MIS Quarterly

RESEARCH NOTE

BEHAVIOR TOWARD NEWCOMERS AND CONTRIBUTIONS TO ONLINE COMMUNITIES[1]

Florian Pethig
Department of Information Systems and Operations Management, Tilburg School of Economics and Management,
Tilburg University, Tilburg, THE NETHERLANDS. {f.pethig@tilburguniversity.edu}

Hartmut Hoehle
Department of Enterprise Systems, Business School,
University of Mannheim, Mannheim, GERMANY. {hoehle@uni-mannheim.de}

Kai-Lung Hui
Department of Information Systems, Business Statistics, and Operations Management, School of Business and Management,
Hong Kong University of Science and Technology, Clear Water Bay, HONG KONG {klhui@ust.hk}

Andreas Lanz
Faculty of Business and Economics, University of Basel,
Basel, SWITZERLAND. {andreas.lanz@unibas.ch}

In this paper, we study whether and how behavior toward newcomers impacts their socialization outcomes in terms of retention and the quality of contributions in online communities. By exploiting a natural experiment on a large deal-sharing platform, we found that an intervention that proactively reminds other community members to be more considerate of newcomers caused newcomer deals to receive 54% more comments and increased the positive sentiment of the comments. The newcomers in the treatment group were 10% more likely than newcomers in the non-treatment group to post another deal, suggesting an increase in retention. However, we did not observe any effect of the intervention on the quality of subsequent contributions. Our evidence suggests that the intervention merely caused a temporary shock to newcomers' first contributions but did not improve their learning or motivate greater effort in subsequent contributions. We draw implications on the design of socialization processes to help communities improve the retention and performance of newcomers.

Keywords: Online communities, user-generated content, newcomers, socialization, natural experiment, difference-in-differences.

Introduction

A central concern of online communities is motivating the sustained contribution of knowledge. Newcomers are an important source of knowledge contribution because they often have a different background, experience, and perspective when compared with existing members. Their knowledge can be of great marginal benefit to online communities (Ransbotham & Kane, 2011; Ren et al., 2016). However, newcomers may also ask questions and make comments that existing members have seen or answered before. They must "learn the ropes," ensuring that they make valuable contributions to integrate with existing members (Kraut et al., 2012). This process of transforming from being an outsider to an insider is called *newcomer socialization* (Louis, 1980). Insiders—existing members of the community—play a pivotal role in this socialization process (Joyce & Kraut, 2006). They shape newcomers' initial interaction experience with the group and directly impact whether newcomers feel liked

[1] Wai Fong Boh was the accepting senior editor for this paper. Ben Liu served as the associate editor.

and accepted by other community members. They also provide valuable guidance for newcomers to learn how to better function in the new environment.

Because existing members affect how newcomers adjust to the new environment, their behavior toward newcomers has important implications for online communities in attracting continuous contributions from new members. The prevailing focus of the literature on newcomer socialization has been on proactive strategies of newcomers to persuade existing members to support them, including membership claims and information seeking behaviors (e.g., Ahuja & Galvin, 2003; Burke et al., 2010), as well as interventions to educate and retain newcomers, including awards (Gallus, 2017), behavioral information (Chen et al., 2010), and collective socialization tactics (Tausczik et al., 2018). An important outstanding question is how online communities can shape the behavior of existing members to facilitate newcomer socialization.

This question is important because newcomers may differ in their propensity to participate positively and actively in the socialization process (Miller & Jablin, 1991). Thus, not all newcomers proactively accustom themselves to existing members. Even if a community can educate newcomers and prompt them to be proactive, they may not comply with the norms or policies hidden in the community. Accordingly, their contributions might face scrutiny, which could demotivate and drive valuable newcomers out of the community (Ren et al., 2016). In practice, online communities often have policies to encourage existing members to be friendly to newcomers.[2] Examples include Mozilla's "Be Kind to Newcomers" and Wikipedia's "Don't Bite the Newcomer" policies. These policies may, however, not be faithfully read by all existing members. If existing members post hostile messages directed toward newcomers, newcomers may decide to leave the community even if the community removes the messages later because the damage has already been inflicted.

In this study, we advance a novel strategy to improve behaviors toward newcomers, i.e., *anticipatory excuses*, where a contributor's newcomer status is announced to other members *before* they engage with their post (Greenberg, 1996; Higgins & Snyder, 1989). This strategy is common in offline contexts: Organizations often use badges to identify new employees in situations where public scrutiny is expected. The idea is to discourage observers from attributing the newcomer's performance to internal factors (e.g., lack of ability), suggesting instead that the performance might be attributable to external factors (e.g., inexperience) for which the newcomer should be excused (Greenberg, 1996). In the context of online communities, we theorize that existing members may see inexperience as a plausible cause for poor contributions (Kelley, 1973). An anticipatory excuse that alerts existing members to a contributor's newness can motivate existing members to discount the role of ability in evaluating the newcomer's contribution.

We are particularly interested in whether anticipatory excuses through a nudge that informs existing members about a contributor's newcomer status can help improve socialization outcomes in terms of retention and contribution quality. We question whether newcomers might translate more positive responses from existing members into sustained and more valuable contributions. The reinforcement literature suggests that this is likely because positive responses can amplify intrinsic motivations—attention from others tends to make people feel good about themselves (Delin & Baumeister, 1994). However, positive outcomes may not occur if newcomers do not acquire relevant knowledge about their environment due to the lack of critical feedback (Wilhelm et al., 2019) or if they have a natural propensity to maintain their initial (low) levels of activity (Panciera et al., 2009).

Here, we use a *newcomer nudge* to study the effects of revealing a contributor's newness on (1) existing members' behavior toward newcomers and (2) newcomers' retention and future contribution quality.[3] Our empirical strategy uses a difference-in-differences (DID) approach, exploiting a natural experiment on a deal-sharing community dedicated to price promotions. The community allows users to post deals and vouchers. Other members can rate the post quality by upvoting or downvoting it or making comments. Since October 20, 2016, the community has displayed a newcomer nudge (hereafter "nudge") above the comment field of a newcomer's *first* post (Figure 1). The nudge is permanently attached to the post and visible even after the contributor has published additional posts. It affects initial newcomer posts only. To our knowledge, the nudge was not announced in advance and the community did not make any other major changes around the time when it was introduced.

Our DID estimation shows that the nudge caused newcomer deals to receive 54% more comments and increased the positive sentiment of the comments, indicating that anticipatory excuses may indeed lead to more enthusiastic responses from existing members. We also found that newcomers socialized with the nudge were 10% more likely to post another deal within 12 months, but the quality of their subsequent contributions did not improve. Hence, the nudge appears to facilitate newcomer retention but does not have a significant or sizable effect on newcomers' contribution quality.

[2] Sample policies discussing the desirable treatment of newcomers are available at https://osf.io/sgmv2.

[3] A nudge is any aspect of a choice architecture (e.g., user interface) that alters people's behavior in a predictable way without restricting the freedom of choice. For more discussion of nudges, see Thaler and Sunstein (2008).

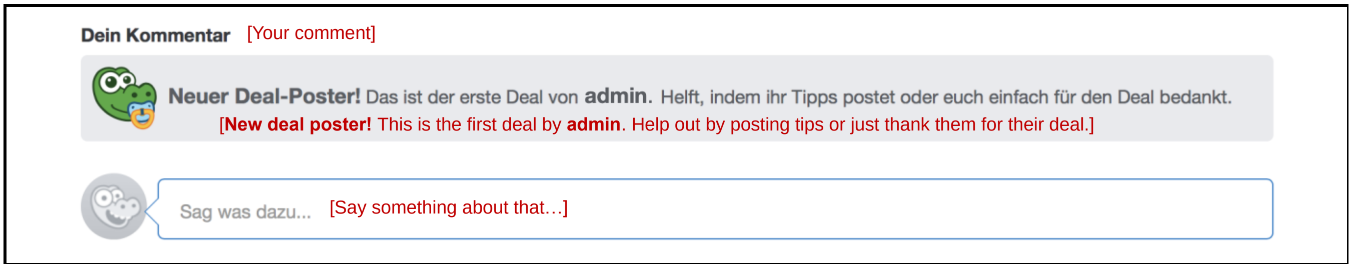


Figure 1. The Nudge

This paper makes three contributions. First, it contributes to the literature on interventions to socialize newcomers in online communities. Many studies have documented positive outcomes of socialization programs, such as collective socialization (e.g., Li et al., 2020; Tausczik et al., 2018). However, the success of these programs often hinges on the extent to which newcomers show good citizenship behaviors after being educated by them. We highlight a different but related approach based on the idea that socialization outcomes are malleable to changes in insiders' behavior as well. This new focus (on insiders instead of newcomers) provides a powerful alternative for online communities to enhance the newcomer socialization process.

Second, this study is the first to examine how a positive distortion of existing members' behavior relates to different aspects of socialization—specifically, newcomer retention versus contribution quality. Prior work has shown that positive responses to newcomers can improve retention through a reinforcement mechanism (Joyce & Kraut, 2006; Phang et al., 2015). However, this paper shows that the benefits of positive reinforcement do not translate to the enhancement of contribution quality. We advance the lack of change in task-relevant knowledge sharing as a plausible reason for why the newcomers' contributions do not improve. The implication is that the positive effects of institutional pressure toward lenient treatment of newcomers are contingent on having the right enabling environment—one that instills task-relevant knowledge in newcomers.

Third, this study contributes to a growing literature on using anticipatory excuses to preempt the negative effects of service failures—for example, by providing trainee badges to inexperienced employees (Flacandji et al., 2023; Greenberg, 1996). Prior research has established the value of signaling employees' inexperience to external customers. This study also shows its effectiveness in interacting with insiders. It suggests that communities can influence content production upstream—*before* the content is published—complementing research that has focused on the moderation of content downstream, i.e., *after* the content has been published (Jiang et al., 2023).

Related Literature

We draw on the literature on newcomer socialization in online communities and anticipatory excuses to frame our contributions.

Newcomer Socialization in Online Communities

Socializing newcomers is central to online community success because newcomers can replace departing members and contribute new knowledge (Ren et al., 2012). Our work is particularly related to two streams of socialization research: (1) how interventions affect socialization outcomes and (2) the influence of insiders' behavior.

First, a growing body of research has studied community interventions to socialize newcomers (Gallus, 2017; Li et al., 2020; Tausczik et al., 2018). These interventions primarily promote newcomers' effective functioning in the new environment to improve their retention and contribution quality. For example, a study of the WikiEd program, where students make Wikipedia edits as a class assignment (Li et al., 2020), showed that newcomers who participated in the program were twice as likely to continue contributing and made higher-quality edits. By contrast, another study showed that an interactive game that helped newcomers accomplish tasks on Wikipedia had no discernible impact on their activities, despite its popularity (Narayan et al., 2017). The literature has shown that the outcomes of socialization interventions are closely tied to newcomers' capabilities and citizenship behavior after onboarding. We depart from this literature by examining whether socialization outcomes can be improved by interventions independent of newcomers' initial behavior. The advantage of our approach is that communities can provide a more positive and consistent new user experience even if newcomers are not yet acquainted with the community. Such an approach has not been tested in the literature.

Second, prior research has examined the role of insiders in the socialization process. Insiders shape the environment in which newcomers try to fit in. Receiving a response to their post can increase newcomers' likelihood to post again because it indicates that the community is positive and receptive (Joyce & Kraut, 2006; Lampe & Johnston, 2005; Zhang et al., 2013). However, research has also suggested that newcomers who receive high-quality answers might reduce contributions, believing that their own contributions are not needed (Yan & Jian, 2017). Aside from their inconclusive findings, these studies are observational, making it difficult to tease out the causal influence of insiders' responses on socialization outcomes.[4] In particular, people interested in the community might have a natural propensity to participate, leading their posts to receive more responses from existing members. To more precisely identify the causal influence of community response on newcomer behavior, it is essential to examine exogenous changes in insiders' behavior. The nudge in this study serves as one such exogenous change.

Anticipatory Excuses and Inexperience

This study is also related to research on *anticipatory excuses* (Higgins & Snyder, 1989)—the attempt to provide an excuse for a performance that has yet to be evaluated (Greenberg, 1996). Whereas retrospective excuses aim to distance the actor as much as possible from a particular performance *after* the act, anticipatory excuses are disseminated *before* the anticipated (poor) performance (Snyder & Higgins, 1988). The goal is to preemptively weaken the link between the actor and a subsequent outcome. Thus, anticipatory excuses are often used when the actors will predictably not meet the performance standard—for example, when they are inexperienced. Research has shown that revealing employee inexperience through a badge or corporate uniform can modify perceptions of service quality (Flacandji et al., 2023; Greenberg, 1996). Greenberg (1996) found that people who asked others to forgive them because they were new to their job were more likely to be excused for poor performance. Flacandji et al. (2023) found that customers who experienced a service failure were more likely to remain loyal to the organization if they encountered an inexperienced employee versus an experienced employee. In such cases, the poor performance tends to be attributed to the employee's inexperience instead of the organization.

The literature reviewed above pertains to encounters between new employees and external customers. In contrast, our study focuses on shifting the behavior of insiders, that is, existing members of a community. Insiders with a long tenure may be more protective of community quality than customers (Ren et al., 2023). Furthermore, since they are devoid of traditional social signals, text-based asynchronous communications in online communities tend to be less personal (Ma & Agarwal, 2007). Whether revealing the inexperience of a newcomer can defuse insiders' dissatisfaction with the newcomer's contributions in an online context is unclear.

Theory and Hypotheses

We draw on attribution theory (Kelley, 1973) to analyze how the newcomer nudge shapes the behavior of existing members toward newcomers. Attribution theory explains how people attribute causes to someone's behavior and the consequences of such attribution (Jones et al., 1987). It posits that the interpretation of others' behavior plays an important role in determining reactions to the behavior. In our setting, the nudge provides a cause (inexperience) for the newcomer's performance. The expected consequence is that other members may respond more leniently by providing (1) more responses and (2) increasing the positive sentiment of responses. Specifically, by introducing the nudge, the platform provides an excuse for newcomers by highlighting their inexperience *before* others respond to them. Similar to a trainee badge (Greenberg, 1996), the nudge can therefore be considered an *anticipatory excuse* (Higgins & Snyder, 1989).

The effectiveness of the anticipatory excuse results from the predictions of the discounting principle (Kelley, 1973). By definition, newcomers have had less exposure to a community than existing members. Therefore, they have fewer opportunities to learn about the community's policies and norms. Since poor performance may be expected when newcomers post for the first time, identifying a community member as a newcomer allows existing members to discount the role of ability as the behavior-correspondent disposition. After being exposed to the nudge, we expect that existing members will accept the newcomer's inexperience as a legitimate explanation for their poor performance, allowing them to attribute it to situational pressure rather than the inherent ability of the person. Hence, we expect such members to be more responsive and forgiving when

[4] Several papers note this shortcoming. Joyce and Kraut (2006, p. 743) note that "ours is not experimental research. Therefore, we cannot definitely say that the empirical relationships shown here ... between getting a reply and posting again, are causal." Yan and Jian (2017, p. 16) note that "this study is not a controlled experiment. So, none of the relationships we have identified is, strictly speaking, causal. However, we have taken measures to make sure our predictors (community response) preceded the outcomes (i.e., future participation)." Zhang et al. (2013, p. 1121) note that "it is likely that some unobserved heterogeneity or omitted variables that influence a member's likelihood of receiving responses from the community also influence his continued participation in the community."

evaluating the contributions of newcomers. Furthermore, the augmentation effect in attribution suggests that if a person can rise above conditions (e.g., inexperience) that would lead them to perform poorly, good performance may be perceived as internally caused, leading to inflated perceptions of good performance (Greenberg, 1996; Kelley, 1973). In other words, if newcomers perform unexpectedly well, their contributions will likely be seen in a particularly positive light.

Taken together, the analysis above points to a positive effect of the nudge on the behavior toward newcomers because (1) existing members may discount the role of ability if newcomers perform poorly, and (2) they may augment the role of ability if newcomers perform well. We expect the nudge to encourage existing members to leave more comments on newcomer contributions and we would expect these comments to be more positive, on average. Therefore, we hypothesize:

H1: *Newcomer contributions will receive more comments after the nudge.*

H2: *The sentiment of comments posted on newcomer contributions will be more positive after the nudge.*

The impact of the nudge on existing members' behavior toward newcomers captures only the initial socialization process. The final socialization outcomes depend on how often the newcomers post and what they post after being socialized into the community. In the following, we therefore focus on the two primary socialization outcomes (e.g., Li et al., 2020)—retention and contribution quality—after the nudge intervention shifts the existing members' responses.

Retention: Positive responses may increase newcomers' future contributions because people tend to repeat actions that lead to positive reinforcements (Joyce & Kraut, 2006). Contributors who perceive themselves to be well-connected in the community are more likely to contribute because they receive acknowledgment from others (Phang et al., 2015). Research offers several theoretical explanations for such reinforcement. One emerges from the finding that an individual's behavior depends on its consequences (Ferster & Skinner, 1957). For example, in a conversation, speakers are more likely to express their opinions when their conversation partners agree with them (Verplanck, 1955). Receiving more responses may also amplify intrinsic motivations because attention from others creates a positive mood and makes people feel good about themselves (Delin & Baumeister, 1994). If the nudge increases newcomers' exposure to positive responses, it may reinforce their decision to stay in the community. In contrast, negative social experiences may lead to alienation. If insiders reject a newcomer, the newcomer may stop asking questions or may leave the community altogether due to fears that they may be "bugging" community members (Miller & Jablin, 1991, p. 97).

Another explanation for positive reinforcement is that individuals reciprocate others' support by paying it forward (Gouldner, 1960). Newcomers might feel indebted and obligated to reciprocate the beneficial resources they have received from existing members (see Joyce & Kraut, 2006, who suggest reciprocity as a mechanism underlying newcomers' information-sharing behavior). Both the desire to receive positive reinforcement and the perceived obligation to reciprocate others' responses support our conjecture that the change in behavior toward newcomers will result in a higher likelihood that they will continue participating in the community. Therefore, we hypothesize:

H3: *The likelihood of retaining newcomers will be higher after the nudge.*

Contribution quality: Whether nudge-induced positive responses from existing members can lead to a higher quality of future contributions is an open question. The literature on socialization has argued that "positive reinforcement induces more learning than negative reinforcement" (Cable & Parsons, 2001, p. 7). Socializing with insiders may help newcomers internalize the community's values. Thus, if the nudge promotes positive interactions between newcomers and existing members, it may help newcomers improve the quality of their subsequent contributions as well.

However, individuals also learn from negative feedback and can use that knowledge to improve their contributions (Wilhelm et al., 2019). Negative feedback is particularly effective in arousing cognitive awareness that leads to adaptation and change. Therefore, the lack of such feedback can lead to quality degradation. As illustrated by a Stack Overflow member, negative feedback can serve as a reminder for newcomers to include missing information: "Yes it is hard for beginners. But I have to admit that the negative feedback helped me to write better questions. At start I was a bit lazy and did not provided [*sic*] enough details and people were downvoting me, but that's ... how I learned to always provide enough details" (Black, 2019).

Lastly, barriers to joining a group and initiation rituals can increase newcomers' commitment and loyalty and motivate them to make high-quality contributions (Kraut et al., 2012). Research suggests that people are more committed to groups when they experience more rigorous initiation processes (Aronson & Mills, 1959) because the actions they must undertake to become part of the group build their selfesteem. Therefore, if the nudge softens the initiation process for newcomers by making existing members act more positively toward them, newcomers might be more likely to remain in the community but may feel less committed to making high-quality contributions (see Kraut et al., 2012). Overall, the existing theories do not point to an unequivocal impact of the nudge on contribution quality. Therefore, we formulate the following competing hypotheses:

H4a: *The quality of newcomers' subsequent contributions will be higher after the nudge.*

H4b: *The quality of newcomers' subsequent contributions will be lower after the nudge.*

Setting

The community of interest we selected for this study is mydealz, a large German consumer-to-consumer community dedicated to sharing, rating, and reviewing deals and vouchers. Similar communities exist in other countries, such as Slickdeals.net in the U.S. and hotukdeals.com in the U.K. Members post deals and vouchers that can be up- or downvoted by others. The net number of votes (upvotes minus downvotes) is called the deal temperature. If a deal receives a temperature above 100, it is "hot," and if it is downvoted to below zero, it is "cold" (Figure 2). Deals are displayed in reverse chronological order. Well-received deals are selected by editors to appear on a highlight page, i.e., the default landing page for visitors. In addition to voting, members can write comments below a deal.

We conducted semi-structured interviews with 18 users to assess the suitability of mydealz for studying behavior toward newcomers (Appendix E). The interviewees observed negative comments directed at newcomers—for example when newcomers make mistakes:

> *They're pretty quick to go after people who are beginners and don't know exactly, okay, what's a good deal now, how do I make the best price comparison, and so on. So, yes, hate comments [on these deals] are usually pouring in very fast.*

The interviewees also shared that other members made fun of newcomer deals that did not offer much savings. Some interviewees believed that new contributors face a lot of scrutiny regarding adherence to the community's policies (e.g., on mydealz, stated prices must always include shipping costs and contributors must conduct a thorough price comparison):

> *I had the feeling that there's always a lot of criticism, that you can't make any mistakes, that you have to pay close attention to the wording and as soon as you somehow have something in there, that it's then immediately noted, criticized, you're ... stoned.*

Such observations motivated mydealz to implement the newcomer nudge. This provided an excellent opportunity to study how online communities can better socialize newcomers because the negative behavior that existed on mydealz discouraged some users from ever posting again.

Data

To analyze the effect of the nudge on newcomer deals and socialization outcomes, we collected historical data from mydealz. In our main analysis, we considered deals posted between July 22, 2016, and January 17, 2017, covering 90 days before and 90 days after the introduction of the nudge. In Appendix A, we describe the data collection and preparation process. The deals cover a broad range of products in multiple categories such as electronics, food and drink, and household and garden. For each deal, we recorded the contributor's username, publication date (*Day*) and hour (*Hour*), title, description, net number of votes (*DealTemp*), number of comments (*NumComments*), number of categories (*NumCategories*), content type (*Content*; 0 = deal and 1 = voucher), and whether it was restricted to a certain location (*LocalDeal*).

We counted the description length in words (*DescLen*) and recorded the commenter's username, day, and comment text to identify its length (*AvgCommentLen*) and sentiment. We measured the average sentiment of the comments using the German sentiment analysis tool provided by Microsoft's Azure Cognitive Services (API version 2021-04-30). Azure Cognitive Services, including its Face API, is well-established and has been used in prior research (Malik et al., 2023). Microsoft's sentiment analysis applies well to texts with more extreme opinions (Pallas et al., 2020), which is typical for online communities. It returns three non-negative sentiment scores for each comment—a positive score (*Positive*), neutral score (*Neutral*), and negative score (*Negative*). The three scores sum to 1. We also collected data from each contributor's public user profile, including the date they joined the community, to compute their tenure in months (*Tenure*). Taken together, we constructed a cross-sectional data set with one row for each deal. Our data set includes all comments written up to the point of data collection. Table 1 presents summary statistics of our data. The deal temperature and description lengths differ markedly between newcomer and non-newcomer deals.

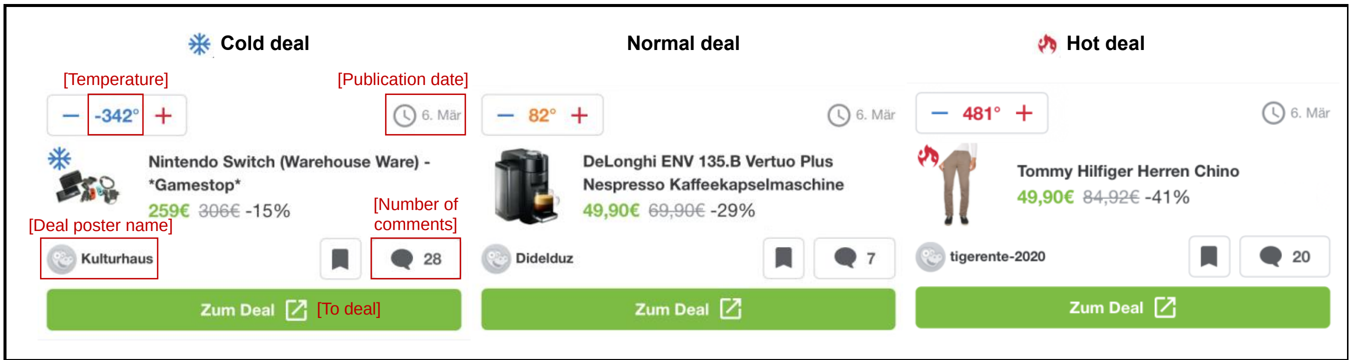


Figure 2. Examples of a Cold, Normal, and Hot Deal

Table 1. Summary Statistics of Newcomer and Non-Newcomer Deals

Variables	Unit	Newcomer deals					Non-newcomer deals					*t*-statistic
		N	**Mean**	***SD***	**Min**	**Max**	***N***	**Mean**	***SD***	**Min**	**Max**	
DealTemp	Degrees	4,952	191.36	465.18	-935	17,746	35,971	290.38	439.32	-1,105	20,668	14.14***
NumComments		4,952	18.47	173.96	0	10,693	35,971	20.07	68.39	0	6,474	0.64
DescLen	Words	4,952	86.02	88.25	0	1,819	35,971	116.74	153.14	0	5,560	20.60***
LocalDeal	Dummy	4,952	0.19	0.40	0	1	35,971	0.14	0.34	0	1	-10.01***
NumCategories		4,952	3.87	2.15	1	13	35,971	3.85	2.09	1	16	-0.67
Content	Dummy	4,952	0.07	0.26	0	1	35,971	0.07	0.25	0	1	-0.51
Tenure	Months	4,952	11.80	17.93	0	108	35,971	33.8	24.58	0	112	76.99***
AvgCommentLen	Words	4,666	19.24	12.06	1	178	34,641	19.64	14.10	0	741	2.05*
Positive	0~1	4,666	0.29	0.17	0	1	34,641	0.29	0.16	0	1	1.84
Neutral	0~1	4,666	0.42	0.18	0	1	34,641	0.42	0.17	0	1	-0.67
Negative	0~1	4,666	0.29	0.16	0	1	34,641	0.28	0.14	0	1	-1.23

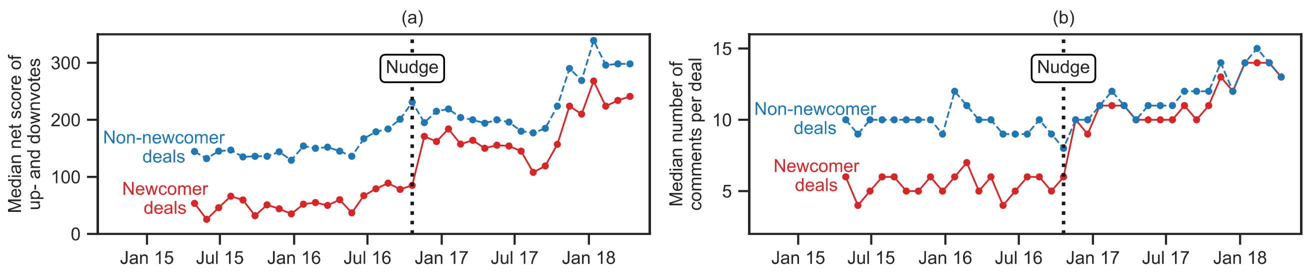
Note: *SD* = standard deviation. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Socialization outcomes: We constructed several measures to evaluate the effect on socialization (see the Effect of the Nudge on Socialization Outcomes section below). Our measure of newcomer retention, *DealPosted*, is a binary indicator of whether users posted any deals within 12 months after their first deal. Our measures of contribution quality are defined as follows: *ΔDealTemp* measures changes in the quality of contributions by subtracting the deal temperature of the first deal from the average temperature of all deals that were posted by the same user within 12 months after the first deal. As alternative measures of quality, we measured the average deal temperature of subsequent deals (*AvgDealTemp*), the average likelihood of users mentioning a price comparison in the descriptions of subsequent deals (*AvgPriceComp*), and, if both the original price and discounted price were available, the average percentage discount (*AvgDiscount*).[5] We analyzed contribution quality for users who had posted at least one deal in the 12 months following the first deal only.

We also measured two outcomes for exploratory purposes. *AvgDescLen* captures the average description length of the subsequent deals posted during the 12 months following a newcomer's first deal. It reflects users' effort to post subsequent deals. *DaysSecDeal* measures the time gap (in days) from the first to the second deal. It reflects users' interest in posting another deal after their inaugural deal.

Lastly, in our analysis of socialization outcomes, we controlled for the badges earned by users (primarily by existing members), which generally reflect their activity levels. Specifically, *BadgeDeal* is a binary indicator denoting whether a user posted at least 10 deals; *BadgeComment* is a binary indicator denoting whether a user posted at least 100 comments; *BadgeVote* is a binary indicator denoting whether a user rated at least 200 deals.

[5] We describe the keyword extraction process for *AvgPriceComp* and *AvgDiscount* in Appendix B.



Note: The unit of observation is 30 days. The vertical dotted line indicates the introduction of the nudge. The shaded area between the two solid lines depicts the window of our main analysis (90 days pre- and post-nudge).

Figure 3. Newcomer vs. Non-Newcomer Deals in the Long Run (1,110 Days)

Empirical Analysis and Results

Effect of the Nudge on Newcomer Deals

Model-Free Evidence

Figure 3 visualizes the long-term effects of the nudge. The plot spans 1,110 days (~3 years) with observations recorded in 30-day intervals. Figure 3a shows the median deal temperatures, which differed substantially between the newcomer and non-newcomer deals before the nudge. The gap narrowed significantly after the nudge. In particular, the median temperature of newcomer deals increased from about 50 to about 150. Figure 3b shows a similar pattern for the number of comments.

Regression Results

The model-free trends in Figure 3 do not account for control variables that might have confounded the nudge effect. To formally test H1 and H2, we used a DID strategy to identify the effect of the nudge on *NumComments* (H1) and the three sentiment scores, *Positive*, *Neutral*, and *Negative* (H2). Our unit of analysis is the deal, with newcomer deals as the treatment group and non-newcomer deals as the control group. We considered the following ordinary least squares (OLS) regression, in which we varied the time windows between 3, 5, 30, and 90 days before and after the newcomer nudge:

$$y_i = \beta_0 + \beta_1 Newcomer_i + \beta_2 Newcomer_i \times After_i + \beta_3 Tenure_i + \gamma_1 X_i + \gamma_2 Day_i + \gamma_3 Hour_i + \varepsilon_i, \quad (1)$$

where y_i variously denotes the log-transformed number of comments (H1) and the sentiment scores of the comments on deal *i* (H2). *Newcomer* is a dummy variable that equals 1 if deal *i* is a newcomer deal and 0 otherwise. As the nudge did not affect deals posted before the policy change, *After* was set to 0 if deal *i* was posted before the introduction of the nudge and 1 otherwise. The coefficient, β_2, of the interaction term *Newcomer* × *After* represents the marginal effect of the nudge on the responses to newcomer deals posted after the policy change. The main effect of *After* was omitted because of collinearity with the day variables. *Tenure* denotes the number of months since the contributor of deal *i* joined the community (fixed at the day of the post).

The control variables, X_i, include deal characteristics, i.e., *LocalDeal*, *Content*, *DescLen*, and *NumCategories*. We included category dummy variables in X_i to account for differences between deal categories. In the sentiment score regressions, we controlled for the average length of comments (*AvgCommentLen*) because comment length may affect content richness and hence the classification accuracy. *Day* and *Hour* are dummy variables to control for the published date and hour of deal *i*. As the deals usually receive the most attention shortly after being posted, both *Day* and *Hour* may affect how others interact with the deals (e.g., deals published at night may attract fewer comments than deals published in the morning). Finally, ε_i captures the random error.

Table 2 shows the regression results with the standard errors, ε_i, clustered by user. Each column in Table 2 corresponds to one of the four time windows, 3 days, 5 days, 30 days, and 90 days before and after the nudge. The left side of Panel A shows that *Newcomer* has a negative relationship with *NumComments*, which indicates that newcomer deals generally received fewer comments than non-newcomer deals. Because the coefficients of the interaction term, *Newcomer* × *After*, are consistently positive and precisely estimated, H1 is supported. The coefficient obtained from the 90-day sample, for instance, is 0.435, indicating that the nudge led to a 54% increase in the number of comments during the first 90 days.[6] Among the control variables, *Tenure* and *DescLen* are positively correlated with *NumComments*, indicating that deals that conveyed more information and were posted by more experienced community members received more attention. Local deals attracted fewer comments than non-local deals, meaning they were of interest to fewer members. *Content* is negatively correlated with *NumComments*, meaning vouchers garnered less discussion than deals. The more categories a deal was assigned to, the more comments it received.

Table 2. Test of H1 and H2: Effect of Nudge on Newcomer Deals

Panel A: Effect of Nudge on Number of Comments and Positive Sentiment								
	log(1+*NumComments*)				***Positive***			
	±3 days	±5 days	±30 days	±90 days	±3 days	±5 days	±30 days	±90 days
Newcomer	-0.120 (0.129)	-0.200** (0.101)	-0.269*** (0.044)	-0.362*** (0.029)	-0.027 (0.019)	-0.034* (0.017)	-0.016** (0.008)	-0.004 (0.005)
Newcomer × *After*	0.322* (0.195)	0.300* (0.158)	0.393*** (0.057)	0.435*** (0.034)	0.038 (0.030)	0.064** (0.028)	0.035*** (0.010)	0.012** (0.006)
Tenure	0.004*** (0.001)	0.005*** (0.001)	0.004*** (0.000)	0.003*** (0.001)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000** (0.000)
log(1+*DescLen*)	0.231*** (0.039)	0.200*** (0.031)	0.181*** (0.017)	0.159*** (0.012)	0.001 (0.006)	0.005 (0.005)	0.012*** (0.002)	0.011*** (0.001)
LocalDeal	-0.389*** (0.099)	-0.309*** (0.073)	-0.371*** (0.035)	-0.372*** (0.024)	-0.008 (0.015)	-0.008 -0.013)	-0.018*** (0.005)	-0.018*** (0.003)
NumCategories	0.077*** (0.015)	0.063*** (0.013)	0.044*** (0.006)	0.037*** (0.004)	0.001 (0.002)	0.001 (0.002)	-0.002* (0.001)	-0.002*** (0.000)
Content	-0.596*** (0.144)	-0.495*** (0.097)	-0.518*** (0.047)	-0.546*** (0.025)	-0.041 (0.027)	-0.014 (0.022)	-0.012 (0.008	0.002 (0.004)
AvgCommentLen					0.002*** (0.000)	0.002*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Time FE (day, hour)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Category FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,109	1,834	11,815	40,923	1,065	1,748	11,234	39,307
Adjusted R-squared	0.144	0.139	0.141	0.139	0.023	0.021	0.031	0.031
Panel B: Effect of Nudge on Neutral Sentiment and Negative Sentiment								
	Neutral				***Negative***			
	±3 days	±5 days	±30 days	±90 days	±3 days	±5 days	±30 days	±90 days
Newcomer	0.013 (0.020)	0.019 (0.020)	0.007 (0.008)	-0.002 (0.005)	0.013 (0.020)	0.015 (0.018)	0.009 (0.007)	0.006 (0.004)
Newcomer × *After*	-0.025 (0.029)	-0.047* (0.027)	-0.021** (0.011)	-0.006 (0.006)	-0.013 (0.026)	-0.017 (0.024)	-0.014 (0.009)	-0.007 (0.005)
Tenure	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)
log(1+*DescLen*)	0.005 (0.006)	-0.003 (0.005)	-0.010*** (0.003)	-0.010*** (0.002)	-0.006 (0.006)	-0.003 (0.004)	-0.001 (0.002)	-0.001 (0.001)
LocalDeal	0.038** (0.017)	0.040*** (0.014)	0.041*** (0.006)	0.035*** (0.003)	-0.029** (0.015)	-0.031*** (0.011)	-0.023*** (0.005)	-0.017*** (0.002)
NumCategories	-0.001 (0.002)	-0.003 (0.002)	0.002** (0.001)	0.002*** (0.000)	0.000 (0.002)	0.002 (0.002)	0.000 (0.001)	0.000 (0.000)
Content	-0.033 (0.031)	-0.022 (0.020)	0.012 (0.008)	0.000 (0.004)	0.074*** (0.028)	0.036* (0.018)	0.000 (0.007)	-0.001 (0.004)
AvgCommentLen	-0.005*** (0.000)	-0.005*** (0.000)	-0.004*** (0.001)	-0.003*** (0.000)	0.003*** (0.001)	0.003*** (0.000)	0.003*** (0.001)	0.002*** (0.000)
Time FE (day, hour)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Category FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,065	1,748	11,234	39,307	1,065	1,748	11,234	39,307
Adjusted R-squared	0.137	0.127	0.108	0.102	0.082	0.072	0.066	0.064

Note: FE = fixed effects. Robust standard errors clustered by user are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

[6] We calculate effect size as exp(0.435) - 1 = 54%.

Table 2 also shows the impact of the nudge on sentiment scores. Comments on newcomer deals became significantly more positive after the nudge (except in the 3-day sample), but we did not observe a consistent and significant effect for negative or neutral sentiment. If anything, both types of sentiment seem to have decreased. These results indicate that the nudge improved the sentiment toward newcomer deals, supporting H2. We used the 90-day window as our preferred estimate.[7]

Parallel Trends

The identification of the treatment effect in DID is based on the parallel trends assumption. In the absence of treatment, the treated and untreated (control) groups should follow a similar trend, i.e., their difference should be relatively stable over time. We added a series of time dummies to capture the relative chronological distance between the observation time and the time when the nudge was introduced:

$$y_i = \beta_0 + \beta_1 Newcomer_i + \sum_{j=-6}^{5} \lambda_j Newcomer_i \times Distance_{ij} + \beta_3 Tenure_i + \gamma_1 X_i + \gamma_2 Day_i + \gamma_3 Hour_i + \varepsilon_i, \quad (2)$$

where *Distance* is a dummy variable indicating the relative chronological distance j from the policy change using a 15-day time window. Equation (2) is similar to Equation (1), with *Newcomer* × *After* replaced by a set of dummy variables, *Newcomer* × *Distance*. The coefficients λ_j help identify whether a pre-treatment trend existed and how the effect dynamically evolved after the new policy. We estimated Equation (2) with j ranging from -6 to 5, which evenly divides the 180 days of our main analysis into 12 periods. We set the first time period (j = -6) as the baseline by normalizing the coefficient of that time period to zero.

Table 3 presents the estimation results. None of the pre-treatment coefficients of *Newcomer* × *Distance* are statistically different from zero. By contrast, all post-treatment coefficients for *NumComments* are statistically significant and positive. Three coefficients in the post-treatment periods of the sentiment regressions are marginally significant at $p < 0.1$ (*Positive* at j = 0, *Negative* at j = 1 and 2). These results suggest that changes in the number of comments and sentiment scores occurred only after the policy change and that there were no spurious or erroneous associations.

Table 3. Dynamic Treatment Effect

	log(1+*NumComments*)		*Positive*		*Neutral*		*Negative*	
	(1)		(2)		(3)		(4)	
Newcomer	-0.337***	(0.067)	-0.004	(0.012)	-0.012	(0.012)	0.015	(0.011)
Newcomer × *Distance*$_{-5}$	-0.010	(0.095)	0.028	(0.017)	-0.007	(0.017)	-0.021	(0.015)
Newcomer × *Distance*$_{-4}$	-0.099	(0.094)	-0.004	(0.016)	0.020	(0.017)	-0.016	(0.015)
Newcomer × *Distance*$_{-3}$	-0.132	(0.094)	0.005	(0.016)	0.006	(0.017)	-0.011	(0.015)
Newcomer × *Distance*$_{-2}$	-0.033	(0.090)	-0.008	(0.016)	0.012	(0.017)	-0.004	(0.015)
Newcomer × *Distance*$_{-1}$	0.122	(0.093)	-0.017	(0.015)	0.023	(0.017)	-0.006	(0.015)
Newcomer × *Distance*$_{0}$	0.338***	(0.089)	0.025*	(0.015)	-0.012	(0.016)	-0.013	(0.014)
Newcomer × *Distance*$_{1}$	0.524***	(0.088)	0.021	(0.015)	0.004	(0.015)	-0.025*	(0.014)
Newcomer × *Distance*$_{2}$	0.361***	(0.079)	0.015	(0.014)	0.006	(0.014)	-0.021*	(0.012)
Newcomer × *Distance*$_{3}$	0.444***	(0.077)	0.011	(0.013)	0.007	(0.014)	-0.018	(0.012)
Newcomer × *Distance*$_{4}$	0.389***	(0.081)	-0.002	(0.014)	0.009	(0.014)	-0.007	(0.013)
Newcomer × *Distance*$_{5}$	0.415***	(0.080)	0.010	(0.014)	0.002	(0.014)	-0.012	(0.013)
Time FE (day, hour)	Yes		Yes		Yes		Yes	
Category FE	Yes		Yes		Yes		Yes	
Control variables	Yes		Yes		Yes		Yes	
Observations	40,923		39,307		39,307		39,307	
Adjusted R-squared	0.139		0.031		0.102		0.064	

Note: FE = fixed effects. Robust standard errors clustered by user are in parentheses. Control variables include *Tenure*, log(1+*DescLen*), *LocalDeal*, *NumCategories*, and *Content*. Columns 2-4 additionally include *AvgCommentLen*. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

[7] Our preference for the 90-day window is based on prior work (Foerderer et al., 2018) and the fact that the platform had implemented a new badge system three months before the nudge. Although the use of shorter windows produces significant estimates, the rapid decrease in sample size may affect the precision of the estimates.

Table 4. Testing for SUTVA and Compositional Changes								
	log(1+*NumComments*)		*Positive*		*Neutral*		*Negative*	
	SUTVA	Comp. Changes	SUTVA	Comp. Changes	SUTVA	Comp. Changes	SUTVA	Comp. Changes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NumTreatedDeals	-0.002		0.000		0.001		-0.001	
	(0.006)		(0.001)		(0.001)		(0.001)	
SecondDeal		-0.194*		0.002		-0.010		0.008
		(0.103)		(0.018)		(0.019)		(0.014)
SecondDeal × *After*		0.168		-0.013		0.011		0.001
		(0.120)		(0.021)		(0.022)		(0.018)
TimeFE (day, hour)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Category FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	20,467	35,966	19,785	34,636	19,785	34,636	19,785	34,636
Adjusted R-squared	0.131	0.130	0.032	0.030	0.088	0.099	0.063	0.066

Note: SUTVA = stable unit treatment value assumption; FE = fixed effects. Robust standard errors clustered by user are in parentheses. Control variables include *Tenure*, log(1+*DescLen*), *LocalDeal*, *NumCategories*, and *Content*. Columns 3-8 additionally include *AvgCommentLen*. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Spillover of the Policy Change

We needed to rule out the possibility, known as the stable unit treatment value assumption (SUTVA), that the nudge could have affected non-newcomer deals (Eckles et al., 2017; Rosenbaum, 2007). To do so, we constructed a proximity-based measure of exposure (Jo et al., 2020). We tested whether the behavior toward non-newcomer deals depended on the number of treated deals posted before a non-newcomer deal. A potential spillover should be more pronounced for non-newcomer deals that directly compete for attention with treated deals. We created the variable *NumTreatedDeals*, which captures the number of newcomer deals published in the 30 minutes prior to a non-newcomer deal. We reestimated Equation (1) by restricting the analysis to non-newcomer deals after the policy change. The results in Table 4 show that the coefficients of *NumTreatedDeals* are not statistically significant (odd columns). Thus, the nudge did not attract comments or lead to a sentiment change for non-newcomer deals.

Compositional Changes

Given that our analysis used a DID design with repeated cross-sections (i.e., different deals posted before and after the nudge), it is important to address possible compositional changes (Athey & Imbens, 2006). First, compositional changes would be less likely to occur in short time windows around the intervention because it likely took some time for members to become aware of the nudge. As shown in Table 2, the results of our analysis are consistent for short and long windows. Second, we compared the second deals of newcomers (which were not treated by the nudge) posted shortly after the first deal. If a compositional change had occurred, these deals would likely have been different because they would have come from newcomers with different characteristics. We restricted our sample to deals posted within one week after the first deal. We chose a short, one-week window to ensure that the second deals were less influenced by newcomer learning. The variable *SecondDeal* equals 1 for a newcomer's second deal posted between 1 and 8 days after the first deal.[8] We removed observations with second deals posted after the policy change but the first deals posted before. We also removed the first newcomer deals to prune the impact of the nudge in this analysis. The results in Table 4 indicate that the second deals did not receive more comments or have different sentiments after the change (even columns). This finding suggests that there is no evidence of compositional change—i.e., the newcomers before and after the policy change did not seem to differ in terms of characteristics.

Robustness Checks

Table 5 reports the robustness checks. In the odd columns, we show that our results are robust after removing deals posted by hyperactive members whose number of deals was more than three standard deviations (*SD*) above the mean (mean = 2.68, *SD* = 9.88). In the even columns, we included deals posted by deleted, banned, or employed members (Appendix A). We excluded *Tenure* as a control because it is only available for members with active profiles during data

[8] We excluded second deals posted within one day of the first deal because we found a number of duplicates or near duplicates among those deals (e.g., in-store promotion of the same local store). Since they often receive fewer comments or are marked as "expired" sooner, including such entries might have introduced noise to our estimation.

collection. For Column 9, we used the percentage of negative words as an alternative operationalization of sentiment (Shen et al., 2015).[9] All of these estimations produced results consistent with H1 and H2—i.e., the nudge aroused more responses and more positive sentiment on the newcomer deals.[10]

In Appendix C, we show that existing members changed their behavior because of the anticipatory excuse provided by the platform rather than newcomers. We made this inference by leveraging the fact that some contributors revealed their newcomer status themselves and asked for forgiveness when posting the deals. We found that the nudge had a stronger influence than newcomer self-disclosure—i.e., the nudge had a robust positive effect on the number of comments and their sentiment even after controlling for the dissemination of the anticipatory excuse by newcomers themselves.

Effect of the Nudge on Socialization Outcomes

To test H3 and H4a-H4b, we considered the effect of the nudge on newcomer retention and newcomer contribution quality. We modified Equation (1) and dropped the deal characteristics (X_i) and hour dummies of the first deal (*Hour*) because they are unlikely to account for differences in continuous user engagement. In addition to *Tenure*, which was used in Equation (1), we included a set of badges that users had earned before the focal deal to better capture users' motivation to contribute:

$$y_i = \beta_0 + \beta_1 Newcomer_i + \beta_2 Newcomer_i \times After_i + \beta_3 Tenure_i + \gamma_1 Badges_i + \gamma_2 Day_i + \varepsilon_i, \quad (3)$$

where y_i denotes retention or the quality of contributions. In contrast to the main analysis, we modified the control group to capture changes at the user level. Specifically, because non-newcomers may have posted multiple deals in each period, we only selected each non-newcomer's first post in the pre-nudge and post-nudge periods. We considered these deals the "first deals" of non-newcomers and used their posting dates as the start of the 12-month time frame.

Retention

To test H3, we examined the effect of the nudge on retention, measured by a binary indicator, *DealPosted*, of whether a user posted another deal within 12 months after the first deal. We estimated the effect of the policy change on this outcome using a linear probability model (LPM). The results are shown in Column 1 of Table 6. Because we found that newcomers in the post-nudge period were significantly more likely to post a deal in the 12 months after the first deal, H3 is supported. On average, the nudge increased newcomer retention by 3.7 percentage points compared to non-newcomers. Over the pre-nudge period, the probability of a newcomer returning within 12 months was 38%. We related the DID coefficient to the baseline probability by dividing 0.037 by 0.38, which suggests a change of 9.7%. In Appendix D, we present the same analysis for (1) the volume of deals and (2) comments. The results are consistent.[11]

Table 5. Robustness Checks

	log(1+*NumComments*)		*Positive*		*Neutral*		*Negative*		*PercNegWords*
	Outliers removed	All users included	Outliers removed	All users included	Outliers removed	All users included	Outliers removed	All users included	Main model
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Newcomer	-0.376***	-0.443***	-0.001	-0.007*	-0.003	-0.002	0.004	0.010**	0.312***
	(0.027)	(0.033)	(0.005)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.088)
Newcomer × After	0.445***	0.447***	0.012**	0.012**	-0.006	-0.003	-0.006	-0.009*	-0.267**
	(0.032)	(0.033)	(0.006)	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)	(0.106)
Time FE (day, hour)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Category FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33,541	52,365	32,179	50,350	32,179	50,350	32,179	50,350	39,307
Adjusted *R*-squared	0.148	0.134	0.030	0.030	0.099	0.104	0.062	0.066	0.010

Note: FE = fixed effects. Robust standard errors clustered by user are in parentheses. Control variables include log(1+*DescLen*), *LocalDeal*, *NumCategories*, and *Content*. In addition, odd columns include *Tenure* and Columns 3-8 include *AvgCommentLen*. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

[9] We describe the derivation of percentage of negative words, *PercNegWords*, in Appendix B.

[10] In Table B2 in Appendix B, we show that the results are largely consistent when using a dictionary-based sentiment analysis on a subset of the comments translated into English.

[11] The results in Table 6 are robust to using a numerical variable, *NumPriorComments*, instead of *BadgeComment*. *NumPriorComments* denotes the number of comments posted by a user prior to posting deal *i*.

Table 6. Test of H3 and H4a-H4b: Effect of Nudge on Retention, Quality, and Motivation

	Retention	Quality				Motivation	
	DealPosted	*ΔDealTemp*	*AvgDealTemp*	AvgPriceComp	AvgDiscount	AvgDescLen	DaysSecDeal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Newcomer	-0.230***	0.048	-0.557***	-0.005	-0.003	-0.077***	8.974**
	(0.013)	(0.200)	(0.125)	(0.016)	(0.011)	(0.028)	(3.810)
Newcomer × After	0.037**	-0.695***	0.185	-0.011	-0.011	-0.006	2.994
	(0.016)	(0.235)	(0.154)	(0.020)	(0.014)	(0.036)	(4.929)
Time FE (day)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	19,153	11,757	11,757	11,757	5,205	11,757	11,757
Adjusted *R*-squared	0.114	0.017	0.033	0.003	0.014	0.034	0.044

Note: Δ*DealTemp*, *AvgDealTemp*, and *AvgDescLen* are log-transformed. Control variables include *Tenure*, *BadgeDeal*, *BadgeComment*, and *BadgeVote*. Robust standard errors clustered by user are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Contribution Quality

To test the competing H4a and H4b, we analyzed how the quality of the subsequent contributions of newcomers changed compared to their first contribution. We considered the change in quality, *ΔDealTemp*, using the sample of newcomers who posted another deal within 12 months after the first deal. We observed a statistically significant effect for *ΔDealTemp* (1.152 vs. 0.049, $t(2{,}024) = 5.238$, $p < 0.001$). Before the nudge, newcomers' subsequent deals received, on average, more upvotes than the first deal, indicating that newcomers improved over time. Surprisingly, after the intervention, *ΔDealTemp* was almost zero and much lower than in the pre-nudge period. We formally conducted the analysis including non-newcomers who posted another deal within 12 months as a control group in estimating Equation (3). The results in Column 2 of Table 6 show that the newcomers' second deal indeed had a lower temperature than their first deal, relative to non-newcomers after the nudge. What caused such a relative drop in the quality of the subsequent deal?

One explanation for the decline in *ΔDealTemp* is that without the nudge, existing members were less likely to discount the role of ability. Hence, they were more critical of newcomers' first deals, leading to the lower temperature of such deals and hence the larger *ΔDealTemp* before the nudge. This effect, due to anticipatory excuses preceding newcomers' first deals, would be absent starting from the second deal onwards. Accordingly, we expect the temperature of deals posted after the first deal, *AvgDealTemp*, to be similar before and after the nudge. Column 3 of Table 6 shows that *AvgDealTemp* indeed remained unchanged. Furthermore, Columns 4 and 5 show no significant differences using alternative measures of deal quality, i.e., the likelihood of users mentioning a price comparison in their deal description (*AvgPriceComp*) and the average percentage discount (*AvgDiscount*). Collectively, these results do not support H4a or H4b. Instead, the net quality of subsequent deals by newcomers is similar before and after the nudge, supporting the explanation that the change in *ΔDealTemp* can be attributed to the absence of the nudge on subsequent deals. In Appendix E, we offer qualitative evidence in support of this explanation.

We now explore other explanations for why newcomers did not surpass the quality of their first deals. The lenient feedback induced by the nudge may have suppressed the motivation of newcomers to learn—since they did not have to work hard to get accepted into the group, they devoted less effort to subsequent deals. To identify a reduction in the motivation of newcomers after the nudge, we compared the average deal description length of the subsequent deals (*AvgDescLen*) and the time gap between the first and second deal (*DaysSecDeal*). The former reflects the effort put into subsequent deals. The latter indicates newcomers' general level of motivation to contribute. Columns 6 and 7 of Table 6 show no significant coefficients for the DID estimators of *AvgDescLen* and *DaysSecDeal*, suggesting that their effort and motivation had not changed.

The lenient behavior toward newcomers might have reduced the information quality of the comments by causing newcomers to learn less, reducing their chances of translating their experience into more successful posts in the future. We used several machine learning classifiers to analyze how the helpfulness, usefulness, and informativeness of the comments changed after the nudge (Appendix F). The results suggest that the nudge did not reduce the percentage of helpful, useful, or informative comments on newcomer deals. Furthermore, we tested whether newcomers in greater need of learning, e.g., those with a short tenure or few prior comments at the time of their first post, experienced a more pronounced decline in *ΔDealTemp* between their first and subsequent deals. We found that the decline exists for both experienced and inexperienced newcomers (Table 7). These results suggest that learning suppression is unlikely to explain our findings.

Table 7. Change in Deal Temperature by Newcomer Experience

	(1) Low-experience newcomers			(2) High-experience newcomers		
Measure	**Pre**	**Post**	***t*-statistic**	**Pre**	**Post**	***t*-statistic**
Cumulative Comments						
ΔDealTemp (log)	1.251	-0.052	-3.702***	1.080	0.121	-3.704***
Observations	315	531		431	747	
Tenure						
ΔDealTemp (log)	1.045	0.007	-3.317***	1.248	0.093	-4.703***
Observations	352	651		394	627	

Note: The sample was split by the median into low- and high-experience newcomers. *** $p < 0.01$.

Implications and Conclusion

Online communities face high turnover, particularly among newcomers. This paper is one of the first empirical studies on how an exogenous shock in existing members' behavior affects newcomer socialization outcomes in a deal-sharing community. By exploiting a natural experiment, we show that an intervention that proactively reminded people to be more considerate of newcomers caused newcomer deals to receive more comments (H1) with a more positive sentiment (H2). Consistent with H3, we found that newcomers were more likely to post another deal after the nudge, suggesting improved newcomer retention. However, since the nudge did not affect the quality of newcomers' subsequent contributions, neither H4a nor H4b is supported.

Community Response and Socialization Outcomes

The positive impact of the intervention on retention suggests that interacting with other members positively reinforces continued participation. We found that newcomers in the post-nudge period were 10% (4 percentage points) more likely to post a deal in the 12 months following their first deal. How does this effect compare with other interventions? Two recent interventions on Wikipedia serve as good references. Gallus (2017) found a 13% (4 percentage points) increase in retention in the month after newcomers received a symbolic award. Li et al. (2020) found that newcomers who edited Wikipedia as part of the WikiEd program had a 51.2% reduction in the risk of dropping out one year after the end of the course compared to editors in the matched control group. However, the difference was only 2.1 percentage points due to the low probability that users would still be editing after one year (2.1% in the control group, 4.2% in the treatment group). Obviously, the contexts and time windows of these studies are different from ours. However, these interventions aimed at newcomers appear to have a marginal effect that is similar to our study, which is aimed at existing members. We believe encouraging existing members to be more friendly is a promising low-cost alternative strategy to retain newcomers in an online community.

We also compared our effect size with an observational study that examined the role of insiders in the socialization process. Joyce and Kraut (2006) found that newcomers who received a response were 12 percentage points more likely to post in the community again. The coefficient is about three times larger than our estimate obtained from an exogenous shock (0.124 compared with 0.037).[12] This discrepancy could arise from endogenous responses—i.e., some newcomers may have a stronger propensity for remaining in a community and interacting with insiders. Our setting of an exogenous natural experiment better controls for such endogenous responses.

Our finding informs the broader tension of whether active and committed community members are born or made—particularly through their interaction with existing members (e.g., Panciera et al., 2009). We contribute new empirical evidence that feeling socially accepted by insiders can make newcomers more likely to return, regardless of their intrinsic propensity to participate. However, this effect is likely to be smaller than that reported in observational studies.

Despite better retention, the initial interaction does not necessarily affect the quality of subsequent contributions. Studies have suggested that newcomers learn through both positive reinforcement (Cable & Parsons, 2001) and negative experiences (Wilhelm et al., 2019). Our results suggest that positive reinforcement does not enhance the quality of newcomers' subsequent contributions. Further analysis in Appendix F shows that while existing members became nicer after the nudge, they did not provide more task-relevant knowledge in their comments. We cannot ascertain if this absence of task-relevant knowledge is the primary cause for the lack of quality improvements, but it seems to be a tenable

[12] Joyce and Kraut (2006, p. 737) state that the coefficient 0.124 corresponds to an increase of 12.4%. The coefficient they obtain using the *dprobit* function in Stata is commonly interpreted as 12.4 percentage points because it represents the marginal effect on the probability of posting again. They mention that "39% of those who failed to receive a reply posted again over the next three months." Thus, we interpret the increase as 12.4 percentage points or 32% (0.124 divided by 0.39).

explanation, as receiving nicer comments means that newcomers may not be further motivated to learn. We suggest that future research should explore whether task-relevant knowledge can help enhance the long-term contribution quality of newcomers.

If task-relevant knowledge can indeed help newcomers enhance their learning and quality, then platform owners should consider how to design interventions to feed such knowledge to newcomers. For example, platform owners could combine a nudge with formal socialization tactics, such as collective socialization (Li et al., 2020). While we found that positive reinforcement encourages newcomers to contribute but does not improve the quality of their contributions, it is possible more extensive treatment could lead to quality improvements. Therefore, platform owners might consider extending the nudge intervention to newcomers' contributions posted within a certain period of time instead of restricting it to the first post only. This would strengthen the positive reinforcement effect and increase the likelihood of creating a lasting impact by fostering a more conducive environment for newcomers to learn to improve their contributions.

Revealing Newcomers in Online Communities

Attribution theory and research on anticipatory excuses suggest that the newcomer nudge may encourage insiders to discount the role of ability if they learn about newcomers' status in online social interactions. Prior research on anticipatory excuses has tested their effectiveness in offline social interactions—for example, using badges or corporate uniforms (Flacandji et al., 2023; Greenberg, 1996). We show that revealing the newcomer status through a nudge in online communities, where interactions are arguably less personal, may serve as a powerful signal for existing members to treat inexperienced individuals more kindly. Interestingly, we found that this effect is stronger if the platform flags the newcomers than it is if newcomers flag themselves. Prior research has not documented any differences between newcomer revelations through self-identification (e.g., in a conversation) versus a standardized badge provided by the organization (e.g., Flacandji et al., 2023). We suspect that this finding is unique to online interactions because observers might find it difficult to judge the credibility of information shared by contributors when they lack reliable social cues (e.g., body language) to verify contributors' claims (Ma & Agarwal, 2007). This distinction has important theoretical implications; it suggests that relying on insiders' intrinsic interest in grooming newcomers may not be as effective in online communities. To better model the newcomer socialization process, we need to establish the theoretical merits of parental measures, such as the newcomer nudge, with the instructional nature as part of the socialization strategy of digital platforms.

Practically, the effectiveness of the newcomer nudge suggests that a simple behavioral intervention can produce significant impacts on the receiving parties (e.g., Gallus, 2017). By influencing the tone of user-generated content, the nudge can complement downstream content moderation initiatives (e.g., Jiang et al., 2023)—if the content submitted to the community is less toxic toward newcomers, platforms can conserve more resources to filter other problematic comments. The newcomer nudge may be especially useful when organizational or community practices or norms are buried in a large repository of information or when tacit knowledge is commonplace in the community. This may particularly be the case for online social networks that focus on knowledge exchange and dissemination. For example, Stack Overflow has introduced a policy similar to the newcomer nudge that flags contributions from new users, arguing that "there are just *too many* nuances to how the system works ... ; we need a safety net" (Post, 2018).

Generalizability to Other Communities

We conclude this paper by discussing the generalizability of its findings to other communities. Although we offer evidence from a deal-sharing community, we believe that our findings are applicable to communities intended for information exchange (see Ridings & Gefen, 2004) and communities requiring contributors to follow specific policies to participate (Kraut et al., 2012). For example, when asking a debugging question on Stack Overflow, users should include a minimal workable example so that other users can reproduce the problem. In such an environment, a newcomer nudge would likely be effective because it would push other members to correct errors or answer questions that they might otherwise ignore. By contrast, our findings may not generalize to communities intended for social support or friendship, such as health support groups or online friendship networks. If people join a community to network with others facing similar situations or to gain emotional support, the community may already be a place where members are inclined to show pro-social behaviors regardless of whether the platform tells them to be nice. In such cases, an intervention that provides protection may be neither necessary nor effective to retain new users. Overall, we encourage future research to replicate this study in different contexts to scrutinize the boundaries of our findings.

Acknowledgments

We thank the senior editor Wai Fong Boh, the associate editor, and the anonymous reviewers for their many comments and helpful suggestions. We also thank participants in seminars and conferences at the University of Cologne, Technical University of Darmstadt, University of Mannheim, 2021 ISMS Marketing Science Conference, City University of Hong Kong, Southern University of Science and Technology, Hong Kong University of Science and Technology, HEC

Paris, Technical University of Munich, Erasmus University, and Tilburg University. We thank Mark-Philipp Decker, Janina Dutzig, Nico Link, and Klara Roth for their helpful research assistance. The work described in this paper was supported by a grant from the Germany/Hong Kong Joint Research Scheme sponsored by the Research Grants Council of Hong Kong and the German Academic Exchange Service (G-HKUST602/22). Florian Pethig received support from the Joachim Herz Foundation.

References

Ahuja, M. K., & Galvin, J. E. (2003). Socialization in virtual groups. *Journal of Management*, *29*(2), 161-185. https://doi.org/10.1016/S0149-2063(02)00213-1

Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *The Journal of Abnormal and Social Psychology*, *59*(2), 177-181. https://doi.org/10.1037/h0047195

Arslan, R. C., Walther, M. P., & Tata, C. S. (2020). formr: A study framework allowing for automated feedback generation and complex longitudinal experience-sampling studies using R. *Behavior Research Methods*, *52*(1), 376-387. https://doi.org/10.3758/s13428-019-01236-y

Athey, S., & Imbens, G. W. (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica*, *74*(2), 431-497. https://doi.org/10.1111/j.1468-0262.2006.00668.x

Black, E. (2019, July 18). Re: What a very bad day at work taught me about building Stack Overflow's community [Blog comment]. *Stack Overflow Blog.* https://stackoverflow.blog/2019/07/18/building-community-inclusivity-stack-overflow/#comment-243110

Burke, M., Kraut, R., & Joyce, E. (2010). Membership claims and requests: Conversation-level newcomer socialization strategies in online groups. *Small Group Research*, *41*(1), 4-40. https://doi.org/10.1177/1046496409351936

Cable, D. M., & Parsons, C. K. (2001). Socialization tactics and person-organization fit. *Personnel Psychology*, *54*(1), 1-23. https://doi.org/10.1111/j.1744-6570.2001.tb00083.x

Chen, Y., Harper, F. M., Konstan, J., & Li, S. X. (2010). Social comparisons and contributions to online communities: A field experiment on movielens. *American Economic Review*, *100*(4), 1358-1398. https://doi.org/10.1257/aer.100.4.1358

Clarke, J., Chen, H., Du, D., & Hu, Y. J. (2020). Fake news, investor attention, and market reaction. *Information Systems Research*, *32*(1), 35-52. https://doi.org/10.1287/isre.2019.0910

Delin, C. R., & Baumeister, R. F. (1994). Praise: More than just social reinforcement. *Journal for the Theory of Social Behaviour*, *24*(3), 219-241. https://doi.org/10.1111/j.1468-5914.1994.tb00254.x

Eckles, D., Karrer, B., & Ugander, J. (2017). Design and analysis of experiments in networks: Reducing bias from interference. *Journal of Causal Inference*, *5*(1), Article 20150021. https://doi.org/10.1515/jci-2015-0021

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. Appleton-Century-Crofts. https://doi.org/10.1037/10627-000

Flacandji, M., Cusin, J., & Lunardo, R. (2023). When and why signaling frontline employee inexperience can prove to be an asset: Effects on consumer forgiveness for service failure. *Psychology & Marketing*, *40*(12), 2728-2742. https://doi.org/10.1002/mar.21897

Foerderer, J., Kude, T., Mithas, S., & Heinzl, A. (2018). Does platform owner's entry crowd out innovation? Evidence from Google Photos. *Information Systems Research*, *29*(2), 444-460. https://doi.org/10.1287/isre.2018.0787

Gallus, J. (2017). Fostering public good contributions with symbolic awards: A large-scale natural field experiment at Wikipedia. *Management Science*, *63*(12), 3999-4015. https://doi.org/10.1287/mnsc.2016.2540

Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review*, *25*(2), 161-178. https://doi.org/10.2307/2092623

Greenberg, J. (1996). "Forgive me, I'm new": Three experimental demonstrations of the effects of attempts to excuse poor performance. *Organizational Behavior and Human Decision Processes*, *66*(2), 165-178. https://doi.org/10.1006/obhd.1996.0046

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, *21*(9), 1263-1284. https://doi.org/10.1109/TKDE.2008.239

Higgins, R. L., & Snyder, C. R. (1989). The business of excuses. In R. A. Giacalone & P. Rosenfeld (Eds.), *Impression management in the organization* (pp. 73-86). Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203763018

Jiang, L. D., Ravichandran, T., & Kuruzovich, J. (2023). Review moderation transparency and online reviews: Evidence from a natural experiment. *MIS Quarterly*, *47*(4), 1693-1708. https://doi.org/10.25300/MISQ/2023/16216

Jo, W., Sunder, S., Choi, J., & Trivedi, M. (2020). Protecting consumers from themselves: Assessing consequences of usage restriction laws on online game usage and spending. *Marketing Science*, *39*(1), 117-133. https://doi.org/10.1287/mksc.2019.1174

Jones, E. E., Kanouse, D. E., Kelley, H. H., Nisbett, R. E., Valins, S., & Weiner, B. (1987). *Attribution: Perceiving the causes of behavior*. Lawrence Erlbaum Associates. https://doi.org/10.2307/1959677

Joyce, E., & Kraut, R. E. (2006). Predicting continued participation in newsgroups. *Journal of Computer-Mediated Communication*, *11*(3), 723-747. https://doi.org/10.1111/j.1083-6101.2006.00033.x

Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, *28*(2), 107-128. https://doi.org/10.1037/h0034225

Kraut, R. E., Burke, M., Riedl, J., & Resnick, P. (2012). The challenges of dealing with newcomers. In R. E. Kraut & P. Resnick (Eds.), *Building successful online communities: Evidence-based social design* (1st ed., pp. 179-230). MIT Press. https://doi.org/10.7551/mitpress/8472.001.0001

Lampe, C., & Johnston, E. (2005). Follow the (slash) dot: Effects of feedback on new members in an online community. In *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work* (pp. 11-20). https://doi.org/10.1145/1099203.109920

Li, A., Yao, Z., Yang, D., Kulkarni, C., Farzan, R., & Kraut, R. E. (2020). Successful online socialization: Lessons from the Wikipedia education program. In *Proceedings of the ACM on Human-Computer Interaction*. https://doi.org/10.1145/3392857

Liu, X., Wang, G. A., Fan, W., & Zhang, Z. (2020). Finding useful solutions in online knowledge communities: A theory-driven design and multilevel analysis. *Information Systems Research*, *31*(3), 731-752. https://doi.org/10.1287/isre.2019.0911

Louis, M. R. (1980). Surprise and sense making: What newcomers experience in entering unfamiliar organizational settings. *Administrative Science Quarterly*, *25*(2), 226-251. https://doi.org/10.2307/2392453

Ma, M., & Agarwal, R. (2007). Through a glass darkly: Information technology design, identity verification, and knowledge contribution in online communities. *Information Systems Research*, *18*(1), 42-67. https://doi.org/10.1287/isre.1070.0113

Malik, N., Singh, P. V., & Srinivasan, K. (2023). When does beauty pay? A large-scale image-based appearance analysis on career transitions. *Information Systems Research*, *35*(4), 1524-1545. https://doi.org/10.1287/isre.2021.0559

Miller, V. D., & Jablin, F. M. (1991). Information seeking during organizational entry: Influences, tactics, and a model of the process. *Academy of Management Review*, *16*(1), 92-120. https://doi.org/10.5465/amr.1991.4278997

Narayan, S., Orlowitz, J., Morgan, J., Hill, B. M., & Shaw, A. (2017). The Wikipedia adventure: Field evaluation of an interactive tutorial for new users. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 1785-1799). https://doi.org/10.1145/2998181.2998307

Natterson-Horowitz, B., & Bowers, K. (2020). *Wildhood: The astounding connections between human and animal adolescents*. Scribner.

Pallas, F., Staufer, D., & Kuhlenkamp, J. (2020). Evaluating the accuracy of cloud NLP services using ground-truth experiments. In *Proceedings of the IEEE International Conference on Big Data* (pp. 341-350). https://doi.org/10.1109/BigData50022.2020.9378188

Panciera, K., Halfaker, A., & Terveen, L. (2009). Wikipedians are born, not made: A study of power editors on Wikipedia. In *Proceedings of the 2009 ACM International Conference on Supporting Group Work* (pp. 51-60). https://doi.org/10.1145/1531674.1531682

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830. http://jmlr.org/papers/v12/pedregosa11a.html

Phang, C. W., Kankanhalli, A., & Tan, B. C. Y. (2015). What motivates contributors vs. lurkers? An investigation of online feedback forums. *Information Systems Research*, *26*(4), 773-792. https://doi.org/10.1287/isre.2015.0599

Post, T. (2018). *Come take a look at our new contributor indicator!* Meta Stack Exchange. https://meta.stackexchange.com/questions/314287

Punyon, J. & Montrose, K. (2020). *The unfriendly robot: Automatically flagging unwelcoming comments.* The Stack Overflow blog. https://stackoverflow.blog/2020/04/09/the-unfriendly-robot-automatically-flagging-unwelcoming-comments/

Ransbotham, S., & Kane, G. C. (2011). Membership turnover and collaboration success in online communities: Explaining rises and falls from grace in Wikipedia. *MIS Quarterly*, *35*(3), 613-627. https://doi.org/10.2307/23042799

Ren, Y., Chen, J., & Riedl, J. (2016). The impact and evolution of group diversity in online open collaboration. *Management Science*, *62*(6), 1668-1686. https://doi.org/10.1287/mnsc.2015.2178

Ren, Y., Harper, F. M., Drenner, S., Terveen, L., Kiesler, S., Riedl, J., & Kraut, R. E. (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

Ren, Y., Zhang, H., & Kraut, R. E. (2023). How did they build the free encyclopedia? A literature review of collaboration and coordination among Wikipedia editors. *ACM Transactions on Computer-Human Interaction*, *31*(1), 1-48. https://doi.org/10.1145/3617369

Ridings, C. M., & Gefen, D. (2004). Virtual community attraction: Why people hang out online. *Journal of Computer-Mediated Communication*, *10*(1). https://doi.org/10.1111/j.1083-6101.2004.tb00229.x

Rosenbaum, P. R. (2007). Interference between units in randomized experiments. *Journal of the American Statistical Association*, *102*(477), 191-200. https://doi.org/10.1198/016214506000001112

Shen, W., Hu, Y. J., & Ulmer, J. R. (2015). Competing for attention: An empirical study of online reviewers' strategic behavior. *MIS Quarterly*, *39*(3), 683-696. https://doi.org/10.25300/MISQ/2015/39.3.08

Sidarenka, U., & Stede, M. (2016). Generating Sentiment Lexicons for German Twitter. In *Proceedings of the Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media* (pp. 80-90).

Snyder, C., & Higgins, R. L. (1988). From making to being the excuse: An analysis of deception and verbal/nonverbal issues. *Journal of Nonverbal Behavior*, *12*, 237-252. https://doi.org/10.1007/BF00987594

Tausczik, Y., Farzan, R., Levine, J., & Kraut, R. (2018). Effects of collective socialization on newcomers' response to feedback in online communities. *ACM Transactions on Social Computing*, *1*(2), 8:1-8:23. https://doi.org/10.1145/3191834

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness.* Yale University Press.

Verplanck, W. S. (1955). The control of the content of conversation: Reinforcement of statements of opinion. The *Journal of Abnormal and Social Psychology*, *51*(3), 668-676. https://doi.org/10.1037/h0046514

Wenninger, H., Krasnova, H., & Buxmann, P. (2019). Understanding the role of social networking sites in the subjective well-being of users: A diary study. *European Journal of Information Systems*, *28*(2), 126-148. https://doi.org/10.1080/0960085X.2018.1496883

Wilhelm, H., Richter, A. W., & Semrau, T. (2019). Employee learning from failure: A team-as-resource perspective. *Organization Science*, *30*(4), 694-714. https://doi.org/10.1287/orsc.2018.1255

Yan, B., & Jian, L. (2017). Beyond reciprocity: The bystander effect of knowledge response in online knowledge communities. *Computers in Human Behavior*, *76*, 9-18. https://doi.org/10.1016/j.chb.2017.06.040

Yin, D., Bond, S. D., & Zhang, H. (2014). Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. *MIS Quarterly*, *38*(2), 539-560. https://doi.org/10.25300/MISQ/2014/38.2.10

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. https://doi.org/10.1287/isre.2013.0485

About the Authors

Florian Pethig is an assistant professor in the Department of Information Systems and Operations Management at Tilburg School of Economics and Management, Tilburg University. He received a Ph.D. in information systems from the University of Mannheim. His research interests include digital platforms, user-generated content, and the societal impact of information technology. His work has appeared or is forthcoming in *MIS Quarterly*, *Journal of the Association for Information Systems*, and *Journal of Business Ethics*. https://orcid.org/0000-0002-6001-2391

Hartmut Hoehle is a professor and Chair of Enterprise Systems at the University of Mannheim, Germany. He received a Ph.D. in information systems from Victoria University of Wellington, New Zealand. His research interests include the design, implementation, and use of enterprise systems. Stemming from his professional experiences gained while working at Deutsche Bank, he is particularly interested in how services and products can be distributed through electronically mediated channels. His work has appeared in *MIS Quarterly*, *Journal of Management Information Systems*, *European Journal of Information Systems*, *Decision Support Systems*, *International Journal of Operations & Production Management*, and other journals. https://orcid.org/0000-0001-8117-0105

Kai-Lung Hui is Elman Family Professor of Business and Chair professor in the Department of Information Systems, Business Statistics, and Operations Management in the School of Business and Management, Hong Kong University of Science and Technology. His research interests include cybercrime economics and policy, information privacy, IT policy, and electronic commerce. His research has been published in scholarly journals including *Information Systems Research*, *Management Science*, *MIS Quarterly*, and *Journal of MIS*, among others. He received his BBA and Ph.D. degrees from the Hong Kong University of Science and Technology. https://orcid.org/0000-0002-7074-1176

Andreas Lanz is an assistant professor of digital marketing analytics in the Faculty of Business and Economics at the University of Basel. Andreas researches social networks and is recognized for his award-winning work on the effectiveness of micro-influence for dissemination. He has published in leading journals including the *Journal of Marketing*, *Journal of Marketing Research*, *MIS Quarterly*, and *Quantitative Marketing and Economics*. https://orcid.org/0000-0003-2643-1960

Appendix A

Data Collection and Preparation

We constructed our data in two rounds. In Round 1, between August and October 2018, we retrieved all deals that were available on mydealz, and our analysis of these deals revealed that the platform broadly adopted the newcomer nudge on October 20, 2016. Initially, we faced the challenge of identifying *exactly* the first deal posted by a new member before the policy change because, apart from the nudge, the first deals did not differ from any other deals. Fortunately, after our first round of data collection, mydealz updated old newcomer deals to show the nudge. In Round 2, in September 2019, we collected the deals again and combined the results of both rounds.

Table A1 outlines our 180-day sample of the two rounds of data collection, divided into six aggregated 30-day intervals. As shown in Table A1, the platform started to gradually test the nudge before October 20. From July 22 to October 19, the nudge was available on fewer than 1% of deals (0.03% to 0.87%). After the intervention was widely introduced, the nudge was available on 11.29% to 12.85% of the deals. As expected, in Round 2, the proportion of deals with the nudge was comparable before and after the intervention. Interestingly, some deals had the nudge in Round 1 but not in Round 2, leading to the assumption that mydealz initially identified some deals as newcomer deals that, in fact, were not truly newcomer deals. Finally, the last column of Table A1 shows that 0.16% to 0.32% of the deals were removed by the platform after Round 1 because they were flagged as spam or duplicates (Figure A1).

To identify suitable newcomer deals (treatment group) and non-newcomer deals (control group), we modified our sample in several ways. First, we removed 119 deals that were removed from the platform after Round 1. Second, we restricted our sample to deals by posters who had not deleted their profile at the time of data collection, had not been banned for violating the rules, and were not employed by the community (e.g., as a deal-hunter, moderator, or administrator). Third, the platform had tested the nudge on select deals before its widespread introduction. We removed these deals from our main analysis to establish clean pre- and post-treatment periods across our treatment and control groups. Fourth, we only kept the newcomer deals for which the nudge was not removed before the second round of data collection to ensure that our treatment group consisted of legitimate newcomer deals. Table A2 shows our final sample, comprising 4,952 newcomer deals in the treatment group and 35,971 non-newcomer deals in the control group. The sum of the bold numbers in Columns 6 and 7 shows the number of newcomer deals, and the sum of Column 8 shows the number of non-newcomer deals.

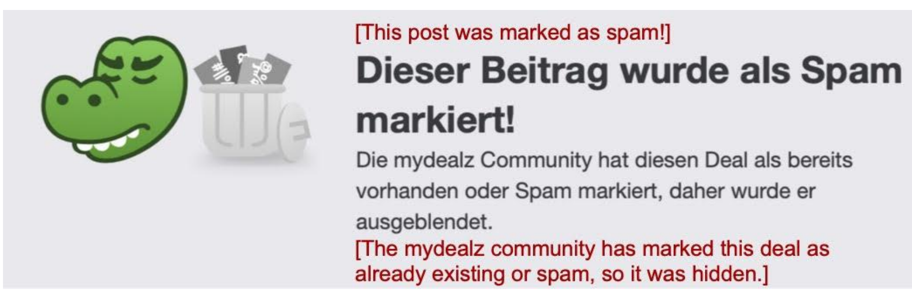


Figure A1. Deal Marked as Spam

Table A1. Data Collection Strategy

Time periods	Round 1 (R1)			Round 2 (R2)			% of deals removed
	Deals	Deals with nudge	% of deals with nudge	Deals	Deals with nudge	% of deals with nudge	
Jul 22 – Aug 20	7,420	2	0.03	7,406	621	8.39	0.19
Aug 21 – Sep 19	7,464	48	0.64	7,452	730	9.80	0.16
Sep 20 – Oct 19	7,930	69	0.87	7,913	799	10.10	0.21
Oct 20 – Nov 18	7,774	878	11.29	7,761	783	10.09	0.17
Nov 19 – Dec 18	13,055	1,677	12.85	13,013	1,542	11.85	0.32
Dec 19 – Jan 17	9,278	1,065	11.48	9,257	978	10.56	0.23

Note: Black Friday was on Nov 25, 2016. Many stores offer highly promoted sales on (and after) this day. Thus, 68% more deals were posted between Nov 19 and Dec 18 compared to the previous 30 days.

Table A2. Sample Selection Procedure

	Deal removed after R1?		Deal posted by deleted, banned, or employed member?		Deal has nudge?			
Time periods	**(1) Yes**	**(2) No**	**(3) Yes**	**(4) No**	**(5) R1 \ R2**	**(6) R2 \ R1**	**(7) R1 ∩ R2**	**(8) No**
Jul 22 – Aug 20	14	7,406	1,726	5,680	0	**584**	1	**5,095**
Aug 21 – Sep 19	12	7,452	1,623	5,829	21	**650**	23	**5,135**
Sep 20 – Oct 19	17	7,913	1,875	6,038	22	**703**	39	**5,274**
Oct 20 – Nov 18	13	7,761	1,833	5,928	86	4	**699**	**5,139**
Nov 19 – Dec 18	42	13,013	2,434	10,579	123	0	**1,409**	**9,047**
Dec 19 – Jan 17	21	9,257	1,992	7,265	76	1	**907**	**6,281**

Note: Numbers highlighted in bold represent the deals included in the main analysis. R1 \ R2 = Nudge present in R1 but not in R2. R2 \ R1 = Nudge present in R2 but not in R1. R1 ∩ R2 = Nudge present in R1 and R2.

Appendix B

Construction of Additional Variables

Price Comparison and Discount

We used the average likelihood with which users mention a price comparison in the descriptions of their subsequent deals, *AvgPriceComp*, and the average percentage discount, *AvgDiscount*, as alternative measures of the quality of subsequent deals. According to the official mydealz community guidelines,[13] deals should include a price comparison so that users can objectively assess their savings potential, allowing us to construct the aforementioned measures of deal quality: (1) if a price comparison is mentioned in the deal description, this indicates that the deal poster was aware of the guidelines and tried to adhere to them by addressing this aspect in his or her post; (2) if the comparison also includes the comparison price, this allows us to calculate the savings potential of a deal.[14] Unfortunately, for the deals in our study period, such data were not available in any kind of structured format. At that time, deals typically only had the discounted price at the top, and the comparison price, if available, was mentioned in the text (Figure B1). Only after our study period did mydealz make it mandatory to enter the comparison price along with the discounted price.

To address this limitation, we use a regular expression (regex) to extract the keywords price comparison (“Preisvergleich”), comparison price (“Vergleichspreis”), their abbreviations (“PVG”, “VGP”), and the names of two platforms for conducting a price comparison in Germany (“Idealo,” “Geizhals”) from the description of all deals. The regex patterns that we used are shown in Table B1. Of the 11,757 newcomers and non-newcomers who posted at least one additional deal within 12 months, 46% mentioned a price comparison in the deal description of the subsequent deals.

To extract the comparison price (and not just mentions of it), we relied on the same regex described above and extracted any euro amount (€|Euro|euro|EUR|eur) that was mentioned in the 100 characters trailing the occurrence of the word “price comparison” (or any variants thereof). The intuition for this approach is demonstrated in Figure B1—the word “price comparison” is typically followed by the actual amount (both circled in red). We divided the deal price (available in a structured form from the field at the top) by the comparison price to obtain the discount percentage. The resulting average discount percentage for subsequent deals was 29.6%.[15]

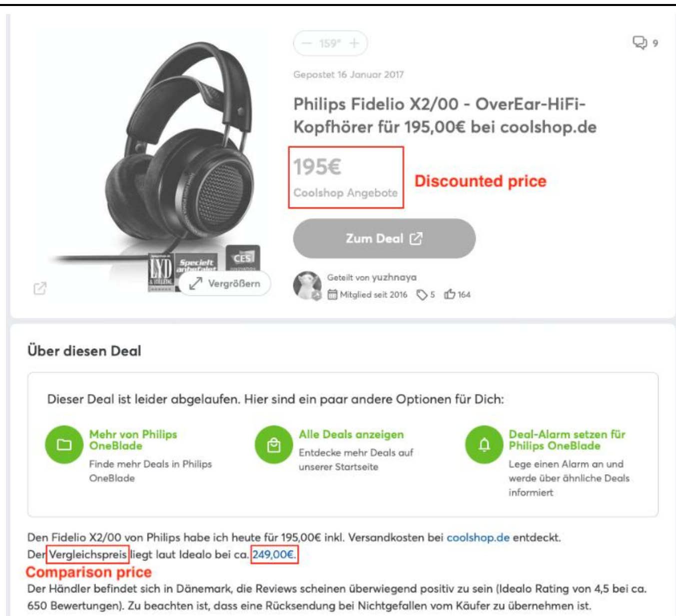


Figure B1. Exemplary Deal From January 2017

[13] See https://help.mydealz.de/help/wie-erstelle-ich-einen-deal (in German).
[14] Mentioning the price comparison does not automatically mean that the price is also mentioned. For example, a deal poster may mention a price comparison if it is not available, e.g., because the deal is a product available only in a certain shop.
[15] We set values to missing where the extracted comparison price was higher than the deal price.

Table B1. List of Regex Patterns to Filter Price Comparisons		
English keyword	**German keyword (abbreviation)**	**Regular expression**
Price comparison	Preisvergleich (PVG)	(p|P)reisvergleich|(p|P)(v|V)(g|G)
Comparison price	Vergleichspreis (VGP)	(v|V)ergleichspreis|(v|V)(g|G)(p|P)
Next best price	Nächster Preis	(n|N)ächste(r?)(p|P)reis
Idealo	Idealo	(i|I)dealo|IDEALO
Geizhals	Geizhals	(g|G)eizhals|GEIZHALS

Percentage of Negative Words

We used the percentage of negative words, *PercNegWords*, as an alternative measure of the sentiment of comments posted for a given deal (e.g., Shen et al., 2015). We tokenized each comment, removed punctuation, and lowercased and matched the words with the negative word list of SentiWS (Remus et al., 2010), a popular German sentiment lexicon, which has been shown to perform particularly well for negative words (Sidarenka & Stede, 2016).[16] To construct the measure of *PercNegWords* at the deal level, we divided the sum of negative words across all comments for deal *i* by the sum of all words across all comments for deal *i*.

Dictionary-Based Sentiment Score

As another robustness check, we conducted a dictionary-based sentiment analysis on an English version of the comments. We used the "deep_translator" package in Python, which provided us with access to the Google Translate API, to translate all comments within 30 days before and after the policy change from German to English.[17] We applied the "sentiment" package in R on the translated comments and obtained one sentiment score for each sentence.[18] We averaged the scores at the deal level to obtain an aggregated score, *SentScore*, and reran Equation (1). The results shown in Table B2 are largely consistent, but smaller in magnitude.

Table B2. Results of the Dictionary-Based Sentiment Analysis			
	SentScore		
	±3 Days	**±5 Days**	**±30 Days**
Newcomer	0.009 (0.014)	-0.001 (0.012)	0.001 (0.006)
Newcomer × After	0.013 (0.021)	0.031* (0.018)	0.012* (0.007)
Tenure	0.000 (0.000)	0.000 (0.000)	0.000** (0.000)
log(1+*DescLen*)	0.002 (0.005)	0.006* (0.004)	0.007*** (0.001)
LocalDeal	0.013 (0.011)	0.016* (0.009)	0.005* (0.003)
NumCategories	0.000 (0.002)	0.000 (0.001)	0.000 (0.001)
Content	0.000 (0.022)	0.015 (0.016)	0.006 (0.005)
AvgCommentLen	0.001 (0.000)	0.000 (0.000)	0.000 (0.000)
Time FE (Day, Hour)	Yes	Yes	Yes
Category FE	Yes	Yes	Yes
Observations	1,065	1,748	11,234
Adjusted *R*-squared	-0.004	0.009	0.006

Note: FE = fixed effects. Robust standard errors clustered by user are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

[16] We added the English word "cold" to the SentiWS negative word list because the word is frequently used to label bad ("cold") deals.
[17] See https://pypi.org/project/deep-translator/.
[18] See https://cran.r-project.org/web/packages/sentimentr/sentimentr.pdf. We manually change the valence of the term "hot" from -0.25 to 0.5 because it reflects a positive sentiment in the context of the mydealz community.

Appendix C

Newcomers Revealing Themselves vs. the Nudge

We separated the identification of the newcomer ("This is the first deal by [user]") versus the suggested treatment of the newcomer ("Help out by posting tips or just thank them for their deal"). We constructed a new independent variable *FirstDealMentioned* that was set to 1 if a newcomer mentioned that it was his or her first deal in the deal description of deal *i* and 0 otherwise.[19] In total, 651 newcomers (13%) disclosed their newcomer status. Thus, there were two treatments, *Newcomer* × *After* and *FirstDealMentioned*. Both variables show an impact if people changed their behavior because the poster was a newcomer. Only the interaction *Newcomer* × *After* (not *FirstDealMentioned*) would have an effect if people changed their behavior because of the platform's instruction instead of the newcomer's status.

The results in Column 1 of Table C1 show that the coefficient of *Newcomer* × *After* is positive and significant for the number of comments. The coefficient of *FirstDealMentioned* is also positive and significant but smaller in magnitude. This finding indicates that self-disclosure may also attract more comments, but to a lesser extent than the nudge. Thus, the nudge may have been more effective in changing behavior, suggesting that people change their behavior because of the platform's instruction. The interaction *FirstDealMentioned* × *After* is not significant, suggesting that self-disclosure in combination with the nudge did not affect the number of comments.

In Column 2, we observe evidence consistent with the idea that the nudge is a more powerful intervention than self-disclosure. Whereas *Newcomer* × *After* had a positive and significant effect on *Positive*, *FirstDealMentioned* and *FirstDealMentioned* × *After* did not. In Columns 3 and 4, we find that *FirstDealMentioned* is negatively related to *Neutral* and positively related to *Negative*, which indicates increased polarization. This effect disappears after the introduction of the nudge.

Overall, this analysis helps to separate the identification of newcomer deals and the suggested treatment of asking established members to be nice to newcomers. *FirstDealMentioned* discloses the newcomer status, but not the message of the nudge. The nudge combines the two. So, including *FirstDealMentioned* × *After* should tease out these two effects. *Newcomer* × *After* should then provide a more precise estimation of the nudging effect.

Table C1. Newcomers Revealing Themselves Versus the Nudge

	log(1+*NumComments*)	*Positive*	*Neutral*	*Negative*
	(1)	(2)	(3)	(4)
Newcomer	-0.381*** (0.030)	-0.004 (0.005)	0.001 (0.005)	0.003 (0.005)
Newcomer × *After*	0.441*** (0.035)	0.013* (0.006)	-0.009 (0.006)	-0.004 (0.005)
FirstDealMentioned	0.140** (0.071)	0.006 (0.012)	-0.025** (0.012)	0.019* (0.011)
FirstDealMentioned × *After*	-0.049 (0.087)	-0.005 (0.015)	0.026* (0.015)	-0.021 (0.013)
Tenure	0.003*** (0.001)	0.000** (0.000)	0.000 (0.000)	0.000*** (0.000)
log(1+*DescLen*)	0.158*** (0.012)	0.011*** (0.001)	-0.010*** (0.002)	-0.001 (0.001)
LocalDeal	-0.372*** (0.024)	-0.018*** (0.003)	0.035*** (0.003)	-0.017*** (0.002)
NumCategories	0.037*** (0.004)	-0.002*** (0.000)	0.002*** (0.000)	0.000 (0.000)
Content	-0.545*** (0.025)	0.002 (0.004)	0.000 (0.004)	-0.001 (0.004)
AvgCommentLen		0.001*** (0.000)	-0.003*** (0.000)	0.002*** (0.000)
Time FE (day, hour)	Yes	Yes	Yes	Yes
Category FE	Yes	Yes	Yes	Yes
Observations	40,923	39,307	39,307	39,307
Adjusted *R*-squared	0.139	0.031	0.102	0.064

Note: FE = fixed effects. Robust standard errors clustered by user are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

[19] We used the following regular expression to extract *FirstDealMentioned* from newcomer deals: (e|E)rste(r?)|1\.?) (d|D)eal (essentially matching "first deal" or "1st deal" in German). As self-disclosure of the newcomer status was only relevant to newcomers, we applied the regular expression only to newcomer deals and set *FirstDealMentioned* to 0 for all non-newcomer deals.}

Appendix D

Effect of Nudge on Alternative Retention Variables

We considered the effect of the nudge on alternative outcomes to measure the retention of newcomers. *NumDealsPosted* is a count measure of the number of deals posted by a user over the 12 months following the first deal. Column 1 of Table D1 shows that the nudge yields a 5% increase in the volume of deals by newcomers compared to non-newcomers. We also considered the effect of the nudge on commenting behavior. We used a binary indicator to determine whether a user posted any comments during the 12 months after the first deal, *CommentPosted*. The results using a linear probability model (LPM) are shown in Column 2 of Table D1. We show evidence that newcomers after the policy change were 7 percentage points more likely to post a comment compared to non-newcomers. We further differentiated whether the comment was posted to a deal posted by the commenter herself or by another community member. The former reflects a revisiting and refinement of their own content, whereas the latter reflects a shift to explore and discuss content generated by the community. These alternative dependent variables have appended suffixes: *Own* refers to a comment on a deal posted by the commenter and *Other* refers to a comment on a deal posted by another community member. In Columns 3 and 4 of Table D1, we present the results of regressions to estimate the decision to comment on a deal posted by the commenter herself, *CommentPostedOwn*, and a deal posted by another user, *CommentPostedOther*. The results indicate a positive effect on *CommentPostedOwn* and *CommentPostedOther*. However, the coefficient is larger for *CommentPostedOwn* than for *CommentPostedOther*. This indicates that newcomers were more likely to comment on their own content than to discuss others' content after the policy change.

Table D1. Effect of Nudge on Alternative Retention Variables

	log(1+*NumComments*)	***CommentPosted***	***CommentPostedOwn***	***CommentPostedOther***
	(1)	**(2)**	**(3)**	**(4)**
Newcomer	-0.398*** (0.021)	0.145*** (0.011)	-0.178*** (0.013)	-0.167*** (0.012)
Newcomer × After	0.046* (0.025)	0.074*** (0.013)	0.101*** (0.015)	0.050*** (0.014)
Tenure	-0.008*** (0.000)	0.001*** (0.000)	0.000** (0.000)	0.002*** (0.000)
BadgeVote	0.092*** (0.020)	0.048*** (0.004)	0.041*** (0.008)	0.086*** (0.006)
BadgeComment	0.167*** (0.021)	0.054*** (0.004)	0.124*** (0.009)	0.112*** (0.006)
BadgeDeal	0.681*** (0.023)	-0.007* (0.004)	0.070*** (0.008)	-0.020*** (0.005)
Time FE (day)	Yes	Yes	Yes	Yes
Observations	19,153	19,153	19,153	19,153
Adjusted *R*-squared	0.190	0.095	0.089	0.147

Note: FE = fixed effects. Robust standard errors clustered by user are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Appendix E

Qualitative Evidence

We interviewed 18 mydealz' members to corroborate our results.[20] The interviews were conducted in November and December 2021 and lasted 23 minutes on average. Participants were recruited through mydealz and deal-related Facebook groups. Six interviewees stated that they felt that the nudge would make them or other members more friendly toward newcomers[21] and that this would sometimes come at the cost of upvoting deals with minor flaws. For example, one interviewee stated:

> *So if it says it's a newcomer, then the user has a bit of a 'puppy license' and I try to take the newcomer somehow in protection and maybe, although it is not such a great deal, still vote hot.*

This notion of a "puppy license" (*Welpenschutz* in German) describes the special status of young puppies, i.e., a "leeway period granted by older members of the group" (Natterson-Horowitz & Bowers, 2020, p. 51). It exists for many animal species—and even humans—and helps new community members explore different behaviors without facing the same consequences as established members. One mydealz' editor also confirmed the efficacy of the nudge, when asked whether people would respond more positively:

> *Yes, definitely yes. So if the deal is really not good, then the user is also informed of that, but if there are minor errors, for example, a price comparison was forgotten, then it is usually just pointed out.*

This quote again confirmed a "lenient" period for newcomers due to the nudge with the result that newcomers are not appraised using the same standards as established members. The qualitative evidence underscores that established members may have been more lenient than justified by the behavior of newcomers. Newcomers would not have been able to achieve the same score in the absence of the nudge.

[20] The interview guide is available at https://osf.io/t7awu.
[21] One additional interviewee stated that she had never noticed the nudge but that she was more lenient toward newcomers who self-disclosed their status.

Appendix F

Detecting Friendly, Helpful, Useful, and Informative Comments

We explore how the content of the comments changed in response to the nudge. We trained several supervised learning algorithms to detect friendly, helpful, useful, and informative comments. To obtain reliable training data, we instructed three research assistants who are native German speakers to label 4,000 comments on the dimensions of friendliness, helpfulness, usefulness, and informativeness.[22] The research assistants were at least somewhat familiar with mydealz. Two had registered user accounts, and one was actively participating on the platform by posting deals and comments. All comments were rated on 9-point semantic differential scales adapted from Wenninger et al. (2019) and Yin et al. (2014). To ensure sufficient variation in the friendliness of the labeled comments, we randomly selected 1,000 comments from deals with a deal temperature less than or equal to zero (25%) and 3,000 comments from deals with a deal temperature above zero (75%). Although only 5.4% of the comments in the data set were from deals with a low deal temperature, highly imbalanced data can be a challenge for machine learning algorithms (He & Garcia, 2009). According to a recent analysis on Stack Overflow (Punyon & Montrose, 2020), only 0.78% of the comments on the platform were labeled as unfriendly and these comments were more likely to be written in response to low-quality questions. Thus, human annotators received more comments from deals with a low net score. We did not expect this choice to influence the variation in helpful, useful, and informative comments because they might have been written in response to a high- or low-quality deal. We also provided human annotators the option to flag suspicious comments, e.g., when they looked truncated or were not understandable.

The annotation process was implemented using formr, an online tool that produces surveys based on comma-separated values (CSV) files (Arslan et al., 2020). We split the annotation task into 20 surveys containing 200 comments each (20 per page) and randomized the order of the comments (per page) to mitigate response-order effects. To match annotations across the surveys, the annotators entered a self-generated identification code at the beginning of each survey. We removed all comments that any of the annotators flagged or did not rate, leaving 3,915 fully annotated comments. We averaged the scores of the annotators and rounded the values to the nearest integer. For our classification task, we were primarily interested in detecting changes in the distribution of (1) friendly comments because they might encourage newcomers to stay and (2) helpful, useful, or informative comments because they might convey important information. Thus, we collapsed the 9-point semantic differential scales into binary scales. We coded friendly, helpful, useful, and informative comments (6, 7, 8, 9) as 1 and unfriendly, not helpful, not useful, and not informative comments (1, 2, 3) as 0. In addition to removing neutral comments with an average rating of 5, we also removed comments with an average rating of 4 because all three annotators rated the majority of the comments as not conveying important information (for a similar argument, see Liu et al., 2020). Including comments with an average rating of 4 led to highly imbalanced classes, which was detrimental to the classification performance. Thus, we did not include these comments in our classification task.

We pre-processed each comment by removing punctuation, digits, single characters, and stop words. The remaining words were lowercased, and the Snowball stemming algorithm was applied to reduce words to their stem.[23] We translated each resulting string into a term-frequency inverse document frequency (tf-idf) representation and applied six commonly used supervised learning algorithms, including gradient boosting, logistic regression, naïve Bayes, neural network, random forest, and support vector machine (Clarke et al., 2020). We adopted the scikit-learn package (Pedregosa et al., 2011) and divided the sample into training data (70%) and test data (30%) to evaluate the performance of each algorithm. Precision, recall, and *f*-measure were our evaluation criteria, and the results are shown in Table F1. The classification performance was in line with recent research that has classified posts in online knowledge communities (Liu et al., 2020).

We selected the machine learning algorithm with the highest *f*-measure (support vector machine for friendly comments, gradient boosting for helpful, useful, and informative comments) to classify all of the remaining comments. According to the resulting machine learning classifications, the percentage of friendly (*PercFriendComments*), helpful (*PercHelpComments*), useful (*PercUseComments*), and informative comments (*PercInfoComments*) were constructed as dependent variables. The results in Table F2 suggest that the nudge did not result in any significant change in the quality of the information provided by the comments.[24]

[22] The annotation instruction (in English) provided to the research assistants is available at https://osf.io/t3fbh.
[23] The Python code for text pre-processing is available at https://osf.io/e3huy.
[24] The resulting sample size is lower because after pre-processing some comments were empty strings and could not be classified.

Table F1. Performance of Machine Learning Algorithms

Measure	Friendly		Helpful		Useful		Informative	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Gradient boosting								
Precision	0.737	0.025	0.765	0.028	0.782	0.034	0.778	0.031
Recall	0.745	0.025	0.668	0.023	0.670	0.025	0.667	0.024
F1	0.737	0.025	**0.700**	0.024	**0.705**	0.027	**0.703**	0.024
Logistic regression								
Precision	0.771	0.031	0.674	0.247	0.694	0.249	0.657	0.237
Recall	0.645	0.043	0.503	0.003	0.503	0.003	0.503	0.004
F1	0.634	0.061	0.471	0.008	0.476	0.008	0.480	0.010
Naive Bayes								
Precision	0.604	0.030	0.479	0.012	0.475	0.010	0.473	0.009
Recall	0.606	0.031	0.454	0.026	0.440	0.022	0.426	0.023
F1	0.603	0.030	0.397	0.015	0.389	0.013	0.384	0.012
Neural network								
Precision	0.740	0.028	0.635	0.032	0.626	0.036	0.643	0.047
Recall	0.700	0.032	0.568	0.016	0.553	0.018	0.544	0.017
F1	0.705	0.034	0.581	0.021	0.563	0.024	0.554	0.026
Random forest								
Precision	0.765	0.024	0.828	0.042	0.858	0.042	0.886	0.040
Recall	0.778	0.024	0.611	0.019	0.592	0.017	0.581	0.018
F1	0.765	0.025	0.647	0.025	0.626	0.025	0.615	0.027
Support vector machine								
Precision	0.782	0.025	0.691	0.022	0.687	0.020	0.695	0.023
Recall	0.796	0.026	0.702	0.024	0.692	0.024	0.691	0.023
F1	**0.777**	0.028	0.696	0.022	0.688	0.019	0.692	0.020

Note: SD = standard deviation. Numbers highlighted in bold represent the highest *f*-measure. The results are based on 100 experiments.

Table F2. Percentage of Friendly, Helpful, Useful, and Informative Comments

	PercFriendComments	***PercHelpComments***	***PercUseComments***	***PercInfoComments***
	(1)	**(2)**	**(3)**	**(4)**
Newcomer	0.011 (0.007)	0.006 (0.004)	0.011*** (0.004)	0.008** (0.004)
Newcomer × *After*	0.000 (0.008)	0.000 (0.004)	-0.004 (0.004)	-0.002 (0.004)
Tenure	0.000*** (0.000)	0.000** (0.000)	0.000** (0.000)	0.000** (0.000)
log(1+*DescLen*)	0.013*** (0.002)	0.016*** (0.001)	0.014*** (0.001)	0.013*** (0.001)
LocalDeal	-0.008** (0.004)	-0.021*** (0.002)	-0.019 (0.002)	-0.017*** (0.002)
NumCategories	0.003*** (0.001)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Content	0.074*** (0.006)	-0.003 (0.003)	0.005* (0.003)	0.003 (0.003)
Time FE (day, hour)	Yes	Yes	Yes	Yes
Category FE	Yes	Yes	Yes	Yes
Observations	39,197	39,197	39,197	39,197
Adjusted *R*-squared	0.018	0.055	0.047	0.050

Note: FE = fixed effects. Robust standard errors clustered by user are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.