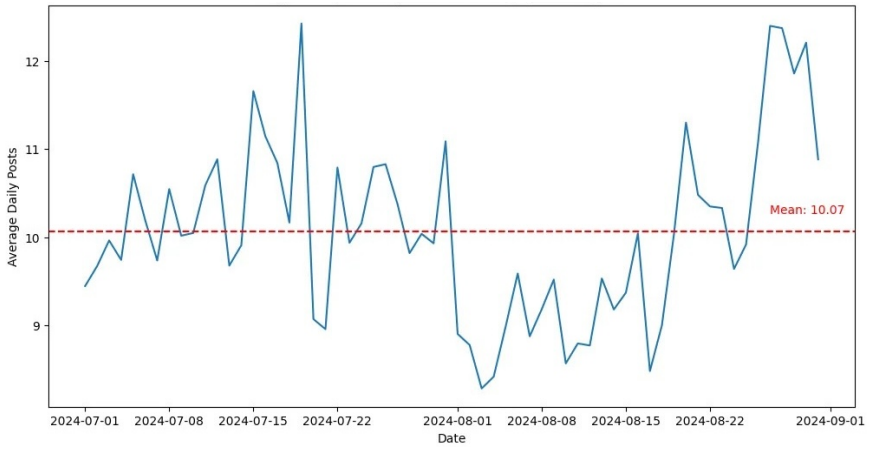
Appendix A: Daily Posts of Influencers

Figure A.1 Average Daily Posts per Influencer



Appendix B: AI Agent's Operation Process, Technology and User Interpretation

B.1. AI Agent's Operation Process

In Weibo's implementation, each AI reply is clearly labeled with an "AI Agent" tag displayed immediately next to the influencer's account name. The AI agent generates and posts replies autonomously without any influencer involvement in the composition, editing, or approval process. Influencer control over the AI agent is highly limited. Influencers cannot adjust the agent's operational parameters, such as response timing or rules determining which users or content the AI agent replies to. These decisions are fully delegated to the autonomous AI system. Influencers can observe AI interactions only after replies have been posted, though we cannot track whether or how closely they monitor these interactions. In summary, influencers in our empirical setting retain only the binary adoption decision, with no control over granular delegation configurations.

B.2. AI Agent's Underlying Technology

Based on publicly available documentation, news reports, and platform statements, the AI agent is powered by *VibeThinker*, a large language model developed by Weibo.[1] According to Weibo's Chief Operating Officer,[2] *VibeThinker*'s base model is trained on high-quality Weibo comments, specifically, those with high reposts, replies, and likes, to ensure quality and representativeness of training data. The model employs reinforcement learning techniques to enhance generalization ability across content domains, enabling effective adaptation to heterogeneous user inputs. For individual influencers, the model also learns their styles from their historical posts. Beyond this high-level description, the platform does not disclose technical details of the AI system.

B.3. User Interpretation of AI Replies

To assess users' understanding of the AI agent and the AI replies, we conducted a survey on Credamo, an online platform widely used in prior research for recruiting Chinese participants (e.g., Hsee and Li 2022; Xu et al. 2023; Zhang et al. 2023). In the survey, we presented participants with an example post published by an influencer, including a user comment and a reply from the influencer's account with an "AI Agent" tag, as shown in Figure B.1 below. Participants were then asked about their understanding of the reply with the

[1] For details, see `https://www.pingwest.com/a/298547`

[2] For details, see `https://cj.sina.com.cn/articles/view/5703921756/v153faf05c001020eng`

AI agent tag. Specifically, they indicated the extent to which they believed the reply came from AI versus from the influencer manually on a 5-point Likert scale.

Among 300 participants, 83% reported believing the reply is generated entirely by AI with minimal to no influencer involvement. This finding suggests that, with the AI agent tag clearly displayed, influencers' followers are likely to correctly attribute AI replies to AI rather than to the influencer's manual efforts. Therefore, it is unlikely that misattribution, i.e., incorrectly believing that the AI replies are written by the influencer, drives our observed effects.

Figure B.1 Survey Stimulus: Influencer Post with an AI Reply

Note. To protect privacy and minimize confounds, we redacted identifying information from the survey stimulus.

Appendix C: Influencers' Adoption of Social AI Agents

C.1. Relationship Between Influencer Characteristics and Adoption Decision

The sample includes only influencers with verified account status, as verification is one of the eligibility criteria for the AI agent feature and is publicly observable. Other eligibility requirements—such as total post views—are not used as filters in our analysis, as this information is only available to the account owner. *RealNameAuth* indicates whether an influencer has real-name authentication. Accounts with 500,000 or more followers are required by Weibo to display their real names; accordingly, this variable serves as a proxy for large-scale influencer status and reflects part of the platform's eligibility criteria.[3]

Table C.1 Adoption of AI Agents by Influencer Characteristics

	DV: AI Agent Adoption
	Logit
AccountTenure	−0.0384
	(0.0401)
LogUpdates	0.1599
	(0.1029)
LogEngagement	−0.0784
	(0.1091)
LogFollowers	0.4265**
	(0.1702)
LogFollowing	0.4237***
	(0.1132)
RealNameAuth	1.4589***
	(0.4234)
Location: Abroad	0.4206
	(0.3847)
Location: Developed Area	−0.0997
	(0.2410)
Commercialized	0.3224
	(0.2344)
Tech Domain	0.4084
	(0.3358)
Male	0.1398
	(0.2534)
AIC	793.9297
BIC	874.8788
Log Likelihood	−384.9648
Deviance	769.9297
Observations	6284

Notes. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

C.2. Influencer's Adoption Rate

While influencer's AI agent adoption rate may appear low, we would like to clarify that it does not mean that only 225 out of the 17,359 influencers on the platform adopted the AI Agent feature. First, the set of eligible influencers is substantially smaller than the full pool of 17,359 influencers. Because the platform does

[3] For details, see `https://restofworld.org/2023/weibo-legal-display-name-influencers/`

not disclose all eligibility criteria, we cannot determine the exact number of eligible influencers. However, one criterion that has been publicly disclosed is verified-account status. Among the 17,359 influencers, only 6,284 have verified accounts, meaning that the potential adopter pool is already less than half of the full population. Additional eligibility criteria, including thresholds based on total post views, are internal metrics accessible only to the platform and account owners. Thus, although we cannot observe all such criteria, it is reasonable to infer that the true eligible population should be considerably smaller than 6,284.

Second, the 225 influencers in our sample represent only those who activated the feature during July and August 2024, the focal period of our study. Adoption outside this window is not captured in this count. Indeed, in later months, adoption continued. For example, we identified 91 additional influencers who activated the feature in January and February 2025. Taken together, these facts indicate that the actual adoption rate is likely much higher than it initially appears, although inherent data limitations prevent us from observing the exact rate.

Appendix D: Summary Statistics and Matching Details

D.1. Summary Statistics of Key Variables

Table D.1 Summary Statistics

Variable	Mean	Std. Dev.	Definition
Outcome Measures			
CommentedPosts	0.29	1.38	Number of posts a user has commented on each day
Treatment			
AIReply	0.02	0.14	Binary variable indicating whether a user has received an AI reply to their comment on the focal post
User Time-Varying Characteristics			
DailyPosts	2.55	8.40	Number of posts published by a user each day
DailyReposts	1.98	7.68	Number of reposts shared by a user each day
User Time-Invariant Characteristics			
LogUserFollowers	4.38	2.49	Log-transformed number of followers the user has
LogUserFollowings	5.89	1.33	Log-transformed number of accounts the user is following
LogUserPosts	6.57	2.44	Log-transformed number of posts ever published by the user
LogUserLikes	4.85	2.69	Log-transformed number of likes received by the user
LogUserComments	4.57	2.84	Log-transformed number of comments received by the user
LogUserReposts	2.81	2.96	Log-transformed number of replies received by the user
PaidUser	0.66	0.48	Binary variable indicating whether a user pays for membership
UserIsFan	0.43	0.50	Binary variable indicating whether a user is a loyal fan of the influencer
UserIsMale	0.46	0.50	Binary variable indicating whether a user is self-identified as male
UserVerified	0.10	0.30	Binary variable indicating whether a user account is verified
Comment Characteristics			
CommentLatency	35.02	79.40	Hours elapsed between post publication and user comment
CommentLength	10.49	17.19	Comment word count
EmojiCount	0.68	1.21	Number of emojis included in the comment
AI Reply Characteristics			
AIReplyContentRelevance	0.6657	0.3009	Relevance of AI reply to user's comment in the context of the original post
AIReplyStyleAlignment	0.4991	0.2646	Alignment of AI reply's style with the influencer's existing style
AIReplyLatency	0.0335	0.1450	Days elapsed between user's comment and AI reply

D.2. Matching Details

For both CEM and PSM, we match users in the treatment and control groups based on a combination of user characteristics and focal comment features. User-level characteristics include the following variables. *LogUserFollowers* and *LogUserFollowings* represent the log-transformed number of a user's followers and followings, respectively. *LogUserPosts* indicates the log-transformed number of posts ever published by the user. *LogUserComments*, *LogUserReposts*, and *LogUserLikes* measure the log-transformed cumulative number of comments, reposts, and likes received by the user. *UserVerified* is a binary indicator for verified accounts, and *UserIsMale* equals 1 if the user is self-identified as male. *UserIsFan* indicates whether the user is a loyal fan of the influencer. *PaidUser* indicates whether the user pays for premium content access.

We also match on focal comment characteristics to account for potential selection on comment-level features. Specifically, we include *CommentLatency*, defined as the number of minutes between the creation time

of the focal post and the time the comment was made, to capture comment promptness. *CommentLength* is the number of characters in the comment. *EmojiCount* captures the number of emojis.

For CEM, we conduct one-to-one matching with replacement. For PSM, we use nearest-neighbor matching based on propensity scores estimated using a generalized linear model (GLM). We implement a one-to-two matching ratio within the same focal post, applying a caliper of 0.2 to ensure close matches in estimated propensity scores. Balance diagnostics for both methods are presented in Table D.2.

Table D.2 Balance Diagnostics: Standardized Mean Differences

	CEM		PSM	
	Original	Matched	Original	Matched
User Characteristics				
LogUserFollowers	0.015	0.019	0.015	-0.016
LogUserFollowings	0.084	-0.007	0.084	-0.054
LogUserPosts	0.047	0.011	0.047	-0.018
LogUserLikes	-0.048	-0.007	-0.048	0.039
LogUserComments	-0.031	-0.014	-0.031	0.020
LogUserReposts	0.019	0.038	0.019	-0.010
PaidUser	-0.004	-0.000	-0.004	0.003
UserIsFan	0.112	-0.000	0.112	-0.100
UserIsMale	0.024	-0.000	0.024	-0.013
UserVerified	-0.001	-0.000	-0.001	0.005
Comment Characteristics				
CommentLatency	0.484	0.061	0.484	0.360
CommentLength	-0.083	0.043	-0.083	0.071
EmojiCount	-0.212	-0.060	-0.212	0.003

Appendix E: Robustness Checks

E.1. Event Study Analysis

A critical assumption of the DID design is that the treatment and control groups would have followed parallel trends in the outcome variable in the absence of treatment (Angrist and Pischke 2009). To check if this assumption holds, we conduct an event study analysis. Specifically, we construct relative time indicators around each user's treatment date, which capture user behavior in event time and allow us to estimate dynamic treatment effects. We define a 17-day window for both the pre-treatment and post-treatment periods. The pre-treatment window provides a sufficient timeframe to test for parallel trends, while the post-treatment window enables us to examine the persistence of treatment effects. Figure E.1 plots the estimated lead and lag coefficients. The pre-treatment coefficients are statistically insignificant, supporting the parallel trends assumption. The post-treatment coefficients are positive and statistically significant throughout the two weeks, indicating that receiving an AI reply increases user commenting behavior with effects persisting for around two weeks. We further evaluate the long-run effects of a single AI reply by extending our post-treatment observation window to six months. We collected historical posts and comments for all focal influencers from July 2024 through December 2024. As shown in Figure E.2, the treatment effect gradually diminishes after around two weeks. We believe this pattern is consistent with the nature of the treatment. A single reply is unlikely to permanently reshape a user's engagement patterns. To the best of our knowledge, the fact that a single reply can sustain increased engagement for two weeks already represents a substantial behavioral effect.

Figure E.1 Event Study Analysis

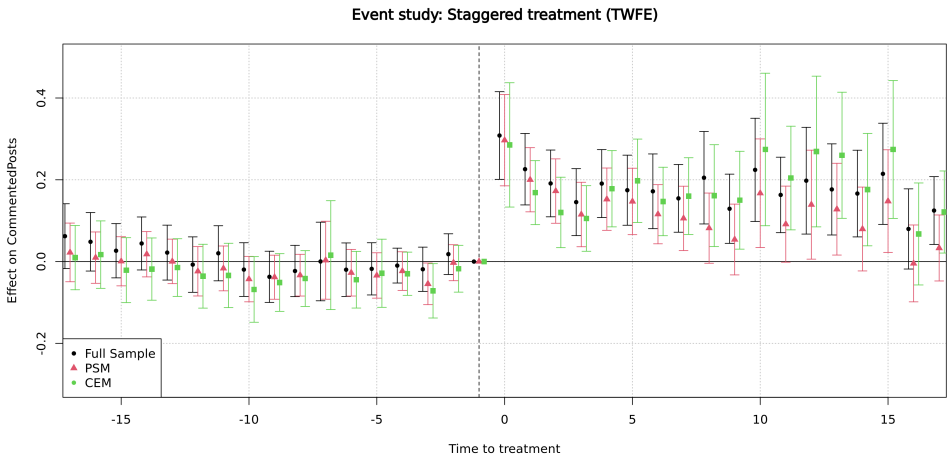
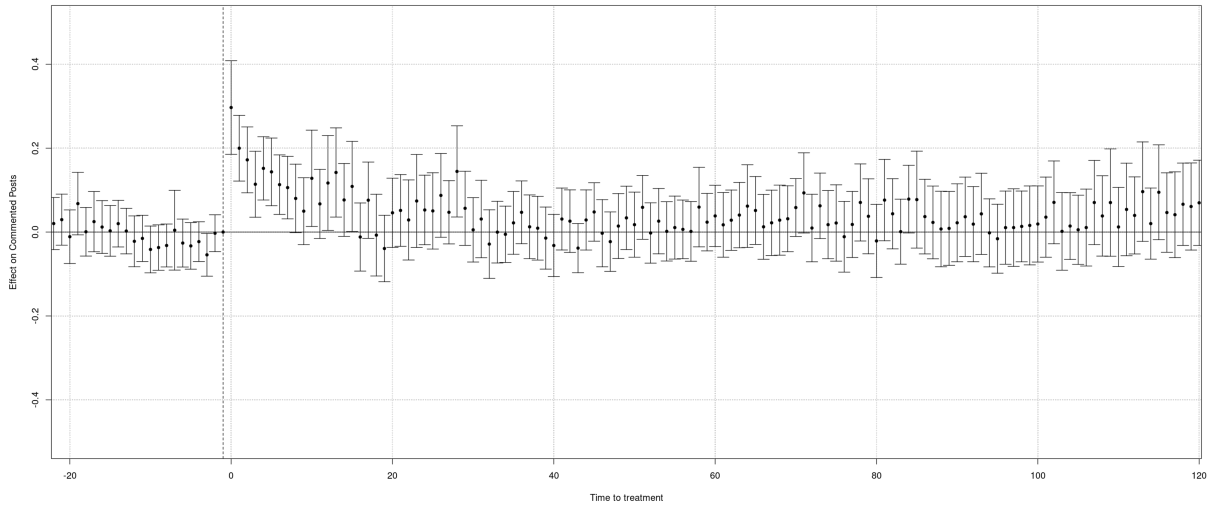


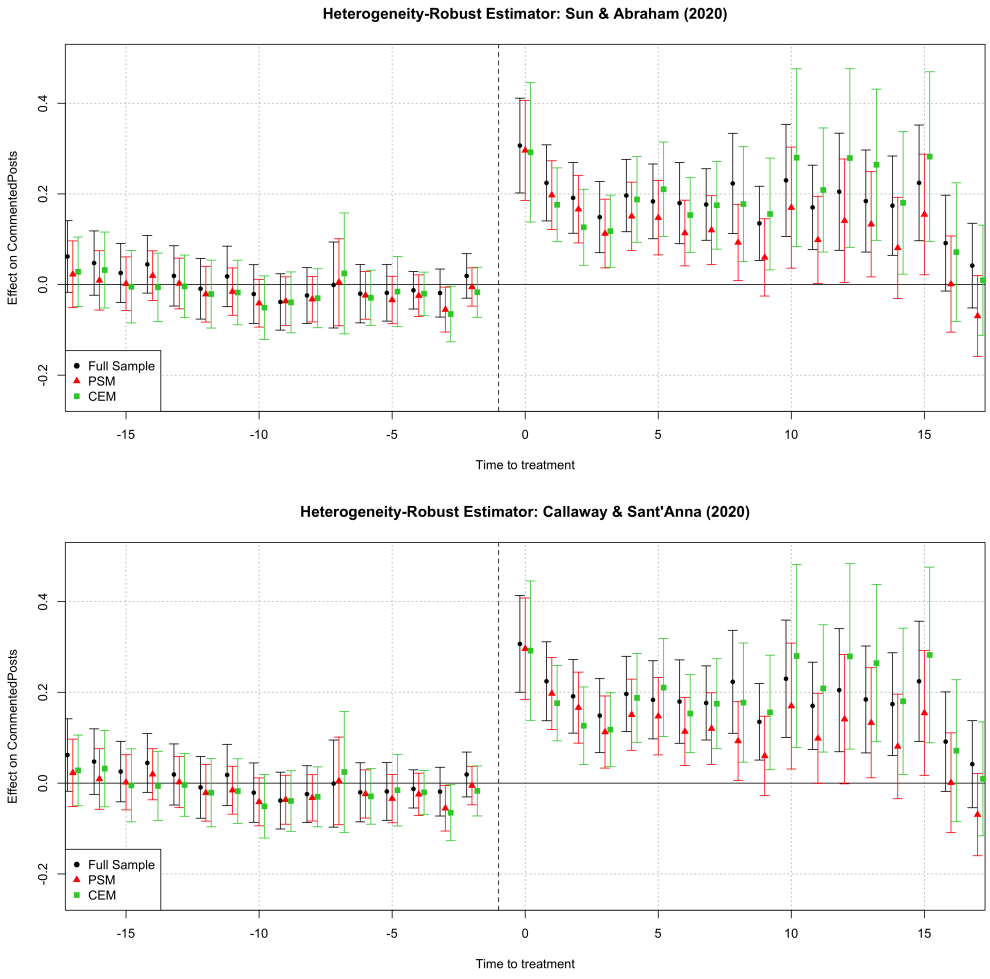
Figure E.2 Event Study Analysis with an Extended Window



E.2. Heterogeneity-Robust Estimators

A well-documented concern with the fixed effects estimator is its potential to yield biased estimates in settings with staggered treatment timing. This bias arises because the fixed effects estimator implicitly uses already-treated units as controls for later-treated units, a structure that can generate so-called "negative weighting" and bias the estimation of treatment effects (Goodman-Bacon 2021). To address this issue, we employ two alternative estimators to accommodate treatment effect heterogeneity in staggered designs. The first estimator is proposed by Callaway and Sant'Anna (2021) and the second estimator is developed by Sun and Abraham (2021). Figure E.3 presents the results using these two estimators. Consistent with our main analysis, both estimators produce positive and statistically significant effects of receiving an AI reply on user commenting behavior.

Figure E.3 Heterogeneity-Robust Estimators



E.3. Alternative Specifications

In the main analysis, we estimate a linear fixed effects model to examine the treatment effect. To account for the fact that the outcome, *CommentedPosts*, is a count variable, we estimate Poisson regressions with user, day, and focal post fixed effects to demonstrate the robustness of our findings across different model specifications. The results are presented in Table E.1. The coefficients of *AIReply* are consistently positive and statistically significant across all columns, aligning with the results in our main analysis.

Table E.1 Count Model

	CommentedPosts					
	Full Sample		PSM		CEM	
	(1)	(2)	(3)	(4)	(5)	(6)
AIReply	0.4128***	0.4003***	0.3809***	0.3645***	0.6352***	0.6216***
	(0.0638)	(0.0639)	(0.0749)	(0.0732)	(0.1193)	(0.1164)
User Controls		✓		✓		✓
User fixed effects	✓	✓	✓	✓	✓	✓
Day fixed effects	✓	✓	✓	✓	✓	✓
Post fixed effects	✓	✓	✓	✓	✓	✓
R^2	0.51	0.51	0.52	0.53	0.51	0.51
Observations	1,403,990	1,403,990	199,950	199,950	88,040	88,040

Notes. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the user level.

The outcome variable used in our main analysis is the number of posts that a user comments on, which is a continuous variable. To assess the robustness of our results across different outcome variable specifications, we construct *LeftComment*, a binary variable indicating whether a user leaves any comments on the influencer's posts each day. Table E.2 presents the results using *LeftComment* as the dependent variable. The coefficients of *AIReply* remain statistically positive across all samples and model specifications, indicating that receiving an AI reply significantly increases the likelihood of leaving comments.

Table E.2 Binary Outcome

	LeftComment					
	Full Sample		PSM		CEM	
	(1) OLS	(2) Logit	(3) OLS	(4) Logit	(5) OLS	(6) Logit
AIReply	0.0338***	0.5296***	0.0359***	0.6017***	0.0535***	0.8107***
	(0.0045)	(0.0629)	(0.0052)	(0.0759)	(0.0064)	(0.0934)
User Controls	✓	✓	✓	✓	✓	✓
User fixed effects	✓	✓	✓	✓	✓	✓
Day fixed effects	✓	✓	✓	✓	✓	✓
Post fixed effects	✓	✓	✓	✓	✓	✓
R^2	0.38	0.32	0.34	0.31	0.33	0.30
Observations	1,403,990	1,403,990	199,950	199,950	88,040	88,040

Notes. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the user level.

E.4. Comment Hour Fixed Effects

We incorporate comment hour fixed effects into the model to control for temporal variations across different hours of the day. As shown in Table E.3 below, the results remain identical to our main estimates.

Table E.3 TWFE Analysis with Comment Hour Fixed Effects

	CommentedPosts					
	Full Sample		PSM		CEM	
	(1)	(2)	(3)	(4)	(5)	(6)
AIReply	0.1074***	0.1046***	0.0982***	0.0962***	0.1630***	0.1598***
	(0.0207)	(0.0206)	(0.0236)	(0.0234)	(0.0329)	(0.0325)
User Controls		✓		✓		✓
User fixed effects	✓	✓	✓	✓	✓	✓
Day fixed effects	✓	✓	✓	✓	✓	✓
Post fixed effects	✓	✓	✓	✓	✓	✓
Comment Hour fixed effects	✓	✓	✓	✓	✓	✓
R^2	0.51	0.52	0.50	0.51	0.47	0.47
Observations	1,403,990	1,403,990	199,950	199,950	88,040	88,040

Notes. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the user level.

E.5. Selection Bias

To assess whether AI agents selectively target specific users or comments, we reverse-engineer their targeting strategy by examining how observable characteristics affect the likelihood of receiving an AI reply. We construct a sample using posts from the first day of influencers' AI Agent activation. The dependent variable is *AIReply*, indicating whether the user received an AI reply to their comment. The independent variables include user characteristics, comment characteristics, and users' prior commenting activities.

For user characteristics, we capture users' profile information, including the number of followers (*LogUserFollowers*), followings (*LogUserFollowings*), total posts (*LogUserUpdates*), and total engagements received (*LogUserEngagements*). We also include indicator variables for whether the user is verified (*UserVerified*), user gender (*UserIsMale*), whether the user is a loyal fan of the focal influencer (*UserIsFan*), and whether the user has a paid membership status (*PaidUser*).

For comment characteristics, we examine several features that might influence the AI agent's selection process. First, we test whether the AI agent targets critical comments, as platforms may strategically prioritize addressing negative feedback to manage influencer reputation. We capture this using sentiment analysis (*NegativeComment*). Second, we examine comment length (*CommentLength*), as longer comments may signal more engaged or invested users whom influencers seek to retain. Third, we assess comment speed (*CommentLatency*), measured as the time elapsed between post publication and commenting, to determine whether the AI agent prioritizes fast comments.

Additionally, other content factors may influence comment selection.[4] Specifically, the platform may program the AI agent to avoid replying to controversial content or to preferentially target comments that are closely related to the context of the original post. To address these concerns, we construct two additional variables: *CommentToxicity* and *CommentPostSimilarity*. First, to capture the possibility that the platform avoids controversial content, we measure the toxicity of each comment using the Perspective API.[5] This API uses pre-trained machine learning models to generate a probability score indicating the likelihood that a text contains toxic content, and it has been widely used in toxicity detection on social media platforms (Lees et al. 2022). Second, to assess whether the AI agent targets comments that are more aligned with the content of the original post, we compute the semantic similarity between the comment and the post. We obtain text embeddings for both the comment and the post using a pre-trained multilingual sentence transformer model.[6] The embeddings of the two texts are then used to calculate the cosine similarity, which serves as an indicator of how closely the comment relates to the original post.

For users' prior commenting activities, we examine users' historical commenting behavior to test whether the AI agent strategically targets more active users. The platform may prioritize replying to frequent commenters to strengthen relationships with highly engaged users. To capture this, we measure *PrevCommentedPosts*, the number of posts the user commented on during multiple pre-periods: 1 day, 3 days, 7 days, 2 weeks, 4 weeks, and 2 months prior to the focal comment. Additionally, the AI agent may prioritize users whose comments historically generate high engagement to stimulate broader interaction. To test this, we incorporate *PrevCommentsPopularity*, measured as the average number of engagements (likes plus replies) received on a user's prior comments within the same set of lookback windows.[7] This allows us to assess whether users who historically generated more popular comments were preferentially targeted.

We then estimate logistic regressions with day fixed effects and focal-post fixed effects, with results reported in Table E.4. Each column corresponds to a different pre-period window (e.g., Column 1 uses 1-day measures; Column 3 uses 7-day measures). Across all windows, both *PrevCommentedPosts* and *PrevCommentsPopularity* remain statistically insignificant predictors of receiving an AI reply. These results suggest that users'

[4] We thank an anonymous reviewer for suggesting these factors.

[5] For details, see `https://perspectiveapi.com/`

[6] For details, see `https://huggingface.co/sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2`

[7] We thank an anonymous reviewer for this suggestion.

prior commenting frequency and prior comment popularity do not significantly influence their probability of receiving an AI reply. Furthermore, none of the user characteristics or comment characteristics significantly affect the probability that a comment or user receives an AI reply. While we can not completely rule out the possibility of targeting based on unobserved factors, the results suggest that AI agents do not appear to target users or comments based on the observable characteristics we examined.

Table E.4 AI Reply Targeting Strategy

	AIReply					
	1 Day	3 Days	7 Days	2 Weeks	4 Weeks	2 Months
	(1)	(2)	(3)	(4)	(5)	(6)
User Characteristics						
LogUserFollowers	0.0071	0.0069	0.0070	0.0074	0.0081	0.0080
	(0.0190)	(0.0189)	(0.0189)	(0.0190)	(0.0190)	(0.0189)
LogUserFollowings	-0.0445	-0.0431	-0.0438	-0.0448	-0.0473	-0.0475
	(0.0358)	(0.0357)	(0.0357)	(0.0356)	(0.0356)	(0.0357)
LogUserPosts	-0.0185	-0.0194	-0.0188	-0.0178	-0.0165	-0.0160
	(0.0347)	(0.0344)	(0.0344)	(0.0345)	(0.0343)	(0.0343)
LogUserEngagement	0.0236	0.0240	0.0237	0.0233	0.0234	0.0232
	(0.0282)	(0.0281)	(0.0281)	(0.0282)	(0.0280)	(0.0280)
UserVerified	-0.0160	-0.0256	-0.0210	-0.0132	-0.0035	-0.0049
	(0.1784)	(0.1789)	(0.1787)	(0.1784)	(0.1778)	(0.1779)
UserIsMale	-0.0237	-0.0229	-0.0232	-0.0234	-0.0204	-0.0196
	(0.0749)	(0.0749)	(0.0750)	(0.0754)	(0.0761)	(0.0761)
UserIsFan	-0.2769	-0.2828	-0.2794	-0.2720	-0.2580	-0.2592
	(0.2146)	(0.2143)	(0.2155)	(0.2191)	(0.2214)	(0.2195)
PaidUser	-0.1236	-0.1257	-0.1243	-0.1220	-0.1186	-0.1192
	(0.0787)	(0.0785)	(0.0782)	(0.0784)	(0.0781)	(0.0780)
Comment Characteristics						
NegativeComment	-0.0523	-0.0509	-0.0514	-0.0518	-0.0483	-0.0508
	(0.1242)	(0.1243)	(0.1243)	(0.1249)	(0.1246)	(0.1246)
CommentLatency	0.0038	0.0038	0.0038	0.0037	0.0037	0.0037
	(0.0062)	(0.0062)	(0.0062)	(0.0062)	(0.0061)	(0.0061)
CommentLength	2.297	2.340	2.327	2.285	2.277	2.411
	(2.726)	(2.725)	(2.727)	(2.721)	(2.730)	(2.704)
CommentPostSimilarity	-0.4745	-0.4732	-0.4728	-0.4700	-0.4632	-0.4603
	(0.6109)	(0.6107)	(0.6116)	(0.6144)	(0.6150)	(0.6145)
CommentToxicity	1.208	1.206	1.207	1.210	1.216	1.210
	(0.9398)	(0.9392)	(0.9380)	(0.9367)	(0.9345)	(0.9379)
User Prior Comments						
PrevCommentsPopularity	0.0014	-0.0034	-0.0056	-0.0175	-0.0823	-0.0544
	(0.0019)	(0.0034)	(0.0071)	(0.0436)	(0.0642)	(0.0496)
PrevCommentedPosts	-0.0272	0.0091	0.0002	-0.0036	-0.0028	-0.0017
	(0.0576)	(0.0265)	(0.0105)	(0.0050)	(0.0024)	(0.0012)
Day fixed effects	✓	✓	✓	✓	✓	✓
Post fixed effects	✓	✓	✓	✓	✓	✓
R^2	0.16	0.16	0.16	0.16	0.16	0.16
Observations	22,645	22,645	22,645	22,645	22,645	22,645

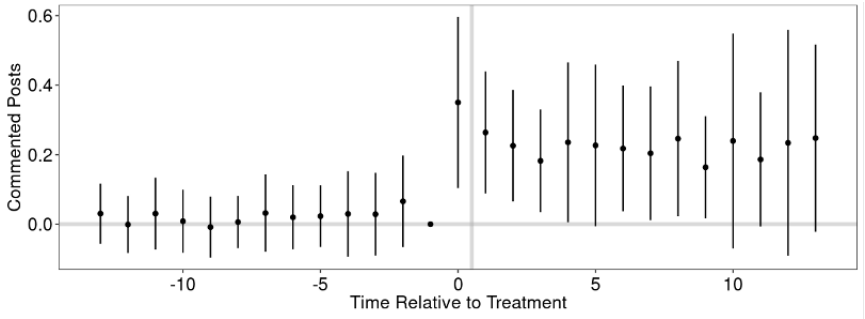
Notes. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the post level.

E.6. Spillover Effects on Control Users

A potential concern is that control users may observe AI replies directed at others, which could influence their own behavior and thereby threaten SUTVA. To empirically alleviate this concern, we conduct three sets of analyses with different alternative sets of control users.

Control Users Never Exposed to Any AI Replies. To ensure control users were never exposed to any AI replies from any influencers, we construct an alternative control group using users who commented on the focal influencer's posts during the first day of May 2024, a period before the platform even launched the AI agent feature at all. Users commenting in this time period could not have been exposed to AI replies from any influencer across the platform. We then implement a stacked DID design (Cengiz et al. 2019) because the focal timing does not overlap with the observation window of the treatment group. Stacking cohorts by relative event time rather than calendar time ensures comparability between treatment and control groups despite differences in focal timing. We set an observation window of two weeks before the focal date and two weeks after. For treatment users, we restrict the observation window to the same pre- and post-periods. Users in the treatment group are assigned to cohorts based on the date they received an AI reply and then stacked with the control users. We then conduct the event study to estimate the treatment effects with user-cohort fixed effects, time-cohort fixed effects, and focal-influencer fixed effects. The result, shown in Figure E.4, demonstrates no significant pre-treatment differences between treatment and control groups, supporting the parallel-trends assumption. In addition, we observe a significant increase in users' commenting frequency after receiving an AI reply, indicating that our main findings remain robust even when restricting control users to those with no opportunity for exposure to AI replies anywhere on the platform.

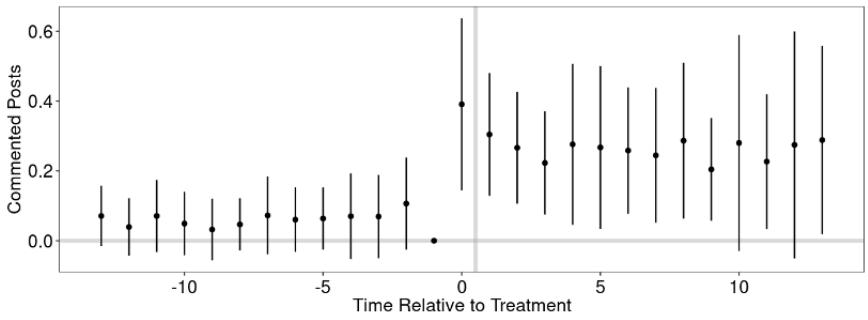
Figure E.4 Stacked DID Result – Control Users from May Sample



Control Users from Prior Day of Agent Activation. An issue with the first set of alternative control users is that their focal timing (May 1) is less tightly aligned to that of the treatment group. To alleviate

this concern, we construct an alternative control group using users who commented on the focal influencer's prior-day posts—specifically, posts published one day before AI agent activation. Because the AI agent was not yet active when these users commented, they could not have been exposed to AI replies from the focal influencer at the time of commenting. Recall that our treatment group is based on posts published on the influencer's first day of activation; thus, using focal comments from the immediately preceding day ensures the control group's commenting timing to closely align that of the treatment group, thereby minimizing, to the best extent possible, concerns related to comment-timing confounders. For this set of control users, we additionally ensure that they never received any AI replies thereafter. Following the approach used in the first set of alternative control users, we also implement a stacked DID design. The event study result, shown in Figure E.5 (prior-day control users), again shows no significant pre-treatment differences between treatment and control groups, and a significant increase in the post-treatment period. Therefore, our main findings still remain robust using this set of alternative control users.

Figure E.5 Stacked DID Result – Control Users from Prior Day of Activation



Control Users Before First AI Reply. To further control for the post-level heterogeneity, we also restrict the control group to a clean subsample consisting only of users whose comments were posted before the first AI reply appeared under the focal post.[8] These users, by construction, had no exposure to AI replies at the time they commented, ensuring that the control group is not contaminated by spillovers from AI replies under the focal post. We then re-estimate our main model using this refined sample. As shown in Table E.5, the coefficients of *AIReply* remain significantly positive, thereby alleviating the concern of potential SUTVA violation.

[8] We thank an anonymous reviewer for suggesting this approach.

Table E.5 TWFE - Control Users Before First AI Reply

	CommentedPosts		
	Full Sample	PSM	CEM
	(1)	(2)	(3)
AIReply	0.1066***	0.0862***	0.1560***
	(0.0206)	(0.0271)	(0.0299)
User Controls	✓	✓	✓
User fixed effects	✓	✓	✓
Day fixed effects	✓	✓	✓
Post fixed effects	✓	✓	✓
R^2	0.56	0.53	0.45
Observations	973,710	119,288	67,766

Notes. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the user level.

Overall, these analyses demonstrate that our findings are robust to alternative control group specifications designed to address potential SUTVA violations. The persistently significant and positive effects are unlikely to be a byproduct of control users witnessing AI interactions with others.

E.7. Effect Persistence After Widespread Adoption

For these 91 later-adopting influencers, we focused on their posts published within the window of January and February of 2025. Treatment users are those who commented and received an AI reply for the first time within the new window, and control users are those who commented on the same posts but did not receive any AI reply. This resulted in 2,413 focal posts, 4,395 treated users, and 22,391 control users. As in our main analysis, we applied PSM and CEM to construct matched samples balanced on observable user characteristics. The results, reported in Table E.6, continue to show a significantly positive effect of receiving an AI reply on subsequent engagement across all samples.

Table E.6 External Validity: TWFE Analysis Using Later-adopting Influencers

	CommentedPosts		
	Full Sample	PSM	CEM
	(1)	(2)	(3)
AIReply	0.0164***	0.0194***	0.0169***
	(0.0027)	(0.0027)	(0.0025)
User fixed effects	✓	✓	✓
Day fixed effects	✓	✓	✓
Post fixed effects	✓	✓	✓
R^2	0.34	0.27	0.23
Observations	1,580,374	758,681	445,922

Notes. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the user level.

E.8. Posts Without Human Replies

One possible concern in our study is that AI replies might appear alongside human replies from influencers themselves, making it difficult to determine whether the observed increase in user commenting behavior is due to AI replies, human replies, or a combination of both. To mitigate this concern, we conduct a subsample analysis using posts with no human replies, therefore isolating the effect of AI replies. Table E.7 presents the results, which show that in the absence of human replies, receiving an AI reply still leads to a significant increase in user commenting behavior.

Table E.7 Subsample Analysis: Posts Without Human Replies

	CommentedPosts		
	Full Sample	PSM	CEM
	(1)	(2)	(3)
AIReply	0.1755***	0.1466***	0.2395***
	(0.0419)	(0.0501)	(0.0625)
User Controls	✓	✓	✓
User fixed effects	✓	✓	✓
Day fixed effects	✓	✓	✓
Post fixed effects	✓	✓	✓
R^2	0.52	0.49	0.46
Observations	707,172	86,490	40,238

Notes. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the user level.

Appendix F: Measurement of AI Reply Content Relevance and Style Alignment

The relevance of an AI reply is operationalized as the extent to which the reply directly addresses the user's comment, given the post content as the context. A reply is considered highly relevant when it responds to the user's comment in a way that is specific, on-point, and tailored to the content of the comment. Conversely, replies that are off topic of the comment are classified as low relevance. This operationalization is guided by our theoretical premise: for AI agents to effectively replicate the influencer's social presence, replies need to demonstrate attentiveness and understanding. Off-topic replies that do not speak to the substance of the user's comment reflect influencers' limited effort in understanding the user, thereby signaling low presence of the influencer. Likewise, generic or templated replies also tend to convey a weaker social presence because they fail to create the impression of a thoughtful or individualized response.

We employ the GPT-4.1 model with few-shot prompting, providing the LLM with few examples from which it could learn generalized rules. We include one example illustrating a high-relevance AI reply (i.e., AI reply that meaningfully addresses the user's comment), one example of medium relevance (i.e., partially to the point but with some abrupt topic shifting), and two examples illustrating low relevance: one involving an off-topic reply (indicating that the AI did not understand the comment) and one involving a templated reply. For each example, we provide a brief rationale for the expected output, enabling the LLM to learn not only the pattern but also the underlying reasoning process. In addition, our prompt follows a structured format, including task description, construct definition, few-shot examples, and explicit input/output specifications, which aligns with best-practice prompting frameworks shown in prior work to improve LLM performance in text-analysis tasks (e.g., Liu et al. 2025).

Furthermore, because we are measuring the extent to which the AI reply addresses user comments well, *AIReplyContentRelevance* should inherently be continuous. We therefore instruct the LLM to produce a relevance score ranging from 0 to 1. To ensure that the model interprets the scoring scale consistently with human judgments, we explicitly specify score ranges that correspond to different levels of relevance. Without such scale anchoring, LLM outputs may tend to be overly centered toward mid-range or extreme values. We therefore specify the intended relevance level of each range (e.g., < 0.4 for low relevance, > 0.6 for high relevance, and 0.4—0.6 for moderate relevance). To further align the model's scoring, we apply the same score ranges when providing few-shot examples. Accordingly, we use the five-part prompting structure, with few-shot examples (including rationales) and explicit score-scale alignment, as listed in Table F.1.

Table F.1 Prompt for AI Reply Content Relevance

Prompt component	Content (Translated to English)
Task Description	You are a Chinese text–processing model specializing in semantic analysis. You will be given a text containing three components: • **Post**: the content of the original post (serves as contextual background) • **Comment**: the user's comment under that post • **Reply**: the reply to that comment Your task is to evaluate how well the Reply addresses the Comment (i.e. reply relevance), taking the Post as contextual information.
Construct Definition (with Scoring Rules)	Reply relevance refers to the extent to which the Reply directly responds to the substance of the Comment, given the Post as context. • **High relevance**: specific, on-point, tailored replies that directly address the user's comment. • **Medium relevance**: replies that partially address the comment but also shift topics somewhat abruptly. • **Low relevance**: replies that do not address the comment, are off-topic, or are generic/templated (e.g., "Thanks for your support!", "Good luck!", "Thanks for following!"). You must output a single score between 0 and 1: • 1 = the reply is fully tailored to and directly addresses the comment. • 0 = the reply is completely unrelated or purely templated. To maintain consistent scale, use the following scoring intervals: • **Low relevance**: $0 \leq \text{score} < 0.40$ • **Medium relevance**: $0.40 \leq \text{score} \leq 0.60$ • **High relevance**: $0.60 < \text{score} \leq 1$
Few-shot Examples	Example 1 (High relevance: > 0.6, close to 1) **Post**: If you had a game account that has been sitting unused for a long time, would you keep it or sell it? #EsportsChat **Comment**: Keep it **Reply**: Keeping it can be a nice way to hold onto memories—who knows, you might suddenly feel like playing again someday! And who knows, maybe it'll even increase in value in the future! **Rationale**: The reply directly engages with the user's answer "keep it," elaborating on the reasons to keep the account in a personalized and contextually meaningful way. → **Expected score**: > 0.6, close to 1 Example 2 (Medium relevance: 0.4–0.6) Post: Why doesn't BlackBerry abandon physical keyboards? Comment: Good evening. Reply: Good evening! The digital world is shining bright—let's talk about BlackBerry! Rationale: The reply briefly responds to the greeting but then abruptly shifts back to the topic. Some relevance but not strong. → Expected score: 0.4–0.6 Example 3 (Low relevance: ¡0.4) Post: I recently fell in love with the new Mi Pocket Photo Printer 1S. Compared with the previous generation, the texture has improved significantly, and its compact size makes it convenient to carry around. It still uses 3-inch photo paper and prints one sheet in about 45 seconds. Priced at 399 RMB—would you buy it? #EVAtechlife #GadgetRecommendations Comment: You do a giveaway? Reply: Sending today's good luck your way! Rationale: The reply is a generic, templated "good luck" message that does not address the user's comment in any substantive or specific way. It shows no evidence of understanding the comment or the post context. → Expected score: < 0.4 Example 4 (Low relevance: ¡0.4) Post: See you in Nanchang this Saturday! Repost for a chance to win 5 sets of our grand-opening limited-edition gift packs! Comment: What's in the gift pack? Reply: Sending you an extra luck buff! Rationale: The user asks a specific informational question ("What's in the gift pack?"), but the reply provides no relevant information, does not address the question, and instead uses a generic, playful phrase. The response shows no evidence of understanding the user's comment and appears templated or pre-set. → Expected score: < 0.4
Input specifications	Input: Post: [post content]; Comment: [comment content]; Reply: [reply content]
Output specifications	Output only a single numerical score between 0 and 1.

We then evaluate the performance of the relevance measure using manual coding. Specifically, we randomly sample 300 AI replies, together with their associated comments and posts. For human coders, it could be difficult to assign precise numerical scores for the perceived extent of relevance. Therefore, when labeling each instance, we use three qualitative categories (high, medium, and low relevance) based on human assessment of how well the AI reply addresses the user's comment in the context of the post. We then map the GPT-generated relevance scores into these same categories using the score ranges (i.e., < 0.4 = low; 0.4—0.6 = medium; > 0.6 = high) and compared these with the human-coded labels. As shown in Table F.2, the GPT-based relevance measure achieves an overall accuracy of 0.94, indicating satisfactory performance.

Table F.2 Performances of GPT-based Relevance Classifier

	Precision	**Recall**	**F_1 Score**	**Support**
Performance by Relevance Level				
Low Relevance	0.98	0.95	0.97	117
Medium Relevance	0.69	0.52	0.59	21
High Relevance	0.94	0.99	0.97	162
Overall Relevance				
Macro Avg	0.87	0.82	**0.84**	300
Weighted Avg	0.94	0.94	**0.94**	300
Accuracy			**0.94**	300

To measure style alignment, we follow a similar approach. We prompt the GPT-4o-mini model to evaluate the stylistic similarity between the AI reply and the influencer's prior content. Specifically, for each influencer, we compile their original posts published in the three days preceding the focal post as the reference content for their existing styles. The complete set of this prior content and the AI reply are input to the model. The GPT model is instructed to evaluate stylistic similarity between the two sets of text, specifically focusing on tone, word choice, sentence structure, and linguistic style, while ignoring semantic content. We apply few-shot learning to guide the model's assessment, providing two illustrative sets of examples.

The first example pair contains two texts that are semantically similar but stylistically different: "The traffic today was awful—I was about to lose my mind!" versus "Today's traffic conditions were undesirable and significantly affected travel efficiency." The model is prompted to recognize that although the meanings are similar in this example, the styles differ, which should result in low style alignment. The second example pair contains two texts with opposite meanings but similar styles: "Wow, my new haircut looks amazing today—I'm so excited!" versus "Ugh, my new haircut looks terrible today—I'm so upset!" Despite the opposite semantic meanings, the model is instructed to recognize the high similarity in tone and expression. These

examples help the model learn to assess stylistic similarity independently of semantic meaning. The model assigns a style alignment score between 0 and 1, with higher values indicating greater stylistic similarity.

Since style alignment can be subtle and therefore difficult for human coders to assess consistently, we validate the robustness of our results using an alternative large language model. Specifically, we replicate the style alignment measurement procedure using the DeepSeek-Chat model, which is well-suited for Chinese language processing due to its extensive Chinese-language training corpus. We then re-estimate the mechanism test using the DeepSeek-based style alignment scores. As shown in Table F.3, the results consistently support the positive effect of style alignment in AI replies on enhancing user engagement.

Table F.3 Mechanism Test of Style Alignment

	CommentedPosts		
	Full Sample	PSM	CEM
	(1)	(2)	(3)
AIReplyStyleAlignment	0.1810***	0.1604***	0.2456***
	(0.0400)	(0.0438)	(0.0576)
User Controls	✓	✓	✓
User fixed effects	✓	✓	✓
Day fixed effects	✓	✓	✓
Post fixed effects	✓	✓	✓
R^2	0.52	0.51	0.47
Observations	1,403,990	199,950	88,040

Appendix G: Additional Heterogeneity Analysis

In our sample, influencers replied to comments themselves on 168 out of 448 posts (37.5%). To examine the heterogeneous treatment effect regarding influencer's human reply, we conducted two subsample analyses based on the sequence of AI reply versus human reply: (1) posts where the influencer's human reply appeared before any AI reply, and (2) posts where the AI agent replied before the influencer. Among the 168 posts that contain both AI replies and human replies, the results are as follows. For posts in which the human reply occurred first, the results in Table G.1 show no significant effect of receiving an AI reply on subsequent user engagement. In contrast, for posts where the AI reply occurred first, the results in Table G.2 indicate a positive and significant effect, with magnitudes larger than those reported in our main analysis.

Table G.1 Subsample Analysis – Human Reply First

	CommentedPosts		
	Full Sample	PSM	CEM
	(1)	(2)	(3)
AIReply	0.0549	0.0475	0.0214
	(0.0390)	(0.0331)	(0.0241)
User Controls	✓	✓	✓
User fixed effects	✓	✓	✓
Day fixed effects	✓	✓	✓
Post fixed effects	✓	✓	✓
R^2	0.52	0.59	0.59
Observations	406,162	32,240	17,980

Notes. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the user level.

Table G.2 Subsample Analysis – AI Reply First

	CommentedPosts		
	Full Sample	PSM	CEM
	(1)	(2)	(3)
AIReply	0.2078***	0.2337***	0.2705***
	(0.0603)	(0.0786)	(0.0973)
User Controls	✓	✓	✓
User fixed effects	✓	✓	✓
Day fixed effects	✓	✓	✓
Post fixed effects	✓	✓	✓
R^2	0.44	0.52	0.41
Observations	43,400	14,880	5,394

Notes. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the user level.

These findings suggest that when users first observe a human reply from the influencer, they are more likely to perceive subsequent AI replies as weaker extensions of the influencer's social presence because the influencer's actual presence has already been established, thereby weakening the extent to which AI-delegated interactions can sustain social presence. These results are consistent with our heterogeneity analysis, where we find that the effect of AI replies is weaker when influencers themselves also reply.

Appendix H: Heterogeneity Analysis with Matched Samples

Table H.1 Heterogeneity Analysis with Matched Samples - CEM

	CommentedPosts					
	(1)	(2)	(3)	(4)	(5)	(6)
AIReply	0.0739***	0.2564***	0.2659***	0.1660***	0.2351***	0.2128***
	(0.0257)	(0.0626)	(0.0514)	(0.0339)	(0.0552)	(0.0411)
AIReply × UserIsFan	0.1512***					
	(0.0492)					
AIReply × InfluencerCommercialized		-0.1688***				
		(0.0623)				
AIReply × TechInfluencer			-0.2220***			
			(0.0501)			
AIReply × InfluencerPrevReplyRate				-0.2143*		
				(0.1232)		
AIReply × InfluencerReplied					-0.1469***	
					(0.0557)	
AIReply × AIReplyPrevalence						-0.0004***
						(7.9×10^{-5})
User Controls	✓	✓	✓	✓	✓	✓
User fixed effects	✓	✓	✓	✓	✓	✓
Day fixed effects	✓	✓	✓	✓	✓	✓
Post fixed effects	✓	✓	✓	✓	✓	✓
R^2	0.47	0.47	0.47	0.47	0.47	0.47
Observations	88,040	88,040	88,040	88,040	88,040	88,040

Notes. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the user level.

Table H.2 Heterogeneity Analysis with Matched Samples - PSM

	CommentedPosts					
	(1)	(2)	(3)	(4)	(5)	(6)
AIReply	0.0493**	0.1745***	0.2130***	0.1040***	0.1847***	0.1699***
	(0.0210)	(0.0502)	(0.0419)	(0.0248)	(0.0474)	(0.0337)
AIReply × UserIsFan	0.0848**					
	(0.0403)					
AIReply × InfluencerCommercialized		-0.1224**				
		(0.0524)				
AIReply × TechInfluencer			-0.2112***			
			(0.0447)			
AIReply × InfluencerPrevReplyRate				-0.3079***		
				(0.0920)		
AIReply × InfluencerReplied					-0.1537***	
					(0.0485)	
AIReply × AIReplyPrevalence						-0.0004***
						(7.14×10^{-5})
User Controls	✓	✓	✓	✓	✓	✓
User fixed effects	✓	✓	✓	✓	✓	✓
Day fixed effects	✓	✓	✓	✓	✓	✓
Post fixed effects	✓	✓	✓	✓	✓	✓
R^2	0.51	0.51	0.51	0.51	0.51	0.51
Observations	199,950	199,950	199,950	199,950	199,950	199,950

Notes. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the user level.

Appendix I: Persistence of Treatment Effect

I.1. Effect Persistence Beyond Users' First AI Rely

To test whether the effect persists beyond the initial hype of receiving an AI reply for the first time, we construct a new sample in which treated users are those who received a *second AI reply*. We further restrict the sample to cases where the second reply occurred at least 5, 10, or 14 days after the first, ensuring that any novelty effect would have subsided. The control group consists of users who commented on the same post where the treatment group received their second AI reply, but never received AI replies themselves. *SecondAIReply* is a binary variable indicating whether a user belongs to the treatment group (i.e., received a second AI reply). Our outcome variable, *CommentedPosts*, measures the number of posts the user commented on in the week following their focal comment. To account for baseline engagement, we control for individual pre-treatment engagement using a seven-day history of commenting behavior before the focal comment.

Results in Table I.1 show that the coefficients for *SecondAIReply* are consistently positive and statistically significant across all time thresholds. These findings undermine the *Hype Effect* explanation and lend support to the social presence mechanism: even after the novelty has worn off, users remain responsive to AI replies. Users interpret the interaction as socially meaningful—consistent with the Social AI Agent serving as a credible extension of the influencer's presence.

Table I.1 Effect Persistence: Users Receiving a Second AI Reply

	CommentedPosts (AI Reply Gap ≥ 5 Days)			CommentedPosts (AI Reply Gap ≥ 10 Days)			CommentedPosts (AI Reply Gap ≥ 14 Days)		
	Full Sample	PSM	CEM	Full Sample	PSM	CEM	Full Sample	PSM	CEM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
SecondAIReply	0.7136***	0.5376***	0.6198***	0.5842**	0.8515**	0.4674*	0.7272**	1.215**	0.8431**
	(0.1559)	(0.1744)	(0.2162)	(0.2218)	(0.3714)	(0.2464)	(0.3516)	(0.5860)	(0.3359)
User Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Comment Date fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Post fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
R^2	0.68	0.78	0.79	0.69	0.80	0.80	0.65	0.78	0.83
Observations	45,788	5,639	4,757	27,941	3,395	2,665	15,821	2,196	1,554

Notes. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. User controls include daily commented posts in each of the seven days preceding the focal comment.

I.2. Effect Persistence After User Familiarization

Recall that our main analysis focuses on posts published on influencers' first day of AI feature activation. To test whether the effect persists after audiences become familiarized with the feature, we shifted our focus to posts published by the same set of focal influencers several months later (November–December 2024). By this time, audiences should have already been familiar with the existence of AI replies and were no longer

experiencing the feature as novel. We identify as the treatment group users who received an AI reply for the first time during this later period, and as the control group users who commented on the same posts as the treatment users but never received any AI reply, either during or prior to this window. In total, we identified 10,179 focal posts, 28,924 treated users, and 94,846 control users. We then replicate the main analysis using this new set of treatment and control users. The results, shown in Table I.2, continue to indicate a significant positive effect of receiving an AI reply on subsequent engagement, suggesting that the observed effect persists long after influencers' initial adoption of the AI agent feature.

Table I.2 Effect Persistence: Using Later Posts from Focal Influencers

	CommentedPosts		
	Full Sample	PSM	CEM
	(1)	(2)	(3)
AIReply	0.0283***	0.0196***	0.0116***
	(0.0012)	(0.0012)	(0.0023)
User fixed effects	✓	✓	✓
Day fixed effects	✓	✓	✓
Post fixed effects	✓	✓	✓
R^2	0.47	0.41	0.21
Observations	7,549,970	3,190,056	367,952

Notes. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses are clustered at the user level.

Appendix J: Measurement of Comment Sentiment

Unlike conventional sentiment classifiers that assess comments in isolation, we leverage GPT-4o-mini to jointly consider the comment and the post content, yielding a more accurate measure of users' sentiment toward the influencer. For example, if an influencer criticizes a product and a user comments, "I also think this product is bad," standard sentiment models may misclassify this agreement as negative, whereas our contextual approach correctly identifies it as supportive.

We prompt the GPT model to evaluate the contextual sentiment of user comments by feeding each comment and its associated post together into the model. We instruct the model to classify the sentiment of each comment as either negative or non-negative, given the post content as the context. To improve accuracy, we adopt a few-shot prompting approach by providing illustrative examples that clarify the difference between contextual sentiment evaluation and traditional sentiment measurement. For instance, when a post criticizes a product and a user comment also expresses criticism, the comment should be classified as non-negative, as it reflects agreement rather than negativity toward the influencer. To validate the classification result, we randomly sampled 500 comments and manually labeled their sentiment as either negative or non-negative. As shown in Table J.1, the GPT-based classifier achieves an overall accuracy of 0.86, demonstrating strong alignment with human judgment.

Table J.1 Performances of GPT-based Sentiment Classifier

	Precision	**Recall**	**F_1 score**	**Support**
Performance by Category				
Non-negative	1.00	0.81	0.90	379
Negative	0.63	0.99	0.77	121
Overall Performance				
Macro Avg	0.81	0.90	**0.83**	500
Weighted Avg	0.91	0.86	**0.86**	500
Accuracy			**0.86**	500

Appendix K: Matching for Non-adopting Influencers

For each adopting influencer in our sample, we identified comparable non-adopting influencers using Weibo's category-specific influencer rankings. Influencers who were in the same content category in the rankings and did not adopt the feature serve as the control pool. We then performed one-to-one matching using CEM on a rich set of account characteristics, including follower count, following count, total posts, total engagement, account tenure, real-name authentication, domestic location, location in developed regions, commercial status, tech-domain indicator, and gender. The balance diagnostics (in Table K.1) show that all standardized mean differences are below 0.1 after matching, indicating good comparability between the two groups. The final matched sample consists of 211 adopting and 211 non-adopting influencers. For these influencers, we collected all posts published between July 1 and August 31, 2024, and constructed an influencer-day panel.

Table K.1 Standardized Mean Differences between Adopting and Non-adopting Influencers

	Original	**Matched**
LogFollowers	-0.0545	-0.0530
LogFollowings	-0.0106	-0.0868
LogPosts	-0.0356	0.0870
LogEngagement	-0.0801	0.0440
AccountTenure	-0.2341	-0.0519
RealNameAuth	0.0850	0.0000
Location: Abroad	0.0221	0.0000
Location: Developed Area	-0.0009	0.0000
Commercialized	-0.0289	0.0000
Tech Domain	0.2456	0.0000
Male	0.0263	0.0000