Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Replying to newcomers enhances engagement with existing users: evidence from the reply network of a social virtual reality (VR) community

Kotaro Hayashi

k.hayashi.shiga@gmail.com

Shiga University

Takuma Tanaka

Shiga University

Research Article

Keywords:

Posted Date: August 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10213118/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Replying to newcomers enhances engagement with existing users: evidence from the reply network of a social virtual reality (VR) community

Kotaro Hayashi[1*] and Takuma Tanaka[1]

[1]Graduate School of Data Science, Shiga University, Hikone, Shiga, 522-8522, Japan.

*Corresponding author(s). E-mail(s): k.hayashi.shiga@gmail.com;

Abstract

Online communities flourish when there is a constant flow of newcomers, but a surge of newcomers can overwhelm existing users, leading to a loss of activity. To address this issue, the present study aimed to examine whether existing users in the Japanese-speaking VRChat community on X increase or decrease their overall activity and engagement with pre-surge interaction partners after replying to newcomers. The surge of newcomers we focused on in this study occurred in August–September 2024. Existing users were defined as users who posted #VRChat始めました ("I started VRChat") in January–February 2024. We defined early partners among existing users using pre-surge reply histories from February–April 2024. We then examined whether users who replied to newcomers in August or September changed their engagement with early partners after the treatment period. After common support trimming, the sample comprised 269 existing users (47 treated and 222 controls) observed over 2,152 user-months. Using propensity score weighting targeting the average treatment effect on the treated (ATT) and difference-in-differences (DID) models with user and month fixed effects, we found post-treatment increases in the total number of posts (0.59, 95% confidence interval: [0.23, 0.95]), the total number of replies (0.69, [0.29, 1.08]), and the total number of replies to early partners (0.30, [0.01, 0.59]), each of which was transformed by $\log(1 + x)$, among treated users relative to weighted controls. These findings indicate that contact with newcomers was followed by increased activity without evidence of reduced engagement with early partners.

Keywords: Online communities, Surge of newcomers, Social VR, VRChat, X, Difference-in-differences, Propensity score weighting

1 Introduction

When an online community experiences a surge of newcomers, existing users' interactions may be redistributed between newcomers and existing interaction partners. On one hand, rapid growth may increase existing users' information load, competition for attention, norm-maintenance efforts, and relational adjustment costs [1–5]. These pressures may degrade existing users' experiences because accommodating newcomers can reduce their interactions with pre-surge partners. On the other hand, newcomers can expand opportunities for interaction. Indeed, recent empirical studies indicate that surges of newcomers do not necessarily degrade online communities [6–8]. Distinguishing these two possibilities is essential for understanding how existing users respond to community growth. However, previous research has predominantly examined newcomer-centered processes, such as entry, norm learning, feedback, and retention or departure, rather than the experiences of existing users [9–15]. Less is known about how existing users adjust their interactions with people they had already interacted with before a newcomer surge. This knowledge gap matters because community growth may be unsustainable if existing users accommodate newcomers by reducing engagement with their pre-surge interaction partners.

This study therefore aimed to examine whether, after a newcomer surge, replying to newcomers displaces interaction with existing users' pre-surge interaction partners or, conversely, increases overall activity with sustained engagement with those partners. Here, we focus on repeatedly contacted pre-surge interaction partners among existing users, which are referred to as early partners. The central question this paper addresses is whether existing users reduce replies to early partners after contact with newcomers.

We examine two competing hypotheses regarding the relationship between replies to newcomers and to early partners. The crowding-out hypothesis posits that, given constraints on reply time and attention, replying to newcomers would be associated with reduced replies to early partners. Conversely, the activation hypothesis posits that replying to newcomers reflects increased overall activity rather than a displacement of replies to early partners, and would therefore be accompanied by increased overall activity and sustained or increased replies to early partners. To distinguish between these hypotheses, early partners must be identified before the surge of newcomers and tracked consistently over time.

Therefore, the present study required a setting in which a newcomer surge could be observable, newcomers identified, and reply histories used to track existing users' interactions with both newcomers and early partners. The Japanese-speaking VRChat community on X, formerly Twitter, provided such a setting. Although VRChat is a social virtual reality (VR) platform, participation announcements and the reply interactions around it are publicly observable on X. The hashtag #VRChat始めました, meaning "I started VRChat," is used to announce entry into VRChat, and posts containing this hashtag increased sharply in August 2024, shortly after the release of a VRChat-related YouTube video by Sutanmi, a popular Japanese streamer and YouTuber, in late July 2024. This increase made it possible to identify the newcomer cohort, and X reply histories allowed us to define pre-surge interaction partners and track subsequent replies to both newcomers and early partners.

We used the August–September 2024 surge to construct a monthly observational comparison between the pre- and post-surge periods. Existing users were defined as users who posted the hashtag #VRChat始めました during January–February 2024, whereas newcomers were defined as users who posted the same hashtag in August or September 2024. Treated users were existing users who replied to at least one August newcomer in August or at least one September newcomer in September; control users

were existing users who did not make such replies. Early partners for an existing user were defined as users to whom the user had repeatedly replied before August. Using this setting, we compared treated users with weighted control users through propensity score weighting targeting the average treatment effect on the treated (ATT) and difference-in-differences (DID) models. This design allowed us to assess whether observed replies to newcomers were followed by reduced engagement with early partners, as predicted by the crowding-out hypothesis, or by increased reply activity that included continued engagement with early partners, as predicted by the activation hypothesis.

The remainder of the paper is organized as follows. In Section 2, we describe the data, research design, measures, weighting procedure, and estimation strategy. In Section 3, we present the results, including balance diagnostics, DID estimates, event study results, and robustness checks. In Section 4, we discuss the study findings, limitations, and directions for future research. Additional diagnostics and sensitivity analyses are provided in the Appendix.

2 Methods

2.1 Study setting and cohort definition

This study investigated the VRChat community of Japanese-speaking users on X. Japanese-speaking users often post self-introductions with the hashtag #VRChat始めました ("I started VRChat") when they start using VRChat. We therefore used posts containing this hashtag as a signal that the poster had started using VRChat. All dates and monthly intervals were defined in Japan Standard Time. Figure 1 summarizes the research design and observation timeline in this study.

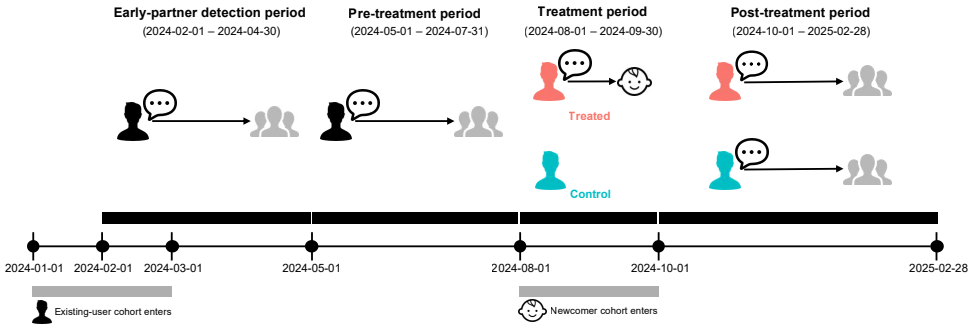


Figure 1: Research design and observation timeline. Early partners were defined using repeated replies from February through April 2024. The DID outcome panel compared the May–July 2024 pre-treatment period with the October 2024–February 2025 post-treatment period, with August–September 2024 used to define treatment status.

To examine existing users' behavior during a surge of newcomers, we defined a newcomer and an existing-user cohort and collected their posts. We defined August–September 2024 as the treatment period because a surge of newcomers was observed in the VRChat-related Japanese discourse on X during these months. This influx followed a VRChat-related YouTube video published by Sutanmi on July 24, 2024 [16]. The magnitude of the influx is visible in the number of VRChat-related X posts shown in Fig. 2. Figure 2a shows that posts containing "VRChat" increased around August 2024, and that most of the increase came from posts for which the language was classified as Japanese in X language metadata. Figure 2b shows that posts containing #VRChat始めました increased from a monthly average of 693 during January–July 2024 to 1,965 in August 2024. On the basis of these descriptive patterns, we used August–September 2024 as the treatment period.

Hashtag posts were collected using the X API full-archive search endpoint, and aggregate volume series were collected using the X API full-archive counts end point [17]. We first constructed the putative newcomer account pool consisting of accounts that posted #VRChat始めました between August 1 and September 30, 2024. In

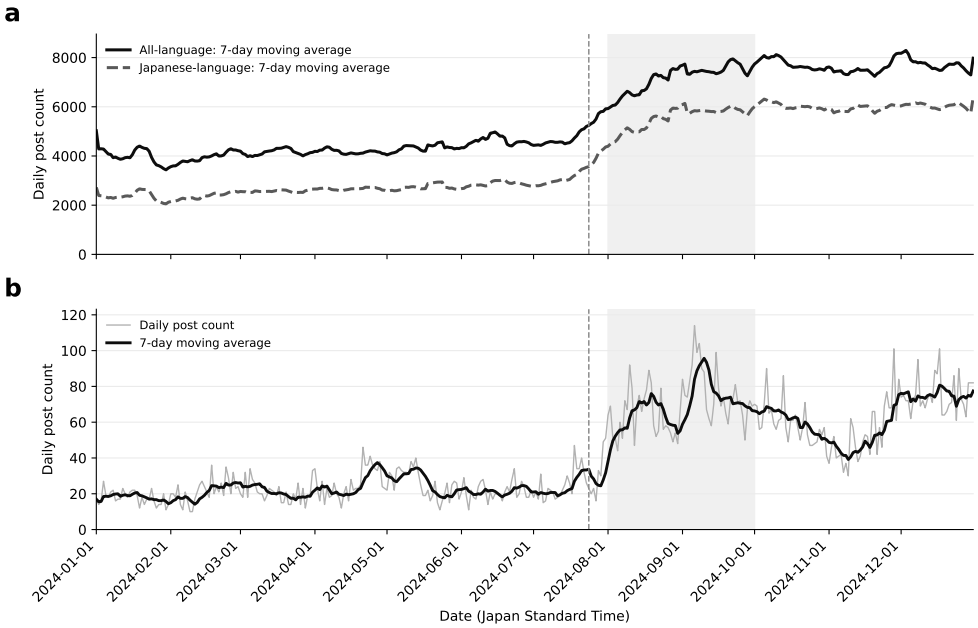


Figure 2: VRChat-related posts in 2024. (a) Seven-day moving average of daily posts containing "VRChat," comparing all language posts with Japanese-language ones according to X language metadata. (b) Daily posts containing #VRChat始めました and the corresponding 7-day moving average. The dashed vertical line indicates the publication date of a VRChat-related YouTube video by Sutanmi in late July 2024.

August, we identified 1,665 posts from 1,240 accounts; in September, we identified 1,793 posts from 1,300 accounts. Manual review of newcomer status was then conducted for accounts in this initial pool that were observed as reply targets of existing users during the treatment period, because replies to these accounts determined treatment assignment. We excluded accounts if prior VRChat use was observed in earlier posts or the profile text available at the time of data collection, or if the account was judged to be primarily engaged in newcomer-oriented advertising or recruitment. The accounts that were not excluded by these criteria constituted the newcomer cohort, and only replies to these accounts were counted as replies to newcomers in the subsequent analyses.

The existing-user cohort was defined as accounts that announced VRChat participation approximately 6 months before the newcomer cohort. Specifically, accounts

that posted #VRChat始めました between January 1 and February 29, 2024, were collected. Accounts meeting the same criteria as those applied in constructing the newcomer cohort were excluded. Figure 2b indicates that January–February 2024, the period from which the existing-user cohort was drawn, did not show a surge comparable to that in August 2024.

For each account in the existing-user cohort, we collected posts from February 2024 through February 2025 using the user timeline end points of the X API [17]. Data collection was restricted to original posts and replies; reposts were not collected and were not used in any analysis. We reconstructed the collected data as a monthly user-level panel. The observation period spanned 13 months, from February 2024 through February 2025. Descriptive statistics for the monthly panel are shown in Table 1. We summarized reply-target diversity using Shannon entropy and reply-target concentration using the Herfindahl–Hirschman Index (HHI) [18, 19].

Table 1: User-level descriptive statistics for monthly measures

Variable	N	Mean	SD	Median	Min	Max
Posting activity						
Posts per month	405	131.29	200.20	57.62	0.00	1168.08
Original posts per month	405	63.24	102.26	28.23	0.00	868.46
Reply posts per month	405	63.71	121.82	19.23	0.00	1016.31
Proportion of active months	405	0.78	0.34	1.00	0.00	1.00
Per-post engagement						
Likes received per original post	383	9.61	25.11	3.37	0.00	286.89
Reposts received per original post	383	0.87	3.52	0.13	0.00	38.53
Replies received per original post	383	0.46	0.87	0.24	0.00	12.54
Reply network						
No. of users replied to per month	405	14.93	21.78	7.46	0.00	148.15
Reply-target diversity, entropy	405	1.45	1.18	1.47	0.00	4.46
Reply-target concentration, HHI	405	0.18	0.14	0.15	0.00	0.72

Note. All observations are averaged to obtain monthly measures within users. Original posts exclude replies, reposts, and quote posts. Proportion of active months is the proportion of months with at least one post. Per-post engagement measures are calculated only for months with original posts and limited to users with at least one such month $N = 383$. Entropy and the HHI are set to 0 in months with no observed reply targets. SD, standard deviation.

2.2 Research design

This study examined whether replying to newcomers was followed by changes in existing users' engagement with early partners. We partitioned the observation period into four distinct periods (Fig. 1): the early-partner detection period (February–April 2024), pre-treatment period (May–July 2024), treatment period (August–September 2024), and post-treatment period (October 2024–February 2025).

During the early-partner detection period, we defined a user's early partners as accounts to whom the user replied in at least 2 of the 3 months. This definition identifies users to whom the existing user repeatedly replied before the newcomer surge

[20–22]. We then used post-treatment replies to these early partners as an observable measure of continued interaction with them. The robustness of this early-partner definition was examined in sensitivity analyses.

The treatment indicator was defined as at least one reply by an existing user to an August newcomer in August or to a September newcomer in September. Existing users who made such replies during the treatment period were classified as treated users, and those who did not were classified as control users. This definition should therefore be interpreted as same-month reply contact with newcomers rather than as precisely timed contact after first-observed hashtag posts. Using this classification, the main DID models compared outcomes in the post-treatment period with those in the pre-treatment period, excluding both the early-partner detection period and the treatment period from the outcome panel.

2.3 Outcome families

We measured post-treatment behavioral differences in two outcome families: overall activity and early-partner engagement. The first family captures whether replying to newcomers was followed by changes in users' overall posting and reply activity. The second family captures whether replying to newcomers was followed by increased or decreased engagement with the early partners defined in Section 2.2. All outcomes were constructed at the user-month level. Definitions are shown in Table 2.

A contacted early partner is an early partner contacted by the user at least once in a given month. We defined replies per contacted early partner as the number of replies to early partners divided by the number of contacted early partners in that month. This outcome was set to 0 when the user replied to no early partners in the month. This measure captured both whether the user contacted any early partner in a given month and how intensively the user replied to the early partners who were contacted.

Table 2: Outcome families

Family	Outcome	Operational definition
Overall activity	Log posts	Monthly count of posts excluding reposts with a $\log(1+x)$ transformation.
	Log replies	Monthly count of reply posts with a $\log(1+x)$ transformation.
Early-partner engagement	Log replies to early partners	Monthly count of replies to the user's early partners with a $\log(1+x)$ transformation.
	Replies per contacted early partner	Replies to early partners divided by the number of early partners contacted in that month; defined as 0 when none were contacted.

2.4 Propensity score weighting

Because the study was based on observational data, treated and control users may differ in regard to pre-treatment activity and reply propensity. In particular, existing users who replied to newcomers may have been more active or more prone to reply before the treatment period. To improve comparability, we used ATT-oriented propensity score weighting [23–25].

The propensity score was defined as the probability that an existing user replied to at least one newcomer during the treatment period. We estimated this probability with a logit model using February–July covariates. The covariates included each user's February–July means for log posts (mean log posts), log replies (mean log replies), unique reply targets (mean unique reply targets), and reply-target entropy (mean reply-target entropy) with log posts and log replies defined as monthly $\log(1+x)$-transformed counts in Table 2. To adjust for February–July activity trends, we also

included user-specific slopes from ordinary least-squares regressions of monthly log replies (trend in log replies) and reply-target entropy (trend in reply-target entropy) on a month index. In addition, we included the $\log(1+x)$-transformed number of early partners identified during the early-partner detection period (log early partners) and the $\log(1+x)$-transformed number of replies sent to early partners in the same period (log early-partner replies).

We retained only users in the common support region where the treated and control propensity score distributions overlapped [26]. ATT weights were assigned as 1 for treated users and $p_i/(1-p_i)$ for control users, where p_i is the estimated propensity score. Control weights were normalized so that the total weight of the control group equaled the total weight of the treated group. We assessed comparability after weighting using the overlap of propensity score distributions and standardized mean differences (SMDs) for all covariates included in the propensity score model [25, 27]. The propensity score and common support summary and control-weight distribution for the main specification are reported in Appendix Tables 4 and 5; exact covariate means and SMDs are reported in Appendix Table 6.

2.5 Estimation strategy

We estimated weighted DID models with normalized ATT weights and used an event study specification to assess pre-treatment divergence and post-treatment dynamics [28–30]. This approach estimates changes among treated users relative to weighted control users after adjustment for observed pre-treatment characteristics. As described above, the main DID panel included only the pre- and post-treatment periods; the early-partner detection period was used to define early partners, and the treatment period was used to define treatment status.

The DID is modeled by:

$$Y_{it} = \alpha_i + \gamma_t + \beta\,(\mathrm{Treat}_i \times \mathrm{Post}_t) + \varepsilon_{it}. \tag{1}$$

Here, Y_{it} is the outcome for user i in month t, α_i is the user fixed effect, and γ_t is the month fixed effect. Treat_i and Post_t are indicators for treated users and the post-treatment period, respectively. The coefficient β captures the difference in pre-to-post changes between treated and control users. Because the surge-related video was published on July 24, 2024, July may contain early responses; we therefore report a drop-July sensitivity analysis. In this specification, July 2024 is excluded from both the general pre-treatment covariate window and the DID outcome panel; the propensity score, common support trimming, ATT weights, covariate balance, and DID model are then re-estimated.

To examine temporal patterns and pre-period differences, we estimated the following event study models:

$$Y_{it} = \alpha_i + \gamma_t + \sum_{k \in \mathcal{K}} \beta_k\,(\mathrm{Treat}_i \times \mathbf{1}\{r_t = k\}) + \varepsilon_{it}, \tag{2}$$

where $\mathcal{K} = \{-3, -1, 2, 3, 4, 5, 6\}$. The coefficients β_k are the event-time coefficients to be estimated. Here, r_t denotes event time in months, with August 2024 coded as $r_t = 0$. August and September 2024 were excluded from the estimation because they belong to the treatment period. June 2024 ($r_t = -2$) is not included in $\mathcal{K}$ because it was used as the reference month.

All models included the user and month fixed effects and used normalized ATT weights unless otherwise noted. Standard errors were clustered at the user level to account for serial dependence within users [31]. The event study was used as a diagnostic complement to the DID estimates rather than as a separate primary estimator.

2.6 Robustness and sensitivity analyses

We assessed robustness by varying the weighting scheme, early-partner criterion, and outcome window. The checks included unweighted DID, 95% ATT-weight capping [32], requiring early partners to appear in all three detection months, and dropping July 2024. Propensity score diagnostics are reported in Appendix Tables 4–6, and robustness and sensitivity estimates are reported in Appendix Table 7.

3 Results

3.1 Sample and covariate balance

Figure 3 shows the sample flow for the weighted DID analysis. The existing-user cohort initially contained 405 users. The propensity score model was estimated among users with pre-treatment covariates and at least one early partner. After common support trimming, the sample contained 269 users, including 47 treated and 222 control users, observed over 2,152 user-months from the pre- and post-treatment periods.

Figure 4 shows the propensity score distributions before and after weighting in the sample. Before weighting, the distribution for treated users was shifted toward higher estimated probabilities of replying to newcomers than that for control users, while still showing common support. After normalized ATT weighting, the weighted control-user distribution more closely matched the treated distribution. The effective sample size (ESS) for the weighted control group was 153.6.

Figure 5 shows absolute SMDs before and after ATT weighting. Before weighting, several covariates showed imbalance. After weighting, covariate imbalance was reduced; the maximum absolute SMD was 0.085 and the mean absolute SMD was 0.044.



Figure 3: Sample flow for the weighted DID analysis. The diagram shows construction of the sample, treatment-indicator assignment, common support trimming, and ATT weighting scheme.

3.2 DID across outcomes

Table 3 reports DID estimates for all four outcomes. For overall activity, the DID estimate was 0.59 for log posts (95% confidence interval [CI]: [0.23, 0.95]) and 0.69 for log replies (95% CI: [0.29, 1.08]). These estimates indicate larger post-treatment increases among treated users relative to weighted control users under the ATT-weighted design.

The early-partner engagement outcomes also showed positive DID estimates. The estimate for log replies to early partners was 0.30 (95% CI: [0.01, 0.59]), and the

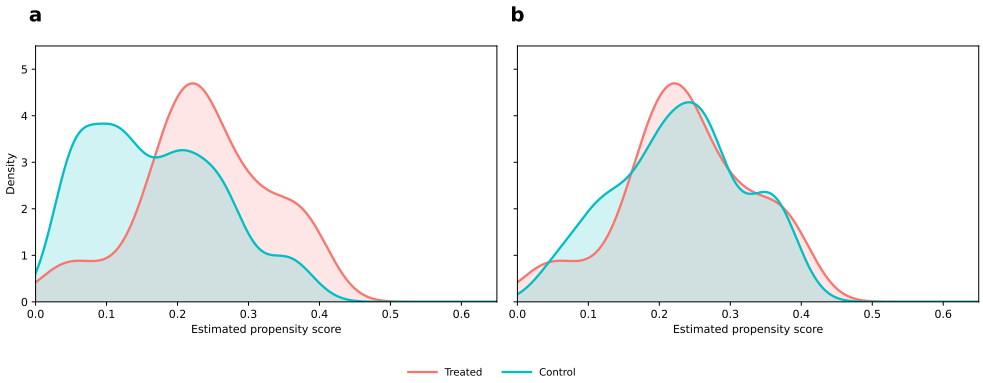


Figure 4: Propensity score distributions before and after ATT weighting. (a) Unweighted sample. (b) ATT-weighted sample.

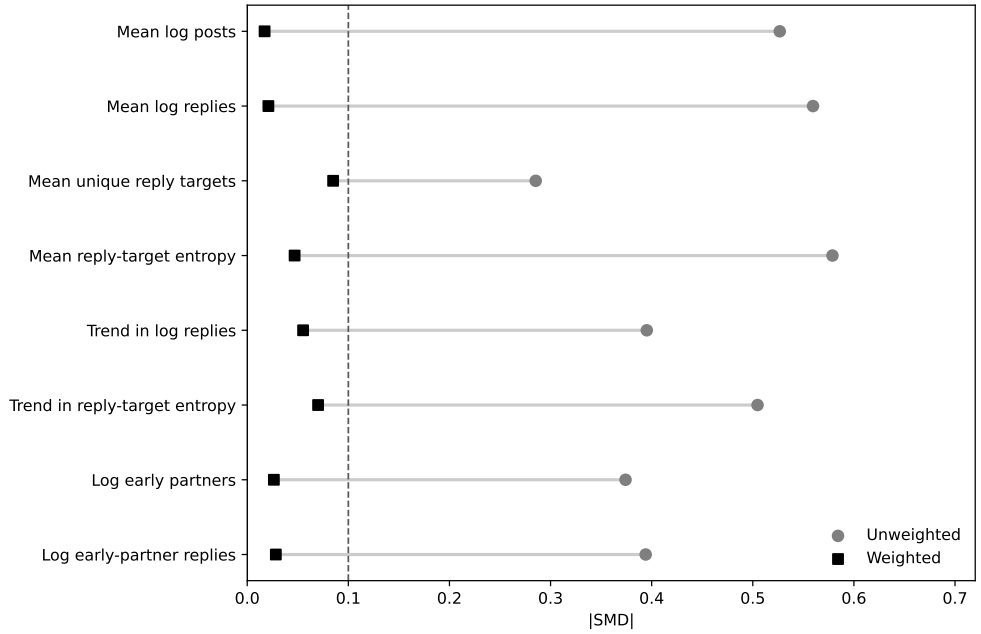


Figure 5: Covariate balance before and after ATT weighting. Points show absolute SMDs; the dashed line indicates 0.10.

estimate for replies per contacted early partner was 1.26 (95% CI: [0.00, 2.51]). Taken together, the outcomes are more consistent with the activation hypothesis than with the crowding-out hypothesis: users who replied to newcomers showed higher overall activity and no evidence of a reduction in absolute replies to early partners.

Table 3: Main DID estimates

Outcome	Estimate	Std. error	95% CI	p-value
Log posts	0.59	0.19	[0.23, 0.95]	0.001
Log replies	0.69	0.20	[0.29, 1.08]	0.001
Log replies to early partners	0.30	0.15	[0.01, 0.59]	0.040
Replies per contacted early partner	1.26	0.64	[0.00, 2.51]	0.049

Note. All estimates use the same sample: 2,152 user-months from 269 users. Standard errors are clustered at the user level.

3.3 Event study

Figure 6 shows event study estimates for all four outcomes. The event study uses June 2024 ($k = -2$) as the reference month and excludes August and September 2024 as the treatment period. Joint tests of the pre-treatment lead coefficients yielded p-values of 0.755 for log posts, 0.576 for log replies, 0.658 for log replies to early partners, and 0.186 for replies per contacted early partner.

Post-period coefficients were jointly different from zero for log posts ($p = 0.008$), log replies ($p < 0.001$), and log replies to early partners ($p = 0.010$), but not for replies per contacted early partner ($p = 0.298$). The event study diagnostics therefore show no clear pre-period divergence for the four main outcomes, although the short pre-period limits the power of lead tests.

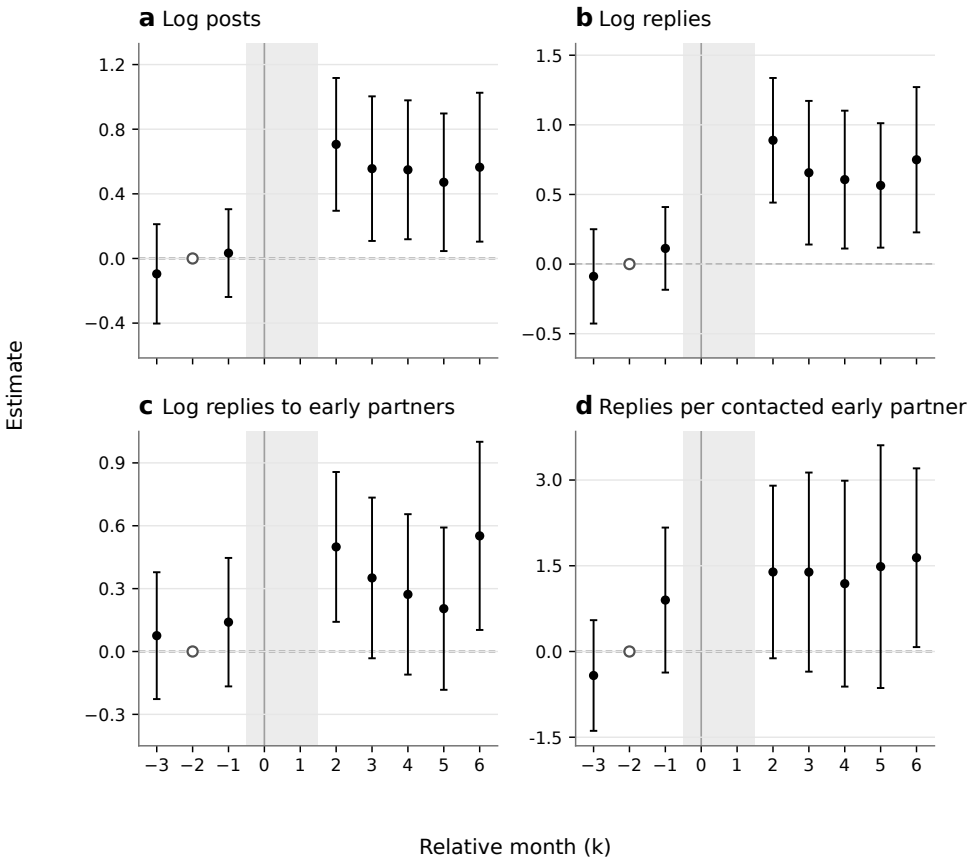


Figure 6: Event study estimates across outcomes. The omitted reference month is June 2024 ($k = -2$), and August–September 2024 are excluded as the treatment period. Points show estimates; bars show 95% confidence intervals.

3.4 Robustness and sensitivity analyses

Figure 7 summarizes the robustness and sensitivity analyses across alternative specifications for the four main outcomes. Overall activity estimates remained positive across unweighted DID, 95% ATT-weight capping, the stricter 3-month early-partner definition, and the drop-July specification. DID estimates for log replies to early partners also remained positive across these checks: 0.05 in the unweighted DID, 0.30 after 95% ATT-weight capping, 0.37 under the stricter 3-month early-partner definition, and 0.34 after dropping July. In the re-estimated drop-July specification, the estimates

were 0.67 for log posts, 0.80 for log replies, 0.34 for log replies to early partners, and 1.51 for replies per contacted early partner; the CI for the last outcome included zero.

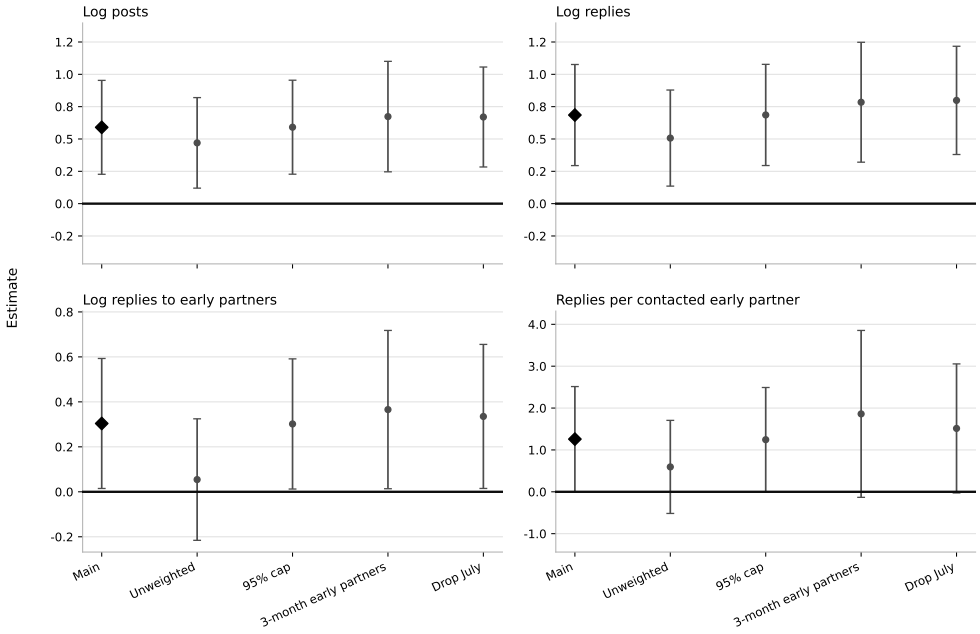


Figure 7: Robustness-check DID estimates across alternative specifications. Points show estimates; bars show 95% confidence intervals. The horizontal line indicates zero.

Additional propensity score diagnostics are reported in Appendix Tables 4–6, and the numerical robustness estimates corresponding to Fig. 7 are reported in Appendix Table 7.

4 Discussion

This study has examined whether replying to newcomers was followed by changes in existing users' overall activity and engagement with early partners in the Japanese-speaking VRChat community on X. The results were more consistent with the activation hypothesis than with the crowding-out hypothesis: users who contacted newcomers showed higher overall activity and no evidence of a reduction in the number of replies to early partners. DID estimates for log replies to early partners were

positive in the ATT-weighted specification and remained so across the main robustness checks, although their precision varied by specification. The overall activity results were the most stable across specifications, while the relationship-specific results provided evidence against a reduction in the number of replies to early partners. However, these results should be interpreted cautiously because the treatment indicator was observational and self-selected. The present findings should therefore be interpreted as ATT-oriented observational evidence among existing users who made observed same-month replies to newcomers rather than conclusive causal evidence that newcomer contact itself changed behavior.

The findings of this study contribute to online-community research by shifting attention from newcomer retention alone to existing user responses. Online communities depend on newcomers, but existing users sustain everyday interaction, tacit practices, and relational continuity. Prior work on newcomer growth and surges has emphasized both the risks of disruption and the possibility that communities adapt to growth [1, 4, 6–8, 33]. The present study provides observational evidence that engagement with early partners can coexist with higher overall activity among existing users who reply to newcomers. This framing is consistent with theories that online-community attachment can be grounded in interpersonal bonds, as well as in group-level identity [34, 35].

The findings of this study also contribute to the analysis of social VR communities as cross-platform communities. Although the focal community was organized around VRChat, the data analyzed here were obtained from X. Participation in social VR communities often extends beyond the VR platform itself. In this setting, X is one of the spaces where VRChat participation is announced and where reply-based relational activity around the community can be observed. This cross-platform framing is consistent with work showing that social VR supports interpersonal relationship building and the maintenance of social relationships beyond a single platform [36–41].

X-based reply behavior therefore provides a window into the broader interactional ecology surrounding a social VR community, even though it does not directly measure interactions inside VRChat.

This study has several limitations. First, the study design was observational in nature. Users who replied to newcomers may have differed from controls in unobserved sociability, community role, free time, or prior relationships. Propensity score weighting and user fixed effects were used to adjust for observed pre-treatment characteristics and time-invariant user heterogeneity, but could not eliminate all unobserved confounding. Relatedly, the main treatment indicator was account-level same-month reply contact in August or September, which does not guarantee that every reply occurred after the newcomers' first hashtag posts. The estimates should therefore be read as ATT-oriented DID estimates for existing users who replied to newcomers in August or September, not as a clean causal effect of precisely timed first contact [42, 43].

Second, the measures are behavioral proxies. Early partners are defined from X replies, not from VRChat friendship status, VRChat co-presence, or subjective relationship reports. A reply target observed in at least 2 of the 3 months in the early-partner detection period is a useful indicator of repeated interaction, but it may represent friendship, casual exchange, event coordination, information seeking, or another form of social connection. More generally, X reply histories do not reveal how close these relationships were, what the exchanges meant to participants, or whether related interaction occurred in VRChat or other channels outside the observed reply network [20, 21, 44]. The newcomer cohort was also defined through self-reported hashtag use. The hashtag #VRChat始めました is a meaningful community signal but may include returning users, users who had already tried VRChat, or promotional posts. Cohort construction attempted to reduce such misclassification, but some ambiguity remains.

Third, the setting and temporal granularity limit generalization. The August 2024 surge of newcomers was a Japanese-speaking community event associated with a VRChat-related YouTube video by Sutanmi [16]. In addition, the results may not be generalizable to other language communities, social VR platforms, or surges generated by different mechanisms. July 2024 may also contain anticipation or early exposure because the associated video was published on July 24, 2024. The drop-July analysis addresses this concern as a robustness check, but future work should use longer pre-treatment periods and higher-frequency daily or weekly panels around late July and August. Finally, the August surge was a community-level event, so both treated and control users were exposed to various changes in the community. The treatment indicator captures observed replies to newcomers rather than exposure to the surge itself.

Future work can extend this study in at least three directions. First, it should distinguish the quality and directionality of contact with newcomers. A reply may involve welcoming, guidance, casual conversation, event invitations, technical support, and other forms of interaction. The present study measured existing users' replies to newcomers rather than replies that existing users received from newcomers; future work should examine reciprocal exchange and newcomer-initiated replies as distinct mechanisms. Second, future research should examine heterogeneity. Users with high baseline activity, many early partners, or community-organizing roles may respond differently to surges of newcomers. Third, future work should connect X-based behavior with VRChat experiences through surveys, interviews, or platform-level data. Such mixed-methods evidence could clarify the mechanisms through which replying to newcomers relates to engagement with early partners and connect individual-level responses to broader questions about online-community growth and survival [45].

Abbreviations

ATT, average treatment effect on the treated; CI, confidence interval; DID, difference-in-differences; ESS, effective sample size; HHI, Herfindahl–Hirschman Index; SD, standard deviation; SMD, standardized mean difference; VR, virtual reality.

Declarations

Availability of data and materials. The raw X posts, platform-specific user identifiers, dyadic reply records, and derived user-level and dyadic datasets are not publicly available. Public release is restricted because these materials contain X Content, platform-specific identifiers, and social media traces that may create privacy and re-identification risks, and because redistribution is constrained by X's Developer Agreement and Developer Policy, ethics approval conditions, and institutional requirements. Analysis code that does not contain X Content, user identifiers, API credentials, or non-redistributable derived data may be made available from the corresponding author upon reasonable request. Requests for any additional materials will be considered only when they are compatible with X's terms, institutional approval, and ethics review.

Competing interests. The authors declare no competing interests.

Funding. This work was supported by the Telecommunications Advancement Foundation under the research project "Formation of Social Networks in Virtual Reality Worlds and Their Effects on Subjective Well-Being." The funder had no role in the study design, data collection, analysis, interpretation, or writing of the manuscript.

Authors' contributions. KH designed the study, collected the data, performed the analysis, prepared the tables and figures, and drafted the manuscript. TT supervised the study and critically revised the manuscript. Both authors read and approved the final manuscript.

Acknowledgements. The authors used AI-assisted tools to support translation, language revision, code-related editing, and manuscript-structure review. All AI-assisted output was reviewed and revised by the authors, who take full responsibility for the final manuscript. The authors thank FORTE Science Communications for English language editing.

Ethics approval and consent to participate. Ethics approval was obtained from the author's institutional review board for the analysis of publicly accessible social media traces (B250009). This study used only publicly accessible X posts and excluded private or protected accounts from analysis. The manuscript reports only aggregate results and does not quote identifiable user posts. Although the source posts were publicly accessible, the collected X records and derived datasets are not redistributed because they contain platform-specific identifiers, profile information, and dyadic reply records subject to platform terms, institutional requirements, and the approved ethics protocol. Direct identifiers were replaced with study-specific pseudonyms for analysis, and the data were stored in access-restricted local research storage. Individual consent was not obtained because the study analyzed publicly accessible social media traces under institutional ethics review.

Consent for publication. Not applicable. The manuscript does not include identifiable quotations or personal information from individual users.

Appendix: Supplementary diagnostics and robustness checks

This Appendix reports supplementary diagnostics for the main specification and numerical estimates for the robustness checks summarized in Fig. 7. The first subsection reports propensity score overlap, control-weight diagnostics, and covariate

balance. The second subsection reports the numerical DID estimates corresponding to the robustness checks shown in Fig. 7.

Propensity score diagnostics

Table 4: Propensity score and common support diagnostics for the main specification

Specification	Users	Treated	Control	Control ESS	Max |SMD|	Mean |SMD|
Main after trimming	269	47	222	153.6	0.085	0.044
Unweighted sample	269	47	222	222.0	0.579	0.452

Table 5: Control-weight distribution after common support trimming

Users	Weight sum	ESS	Median	P90	P95	Max
222	47.0	153.6	0.183	0.398	0.529	0.642

Table 6: Covariate balance before and after ATT weighting in the main specification

	Before weighting			After weighting		
Covariate	Treated	Control	|SMD|	Treated	Control	|SMD|
Mean log posts	4.900	4.247	0.527	4.900	4.919	0.017
Mean log replies	3.945	3.203	0.560	3.945	3.971	0.021
Mean unique reply targets	24.979	18.654	0.285	24.979	26.962	0.085
Mean reply-target entropy	2.374	1.848	0.579	2.374	2.335	0.047
Trend in log replies	0.136	0.006	0.395	0.136	0.154	0.055
Trend in reply-target entropy	0.121	0.009	0.505	0.121	0.106	0.070
Log early partners	2.367	2.007	0.374	2.367	2.392	0.026
Log early-partner replies	4.467	3.869	0.394	4.467	4.509	0.028

Note. After weighting, treated means are unchanged because treated users receive weight 1 under ATT weighting. After-weighting control means use normalized ATT weights.

Robustness estimates

Table 7: Robustness-check DID estimates for the four main outcomes

Specification	Log posts	Log replies	Log replies to early partners	Replies per contacted early partner	Users	Observations
Main	0.59 [0.23, 0.95]	0.69 [0.29, 1.08]	0.30 [0.01, 0.59]	1.26 [0.00, 2.51]	269	2152
Unweighted DID	0.47 [0.12, 0.82]	0.51 [0.14, 0.88]	0.05 [−0.22, 0.32]	0.59 [−0.52, 1.71]	269	2152
95% weight cap	0.59 [0.23, 0.95]	0.69 [0.29, 1.08]	0.30 [0.01, 0.59]	1.24 [0.00, 2.49]	269	2152
3-month early-partner criterion	0.67 [0.25, 1.10]	0.78 [0.32, 1.25]	0.37 [0.01, 0.72]	1.86 [−0.13, 3.85]	188	1504
Drop-July	0.67 [0.28, 1.06]	0.80 [0.38, 1.22]	0.34 [0.01, 0.66]	1.51 [−0.03, 3.05]	270	1890

Note. Entries are DID estimates with 95% confidence intervals in brackets. The drop-July specification omits July 2024 from the general pre-treatment covariate window and DID outcome panel, then re-estimates the propensity score, common support trimming, ATT weights, balance diagnostics, and DID model.

References

[1] Butler, B.S.: Membership size, communication activity, and sustainability: a resource-based model of online social structures. Information Systems Research **12**(4), 346–362 (2001) https://doi.org/10.1287/isre.12.4.346.9703

[2] Jones, Q., Ravid, G., Rafaeli, S.: Information overload and the message dynamics of online interaction spaces: a theoretical model and empirical exploration. Information Systems Research **15**(2), 194–210 (2004) https://doi.org/10.1287/isre.1040.0023

[3] Kraut, R.E., Resnick, P.: Building Successful Online Communities: Evidence-Based Social Design. MIT Press, Cambridge (2012). https://doi.org/10.7551/mitpress/8472.001.0001

[4] Halfaker, A., Geiger, R.S., Morgan, J.T., Riedl, J.: The rise and decline of an open collaboration system: how Wikipedia's reaction to popularity is causing its decline. American Behavioral Scientist **57**(5), 664–688 (2013) https://doi.org/10.1177/0002764212469365

[5] Masson, Z., Parmentier, G.: Drivers and mechanisms for online communities performance: a systematic literature review. European Management Journal **41**(4), 590–606 (2023) https://doi.org/10.1016/j.emj.2022.08.005

[6] Lin, Z., Salehi, N., Yao, B., Chen, Y., Bernstein, M.S.: Better when it was smaller? Community content and behavior after massive growth. Proceedings of the International AAAI Conference on Web and Social Media **11**(1), 132–141 (2017) https://doi.org/10.1609/icwsm.v11i1.14884

[7] Guo, Q., Shi, C., Yin, Z., Liu, C., Ma, X.: Exploring the effects of event-induced sudden influx of newcomers to online pop music fandom communities: content, interaction, and engagement. Proceedings of the ACM on Human-Computer Interaction **7**(CSCW2), 1–24 (2023) https://doi.org/10.1145/3610063

[8] Atcheson, A., Koshy, V., Karahalios, K.: Not what it used to be: characterizing content and user-base changes in newly created online communities. In: Proceedings of the CHI Conference on Human Factors in Computing Systems, pp. 1–12 (2024). https://doi.org/10.1145/3613904.3642769

[9] Arguello, J., Butler, B.S., Joyce, E., Kraut, R., Ling, K.S., Rosé, C., Wang, X.: Talk to me: foundations for successful individual-group interactions in online communities. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 959–968 (2006). https://doi.org/10.1145/1124772.1124916

[10] Choi, B., Alexander, K., Kraut, R.E., Levine, J.M.: Socialization tactics in Wikipedia and their effects. In: Proceedings of the ACM 2010 Conference on Computer Supported Cooperative Work, pp. 107–116 (2010). https://doi.org/10.1145/1718918.1718940

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

[12] Steinmacher, I., Graciotto Silva, M.A., Gerosa, M.A., Redmiles, D.F.: A systematic literature review on the barriers faced by newcomers to open source software projects. Information and Software Technology **59**, 67–85 (2015) https://doi.org/10.1016/j.infsof.2014.11.001

[13] Morgan, J.T., Halfaker, A.: Evaluating the impact of the Wikipedia Teahouse on newcomer socialization and retention. In: Proceedings of the 14th International Symposium on Open Collaboration, pp. 1–7 (2018). https://doi.org/10.1145/3233391.3233544

[14] Seering, J., Hammer, J., Kaufman, G., Yang, D.: Proximate social factors in first-time contribution to online communities. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–14 (2020). https://doi.org/10.1145/3313831.3376151

[15] Pethig, F.T., Hoehle, H., Hui, K.-L., Lanz, A.: Behavior toward newcomers and contributions to online communities. MIS Quarterly **49**(2), 731–758 (2025) https://doi.org/10.25300/MISQ/2024/17759

[16] Sutanmi: VRChatで偶然出会った謎の生命体に、腹筋を壊されるスタンミじゃばん【VRChat】. YouTube. Published 24 July 2024. Accessed 12 May 2026 (2024). https://www.youtube.com/watch?v=J4zY0IxCyA4

[17] X Corp: X API documentation. X Developer Platform. Accessed 12 May 2026 (n.d.). https://docs.x.com/x-api

[18] Shannon, C.E.: A mathematical theory of communication. Bell System Technical Journal **27**(3), 379–423 (1948) https://doi.org/10.1002/j.1538-7305.1948.tb01338.x . See also 27(4):623–656

[19] Kwoka, J.E. Jr.: The Herfindahl index in theory and practice. The Antitrust Bulletin **30**(4), 915–947 (1985) https://doi.org/10.1177/0003603X8503000405

[20] Marsden, P.V., Campbell, K.E.: Measuring tie strength. Social Forces **63**(2), 482–501 (1984) https://doi.org/10.1093/sf/63.2.482

[21] Jones, J.J., Settle, J.E., Bond, R.M., Fariss, C.J., Marlow, C., Fowler, J.H.: Inferring tie strength from online directed behavior. PLOS ONE **8**(1), 52168 (2013) https://doi.org/10.1371/journal.pone.0052168

[22] Pan, W., Shen, C., Feng, B.: You get what you give: understanding reply reciprocity and social capital in online health support forums. Journal of Health Communication **22**(1), 45–52 (2017) https://doi.org/10.1080/10810730.2016.1250845

[23] Rosenbaum, P.R., Rubin, D.B.: The central role of the propensity score in observational studies for causal effects. Biometrika **70**(1), 41–55 (1983) https://doi.org/10.1093/biomet/70.1.41

[24] Stuart, E.A.: Matching methods for causal inference: a review and a look forward. Statistical Science **25**(1), 1–21 (2010) https://doi.org/10.1214/09-STS313

[25] Austin, P.C., Stuart, E.A.: Moving towards best practice when using inverse probability of treatment weighting using the propensity score to estimate causal treatment effects in observational studies. Statistics in Medicine **34**(28), 3661–3679 (2015) https://doi.org/10.1002/sim.6607

[26] Crump, R.K., Hotz, V.J., Imbens, G.W., Mitnik, O.A.: Dealing with limited overlap in estimation of average treatment effects. Biometrika **96**(1), 187–199 (2009) https://doi.org/10.1093/biomet/asn055

[27] Austin, P.C.: Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity score matched samples. Statistics in Medicine **28**(25), 3083–3107 (2009) https://doi.org/10.1002/sim.3697

[28] Abadie, A.: Semiparametric difference-in-differences estimators. Review of Economic Studies **72**(1), 1–19 (2005) https://doi.org/10.1111/0034-6527.00321

[29] Bertrand, M., Duflo, E., Mullainathan, S.: How much should we trust differences-in-differences estimates? Quarterly Journal of Economics **119**(1), 249–275 (2004) https://doi.org/10.1162/003355304772839588

[30] Freyaldenhoven, S., Hansen, C., Shapiro, J.M.: Pre-event trends in the panel event study design. American Economic Review **109**(9), 3307–3338 (2019) https://doi.org/10.1257/aer.20180609

[31] Cameron, A.C., Miller, D.L.: A practitioner's guide to cluster-robust inference. Journal of Human Resources **50**(2), 317–372 (2015) https://doi.org/10.3368/jhr.50.2.317

[32] Lee, B.K., Lessler, J., Stuart, E.A.: Weight trimming and propensity score weighting. PLOS ONE **6**(3), 18174 (2011) https://doi.org/10.1371/journal.pone.0018174

[33] Kiene, C., Monroy-Hernández, A., Hill, B.M.: Surviving an "Eternal September": how an online community managed a surge of newcomers. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pp. 1152–1156 (2016). https://doi.org/10.1145/2858036.2858356

[34] Ren, Y., Kraut, R., Kiesler, S.: Applying common identity and bond theory to design of online communities. Organization Studies **28**(3), 377–408 (2007) https://doi.org/10.1177/0170840607076007

[35] Ren, Y., Harper, F.M., Drenner, S., Terveen, L., Kiesler, S., Riedl, J., Kraut, R.E.: Building member attachment in online communities: applying theories of group identity and interpersonal bonds. MIS Quarterly **36**(3), 841–864 (2012) https://doi.org/10.2307/41703483

[36] Freeman, G., Acena, D.: Hugging from a distance: building interpersonal relationships in social virtual reality. In: Proceedings of the 2021 ACM International Conference on Interactive Media Experiences, pp. 84–95 (2021). https://doi.org/10.1145/3452918.3458805

[37] Deighan, M.T., Ayobi, A., O'Kane, A.A.: Social virtual reality as a mental health tool: how people use VRChat to support social connectedness and wellbeing. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, pp. 1–13 (2023). https://doi.org/10.1145/3544548.3581103

[38] Rzeszewski, M., Evans, L.: Social relations and spatiality in VR: making spaces meaningful in VRChat. Emotion, Space and Society **53**, 101038 (2024) https://doi.org/10.1016/j.emospa.2024.101038

[39] Leo-Liu, J., Wu-Ouyang, B.: The good, bad, and ugly in social virtual reality: the alternative global village for identity negotiations in VRChat. Social Media + Society **11**(2), 1–13 (2025) https://doi.org/10.1177/20563051251340854

[40] Hide, M., Hatada, Y., Kuzuoka, H., Narumi, T.: "Closer than Real": how social VR platform features influence friendship dynamics. In: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, pp. 1–17 (2025). https:

//doi.org/10.1145/3706598.3714170

[41] Zamanifard, S., Freeman, G.: A surprise birthday party in VR: Leveraging social virtual reality to maintain existing close ties over distance. In: Information for a Better World: Normality, Virtuality, Physicality, Inclusivity, pp. 268–285 (2023). https://doi.org/10.1007/978-3-031-28032-0_23

[42] Sant'Anna, P.H.C., Zhao, J.: Doubly robust difference-in-differences estimators. Journal of Econometrics **219**(1), 101–122 (2020) https://doi.org/10.1016/j.jeconom.2020.06.003

[43] Roth, J., Sant'Anna, P.H.C., Bilinski, A., Poe, J.: What's trending in difference-in-differences? a synthesis of the recent econometrics literature. Journal of Econometrics **235**(2), 2218–2244 (2023) https://doi.org/10.1016/j.jeconom.2023.03.008

[44] Haythornthwaite, C.: Strong, weak, and latent ties and the impact of new media. The Information Society **18**(5), 385–401 (2002) https://doi.org/10.1080/01972240290108195

[45] Tsugawa, S., Niida, S.: Analyzing the effects of social network structure on the growth and survival of online communities in Reddit. IEICE Transactions on Communications **E104.B**(7), 760–769 (2021) https://doi.org/10.1587/transcom.2020CQP0006