018). On the other hand, humans tend to be limited in their scope and speed of processing. *Interfacing* capabilities of bots rely on routinized processes for automated responses that allow for limited interactivity (Salge et al., 2022).ⁱⁱ Bots are well-suited for routine, superficial activities but lack the human rationality and emotion that integrate diverse stimuli for deep, meaningful interaction. Hence, while bots can manage broader simultaneous interactions, they lack the capacity for deeper engagement. This also affects decision-making; bot decisions are more procedural and automated, whereas human decisions are up for interpretation and negotiation within communities (see Table 2).

Table 2: Behavioral Differences between Humans and Bots

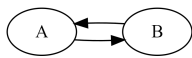
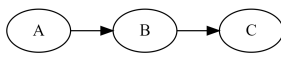
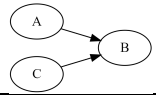
| Behavior | Human participant | Reflexive bot | Human moderator | Supervisory bot |
|--|-------------------|---------------|-----------------|-----------------|
| Attention (awareness capabilities) | Episodic | Constant | Episodic | Constant |
| Reach (computational capabilities) | Local | Broad | Local | Broad |
| Interactivity (interfacing capabilities) | Deeper | Superficial | Negotiated | Procedural |

In summary, although bots are delegates of human participants, the behavior of the bots in achieving these goals differs from the behavior of human participants. Bot behaviors trigger responses from humans, and outcomes of these interactions can spill over to alter patterns of human-to-human social interaction (Neff & Nagy, 2016; Traeger et al., 2020).

2.2 Hypotheses: Bots and Human Social Interaction in Online Communities

Research often conceives of online communities in terms of social networks (Safadi et al., 2021; Wasko et al., 2004; Yan et al., 2016) where interactions among participants form network patterns (Faraj & Johnson, 2011; Surma, 2016). Table 3 shows three such patterns: (1) direct reciprocity ($A \rightarrow B \rightarrow A$), (2) generalized reciprocity ($A \rightarrow B \rightarrow C$), and (3) preferential attachment ($A \rightarrow B \leftarrow C$) (Faraj & Johnson, 2011). We draw on how reflexive and supervisory bots differ from human participants and moderators, to consider how bots may influence network exchange patterns.

Table 3: Key Network Exchange Patterns

| Construct | Direct Reciprocity | Generalized Reciprocity | Preferential Attachment |
|---------------------|---|--|---|
| Illustration |  |  |  |
| Definition | An exchange pattern of participants responding directly to others interacting with them | An exchange pattern of indirect communication that supports generalized exchange | An exchange pattern that reflects a concentration of interaction |
| References | (Faraj & Johnson, 2011; Mashima & Takahashi, 2008; Nowak, 2006; Surma, 2016) | (Faraj & Johnson, 2011; Mashima & Takahashi, 2008; Nowak & Sigmund, 2005; Safadi et al., 2021) | (Barabási & Albert, 1999; Capocci et al., 2006; Faraj & Johnson, 2011; Johnson et al., 2014; Lu et al., 2013) |

2.2.1 Reflexive Bots and Direct Reciprocity

Exchange follows a set of norms, foremost among which is reciprocity (Cropanzano & Mitchell, 2005). Expectations of reciprocity, or mutual interchange, are predicated on fairness in interaction, which forms the basis for sociality (Mashima & Takahashi, 2008; Nowak, 2006). *Direct reciprocity* undergirds social exchange in online communities (Faraj & Johnson, 2011; Surma, 2016) and drives participation and sharing in online communities (Yan et al., 2016). Direct reciprocity is premised on the expectation that a human's cooperation with another will be returned (Nowak, 2006). Direct reciprocity is dyadic – a participant will act by responding to or providing resources in return for or anticipation of another participant's actions (Faraj & Johnson, 2011). Many participants contribute to online communities expecting others to reciprocate and return their help, and there is extensive evidence of direct reciprocity in online communities (Kathan et al., 2015; Wang et al., 2015).

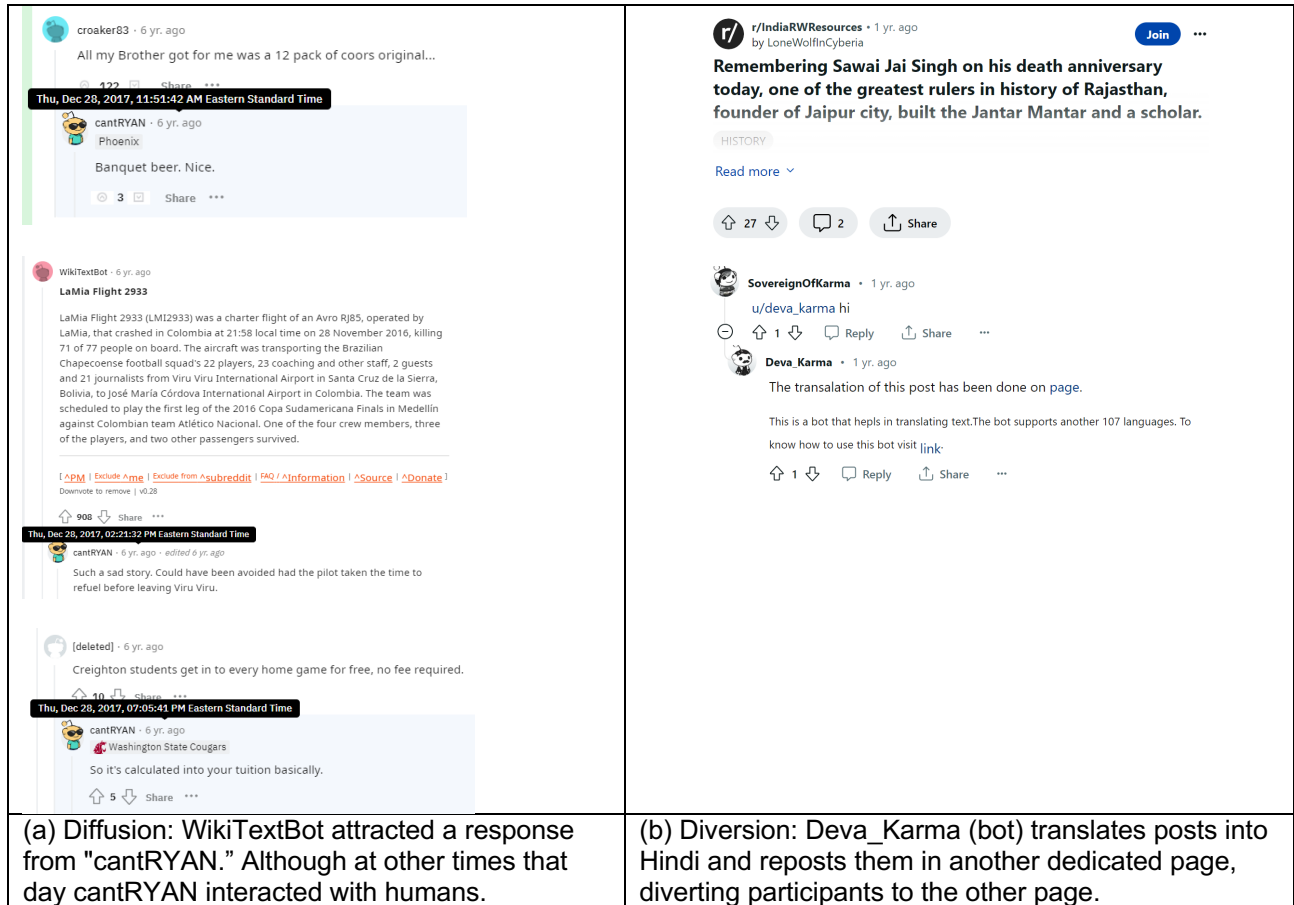


Figure 2: Reflexive bots and direct reciprocity: examples of diffusion and diversion

Reflexive bots are programmed by humans for a variety of purposes (Long et al., 2017). Due to their constant attention and broad reach, reflexive bots can diffuse the episodic attention of human participants across broader areas and divert human attention away from particular interactions. Figure 2 provides examples of these mechanisms of diffusion and diversion: (a) diffusion involves presenting stimuli to humans that spread their attention across a greater number of domains; (b) diversion describes how bots can guide human attention away from particular discussions toward new ones. As bots broadly participate in many different discussions, they introduce contributions that would not have been otherwise included. Humans, therefore, have more information presented to them. Human attention can be diffused as a result, because these bot activities may pull human attention across topics within a discussion, and may divert this attention to different discussions. Because human attention is scarce and episodic, attention to information presented by bots is likely

to be at the expense of deep interaction among dyadic sets of human participants. Reflexive bots can be seen as a particular type of super-participants who drive an increasing share of discussions (Johnson et al., 2014) but do so at the expense of deep, reciprocal dyadic discussions. As a result, human participants will be drawn to more discussions with more participants but less reciprocity in existing discussions. Therefore, we hypothesize:

H1: Reflexive bot activity is associated with less direct reciprocity between humans in an online community.

2.2.2 Reflexive Bots and Generalized Reciprocity

Reciprocity, when not directed, involves interaction with the community overall. Participants expect that others will act in kind and that their contributions will be rewarded by the community (Levine & Prietula, 2014; Surma, 2016). This indirect reciprocity results in more generalized social exchange across the community as a whole (Mashima & Takahashi, 2008). When someone helps another person, that other person will be inclined to help a third person, or someone else will be inclined to help the first person, in what is often referred to as downstream and upstream indirect reciprocity, respectively (Mashima & Takahashi, 2008; Nowak & Sigmund, 2005). Indirect or generalized reciprocity is a common feature of online communities. When participants find value in the community, they tend to contribute to the community in order to encourage engagement and strengthen the community (Constant et al., 1996; Cross & Sproull, 2004; Wasko et al., 2004). Hobby, leisure, and social support communities are largely comprised of altruistic generalized reciprocity (Goh et al., 2016). Even production communities, which are goal-oriented and focused on producing artifacts (e.g., articles, software), inevitably involve generalized reciprocity. In such communities, expert participants often provide help without the expectation of dyadic reciprocation (Constant et al., 1996; Wasko & Faraj, 2005).


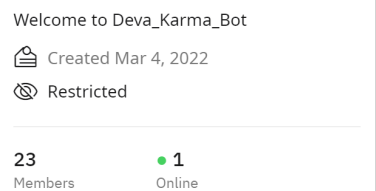
| | |
|---|--|
|  <p>tootlymegapoopoo · 6 yr. ago</p> <p>The aircraft was transporting the Brazilian Chapecoense football squad's 22 players, 23 coaching and other staff, 2 guests and 21 journalists from Viru Viru</p> <p>Is it odd that there were almost as many journalists as players? Forgive my ignorance of S.American football, but does this make sense?</p> <p>2 ↑ ↓ Share ...</p> <p>[deleted] · 6 yr. ago</p> <p>It was the cup final, Chapecoense were the underdogs and caught a lot of attention (Atlético Nacional is a powerhouse of South American football), so there were many journalists from all the continent traveling to cover the match. This includes the Brazilian Fox Sports staff, who were going to do the official broadcasting.</p> <p>The flight was private, so it seemed like a good idea sharing the costs between the club and the press.</p> <p>6 ↑ ↓ Share ...</p> <p>gunsof · 6 yr. ago</p> <p>It was their first time playing in the final of a tournament that big, they'd recently come up from a lower league. It was a huge deal, they were kind of like the Leicester City of Brazil.</p> <p>1 ↑ ↓ Share ...</p> |  <p>Welcome to Deva_Karma_Bot</p> <p>Created Mar 4, 2022</p> <p>Restricted</p> <p>23 Members</p> <p>1 Online</p> |
| <p>(a) Within discussion: WikiTextBot's contribution (Figure 2a above) sparked further discussion about new topics not mentioned in the original post.</p> | <p>(b) Across discussions: Posts translated by Deva_Karma (Figure 2b) attracted 22 human participants to the other subreddit.</p> |

Figure 3: Reflexive bots and generalized reciprocity: within and across discussions

Although reflexive bots will diffuse human attention resulting in less dyadic interactions with other humans, they will also drive human interaction through their reach across the entire community. Their contributions often provide new information that can elicit reactions from humans that were not previously interacting with each other. As such, although reflexive bots diffuse attention, they also provide occasions for general interaction. Further, by diverting attention to new areas, this can attract attention generally to those areas. Figure 3 provides two illustrations of how reflexive bots can attract interactions through two mechanisms – contributing content to existing discussions and starting new discussions: (a) within a discussion, bots can introduce new topics that capture the attention of a wider audience; (b) across discussions, bots' participation can attract human participants who may not have previously engaged in discussions. Reflexive bots amplify diverse perspectives across the community in a way humans do not. Through their broad reach, reflexive bot activity can inject variety into the discourse of the community, introduce different perspectives more broadly, and generate serendipitous occasions for human interaction.

Essentially, this activity provides occasions for broad social exchange among humans. Because reflexive bots operate continuously, they present opportunities for interaction between humans who otherwise would not have connected. These discussions can lead to a greater engagement with the community, which can lead to more interaction, thus strengthening the community (Constant et al., 1996; Wasko et al., 2004) – resulting in greater generalized reciprocity. Therefore, we hypothesize:

H2: Reflexive bot activity is associated with more generalized reciprocity among humans in an online community.

2.2.3 Supervisory Bots and Preferential Attachment

Participants with a central role in the community – typically resulting from a history of both direct and generalized reciprocity with others – enjoy what is referred to as preferential attachment (Capocci et al., 2006; Newman, 2001). Preferential attachment refers to the tendency of a few nodes in the network to generate a large number of interactions (Johnson et al., 2014), since the probability by which a new node attaches to an existing node depends on the existing node's current connections (Barabási & Albert, 1999). In online communities, there is a subset of humans who drive a large share of the interactions (Krishnamurthy, 2002; Maillart et al., 2008) and typically have a degree of popularity as a result, which means new members of the community will interact with them – thus exhibiting preferential attachment (Capocci et al., 2006; Lu et al., 2013).

Preferential attachment is often reflected in a concentration of communication in the community (Faraj & Johnson, 2011, p. 1467). Both informal and formal leaders hold significant influence over the community. This is especially true for moderators, who play a crucial role in resolving conflicts and disputes (Kane et al., 2014) – members often turn to leaders for support and assistance (Johnson et al., 2015). Therefore, we anticipate a preferential attachment towards community leaders.

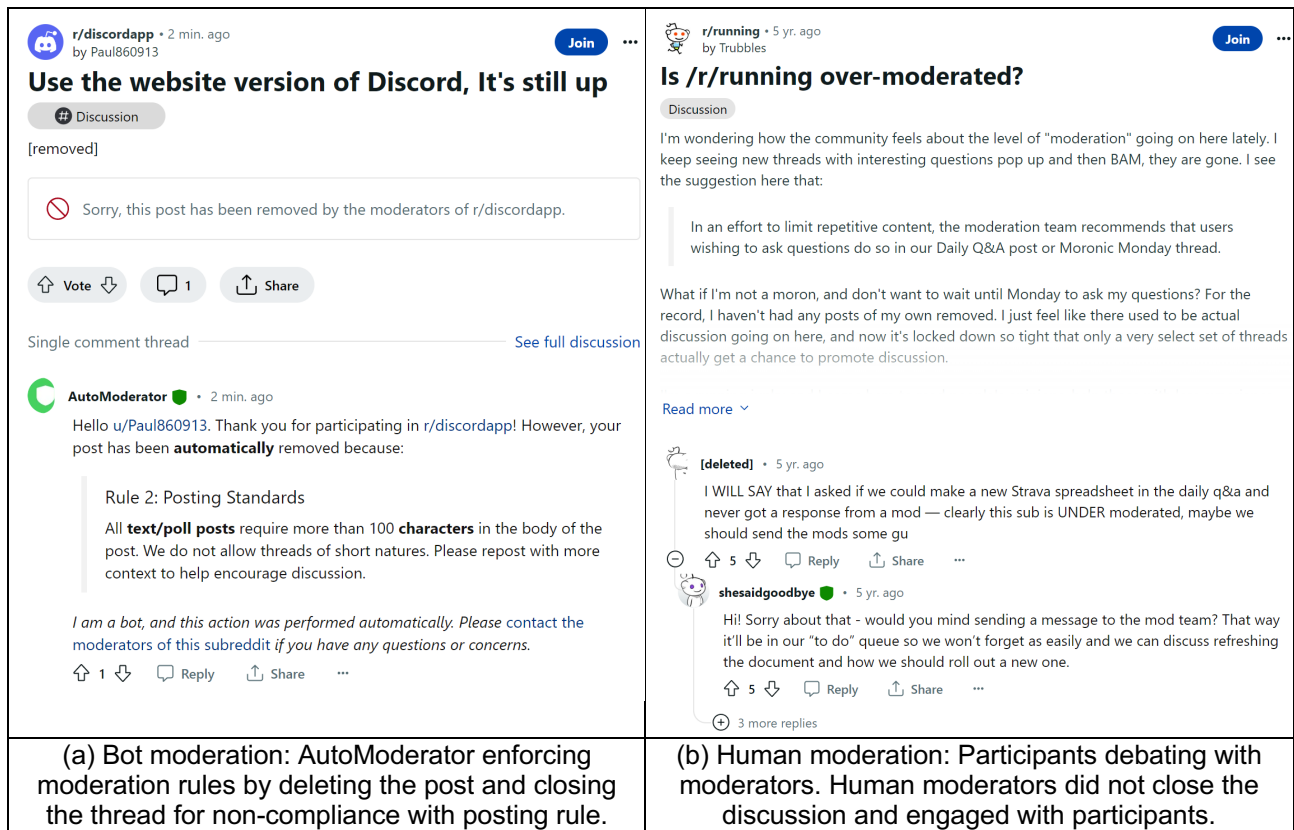


Figure 4: Supervisory bots and preferential attachment

Many tasks delegated to supervisory bots are traditionally handled by community leaders and moderators who take a central role in the community. However, because supervisory bots typically do not allow for much interactivity – they are executing procedures typically in the form of conditional statements – they do not allow direct dispute. A human would always be subject to inquiry and appeal, but supervisory bots generally are not. Figure 4 illustrates this contrast. In (a), with bot moderation, the AutoModerator bot made a decision to remove and close a discussion for violating the community's norms. There was no opportunity for further exchange. However, in (b), with a human moderator, participants engaged in a lively debate about norms, exchanging views with human moderators and other participants. Thus, bot supervision is simply procedural, whereas human moderation is often open to questioning and bargaining – it is negotiated. Supervisory bots encode and enforce moderation (David & Rullani, 2008; Geiger, 2014). Procedural decisions reduce the need for humans to coordinate with each other and negotiate each other's decisions (Hukal et al.,

2019). Humans simply do not interact with supervisory bots like they interact with human leaders. As a result, supervisory bots will reduce human-to-human preferential attachment because there will be less reason to interact with community leaders. This leads to the following hypothesis:

H3: Supervisory bot activity is associated with decreased preferential attachment among humans in an online community.

3 METHODS

To test our hypotheses, we studied Reddit, an online community platform that allows web users to share and discuss a wide range of topics. Reddit hosts some of the most diverse and active online communities in the English-speaking world. Participation is organized in communities devoted to specific topics (e.g., politics, technology, sports) referred to as subreddits. Participants, or “Redditors,” post links, texts, or images to be commented on and discussed by other participants. Redditors provide feedback to each other with upvotes or downvotes.

Data on Reddit activity is collected by the PushShift projectⁱⁱⁱ and posts and comments from 2005 through 2019 are publicly available on Google BigQuery. However, bot activity on Reddit is relatively low compared to human activity. We designed a sampling strategy to ensure that we identified subreddits with enough bot activity to analyze the effects of bots over time. We designed inclusion criteria at the subreddit level as follows: we selected subreddits with total comment activity (between 2005 and 2019) between 16,000 and 250,000 comments and a total proportion of bot activity among those comments between 3% and 30%. This leaves us with subreddits that are generally active and have a small to moderate amount of bot activity. With this strategy, 1074 subreddits fit our inclusion criteria. We examined each subreddit every month and organized the data in a panel. The final sample includes 69,336 subreddit months of activity (almost 70 million posts) from April 2007 to December 2019, with a median of 66 months per subreddit.

3.1 Bot Operationalization

Reddit has two types of bots: *user* and *moderator* bots, which correspond to the concepts of reflexive and supervisory bots, respectively. User bots are created by participants. User bots post and comment just as human users do, typically in response to specific keywords or phrases, and have the privileges associated with other human participants. There are many user bots, each coded to perform a different action. For example, WikiTextBot listens for any posted Wikipedia links and replies with a summary of the linked article (Figure 1), and Deva_Karma waits for any user mentions in a post to translate it into Hindi (Figure 2).^{iv}

Reddit participants can design user bots to automatically interact with users on Reddit through the Reddit API.^v Bots have usernames just as human users do on Reddit and can access all Reddit activity. Because user bots do not need to be explicitly identified as such on Reddit, we devised a way to identify them. We relied on a convention among Reddit participants to reply to posts from bots with the phrase “good bot” or “bad bot” if the participant approves or disapproves of the bot’s action, respectively. We used an online ranking of bots using this information to identify the best bots based on the number of times users commented “good bot” or “bad bot” on the bot’s activities.^{vi} The online ranking score is the lower bound of the Wilson score (Wilson, 1927) confidence interval of the proportion of “good bot” votes.^{vii} This feature gave us a metric for identifying the most popular bots. We used this list as our list of bots. We focused on the top 2,000 ranked bots according to this metric. Beyond this, we found that entries were either very inactive with very few posts or were false positives (humans that had received a “good/bad bot” reply to one of their posts).

Moderator bots are delegated by community moderators to scan new posts and comments for flagged content, identify and remove spammers, etc. Reddit offers a class of moderator bots under the AutoModerator functionality. In any subreddit, there can be only up to one AutoModerator

instance. This instance of AutoModerator is customized and deployed at the subreddit level by human subreddit moderators based on how they would like activity to be monitored on their subreddit.^{viii} By granting the bot access to subreddit moderator privileges, the bot can enforce predefined subreddit rules automatically (Jhaver et al., 2019). AutoModerator’s status for each subreddit is available publicly, so we were able to identify if AutoModerator was active for each subreddit and if so, the month it was activated, as well as view all of its comments. In our data, 947 of the 1074 communities had AutoModerator activated during the observation period.

User and moderator bots are delegated by community participants and moderators, respectively. Both types of bots are triggered by human activity, with moderator bots having more decision-making latitude beyond merely posting or commenting like user bots. As such, we consider user and moderator bots to be instances of reflexive and supervisory bots within Reddit.



Figure 5: Between January 2014 (left) and June 2018 (right), bot activity on the Football subreddit increased by a factor of ten. The human user network graphs illustrate a change in how human users interacted in the subreddit associated with increased bot activity

3.2 Network Exchange Operationalization

Conversations in Reddit are structured as nested threaded discussions. Each thread has a starting post, and then participants can comment on the post or comment on others’ comments. In this work, we focus on the effects of bots on human interaction, and thus, we model interaction as a directed network where an edge connecting two participants represents a reply from the first participant to the second (see Figure 5 for an example). We are interested in the influence of bots on human-to-

human interaction, and therefore, we do not include bots in our networks. We only include edges representing replying comments, where both nodes are human participants.

Network Exchange Patterns:^{ix} The reciprocity score is a well-established social network measure for quantifying dyadic reciprocal exchange. It measures the ratio of bidirectional links to the total number of links in a directed network (Wasserman & Faust, 1994, p. 124). A larger value indicates that a given directed edge in the network is more likely to have an associated edge pointing in the opposite direction. In our case, higher reciprocity indicates that if participant A has replied to one of participant B's posts/comments, then it is more likely that participant B has also replied to one of participant A's posts/comments. To operationalize generalized reciprocity, we use the density of the directed network calculated as the proportion of connected participant pairs over all possible dyads in the network (Wasserman & Faust, 1994, Chapter 5). A higher-density network is indicative of more generalized exchange with many interactions among participants.

Preferential attachment increases the incoming and outgoing interactions and thus correlates with node centrality measures.^x Thus, we measure preferential attachment using Freeman's (1979) measure of network centralization as the sum of differences between the in-degree centrality of the most central node and all other nodes, divided by the same sum calculated on a star graph with the same number of nodes. This measurement provides the extent to which the observed network conforms to the ideal type of a centralized network (Borgatti & Everett, 2006) in which few nodes enjoy most ties as expected when preferential attachment is prevalent (Faraj & Johnson, 2011).

Bot Activity: We operationalize bot activity by considering the propensity of their activity in the community. Activity is measured as the ratio of bot comments to total comments in the subreddit-month. We calculate this ratio for user bots and the moderator bots to operationalize user bot activity and moderator bot activity, respectively.

Table 4: Theoretical Constructs and their Operationalization

| Construct | Operationalization |
|-----------------------------|---|
| | Outcomes |
| (1) Direct reciprocity | Reciprocity score of the human-human interaction network of the subreddit-month |
| (2) Generalized reciprocity | Density of the human-human interaction network of the subreddit-month |
| (3) Preferential attachment | In-degree centralization of the human-human interaction network of the subreddit-month |
| Predictors | |
| (4) User bot activity | The number of user bot comments divided over the total number of comments of both bots and humans in the subreddit-month |
| (5) Moderator bot activity | The number of moderator bot comments divided over the total number of comments of both bots and humans in the subreddit-month |
| Controls | |
| (6) Community activity | Number of Comments in the subreddit-month |
| (7) Community moderation | The number of human Moderators in the subreddit-month |
| (8) Community membership | The number of human participants in the subreddit-month |

Table 5: Descriptive Statistics and Correlation Matrix, 71,173 observations in 1,074 groups
Pearson correlations in the bottom triangle, Spearman correlations in the top triangle, * p<0.05

| | Mean | Std. Dev. | Minimum | p25 | Median | p75 | Maximum |
|-----------------------------|--------|-----------|---------|-------|--------|-------|---------|
| (1) Direct reciprocity | .42 | .16 | 0 | .33 | .43 | .52 | 1 |
| (2) Generalized reciprocity | .042 | .11 | 0 | .0049 | .011 | .028 | 1.50 |
| (3) Preferential attachment | .13 | .16 | 0 | .053 | .089 | .15 | 2 |
| (4) User bot activity | .015 | .041 | 0 | 0 | .0013 | .0095 | .98 |
| (5) Moderator bot activity | .046 | .088 | 0 | 0 | 0 | .056 | .97 |
| (6) Community activity | 1143.9 | 2179.5 | 1 | 191 | 533 | 1300 | 100547 |
| (7) Community moderation | 3.07 | 8.78 | 0 | 1 | 2 | 4 | 363 |
| (8) Community membership | 227.8 | 366.9 | 1 | 44 | 118 | 273 | 16389 |

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|-----|--------|--------|--------|--------|--------|-------|-------|-------|
| (1) | 1 | .14* | .26* | -.062* | -.022* | .1* | .042* | .024* |
| (2) | .019* | 1 | .45* | -.33* | -.28* | -.84* | -.29* | -.95* |
| (3) | -.081* | .53* | 1 | -.25* | -.18* | -.17* | -.18* | -.32* |
| (4) | -.05* | -.019* | -.031* | 1 | .18* | .35* | .3* | .33* |
| (5) | -.017* | -.062* | -.065* | -.068* | 1 | .34* | .3* | .28* |
| (6) | .061* | -.15* | -.059* | .049* | .081* | 1 | .31* | .93* |
| (7) | .025* | -.018* | -.018* | .096* | .057* | .053* | 1 | .32* |
| (8) | -.003 | -.2* | -.16* | .017* | .03* | .79* | .068* | 1 |

Controls: Controls include the overall number of comments in the subreddit/month, the total number of human moderators, and the number of human participants to measure community

activity, community moderation, and community membership, respectively. Table 4 shows constructs and measures, and Table 5 presents descriptive statistics and correlations. All variables are per subreddit-month.

4 RESULTS

Our main estimation approach is the panel vector autoregression (PVAR) model (Abrigo & Love, 2016; Holtz-Eakin et al., 1988), which allows us to examine the relationship between bot activity and network exchange. Additionally, the PVAR model can establish Granger causality. We adopt the reduced form of the PVAR model where the dependent variable is endogenous and, hence a linear function of its lagged value, the lagged independent variables of network characteristics, and the lagged control variables. We used a lag of 1 based on lag-order selection statistics for lags 1 to 4. The overall coefficient of determination increases with more lags, but the increases were minimal beyond lag 1 (0.1% to 0.8%). We also account for unobserved heterogeneity of individual communities and temporal trends by introducing the individual-specific effect and time-fixed effects. We employ robust standard errors to mitigate heteroskedasticity. The model is as follows; results are summarized in Table 6.

$$DV_{s,m} = \beta_1 DV_{s,m-1} + \beta_2 \text{Community activity}_{s,m-1} + \beta_3 \text{Community moderation}_{s,m-1} + \beta_4 \text{Community membership}_{s,m-1} + \beta_5 \text{User bot activity}_{s,m-1} + \beta_6 \text{Moderator bot activity}_{s,m-1} + \alpha_s + u_{sm},$$

(where s and m refer to subreddit and month)

We found support for H1 ($\beta_5 = -.679^{***}$). The reflexive activity of user bots is associated with less human direct reciprocity. Because bot activity is measured with the percentage of posting by bots and network reciprocity varies from 0 to 1, this relationship can be interpreted as a decrease of 0.679% in direct reciprocity associated with a 1% increase in user bots' activity. H2 is also supported. User bots' activity led to increased generalized reciprocity ($\beta_5 = 0.071^*$). Because density in a directed network ranges from 0 to 2, the relationship can be interpreted as an increase of 3.55% associated with an increase of 1% in user bots' activity. Finally, supporting H3, supervisory

bots were negatively associated with preferential attachment given decreased network in-degree centralization ($\beta_6 = -.153^{***}$). Because centralization varies from 0 to 2, this relationship can be interpreted as a decrease of 7.65% in centralization associated with a 1% increase in moderator bot activity.

Table 6: PVAR results (Robust standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, * $p < 0.001$)**

| DV _{s,m} = | Direct reciprocity _{s,m} | | Generalized reciprocity _{s,m} | | Preferential attachment _{s,m} | |
|---|-----------------------------------|--------|--|--------|--|--------|
| β_1 DV _{s,m-1} | .097*** | (.027) | .643*** | (.032) | .234*** | (.029) |
| β_2 Community activity _{s,m-1} | .000* | (.000) | -.000 | (.000) | .000* | (.000) |
| β_3 Community moderation _{s,m-1} | -.000 | (.001) | .000* | (.000) | .000 | (.000) |
| β_4 Community membership _{s,m-1} | -.000 | (.000) | -.000 | (.000) | -.000** | (.000) |
| β_5 User bot activity _{s,m-1} | -.679** | (.258) | .071* | (.034) | -.080 | (.110) |
| β_6 Moderator bot activity _{s,m-1} | -.064 | (.035) | -.003 | (.006) | -.153*** | (.023) |
| Hansen's J χ^2 statistic | 431.0*** | | 446.3*** | | 478.3*** | |

4.1 Causal Interpretation

The PVAR model can establish a less strict form of causality referred to as Granger causality (Granger, 1988). In this form of causality, the focus is on whether one time-series variable can be used to predict the future outcomes of another time-series variable. Formally, we say that the time series X Granger-causes the time series Y if the past values of X can be used for the prediction of the future values of Y. Following the PVAR estimation, we perform pairwise Granger causality tests (Abrigo & Love, 2016) between the dependent variables (network exchange patterns) and the key independent variables of user bot activity and moderator bot activity. The results of these tests are detailed in Appendix B. However, they conform with the findings of PVAR (Table 6). Together, these findings suggest that we can interpret the results in a Granger-causal way. Past bot activity can predict a change in future human-human interaction.

4.2 Robustness Checks

We conducted a series of robustness checks to measure the impact of various operational decisions on our results (Appendix A). Our checks cover how we identify bots and how many bots we include, inclusion criteria for subreddits, and accounting for false negatives in our data (i.e., bots that are not identified as bots and mistakenly considered humans). Specifically, we account for

changing network dynamics (A.1.1), account for false negatives in bot identification (A.1.2 and A.1.3), exclude human moderators from the networks (A.1.4), and use a different source for bot identification (A.1.5). We then increase (A.1.6) and decrease (A.1.6) the bot detection thresholds, and apply more strict inclusion criteria for bots (A.1.8 and A.1.9) We also exclude outlier communities (A.1.10) and employ alternative measurements for the dependent variables (A.2). The results are summarized in Table 7. A detailed discussion is available in Appendix A.

Table 7: Summary of robustness checks

| Analysis | Hypothesis 1 | Hypothesis 2 | Hypothesis 3 |
|--|------------------------------|----------------------------------|----------------------------------|
| Main | Negative, significant | Positive, significant | Negative, significant |
| <i>Robustness Tests in Appendix A</i> | | | |
| A.1.1 | Negative, significant | Positive, significant | Negative, significant |
| A.1.2 | Negative, significant | Positive, significant | Negative, significant |
| A.1.3 | Negative, significant | Positive, significant | Negative, significant |
| A.1.4 | Negative, significant | Positive, significant | Negative, significant |
| A.1.5 | Positive, significant | Negative, not significant | Negative, not significant |
| A.1.6 | Negative, significant | Positive, significant | Negative, significant |
| A.1.7 | Negative, significant | Positive, significant | Negative, not significant |
| A.1.8 | Negative, significant | Negative, not significant | Negative, significant |
| A.1.9 | Negative, significant | Positive, significant | Negative, significant |
| A.1.10 | Negative, significant | Positive, significant | Negative, significant |
| A.2 | Negative, significant | Positive, significant | Negative, significant |

After replacing our bot list with a different list (A.1.5), we discovered that our results were not statistically significant. We believe that this particular list is inadequate, particularly because it is no longer accessible on the webpage from which we obtained it. Despite this, we have included this check in our results for the sake of comprehensiveness. When we decrease the number of bots being examined from 2000 to 1000 (A.1.7), we discover that H3 is no longer supported, while H1 and H2 remain supported. This is likely due to the fact that the AutoModerator bot is ranked 1849 in our list of bots. Therefore, when we limit our analysis to 1000 bots, AutoModerator is no longer considered a bot and is instead included in the human-human interaction network. The inclusion of this bot specifically impacts H3, as it is the primary supervisory bot on the Reddit platform (although its behavior can be customized for each subreddit that chooses to use it). By excluding AutoModerator from our bot analysis, we are also excluding the main mechanism for influencing preferential

attachment (H3). When we limit the number of subreddits being considered (A.1.8), we find that H2 is no longer supported. The changes in network density are not evident, and we attribute this to the fact that networks with higher bot activity will also experience a greater change in density. This assumption is validated when we loosen the inclusion criteria to a more moderate restriction in our main analysis (A.1.9).

5 DISCUSSION

Reddit was valued at \$10 billion in August 2021.^{xi} This valuation reflects the value of human participation and contribution to online communities (Baldwin & von Hippel, 2011; Faraj et al., 2016). Maintaining the health of online communities, retaining their members, and sustaining their contribution is therefore an important goal of practical and scholarly interest (Butler, 2001; Faraj et al., 2011; Kraut & Resnick, 2011). Yet, the landscape of online communities is changing. While the moderation of online content is extremely important for the functioning of online communities (Skousen et al., 2020), the growth of participation and the quantity of offensive and inappropriate content stresses the capacities of human moderators to function (Cheng et al., 2015). Many online communities deploy some form of bot to assist in their moderation (Halfaker & Taraborelli, 2015; Hukal et al., 2019), and many online communities enable user bots as well (Geiger, 2014, 2016). However, existing work does not address how this abundance of bot activity affects human participation and social interaction. In this research note, we take an initial step in this direction. This is particularly important given the rise of generative artificial intelligence instantiated through chatbots. In order to navigate the future where bots possess greater capabilities, it is essential to have a solid theory of their foundations. There are many different types of bots, and their capabilities are a moving target. While advancements in artificial intelligence are drawing more attention to bots, many mainstream bots are still largely rule-based.ⁱ Even in many situations that

involve artificially intelligent bots, their participation in online communities is triggered and enacted according to conditions, much like rule-based bots.

In our study, rule-based bots had considerable influence on human interaction. Reflexive bot activity decreases direct reciprocity, and increases generalized reciprocity, while supervisory bot activity decreases preferential attachment. This shows that the potential impact of including bots in online social environments can be significant even when they are relatively simple. Although various studies in IS and human-computer interaction examined bots, with few exceptions (e.g., Salge et al., 2022), not many studies theorized about bots. Our work contributes to this nascent domain to unpack how bots can influence human social interaction by drawing on network exchange theory.

First, bots are delegated by other humans to achieve their goals. However, in achieving these goals, bots participate and engage with human participants, and their participation differs from other humans because of their different attention, reach, and interactivity. These differing behaviors (to reach the same goals) can create unintended consequences on other human participants' behavior. This observation emphasizes the potential for interactive technology, designed with a specific goal in mind, to have unintended consequences due to its interactive nature. This contributes to the ongoing discourse on the unintended impacts of technology (Aanestad et al., 2021; Rahwan et al., 2019). Second, scholarly work in IS has examined the materiality and affordances of online communities (Faraj et al., 2011; Ferguson & Soekijad, 2016; Levina & Arriaga, 2014). This body of work examines how the design of the community platform affords and constrains the actions of its members. Our work contributes to extending this work to show that such affordances extend beyond the static element of community design (e.g., Vaast et al., 2017) to interactive, agentic components such as bots. Third, recent work examining the new wave of technologies such as AI and ML cautions against both the intended and the unintended consequences of the substitution of human

activity and human labor (Berente et al., 2021; Brynjolfsson et al., 2021). Our work shows that understanding such substitution extends beyond who is partaking in the activity. While substituting for human moderator activity, supervisory bots end up changing how other participants perceive such moderation. Such a finding suggests that substitution does not necessarily imply equivalence. This conclusion calls for examining the spillover effects of agentic IS beyond their intended activity (e.g., Lebovitz et al., 2021; Traeger et al., 2020).

5.1 Practical Contributions

During Elon Musk’s \$44 billion takeover attempt of Twitter in 2022, bots were one of the contentious issues; this example exemplifies the controversy and confusion around the role of bots.^{xii} Bots can be good or bad, and it often depends on the perspective. Our work provides practical contributions to stakeholders in online communities. For instance, community managers are interested in understanding the implications of enabling bots. If bots increase human-human interaction, then introducing bots can be of potential benefit, given that human interaction and contribution create value. However, our work cautions against such a simplistic conclusion, given that reflexive bots, while they increase generalized reciprocity, decrease direct reciprocity at the same time. This suggests a trade-off between promoting the breadth and depth of interaction through bots. Whereas promoting the breadth of interaction can be beneficial, it also risks marginalizing deep reciprocal interactions and, as a result, trivializing the community into a promotional space (Bulgurcu et al., 2018).

Further, given the growth of participation, the value of moderation work, and the scarcity of moderation resources, it is not surprising that many online communities are implementing supervisory bots (Halfaker & Riedl, 2012). Moderation tasks are extremely valuable (Skousen et al., 2020). For instance, it is estimated that in 2020, Reddit moderators worked for at least 466 daily hours. This is the equivalent of an estimated 3.4 million USD per year (Li et al., 2022). Our research

points out the impact of supervisory bots in weakening the centrality of human participants as well as potentially decreasing the depth of interaction in communities without the associated increase in breadth of interaction that we found in other reflexive bots. These results caution practitioners to implement widespread supervisory bots carefully.

5.2 Limitations and Future Work

In this study, we focused on the association between bots and human-to-human interaction. Without a prior theoretical foundation, the “detective work” for underlying mechanisms that can be later confirmed or refuted is valuable (Puranam, 2018, pp. 159–161). The intuition of our detective work in this research is reflected in emerging research that is being conducted on how humans interact with each other around autonomous agents (e.g., Traeger et al., 2020). Characteristics of the autonomous agent influence human-to-human social interaction in ways that are different from other technical elements of platform infrastructures.

In this work, we considered rule-based bots that perform specific roles coded by their developers. Today, many bots in operation are rule-based. They execute relatively simple instructions at scale to improve the user experience through efficiency and broad coverage. However, the next phase in automated bots is a phase of “learning bots.” Learning bots, powered by machine learning techniques and generative AI, will be data-driven bots and thus have further implications in exhibiting complex behavior and altering human behavior in more unexpected ways, like the example of Microsoft Tay demonstrates (Salge & Berente, 2017). For example, an “intelligent bot” could lead human participants to focus on topics that need critical thinking, stimulating the discussions among human participants and more reciprocity as a result. Future work includes testing the limit of our findings with rule-based bots as well as developing our understanding of how bots continue to alter and shape human interaction in online environments.

Second, beyond the three mainstream exchange patterns of direct reciprocity, generalized reciprocity, and preferential attachment, future should examine other mechanisms, such as triadic closure and local search (Jackson & Rogers, 2007; Kossinets & Watts, 2006). Third, the scope of this work can be expanded to provide a better understanding of the roles of the participants as they engage with the bots. Such understanding can help community managers save time using these bots without decreasing the overall experience of the participants. Finally, bots may have an impact on users' social interactions outside the focal subreddits where they encountered the bots. This spillover effect can be an interesting extension of this work.

5.3 Summary and Conclusion

In this research note, we took a step toward theoretically unpacking how the proliferation of bots in online communities can influence human-to-human social interaction. Specifically, and notwithstanding advances in machine learning, we focus on two types of bots with relatively limited decision-making latitude (Baird & Maruping, 2021), reflexive and supervisory bots. We theorize how these two types of bots are associated with changes in network exchange patterns of direct reciprocity, generalized reciprocity, and preferential attachment. Future work can draw on this as a basis for understanding the role of different forms of AI on human social interaction.

6 REFERENCES

- Aanestad, M., Kankanhalli, A., Maruping, L., Pang, M.-S., & Ram, S. (2021). Special Issue — Call for Papers — Digital Technologies and Social Justice. *MIS Quarterly*.
- Abrigo, M. R. M., & Love, I. (2016). Estimation of panel vector autoregression in Stata. *Stata Journal*, 16(3), 778–804. <https://doi.org/10.1177/1536867x1601600314>
- Baird, A., & Maruping, L. M. (2021). The Next Generation of Research on IS Use: A Theoretical Framework of Delegation to and from Agentic IS Artifacts. *MIS Quarterly*, 45(1), 315–341. <https://doi.org/10.25300/MISQ/2021/15882>
- Baldwin, C. Y., & von Hippel, E. (2011). Modeling a Paradigm Shift: From Producer Innovation to User and Open Collaborative Innovation. *Organization Science*, 22(6), 1399–1417. <https://doi.org/10.1287/orsc.1100.0618>
- Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509–512.
- Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. Yale University Press.
- Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing Artificial Intelligence. *MIS*

- Quarterly*, 45(3).
- Borgatti, S. P., & Everett, M. G. (2006). A Graph-theoretic perspective on centrality. *Social Networks*, 28(4), 466–484. <https://doi.org/10.1016/j.socnet.2005.11.005>
- Brynjolfsson, E., Wang, C., & Zhang, X. (2021). The economics of IT and digitization: Eight questions for research. *MIS Quarterly*, 45(1), 473–477.
- Bulgurcu, B., Van Osch, W., & Kane, G. C. (2018). The Rise of the Promoters: User Classes and Contribution Patterns in Enterprise Social Media. *Journal of Management Information Systems*, 35(2), 610–646. <https://doi.org/10.1080/07421222.2018.1451960>
- Butler, B. S. (2001). Membership Size, Communication Activity, and Sustainability: A Resource-Based Model of Online Social Structures. *Information Systems Research*, 12(4), 346–362. <https://doi.org/10.1287/isre.12.4.346.9703>
- Capocci, A., Servedio, V. D. P., Colaiori, F., Buriol, L. S., Donato, D., Leonardi, S., & Caldarelli, G. (2006). Preferential attachment in the growth of social networks: The internet encyclopedia Wikipedia. *Phys. Rev. E*, 74(3), 36116. <https://doi.org/10.1103/PhysRevE.74.036116>
- Cheng, J., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2015). Antisocial Behavior in Online Discussion Communities. *Proceedings of the Ninth International Conference on Web and Social Media*, 61–70. <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/view/10469>
- Clément, M., & Guittou, M. J. (2015). Interacting with bots online: Users' reactions to actions of automated programs in Wikipedia. *Computers in Human Behavior*, 50, 66–75. <https://doi.org/10.1016/j.chb.2015.03.078>
- Constant, D., Sproull, L., & Kiesler, S. (1996). The Kindness of Strangers: The Usefulness of Electronic Weak Ties for Technical Advice. *Organization Science*, 7(2), 119–135. <https://doi.org/10.1287/orsc.7.2.119>
- Cropanzano, R., & Mitchell, M. S. (2005). Social exchange theory: An interdisciplinary review. *Journal of Management*, 31(6), 874–900.
- Cross, R., & Sproull, L. (2004). More Than an Answer: Information Relationships for Actionable Knowledge. *Organization Science*, 15(4), 446–462. <http://www.jstor.org/stable/30034748>
- David, P. A., & Rullani, F. (2008). Dynamics of innovation in an open source collaboration environment: lurking, laboring, and launching FLOSS projects on SourceForge. *Industrial and Corporate Change*, 17(4), 647–710.
- Delkhosh, F., Gopal, R. D., Patterson, R. A., & Yaraghi, N. (2023). Impact of Bot Involvement in an Incentivized Blockchain-Based Online Social Media Platform. *Journal of Management Information Systems*, 40(3), 778–806. <https://doi.org/10.1080/07421222.2023.2229124>
- Ellison, N. B., & boyd, D. (2013). Sociality through Social Network Sites. In *The Oxford Handbook of Internet Studies* (pp. 151---172). <https://doi.org/10.1093/oxfordhb/9780199589074.001.0001>
- Faraj, S., Jarvenpaa, S. L., & Majchrzak, A. (2011). Knowledge Collaboration in Online Communities. *Organization Science*, 22(5), 1224–1239. <https://doi.org/10.1287/orsc.1100.0614>
- Faraj, S., & Johnson, S. L. (2011). Network Exchange Patterns in Online Communities. *Organization Science*, 22(6), 1464–1480. <https://doi.org/10.1287/orsc.1100.0600>
- Faraj, S., von Krogh, G., Monteiro, E., & Lakhani, K. R. (2016). Online Community as Space for Knowledge Flows. *Information Systems Research*, 7047, 1–17. <https://doi.org/10.1287/isre.2016.0682>
- Ferguson, J. E., & Soekijad, M. (2016). Multiple interests or unified voice? Online communities as intermediary spaces for development. *Journal of Information Technology*, 31(4), 358–381. <https://doi.org/10.1057/jit.2015.25>
- Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots.

- Communications of the ACM*, 59(7), 96–104. <https://doi.org/10.1145/2818717>
- Freeman, L. C. (1979). Centrality in social networks conceptual clarification. *Social Networks*, 1(3), 215–239.
- Geiger, R. S. (2014). Bots, bespoke, code and the materiality of software platforms. *Information Communication and Society*, 17(3), 342–356. <https://doi.org/10.1080/1369118X.2013.873069>
- Geiger, R. S. (2016). Bot-based collective blocklists in Twitter: the counterpublic moderation of harassment in a networked public space. *Information Communication and Society*, 19(6), 787–803. <https://doi.org/10.1080/1369118X.2016.1153700>
- Goh, J. M., Gao, G. (Gordon), & Agarwal, R. (2016). The Creation of Social Value: Can an Online Health Community Reduce Rural-Urban Health Disparities? *MIS Quarterly*, 40(1), 247–263. <https://doi.org/10.25300/MISQ/2016/40.1.11>
- Granger, C. W. J. (1988). Causality, cointegration, and control. *Journal of Economic Dynamics and Control*, 12(2), 551–559.
- Halfaker, A., Geiger, R. S., Morgan, J. T., & Riedl, J. (2013). The Rise and Decline of an Open Collaboration System. *American Behavioral Scientist*, 57(5), 664–688. <https://doi.org/10.1177/0002764212469365>
- Halfaker, A., & Riedl, J. (2012). Bots and cyborgs: Wikipedia’s immune system. *Computer*, 45(3), 79–82.
- Halfaker, A., & Taraborelli, D. (2015). *Artificial intelligence service “ORES” gives Wikipedians X-ray specs to see through bad edits*. Wikimedia Blog. <https://blog.wikimedia.org/2015/11/30/artificial-intelligence-x-ray-specs/>
- Holtz-Eakin, D., Newey, W., & Rosen, H. S. (1988). Estimating vector autoregressions with panel data. *Econometrica: Journal of the Econometric Society*, 1371–1395.
- Hukal, P., Berente, N., Germonprez, M., & Schechter, A. (2019). Bots Coordinating Work in Open Source Software Projects. *Computer*, 52(9), 52–60. <https://doi.org/10.1109/MC.2018.2885970>
- Jackson, M. O., & Rogers, B. W. (2007). Meeting Strangers and Friends of Friends: How Random Are Social Networks? *American Economic Review*, 97(3), 890–915. <https://doi.org/10.1257/aer.97.3.890>
- Jennings, N. R., Sycara, K., & Wooldridge, M. (1998). A roadmap of agent research and development. *Autonomous Agents and Multi-Agent Systems*, 1(1), 7–38.
- Jhaver, S., Birman, I., Gilbert, E., & Bruckman, A. (2019). Human-Machine Collaboration for Content Regulation. *ACM Transactions on Computer-Human Interaction*, 26(5), 1–35. <https://doi.org/10.1145/3338243>
- Johnson, S. L., Faraj, S., & Kudaravalli, S. (2014). Emergence of Power Laws in Online Communities: The Role of Social Mechanisms and Preferential Attachment. *MIS Quarterly*, 38(3), 795–808.
- Johnson, S. L., Safadi, H., & Faraj, S. (2015). The Emergence of Online Community Leadership. *Information Systems Research*, 26(1), 165–187. <https://doi.org/10.1287/isre.2014.0562>
- Kane, G. C., Johnson, J., & Majchrzak, A. (2014). Emergent Life Cycle: The Tension Between Knowledge Change and Knowledge Retention in Open Online Coproduction Communities. *Management Science*, 60(12), 3026–3048. <https://doi.org/10.1287/mnsc.2013.1855>
- Kathan, W., Hutter, K., Füller, J., & Hautz, J. (2015). Reciprocity vs. Free-Riding in Innovation Contest Communities. *Creativity and Innovation Management*, 24(3), 537–549. <https://doi.org/10.1111/caim.12107>
- Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, 14(1), 366–410. <https://doi.org/10.5465/annals.2018.0174>
- Kossinets, G., & Watts, D. J. (2006). Empirical analysis of an evolving social network. *Science*

- (New York, N.Y.), 311(5757), 88–90. <https://doi.org/10.1126/science.1116869>
- Kraut, R. E., & Resnick, P. (2011). Building successful online communities: Evidence-based social design. In *Building successful online communities: Evidence-based social design*. Mit Press.
- Krishnamurthy, S. (2002). Cave or community?: An empirical examination of 100 mature open source projects. *First Monday*, 7(6).
- Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, 359(6380), 1094–1096. <https://doi.org/10.1126/science.aao2998>
- Lebovitz, S., Levina, N., & Lifshitz-Assa, H. (2021). Is AI Ground Truth Really True? The Dangers of Training and Evaluating AI Tools Based on Experts' Know-What. *MIS Quarterly*, 45(3), 1501–1526. <https://doi.org/10.25300/MISQ/2021/16564>
- Levina, N., & Arriaga, M. (2014). Distinction and Status Production on User-Generated Content Platforms: Using Bourdieu's Theory of Cultural Production to Understand Social Dynamics in Online Fields. *Information Systems Research*, 25(3), 468–488. <https://doi.org/10.1287/isre.2014.0535>
- Levine, S. S., & Prietula, M. J. (2014). Performance Open Collaboration for Innovation : Principles and Performance. *Organization Science*, 25(5), 1414–1433. <https://doi.org/10.1287/orsc.2013.0872>
- Li, H., Hecht, B., & Chancellor, S. (2022). Measuring the Monetary Value of Online Volunteer Work. *ArXiv*. <http://arxiv.org/abs/2205.14528>
- Long, K., Vines, J., Sutton, S., Brooker, P., Feltwell, T., Kirman, B., Barnett, J., & Lawson, S. (2017). “Could You Define That in Bot Terms”? *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 3488–3500. <https://doi.org/10.1145/3025453.3025830>
- Lu, Y., Jerath, K., & Singh, P. V. (2013). The Emergence of Opinion Leaders in a Networked Online Community: A Dyadic Model with Time Dynamics and a Heuristic for Fast Estimation. *Management Science*, 59(8), 1783–1799. <https://doi.org/10.1287/mnsc.1120.1685>
- Maillart, T., Sornette, D., Spaeth, S., & Von Krogh, G. (2008). Empirical tests of Zipf's law mechanism in open source linux distribution. *Physical Review Letters*, 101(21), 1–4. <https://doi.org/10.1103/PhysRevLett.101.218701>
- Mashima, R., & Takahashi, N. (2008). The emergence of generalized exchange by indirect reciprocity. In *New Issues and Paradigms in Research on Social Dilemmas* (pp. 159–176). Springer.
- Möhlmann, M., Zalmanson, L., Henfridsson, O., & Gregory, R. W. (2021). ALGORITHMIC MANAGEMENT OF WORK ON ONLINE LABOR PLATFORMS: WHEN MATCHING MEETS CONTROL. *MIS Quarterly*, 45(4).
- Nan, N., & Lu, Y. (2015). Harnessing the Power of Self-Organization in an Online Community During Organizational Crisis. *MIS Quarterly*, 39(X).
- Neff, G., & Nagy, P. (2016). Talking to Bots: Symbiotic agency and the case of Tay. *International Journal of Communication*, 10, 17.
- Newman, M. E. J. (2001). Clustering and preferential attachment in growing networks. *Physical Review E*, 64(2), 25102.
- Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*, 314(5805), 1560–1563.
- Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, 437(7063), 1291–1298. <https://doi.org/10.1038/nature04131>
- O'Mahony, S., & Ferraro, F. (2007). The Emergence of Governance in an Open Source Community. *Academy of Management Journal*, 50(5), 1079–1106. <https://doi.org/10.5465/amj.2007.27169153>

- Oh, W., Moon, J. Y., Hahn, J., & Kim, T. (2016). Research Note—Leader Influence on Sustained Participation in Online Collaborative Work Communities: A Simulation-Based Approach. *Information Systems Research*, June, isre.2016.0632. <https://doi.org/10.1287/isre.2016.0632>
- Preece, J., & Maloney-Krichmar, D. (2003). Online communities: focusing on sociability and usability. *Handbook of Human-Computer Interaction*, 596–620.
- Puranam, P. (2018). *The microstructure of organizations*. Oxford University Press.
- Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J. F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A. 'Sandy,' ... Wellman, M. (2019). Machine behaviour. *Nature*, 568(7753), 477–486. <https://doi.org/10.1038/s41586-019-1138-y>
- Ren, Y., Harper, F. M., Drenner, S., Terveen, L., Kiesler, S., Riedl, J., Kraut, R. E., Ren, Harper, Drenner, Terveen, Kiesler, Riedl, & Kraut. (2012). Building member attachment in online communities: Applying theories of group identity and interpersonal bonds. *Mis Quarterly*, 36(3), 841–864. <https://doi.org/10.2307/41703483>
- Rheingold, H. (1993). *The virtual community: Finding connection in a computerized world*. Addison-Wesley Longman Publishing Co., Inc.
- Safadi, H., Johnson, S. L., & Faraj, S. (2021). Who Contributes Knowledge? Core-Periphery Tension in Online Innovation Communities. *Organization Science*, 32(3), 752–775. <https://doi.org/10.1287/orsc.2020.1364>
- Salge, C. A. D. L., & Berente, N. (2017). Is that social bot behaving unethically? *Communications of the ACM*, 60(9), 29–31. <https://doi.org/10.1145/3126492>
- Salge, C. A. D. L., & Karahanna, E. (2018). Protesting Corruption on Twitter: Is It a Bot or Is It a Person? *Academy of Management Discoveries*, 4(1), 32–49.
- Salge, C. A. D. L., Karahanna, E., & Thatcher, J. B. (2022). Algorithmic Processes of Social Alertness and Social Transmission: How Bots Disseminate Information on Twitter. *MIS Quarterly*, 46(1).
- Seering, J., Luria, M., Ye, C., Kaufman, G., & Hammer, J. (2020). It Takes a Village: Integrating an Adaptive Chatbot into an Online Gaming Community. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–13.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63(2), 129.
- Skousen, T., Safadi, H., Young, C., Karahanna, E., Safadi, S., & Chebib, F. (2020). Successful Moderation in Online Patient Communities: Inductive Case Study. *Journal of Medical Internet Research*, 22(3), e15983. <https://doi.org/10.2196/15983>
- Sproull, L., & Arriaga, M. (2012). Online Communities. In *Handbook of Computer Networks* (pp. 898–914). John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118256107.ch58>
- Surma, J. (2016). Social exchange in online social networks. The reciprocity phenomenon on Facebook. *Computer Communications*, 73, 342–346.
- Thomaz, F., Salge, C., Karahanna, E., & Hulland, J. (2020). Learning from the dark web: Leveraging conversational agents in the era of hyper-privacy to enhance marketing. *Journal of the Academy of Marketing Science*, 48(1), 43–63.
- Traeger, M. L., Sebo, S. S., Jung, M., Scassellati, B., & Christakis, N. A. (2020). Vulnerable robots positively shape human conversational dynamics in a human-robot team. *Proceedings of the National Academy of Sciences of the United States of America*, 117(12), 6370–6375. <https://doi.org/10.1073/pnas.1910402117>
- Vaast, E., Safadi, H., Lapointe, L., & Negoita, B. (2017). Social Media Affordances for Connective Action: An Examination of Microblogging Use During the Gulf of Mexico Oil Spill. *MIS*

- Quarterly*, 41(4), 1179–1205. <https://doi.org/10.25300/MISQ/2017/41.4.08>
- Wang, G. A., Liu, X., Wang, J., Zhang, M., & Fan, W. (2015). Examining micro-level knowledge sharing discussions in online communities. *Information Systems Frontiers*, 17(6), 1227–1238. <https://doi.org/10.1007/s10796-015-9566-1>
- Wasko, & Faraj. (2005). Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice. *MIS Quarterly*, 29(1), 35. <https://doi.org/10.2307/25148667>
- Wasko, M., Faraj, S., & Teigland, R. (2004). Collective Action and Knowledge Contribution in Electronic Networks of Practice. *Journal of the Association for Information Systems*, 5(11), 493–513. <https://doi.org/10.17705/1jais.00058>
- Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications* (Vol. 8). Cambridge University Press. <https://doi.org/10.1017/CBO9780511815478>
- Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, 22(158), 209–212.
- Yan, Z., Wang, T., Chen, Y., & Zhang, H. (2016). Knowledge sharing in online health communities: A social exchange theory perspective. *Information & Management*, 53(5), 643–653.

ⁱ The landscape of bots in online communities is set to undergo a transformation with the widespread availability of generative AI technology for use in bots. We argue that it is crucial to gain a solid understanding of bots in their early stages in order to build a theoretical foundation for this phenomenon.

ⁱⁱ Indeed, engaging in a conversation is the benchmark of intelligence according to the Turing test

ⁱⁱⁱ <https://pushshift.io/>

^{iv} Bot accounts are not required to identify themselves as such on Reddit, so there is no definitive count on the number of bots on the platform.

^v <https://www.reddit.com/dev/api/>

^{vi} <https://botrank.pastimes.eu/>

^{vii} <https://botranks.com/about>

^{viii} <https://www.reddit.com/wiki/automoderator/>

^{ix} Network exchange patterns all involve a temporal sequence in which the pattern unfolds (Faraj & Johnson, 2011). In our panel data, the temporal events (i.e., interactions) within each panel (subreddit-month) are combined. As a result, our measures for network exchange patterns reflect the eventual outcomes of these interactions. For instance, a higher density network is indicative of more generalized exchange, while a greater preferential attachment leads to increased centralization.

^x The relationship between preferential attachment and authority is self-reinforcing in OCs. Participants with formal authority (e.g., moderators) have a duty to moderate, and thus their action draws further reaction because decision-making is negotiated constantly. Also, members who interact a lot tend to be endowed with formal or informal authority because online communities are meritocratic (Johnson et al., 2015; O'Mahony & Ferraro, 2007).

^{xi} <https://www.businessofapps.com/data/reddit-statistics/>

^{xii} <https://www.nytimes.com/2022/07/09/technology/elon-musk-twitter-spam-bots.html>