 http://dx.doi.org/10.17287/kmr.2025.54.6.1939

Productivity Spillovers from Generative AI: How ChatGPT Adoption Reshapes Digital Engagement

김하은(주저자) · 박성호(교신저자) · 홍주원(공저자) · 이다혜(공저자)
Haeun Kim(First Author) · Sungho Park(Corresponding Author) · Juwon Hong(Co-Author) · Dahye Lee(Co-Author)

서울대학교 Seoul National University(*haeun047@snu.ac.kr*)
서울대학교 Seoul National University(*spark104@snu.ac.kr*)
서울대학교 Seoul National University(*juwonhong@snu.ac.kr*)
서울대학교 Seoul National University(*leedahye@snu.ac.kr*)

This study demonstrates how generative AI, specifically ChatGPT, influences user engagement across digital platforms beyond direct substitution or functional overlap. Using detailed user-level data from Nielsen Korea's KoreanClick+ panel and a Difference-in-Differences design with Propensity Score Matching, we show that ChatGPT adoption is associated with increased total device usage, with stronger effects observed on PC devices and among users with lighter baseline engagement. Furthermore, users primarily reallocate engagement toward platforms they already frequent, suggesting that generative AI reinforces existing usage patterns rather than leading to the expansion into new platforms. We propose that these indirect behavioral spillovers may be driven, in part, by productivity gains that reduce cognitive effort, a finding consistent with Cognitive Load Theory. These findings highlight the broader impact of generative AI on digital engagement and provide a potential theoretical lens for understanding the underlying mechanisms.

Keyword: Generative AI, ChatGPT, Digital Engagement, Cognitive Load, Spillover

Ⅰ. Introduction

Since the launch of ChatGPT, generative AI has raised new questions about how emerging technologies reshape user behavior across digital platforms. Its wide-ranging capabilities—such as answering technical questions, drafting text, and providing tailored information—pose challenges to incumbent services. For example, expert participation on Stack Overflow has declined following increased reliance on AI-generated programming help (Borwankar et al., 2024; Burtch et al., 2023; Del Rio-Chanona et al., 2023). Educational platforms report reduced engagement as students sub-

Submission Date: 07. 11. 2025 Revised Date: (1st: 10. 29. 2025) Accepted Date: 11. 12. 2025

stitute AI tools for traditional learning materials. Even search engines, long central to digital navigation, are being reexamined as users increasingly consult AI models during their information-seeking processes. In response, researchers have turned their attention to how generative AI impacts related services, exploring whether it replaces or complements them.

Yet this framing raises a broader question: are the behavioral effects of generative AI limited to the platforms it directly resembles? A parallel line of research has documented that tools like ChatGPT enhance productivity across a wide range of tasks, helping users write more efficiently, solve problems faster, and generate creative content at greater scale. These productivity effects are typically treated as separate from platform-level engagement. However, we propose that they may operate as a mechanism driving indirect spillovers: by saving time and reducing cognitive effort, generative AI may increase users' capacity to engage with other digital applications, even those unrelated to its core functions.

We develop this idea using Cognitive Load Theory (Sweller, 1988), which posits that offloading routine or mentally taxing subtasks frees up cognitive resources for other activities. Applied to digital behavior, this suggests that generative AI adoption may allow users to reallocate their time and mental bandwidth toward broader engagement across the digital ecosystem. To examine this possibility, we analyze user-level data from Nielsen Korea's KoreanClick+ panel, which provides comprehensive measures of weekly application usage across both PC and mobile devices. Using a Difference-in-Differences design augmented with Propensity Score Matching, we estimate the causal effect of ChatGPT adoption on overall and app-level digital engagement.

The results provide consistent support for our theoretical framework. First, we find that ChatGPT adoption increases total usage across other applications. Second, this effect is stronger on PCs than on mobile devices, reflecting differences in cognitive demands. Third, light users—who likely have more available cognitive bandwidth—show larger relative increases in usage than heavy users. Fourth, increased engagement is concentrated in applications that were already being actively used, rather than dispersed across the full app portfolio.

By highlighting these indirect, productivity-driven spillover effects, this study extends current understanding of generative AI's behavioral impact. Rather than restricting its influence to direct substitution or functional overlap, we provide preliminary empirical evidence that generative AI may function as a capacity-enhancing technology that reshapes users' broader engagement with digital services.

Ⅱ. Literature Review

2.1 Generative AI and Platform Competition

One stream of research examines how generative AI tools reshape user engagement with incumbent digital platforms. This literature primarily conceptualizes large language models (LLMs) as *information tools* that either substitute for or complement existing services with overlapping core functions, such as search engines and knowledge-sharing communities. For instance, Borwankar et al. (2024) analyzed Stack Overflow's policy change to prohibit ChatGPT-generated content and found a reduction in programming-related discussions. Similarly, Burtch et al. (2023) observed a marked decrease in web traffic and question posting volumes on Stack Overflow following ChatGPT's release. Del Rio-Chanona et al. (2023) reported a 25% drop in activity on Stack Overflow, indicating a substitution effect where users shifted from platform-based knowledge sharing to AI-assisted interactions. In contrast, Hong et al. (2025) show that ChatGPT can act as a complement for commercial search engines: by helping users translate vague information needs into precise queries, it increases search volume and encourages deeper exploration within the purchase funnel. Collectively, these studies capture the *competitive and complementary dynamics* that arise when generative AI enters existing digital ecosystems. They illustrate how generative AI reshapes engagement on platforms that overlap in functionality by disrupting traditional information search processes.

2.2 Generative AI and Cognitive Productivity

The second stream focuses on generative AI as a *productivity technology*, documenting how LLM-based tools enhance productivity across diverse domains including creative, analytical, and writing tasks. Noy and Zhang (2023) find that knowledge workers using ChatGPT write 40% faster and at higher quality, and Zhou and Lee (2024) report that artists using text-to-image models produce 25% more creative output. Similarly, Wu et al. (2025) demonstrate that AI collaboration improves performance in creative problem-solving tasks. These productivity gains arise partly because generative AI allows users to offload cognitively taxing subtasks to the model. Direct evidence of this mechanism comes from Kosmyna et al. (2025), who use neurocognitive measurements and find reduced neural markers of cognitive effort during AI-assisted writing, indicating genuine mental offloading.

Studies report empirical evidence that the freed-up cognitive capacity can then be reinvested into higher-value or more complex activities: Brynjolfsson et al. (2025) find that customer service agents using AI resolve issues faster and redirect their attention to more sophisticated tasks, while Hoffmann et al. (2024) show that developers employing AI-assisted coding tools shift focus from routine programming to creative problem-solving. Extending this logic to consumer contexts, Zheng et al. (2023) show that a generative-AI-powered query recommender system reduces cognitive load, enabling users to refine their searches and expand their purchase consideration sets.

Cognitive Load Theory (Sweller, 1988) provides a theoretical framework for understanding these effects. CLT posits that working memory has limited capacity, and performance is constrained by cognitive load (Hyun and Choi, 1994). To reduce such load, people often rely on cognitive offloading, which is utilizing physical actions or tools to alleviate cognitive demands involved in a task and free up cognitive resources. These resources can then be reallocated to other activities (Risko and Gilbert, 2016; Perkins, 2023). When cognitive load is reduced, more working memory resources are available for higher-order processing, enabling deeper understanding and more complex task performance (Sweller et al, 1998). Generative AI functions as a form of cognitive offloading, automating mentally demanding subtasks and freeing cognitive resources for tasks requiring higher-order processing.

2.3 Research Gap and Contribution

Two streams of research frame our study. The first examines how generative AI tools reshape user engagement with incumbent digital platforms. These studies capture the *competitive and complementary dynamics* that arise among incumbent digital platforms because of generative AI. The second stream focuses on generative AI as a *productivity technology*. These studies primarily center on *within-task effects* of generative AI—how AI adoption changes output quality, speed, or learning in the same domain of work.

Our study bridges these two streams by examining how productivity gains from generative AI reverberate beyond the focal task or platform. Rather than focusing on substitution or complementarity among services with overlapping functions, we investigate how the efficiency benefits of AI adoption may spill over into users' broader digital behavior. Specifically, we explore whether the time and effort saved through AI-assisted tasks lead users to engage differently with other applications that are not directly related to the AI tool itself. In doing so, we highlight a new dimension of generative AI's impact—its potential to reshape overall patterns of digital engagement—thus extending both the substitution–complementarity literature and the productivity literature toward a more holistic understanding of AI's behavioral consequences.

This perspective leads to our central research question: *Does ChatGPT adoption reshape users' engagement with other digital applications by expanding their capacity through productivity spillovers?*

III. Hypotheses Development

Generative AI tools like ChatGPT have been shown to substantially enhance individual productivity across a variety of knowledge and creative tasks. Noy and Zhang (2023) demonstrate that knowledge workers using ChatGPT produce written outputs 40% faster and of higher quality. Zhou and Lee (2024) show that artists leveraging text-to-image models generate more creative content within the same time frame, and Brynjolfsson et al. (2025) find that AI-assisted customer service agents resolve issues more efficiently while maintaining or improving quality. These findings converge on the idea that generative AI reduces the cognitive load associated with routine or low-level subtasks, freeing mental bandwidth and time. Importantly, Kosmyna et al. (2025) provide direct neurophysiological evidence for this mechanism: their study shows that AI-assisted writing lowers neural markers of cognitive effort, indicating genuine mental offloading during task performance. Drawing on theories of cognitive resource allocation (Sweller, 1988; Paas et al, 2003), such relief can produce behavioral spillovers —when cognitive friction in one domain decreases, users may reallocate the saved resources toward other meaningful digital activities. Therefore:

H1: ChatGPT adoption increases overall usage of other digital applications.

While the overall spillover mechanism may apply broadly, the extent of its impact likely varies across device types due to their differing cognitive demands and usage contexts. Mobile and PC platforms differ fundamentally in their typical use contexts and cognitive demands (Chung et al., 2015). Research shows that mobile interactions tend to be brief, goal-specific, and constrained by limited screen size and situational context (Karlson et al., 2009; Ghose et al., 2013; Lee and You, 2023). PC use, in contrast, supports longer sessions, multitasking, and complex workflows that require higher levels of sustained attention and information processing (Mark et al., 2015). According to Cognitive Load Theory, the benefit of tools that reduce mental effort should be greater in more cognitively demanding settings. Since ChatGPT assists with synthesis, drafting, and problem-solving—activities that are more prevalent and complex in PC workflows—we expect stronger productivity spillovers in PC environments. Therefore:

H2: The positive effect of ChatGPT adoption on digital engagement is stronger on PCs than on mobile devices.

The magnitude of productivity-driven spillovers is also likely to depend on users' baseline engagement levels. According to resource allocation and technology adoption theories (Venkatesh et al., 2012; Davis et al., 1989), the benefits of productivity-enhancing tools are often most pronounced for users with greater unused capacity—those who were previously constrained by time, cognitive resources, or motivation. In contrast, heavy users who already spend significant time on digital platforms may experience diminishing marginal returns from further efficiency gains. This asymmetry is also consistent with the "law of limited attention" (Simon, 1996), suggesting that once cognitive and temporal resources are fully utilized, additional savings yield less behavioral change. Therefore:

H3: ChatGPT adoption leads to larger relative increases in engagement among light users than among heavy users.

Finally, even when users free up time and mental bandwidth, they are unlikely to distribute these gains uniformly across all activities. Behavioral research shows that individuals tend to reinvest freed-up resources into familiar or intrinsically rewarding activities, reinforcing rather than reshaping established habits (Oulasvirta et al., 2012; LaRose, 2010). In digital contexts, this selective reinforcement implies that spillover engagement will concentrate on platforms users already value most. Cognitive and motivational theories likewise suggest that attention allocation tends to follow preexisting preference hierarchies (Kahneman, 1973). Accordingly, we expect productivity spillovers to amplify existing engagement patterns rather than redirect users to entirely new domains. Accordingly:

H4: The positive impact of ChatGPT adoption on usage is stronger for applications that users previously engaged with more actively.

Ⅳ. Empirical Analysis

4.1 Data

Our analysis draws on user-level behavioral data from Nielsen Korea's KoreanClick+ panel, a demographically balanced sample of South Korean internet users. The dataset spans from August 2022 to September 2023 and provides weekly records of digital activity across both PC and mobile platforms. Panel members are selected through stratified random sampling and include detailed demographic data such as gender, age, job, income, region, education, occupation, and marital status.

Digital behavior is captured through log-tracking software installed on participants' devices, which records all websites and applications accessed. These logs are categorized by Nielsen Korea into 15 high-level content groups (e.g., Social Media, Business, Finance), and each record contains the user ID, content type, week of activity, and duration of use. From this dataset, we construct two main outcome variables: total weekly time (in minutes) spent on PCs and on mobile devices.

We begin with a sample of 9,028 individuals. To ensure comparability between users with and without exposure to ChatGPT, we restrict the sample to individuals who had used either a PC or mobile device during the pre-treatment period. This yields a final analytic sample of 8,736 users. Table 1 presents the summary statistics of the main variables.

4.2 Propensity Score Matching

To estimate the causal impact of ChatGPT adoption, we employ a two-step identification strategy: Propensity Score Matching (PSM) followed by a Difference-in-Differences (DiD) analysis. Our initial sample includes 8,736 panelists, among whom 802 individuals engaged with ChatGPT at least once during the observation period. We classify these individuals as the treatment group ("adopters"), while the remaining 7,934 users who never interacted with ChatGPT serve as the control group ("non-adopters").

To improve comparability between these two groups, we first estimate the probability of ChatGPT adoption based on pre-treatment characteristics. The first set of covariates reflects users' baseline digital behavior, including total time spent on PCs and mobile devices, as well as time allocated to specific content categories. Nielsen Korea classifies all applications and websites into 15 broad functional groups—such as "Entertainment" (e.g., streaming services, ticketing platforms), "Education" (e.g., online learning tools), "Organizations" (e.g., government and nonprofit sites), "Portal", "Social Media" (e.g., messaging apps, online communities), "Utilities" (e.g., calendar, weather, file storage), and "Other" (miscellaneous services). We measure time spent in each category separately by device type.

The second set of covariates includes users' demographic information such as gender, age, income, region, education, and marital status. Based on these variables, we conduct one-to-one nearest-neighbor matching without replacement, using a caliper width of 0.0001 to ensure close matches. This yields a final matched sample of 574 adopters and 574 corresponding non-adopters. As shown in Table 2, 〈Table 1〉 Summary Statistics

Variable	Mean	S.D.	Minimum	Maximum
Demographic attributes				
Age	years 7 - 18 (5.04%), years 19 - 29 (11.03%), years 30 - 39 (14.23%), years 40 - 49 (26.03%), years 50 - 59 (22.25%), years 60 - 69 (16.85%), years 70 - 79 (4.57%)			
Gender	Male (50.84%), Female (49.16%)			
Income	USD 725 or less (2.63%), between USD 725 and USD 2,175 (17.36%), between USD 2,175 and USD 3,625 (38.54%), USD 3,625 or more (41.46%)			
Region	Seoul/Incheon/Gyeonggi (45.39%), Gyeongsang (26.83%), Jeolla or Jeju (12.76%), Chungcheong or Gangwon (15.02%)			
Level of Education	University graduates (62.67%), College and graduate students (6.49%), High school graduates (25.37%), Students attending high school or lower-level schools (5.47%)			
Marital status	Married (68.05%), Single (31.95%)			
Weekly device usage (minutes)				
Total usage	2,105.28	1,661.58	0	19,332.48
Mobile usage	2,011.42	1,639.56	0	19,332.48
PC usage	93.86	237.15	0	6,678.12
Weekly usage by Service Category (minutes)				
Business	1.94	18.34	0	1,864.38
Communication	114.12	244.33	0	5,354.17
Education	11.36	52.15	0	4,100.77
Entertainment	610.74	855.53	0	12,471.42
Finance	84.73	211.70	0	7,359.27
Game	150.39	417.78	0	9,106.43
Leisure	4.69	40.61	0	3,638.18
Lifestyle	193.72	352.42	0	7,918.08
News	4.45	30.54	0	2,206.43
Organization	4.99	25.78	0	1,682.40
Portals	39.62	118.98	0	4,290.00
Shop	72.54	140.15	0	4,061.98
Social media	189.76	234.90	0	4,107.75
Utility	620.55	572.74	0	7,552.37
Other	1.68	30.31	0	3,421.42

〈Table 2〉 Covariate Balance Before and After Matching

Variables	Unmatched			Matched		
	TG	CG	p 〉|t|	TG	CG	p 〉|t|
Time spent by users during pre-treatment periods						
Total usage	2,292.972	2,051.215	0.000	2,206.401	2,194.592	0.908
Average education usage	19.035	10.330	0.000	14.860	12.964	0.447
Average entertainment usage	812.389	740.145	0.051	806.299	802.928	0.956
Average organization usage	8.535	4.291	0.000	5.602	6.379	0.520
Average portals usage	99.029	35.204	0.000	55.705	46.389	0.160
Average social media usage	334.568	294.375	0.003	316.339	324.581	0.695
Average utility usage	657.320	624.012	0.123	647.821	679.070	0.412
Average others usage	2.225	1.785	0.706	1.722	1.437	0.626
Demographic variables						
Male	0.631	0.496	0.000	0.578	0.561	0.552
Age_7_to_18	0.064	0.049	0.072	0.068	0.068	1.000
Age 19 to 29	0.151	0.106	0.000	0.141	0.167	0.221
Age 30 to 39	0.200	0.137	0.000	0.174	0.192	0.446
Age 50 to 59	0.182	0.227	0.004	0.206	0.181	0.296
Age_60 to 69	0.099	0.176	0.000	0.118	0.115	0.854
Age 70 to 79	0.010	0.049	0.000	0.012	0.005	0.204
Income USD 725 or less	0.017	0.027	0.100	0.017	0.014	0.635
Income USD 725 - 2,175	0.176	0.173	0.865	0.171	0.169	0.937
Income USD 2,175 - 3,625	0.357	0.388	0.079	0.375	0.361	0.625
Region Gyeongsang	0.254	0.270	0.350	0.268	0.251	0.501
Region Jeolla or Jeju	0.105	0.130	0.042	0.111	0.110	0.925
Region Chungcheong or Gangwon	0.135	0.152	0.197	0.148	0.141	0.737
College and graduate students	0.095	0.062	0.000	0.080	0.110	0.087
Hight school graduates	0.138	0.265	0.000	0.160	0.164	0.873
Students up to high school	0.064	0.054	0.246	0.070	0.068	0.907
Married	0.631	0.496	0.000	0.610	0.632	0.429

the matching procedure achieves strong covariate balance, supporting the validity of our subsequent DiD analysis.

4.3 Difference in Differences

To estimate the causal impact of ChatGPT adoption on digital engagement, we apply a two-way fixed effects (TWFE) Difference-in-Differences (DiD) model to the matched sample. Our outcome variable, Y_{it}, represents the total weekly time (in seconds) that user i spends using PCs and mobile devices in week t. To reduce skewness and accommodate zero values, we apply a log(x + 1) transformation to all dependent variables.

Our key explanatory variable is $Treatment_{it}$, an indicator equal to 1 if user i has adopted ChatGPT by week t, and 0 otherwise. The model includes individual fixed effects θ_i to account for time-invariant user characteristics, and week fixed effects ϕ_t to control for temporal factors common to all users. To account for serial correlation, we cluster standard errors at the individual level.

β_1 captures the average treatment effect of ChatGPT adoption on digital usage. To test for heterogeneity in line with Hypothesis 3, we extend the model using a Difference-in-Difference-in-Differences (DDD) specification. In this extension, we interact the treatment indicator with a moderator variable: **Heavy Device User**, defined as users whose average weekly device usage (across both PC and mobile) exceeds the sample median. This allows us to assess whether the effects of ChatGPT adoption differ systematically between heavy and light users.

$$\log(Y_{it}+1) = \beta_0 + \beta_1 Treatment_{it} + \theta_i + \phi_t + \varepsilon_{it} \tag{1}$$

$$\begin{aligned}\log(Y_{it}+1) = \beta_0 &+ \beta_1 Treatment_{it} \\ &+ \beta_2 Treatment_{it} \times Heavy\ Device\ User_i \\ &+ \theta_i + \phi_t + \varepsilon_{it}\end{aligned} \tag{2}$$

Ⅴ. Results

Table 3 presents the average treatment effects of ChatGPT adoption on device usage time. Column 1 demonstrates that adoption

〈Table 3〉 Effect of ChatGPT Adoption on Device Usage Time

	(1)	(2)	(3)
Dependent Variable	Total Usage Time	Mobile Device Total Usage Time	PC Device Total Usage Time
Treatment Effect	0.257*** (0.085)	0.234** (0.101)	0.477*** (0.067)
User Fixed Effect	Yes	Yes	Yes
Week Fixed Effect	Yes	Yes	Yes
R-square	0.5604	0.6121	0.5606
Observations	60,844	60,844	60,844

standard errors clustered at the user level are used.
*p〈0.1; **p〈0.05; ***p〈0.01

of ChatGPT significantly increases overall usage of other digital applications, in support of hypothesis 1. Column 2 and 3 support Hypothesis 2—the impact of ChatGPT adoption is higher for application usage on PC devices. Specifically, average mobile usage rises by approximately 23.4%, while PC usage increases by 47.7%. Usage of both devices increases by 25.7%. All results are statistically significant at the 5% level.

Next, we analyze the heterogeneous treatment effect of ChatGPT adoption by users' baseline device usage (Table 4). ChatGPT adoption leads to larger relative increases in engagement among light users than among heavy users, thus supporting hypothesis 3.

〈Table 4〉 Heterogeneous Effect by Baseline Device Usage

Dependent Variable	Total Usage Time
Treatment Effect	0.535*** (0.132)
Treatment Effect × Heavy Device user	-0.572*** (0.134)
User Fixed Effect	Yes
Week Fixed Effect	Yes
R-square	0.5622
Observations	60,844

standard errors clustered at the user level are used.
*p〈0.1; **p〈0.05; ***p〈0.01

〈Table 5〉 Weekly Average Category Usage in the Pre-Adoption Period of ChatGPT

	Category	Average Weekly Usage Time (Unit: Minute)
High Usage Category	Utility	626.426
	Entertainment	587.654
	Social media	192.673
	Lifestyle	188.627
	Game	156.346
Middle Usage Category	Communication	113.879
	Finance	74.297
	Shop	65.392
	Portals	41.568
	Education	11.664
Low Usage Category	Organization	5.049
	News	4.737
	Leisure	4.404
	Business	2.171
	Other	1.837

〈Table 6〉 Effect of ChatGPT Adoption on Device Usage by Pre-Adoption Category Level

Dependent Variable	High Usage Category Time	Middle Usage Category Time	Low Usage Category Time
Treatment Effect	0.264*** (0.079)	0.208*** (0.065)	0.188*** (0.041)
User Fixed Effect	Yes	Yes	Yes
Week Fixed Effect	Yes	Yes	Yes
R-square	0.5861	0.6213	0.5580
Observations	60,844	60,844	60,844

standard errors clustered at the user level are used.
*p 〈 0.1; **p 〈 0.05; ***p 〈 0.01

Lastly, we examine the effect of ChatGPT adoption on device usage by pre-adoption category level. As shown in Table 5, we classify categories into high, middle, and low usage groups based on the average weekly usage time during the pre-adoption period of ChatGPT. Table 6 reports the average treatment effects of ChatGPT adoption on device usage time by each category. Results support Hypothesis 4 —the positive impact of ChatGPT adoption on usage is stronger for applications that users previously engaged with more actively. Specifically, high usage category time rises by approximately 26.4%, while middle and low usage category time increases by 20.8% and 18.8%, respectively.

Ⅵ. Robustness Checks

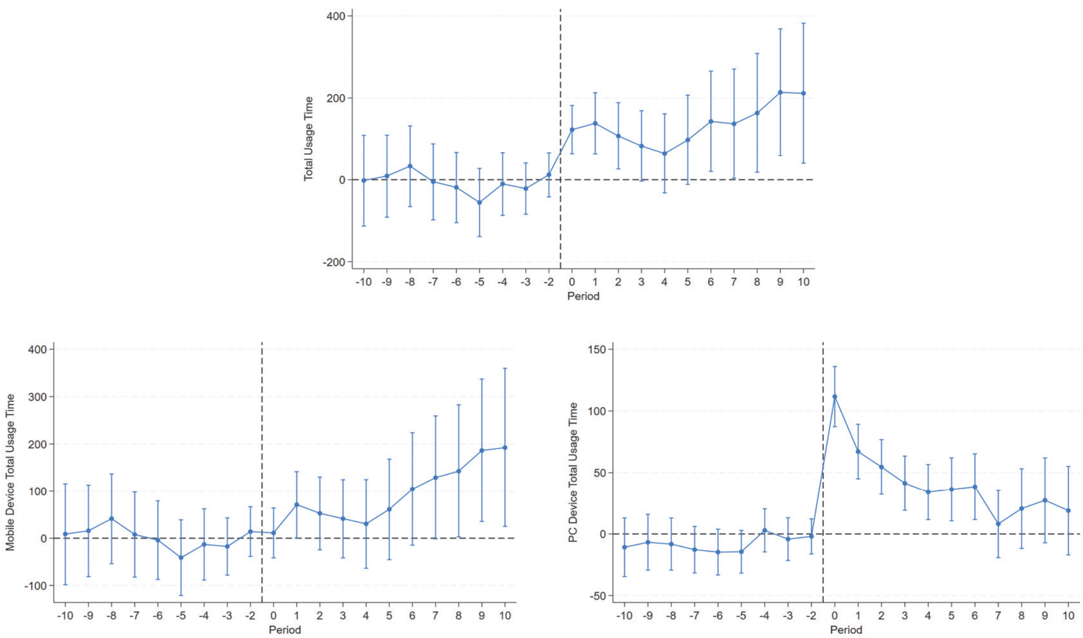
To ensure the credibility of our causal estimates, we conduct a comprehensive set of robustness checks addressing the identification strategy, alternative specifications, and potential threats to internal validity. These analyses demonstrate that the observed effects of ChatGPT adoption on commercial search behavior and downstream engagement are not driven by spurious trends, selection bias, or model artifacts.

6.1 Parallel trends assumption

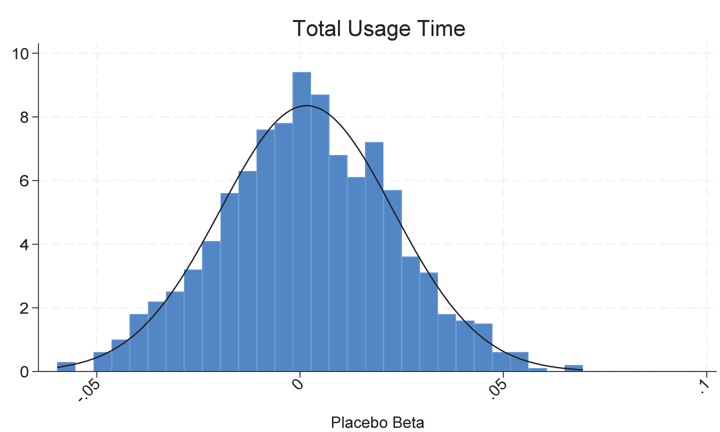
Our identification strategy relies on the parallel trends assumption of the DiD framework. An event-study specification (Angrist and Pischke, 2009), which replaces the treatment dummy with a series of time relative indicators, shows no evidence of pre-treatment divergence and a marked increase following adoption (Figure 1).

6.2 Placebo Simulation

We randomly assign a “placebo” treatment



〈Figure 1〉 Parallel Trend Assumption Check



〈Figure 2〉 Distribution of Placebo DiD Estimate for Total Usage Time

to periods before the actual treatment event and estimate the treatment effect for each of 1,000 draws. Figure 2 shows the kernel density of these 1,000 placebo ATTs, where the x-axis represents the beta estimate values, while the y-axis shows the corresponding probability densities. Distribution of placebo treatment effects tightly centered around zero, which indicates no systematic pre-treatment effect.

6.3. Alternative Matching Methods

We replace propensity score matching with 1) coarsened exact matching (CEM) and 2) inverse probability of treatment weighting (IPTW). First, the CEM approach provides an alternative method of balancing covariates between treated and control groups by grouping observations into coarsened strata before matching. Second, IPTW method retains the entire sample for estimation (Zhang et al., 2021). We compute the weight of user i as follows:

$$w_i = \frac{T_i}{PS_i} + \frac{1 - T_i}{1 - PS_i}$$

where T_i is the binary indicator for ChatGPT adopters. PS_i is propensity score calculated from the logit regression. The results presented in Table 7 are consistent with the findings from the main analyses.

〈Table 7〉 Robustness Checks of Table 4: Alternative Matching Methods

	Method A: CEM		
Dependent Variable	Total Usage Time	Mobile Device Total Usage Time	PC Device Total Usage Time
Treatment Effect	0.284*** (0.085)	0.307*** (0.106)	0.407*** (0.070)
User Fixed Effect	Yes	Yes	Yes
Week Fixed Effect	Yes	Yes	Yes
R-square	0.5627	0.5974	0.4825
Observations	378,473	378,473	378,473
	Method B: IPTW		
Dependent Variable	Total Usage Time	Mobile Device Total Usage Time	PC Device Total Usage Time
Treatment Effect	0.245*** (0.066)	0.203** (0.084)	0.487*** (0.059)
User Fixed Effect	Yes	Yes	Yes
Week Fixed Effect	Yes	Yes	Yes
R-square	0.5530	0.6046	0.5393
Observations	463,008	463,008	463,008

standard errors clustered at the user level are used.
*p 〈 0.1; **p 〈 0.05; ***p 〈 0.01

6.4 Alternative Control Group

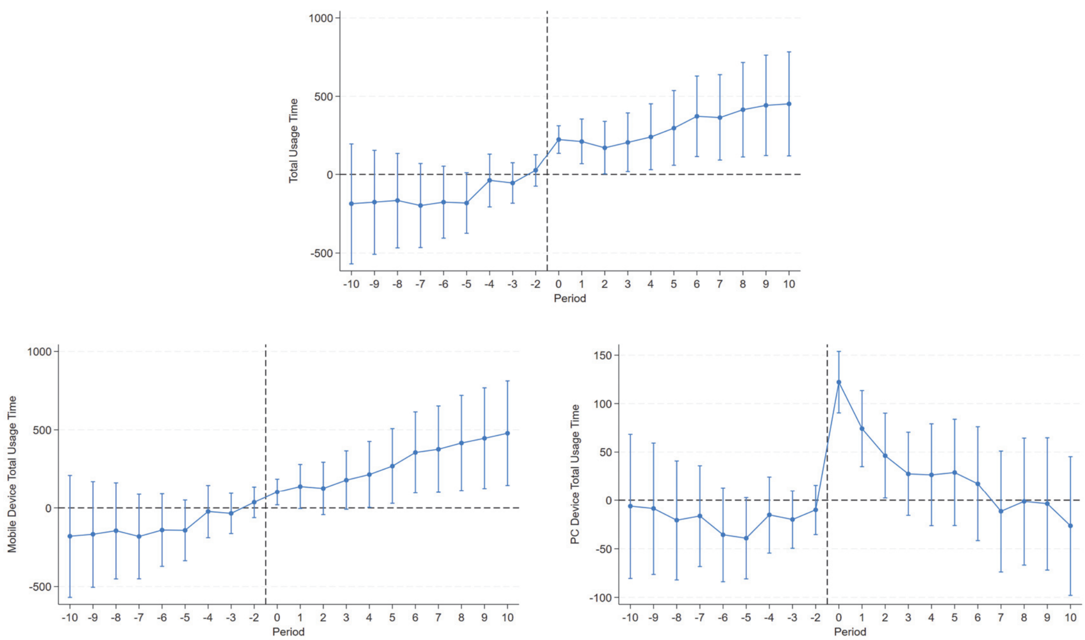
We test the robustness of our findings using two alternative control group specifications. Specifically, we first compare the early adopters to the late adopters of ChatGPT, leveraging a natural staggered rollout to account for unobserved time-invariant differences between adopters and never-adopters. Early adopters are defined as those who adopted ChatGPT prior to mid-May 2023 (i.e., before the median adoption week), while the remaining adopters constitute the control group. We conduct one-to-one nearest-neighbor propensity score matching without replacement, using a caliper of 0.01. Pre-adoption PC usage time and gender are used as matching covariates. This results in a matched sample of 239 treated users and 239 corresponding control users. The resulting DiD estimates (Panel A of Table 8) are consistent in sign, magnitude, and statistical significance with our main results. We validate the parallel trends of the matched sample using pre-treatment event studies (Figure 3).

We also compare the ChatGPT adopters to those who used generative AI apps (e.g., image generators, conversational chatbots)

〈Table 8〉 Robustness Checks of Table 4: Alternative Control Groups

	Panel A: Later ChatGPT Adopters		
Dependent Variable	Total Usage Time	Mobile Device Total Usage Time	PC Device Total Usage Time
Treatment Effect	0.337*** (0.104)	0.215*** (0.126)	1.035*** (0.135)
User Fixed Effect	Yes	Yes	Yes
Week Fixed Effect	Yes	Yes	Yes
R-square	0.6827	0.7750	0.5139
Observations	9,250	9,250	9,250
	Panel B: Non-ChatGPT Generative AI App Users		
Dependent Variable	Total Usage Time	Mobile Device Total Usage Time	PC Device Total Usage Time
Treatment Effect	0.263*** (0.082)	0.222** (0.100)	0.456*** (0.064)
User Fixed Effect	Yes	Yes	Yes
Week Fixed Effect	Yes	Yes	Yes
R-square	0.5025	0.5668	0.5298
Observations	66,568	66,568	66,568

standard errors clustered at the user level are used.
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

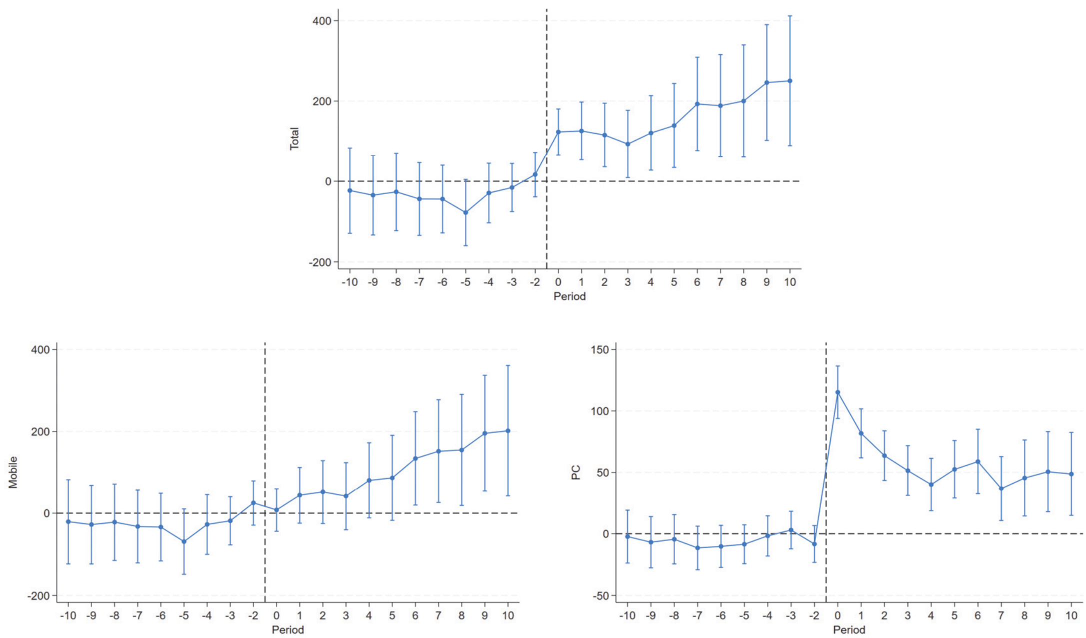


〈Figure 3〉 Parallel Trend Assumption Check with Later ChatGPT Adopters as the Alternative Group

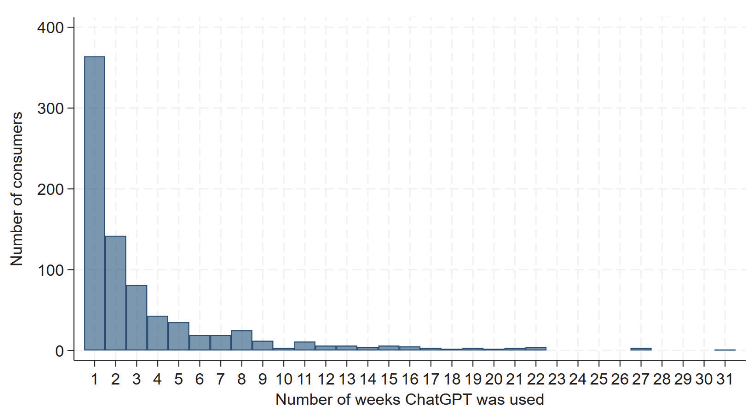
other than ChatGPT. We conduct one-to-one nearest-neighbor propensity score matching without replacement, using a caliper of 0.0001. We use pre-adoption total usage time, PC usage time, gender, age, and education as matching covariates. This results in a matched sample of 628 treated users and 628 corresponding control users. The resulting DiD estimates (Panel B of Table 8) are not statistically different from the main analysis. We validate the parallel trends of the matched sample using pre-treatment event studies (Figure 4).

6.5 Alternative Treatment Group

We examine the robustness of our results by using an alternative treatment group with active users of ChatGPT. As shown in figure 5, 364 out of 802 adopters used the app only once during the sample period. Accordingly, we define active users as those who used ChatGPT at least twice. We apply one-to-one nearest-neighbor propensity score matching without replacement, using a caliper of 0.0001. We include pre-adoption PC and mobile usage time as matching covariates. Our final sample includes 342 users in the treatment group and 342 in the control group after



〈Figure 4〉 Parallel Trend Assumption Check with Non-ChatGPT Generative App Users as the Alternative Group



〈Figure 5〉 Distribution of Consumers by Number of Weeks Using ChatGPT

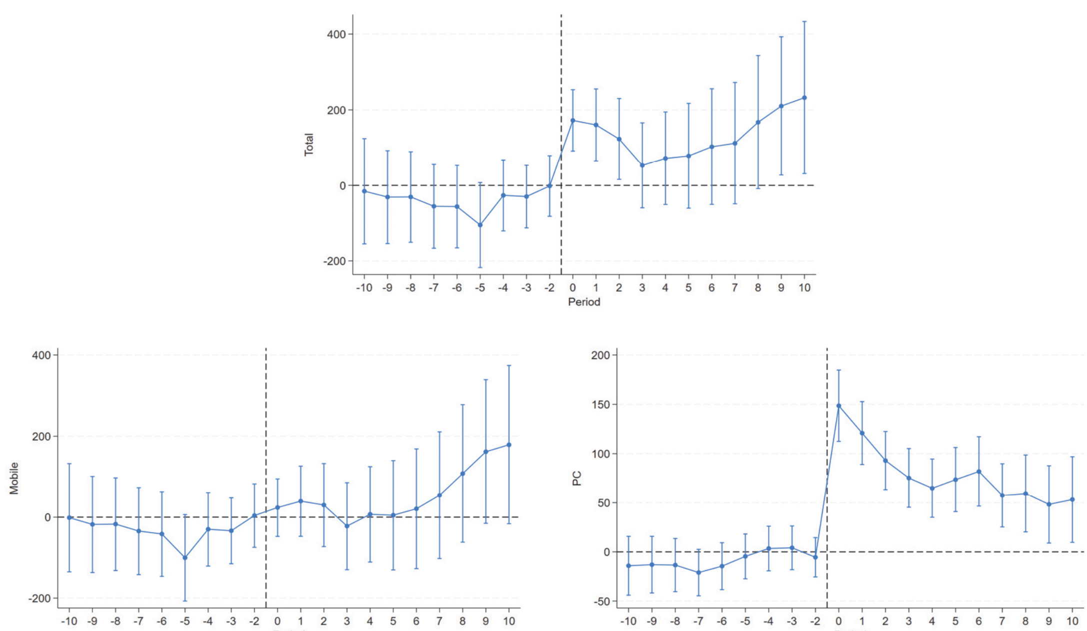
matching. We show that our results are robust (Table 9). We provide evidence for the parallel trends assumption using pre-treatment event studies (Figure 6).

We also analyze a continuous variable measuring ChatGPT usage time, capturing varia-

〈Table 9〉 Robustness Checks of Table 4: Alternative Treatment Group

Dependent Variable	Total Usage Time	Mobile Device Total Usage Time	PC Device Total Usage Time
Treatment Effect	0.308*** (0.096)	0.243** (0.120)	0.718*** (0.090)
User Fixed Effect	Yes	Yes	Yes
Week Fixed Effect	Yes	Yes	Yes
R-square	0.5470	0.6135	0.5336
Observations	36,252	36,252	36,252

standard errors clustered at the user level are used.
*p〈0.1; **p〈0.05; ***p〈0.01



〈Figure 6〉 Parallel Trend Assumption Check of Alternative Treatment Group

tion in engagement intensity. Specifically, we construct each user's ratio of weekly ChatGPT usage time to their maximum weekly ChatGPT usage time. The results presented in Table 10 are consistent with the findings from the main analyses. Users engage more with digital platforms in weeks when they use ChatGPT more intensely. This pattern further supports the interpretation that generative AI use complements, rather than crowds out, traditional digital activities.

〈Table 10〉 Robustness Checks of Table 4: Treatment Intensity

Dependent Variable	Total Usage Time	Mobile Device Total Usage Time	PC Device Total Usage Time
Treatment Effect	0.089*** (0.002)	0.006*** (0.001)	0.084*** (0.002)
User Fixed Effect	Yes	Yes	Yes
Week Fixed Effect	Yes	Yes	Yes
R-square	0.1205	0.0312	0.1176
Observations	463,008	463,008	463,008

standard errors clustered at the user level are used.
*p 〈 0.1; **p 〈 0.05; ***p 〈 0.01

6.6 Alternative Outcome Variable

To further assess the robustness of our results, we use an alternative outcome variable that excludes ChatGPT usage. We conduct one-to-one nearest-neighbor propensity score matching without replacement, using a caliper of 0.01. Pre-adoption PC and mobile usage time are used as matching covariates. This results in a matched sample of 607 treated users and 607 corresponding control users. The resulting DiD estimates (Table 11) are consistent in sign, magnitude, and statistical significance with our main results. We validate the parallel trends of the matched sample using pre-treatment event studies (Figure 7).

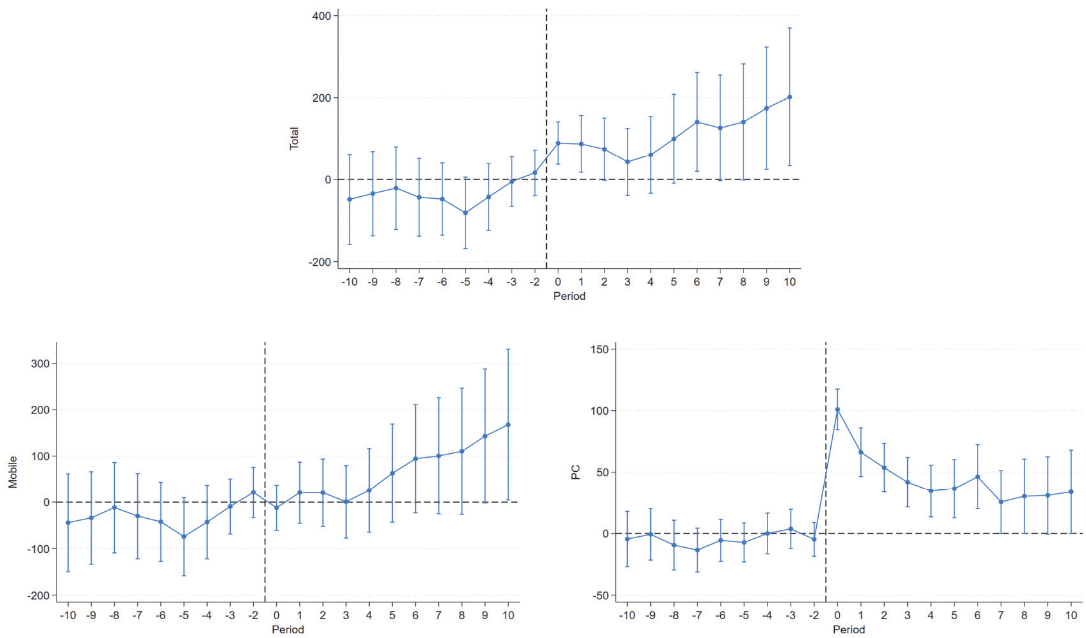
6.7 Alternative Estimator

We use an alternative estimator developed by Callaway and Sant'Anna (2021), to address concerns that DiD estimates may suffer

〈Table 11〉 Robustness Checks of Table 4: Alternative Treatment Group

Dependent Variable	Total Usage Time Excluding ChatGPT	Mobile Device Total Usage Time Excluding ChatGPT	PC Device Total Usage Time Excluding ChatGPT
Treatment Effect	0.241*** (0.080)	0.209** (0.097)	0.391*** (0.066)
User Fixed Effect	Yes	Yes	Yes
Week Fixed Effect	Yes	Yes	Yes
R-square	0.5440	0.6047	0.5395
Observations	64,342	64,342	64,342

standard errors clustered at the user level are used.
*p 〈 0.1; **p 〈 0.05; ***p 〈 0.01



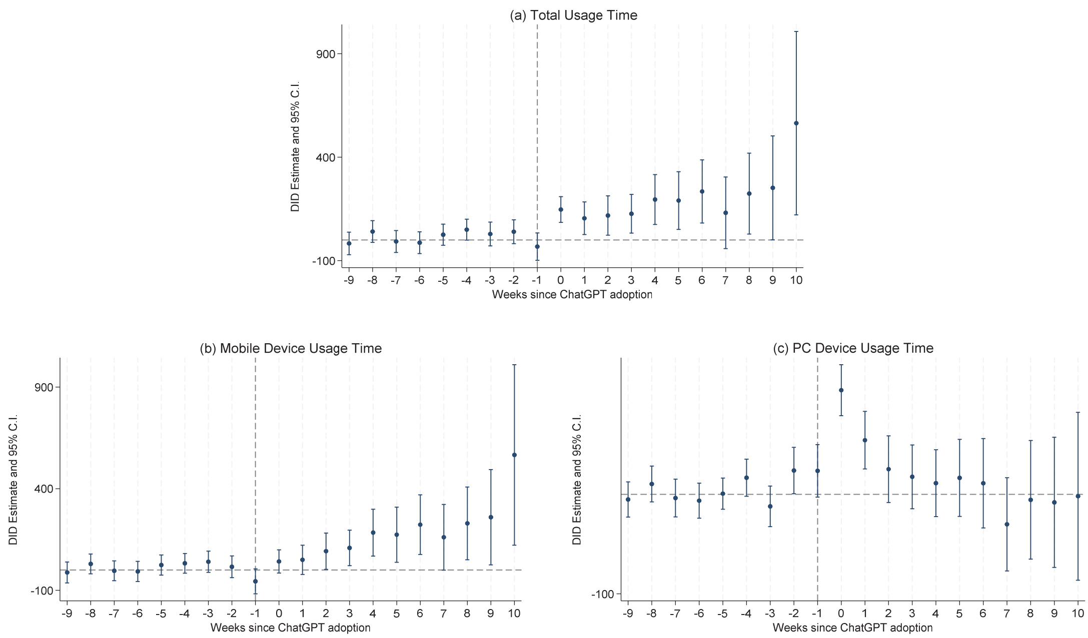
〈Figure 7〉 Parallel Trend Assumption Check of Alternative Treatment Group

from Goodman‐Bacon weighting bias. As shown in Figure 8, Group-specific ATTs using Callaway and Sant'Anna (2021) event-study methods show consistent results.

6.8 Instrumental Variable Approach

To address endogeneity concerns, we employ an instrumental variables (IV) estimation strategy with a time-varying instrument. Specifically, we construct an indicator for the adoption of any conversational generative AI application other than ChatGPT. This instrument is expected to be correlated with ChatGPT adoption, as users who adopt other generative AI tools are generally more receptive to experimenting with new AI technologies. At the same time, it plausibly satisfies the exclusion restriction: during the study period, the alternative conversational AI applications observed in our data—such as SimSimi and Replika—functioned primarily as virtual companions rather than productivity or information tools comparable to ChatGPT. We restrict our analysis to mobile, as comparable application data is unavailable for PC usage.

In the first stage, we regress ChatGPT adoption on the adoption of other conversational AI applications. In the second stage, we estimate the causal effect of instrumented



〈Figure 8〉 Parallel Trend Assumption Check of Alternative Estimator

〈Table 12〉 First-Stage IV Estimate of ChatGPT Adoption

Dependent Variable	ChatGPT Adoption
Adoption of Conversational Generative AI Apps Other Than ChatGPT	0.076*** (0.019)
User Fixed Effect	Yes
Week Fixed Effect	Yes
R-square	0.4726
Observations	463,008

standard errors clustered at the user level are used.
*p〈0.1; **p〈0.05; ***p〈0.01

ChatGPT adoption on users' total device usage time. The results, presented in Table 12 and 13, indicate a strong and positive association between ChatGPT adoption and device usage. This evidence suggests that users engage more actively with digital platforms in weeks when they use ChatGPT more intensively, reinforcing the interpretation that generative AI adoption complements, rather than substitutes for, traditional digital activities.

〈Table 13〉 Second-Stage IV Estimates of ChatGPT Adoption Effects

Dependent Variable	Total Usage Time	Mobile Device Total Usage Time	PC Device Total Usage Time
Treatment Effect	5.565*** (1.532)	5.315*** (1.711)	2.099* (1.084)
User Fixed Effect	Yes	Yes	Yes
Week Fixed Effect	Yes	Yes	Yes
R-square	0.5593	0.5896	0.5791
Observations	463,008	463,008	463,008

standard errors clustered at the user level are used.
*p〈 0.1; **p〈 0.05; ***p〈 0.01

Ⅶ. Discussion

This study examines how generative AI—specifically ChatGPT—may influence digital engagement beyond direct substitution or functional overlap. Recent empirical work shows that generative AI substantially enhances individual productivity across a variety of knowledge and creative tasks: knowledge workers write faster and at higher quality (Noy and Zhang, 2023), artists produce more creative outputs with text-to-image models (Zhou and Lee, 2024), and customer service agents resolve issues more efficiently while maintaining quality (Brynjolfsson et al, 2025). Importantly, Kosmyna et al. (2025) provide direct neurophysiological evidence underlying this productivity increase, showing that AI-assisted writing lowers neural markers of cognitive effort, indicating genuine mental offloading. Drawing on theories of cognitive resource allocation (Sweller, 1988; Paas et al, 2003), such relief can produce behavioral spillovers—when cognitive friction in one domain decreases, users may reallocate the freed resources toward other meaningful digital activities. Building on this, we hypothesize that productivity improvements associated with generative AI can reduce users' cognitive effort and time expenditure, potentially enabling expanded engagement across a broad range of digital platforms.

Our contribution lies in examining generative AI's broader impact on digital behavior. Prior research has largely focused on how generative AI reshapes engagement with functionally overlapping platforms, such as search engines or Q&A communities, highlighting substitution or complementarity dynamics among incumbent digital services. In contrast, our study investigates how the efficiency benefits of AI adoption may lead users to engage differently with other applications that are not directly related to the AI tool itself. In doing so, we highlight a new dimension of generative AI's impact—its potential to reshape overall patterns of digital engagement—thus extending both the substitution–complementarity literature toward a more holistic understanding of AI's behavioral consequences.

Our empirical findings are consistent with this framework. ChatGPT adoption is associated with increases in total device usage, indicating that it may serve not only as a task-specific aid but also as a capacity-enhancing technology. The effect is more pronounced on PCs, where user activities often involve higher cognitive complexity. Additionally, users with lighter baseline engagement tend to exhibit larger relative increases, suggesting that those with more available cognitive bandwidth may be better positioned to benefit from AI-driven productivity enhancements. Rather than expanding usage into unfamiliar applications, users appear to reallocate their freed resources toward platforms they already frequent, reinforcing existing habits rather than substantially changing usage patterns.

These findings suggest several practical implications for platform managers. First, generative AI might function more as an engagement amplifier than a direct competitor. While concerns exist that AI tools could reduce time spent on incumbent services, our results indicate that ChatGPT adoption is associated with broader digital engagement, especially in contexts involving complex tasks such as PC-based workflows. For example, Microsoft's integration of ChatGPT-powered Copilot within Office 365 applications has been linked to increased user productivity and time spent within its ecosystem (Microsoft, 2025). Embedding AI in this way may help platforms improve user retention and satisfaction by streamlining workflows.

Second, managers may benefit from focusing AI integration on platforms and use cases characterized by higher cognitive loads. The stronger spillover effects observed on PCs align with the notion that productivity benefits are amplified in settings requiring substantial mental effort. Platforms supporting professional, knowledge-intensive, or creative work—such as Salesforce's Einstein AI embedded in desktop CRM tools—demonstrate how automating routine tasks and augmenting decision-making can be associated with increased platform usage and worker productivity (Salesforce, 2025). Strategic AI investments in these areas could offer meaningful returns.

Finally, embedding AI within existing workflows may deepen user engagement by reinforcing core usage habits. Given that users primarily reinvest their freed cognitive capacity into familiar applications, platforms may see greater benefit from enhancing habitual use rather than pursuing aggressive diversification. Spotify's AI-driven personalized recommendations illustrate this approach by maintaining user engagement within well-established content preferences, thereby supporting loyalty (Spotify, 2025). AI enhancements that enrich core experiences contribute to improved user retention and lifetime value.

Taken together, these insights highlight potential pathways through which generative AI may influence digital engagement. While further research is needed to fully understand these dynamics, managers may consider how AI integration can support existing user behaviors and focus on contexts where productivity gains are most likely to translate into broader platform use.

Despite the contributions of this study, several limitations should be noted. First, our results are based on the early stage of ChatGPT adoption, when the adoption rate was relatively low (9.8%), and the sample size was modest. This may limit the generalizability of our findings, particularly to contexts with higher or more sustained generative AI usage. Second, while our theoretical framework draws on Cognitive Load Theory to propose a mechanism linking productivity gains to cross-platform engagement, we cannot directly measure cognitive load or the reallocation of cognitive resources with the available data. As a result, we present this mechanism as a plausible explanatory pathway, grounded in prior research, rather than a directly tested causal process. These limitations clarify the scope and interpretive boundaries of our study and point to directions for future research.

References

Angrist, J. D. and Pischke, J. S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, Princeton.

Borwankar, S., Khern-am-nuai, W., and Kannan, K. N. (2024). "Unraveling the Impact: An Empirical Investigation of ChatGPT's Exclusion from Stack Overflow," SSRN Working Paper, https://ssrn.com/abstract=4481959, retrieved July 2025.

Brynjolfsson, E., Li, D., and Raymond, L. R. (2025). "Generative AI at Work," *The Quarterly Journal of Economics*, 140(2), pp.889-942.

Burtch, G., Lee, D., and Chen, Z. (2023). "The consequences of generative AI for UGC and online community engagement," SSRN Working Paper, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4521754, retrieved November 2025.

Callaway, B. and Sant'Anna, P. H. (2021). "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 225(2), pp.200-230.

Chung, K., Noh, M., and Rui, N. (2015), "An Empirical Examination of the Continuance Intention to use on Mobile Communication Service in Korea and China," *Korean Management Review*, 44(1), pp. 215-246.

Davis, F. D., Bagozzi, R. P., and Warshaw, P. R. (1989). "User acceptance of computer technology

: A comparison of two theoretical models," *Management science*, 35(8), pp.982-1003.

del Rio-Chanona, M., Laurentsyeva, N., and Wachs, J. (2023). "Are Large Language Models a Threat to Digital Public Goods? Evidence from Activity on Stack Overflow," *arXiv* preprint arXiv:2307.07367, retrieved July 2025.

Ghose, A., Goldfarb, A., and Han, S. P. (2013). "How Is the Mobile Internet Different? Search Costs and Local Activities," *Information Systems Research*, 24(3), pp.613-631.

Hoffmann, M., Boysel, S., Nagle, F., Peng, S., and Xu, K. (2024). "Generative AI and the Nature of Work," CESifo Working Paper Series, No. 11479.

Hong, J., Yoon, S., Park, S., and Han, S. P. (2025). "How User Adoption of ChatGPT Influences Commercial Search Patterns in Traditional Search Engines," *Academy of Management Proceedings*, 2025(1), p.18602.

Hyun, Y. and Choi, S. (1994), "Information Overload and Cognitive Structure," *Korean Management Review*, 23(3), pp.1-22.

Kahneman, D. (1973). "Attention and Effort," Englewood Cliffs, New Jersey: Prentice-Hall.

Karlson, A., Iqbal, S. T., Meyers, B. R., Ramos, G., Lee, K., and Tang, J. (2009). "Mobile Taskflow in Context: A Screenshot Study of Smartphone Usage," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*, pp.2009-2018.

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., and Maes, P. (2025). "Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task," arXiv preprint arXiv:2506.08872.

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X. H., Beresnitzky, A. V., Braunstein, I. and Maes, P. (2025). "Your brain on chatgpt: Accumulation of cognitive debt when using an ai assistant for essay writing task," *arXiv preprint arXiv:2506.08872*.

LaRose, R. (2010). "The problem of media habits," *Communication theory*, 20(2), pp.194-222.

Lee, J. and You, J. (2023). "Effect of Reviewer Writing Device Type (Mobile vs. PC) on Linguistic Style of Hotel Reviews: Focusing on Foreign Traveler Reviews of Hotels Located in Seoul," *Korean Management Review*, 52(3), pp.735-753.

Mark, G., Iqbal, S., Czerwinski, M., and Johns, P. (2015), "Focused, aroused, but so distractible : Temporal perspectives on multitasking and communications," in Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW 2015), Vancouver, Canada, pp. 903-916.

Microsoft. "Fiscal Year 2023 Fourth Quarter Earnings Results," https://www.microsoft.com/en-us/investor/events/fy-2023/earnings-fy-2023-q4 , retrieved November 2025.

Noy, S. and Zhang, W. (2023). "Experimental evidence on the productivity effects of generative artificial intelligence," *Science*, 381(6654), pp.187-192.

Oulasvirta, A., Rattenbury, T., Ma, L., and Raita, E. (2012). "Habits Make Smartphone Use More Pervasive," *Personal and Ubiquitous Computing*, 16(1), pp.105-114.

Paas, F., Renkl, A., and Sweller, J. (2003). "Cognitive load theory and instructional design: Recent developments," *Educational Psychologist*, 38 (1), pp.1-4.

Perkins, M. (2023). "Academic Integrity Considerations of AI Large Language Models in the Post-Pandemic Era: ChatGPT and Beyond," *Journal of University Teaching and Learning Practice*, 20(2), pp.1-24.

Risko, E. F. and Gilbert, S. J. (2016). "Cognitive Offloading," *Trends in Cognitive Sciences*, 20(9), pp.676-688.

Salesforce. "Agentic AI's Impact on the Workforce," https://www.salesforce.com/news/stories/agentic-ai-impact-on-workforce-research/, retrieved October 2025.

Simon, H. A. (1996). "Designing organizations for an information-rich world," *International Library of Critical Writings in Economics*, 70, pp.187-202.

Spotify. "How Spotify Uses Design To Make Personalization Features Delightful," https://newsroom.spotify.com/2023-10-18/how-spotify-uses-design-to-make-personalization-features-delightful/, retrieved October 2025.

Sweller, J. (1988). "Cognitive load during problem solving: Effects on learning," *Cognitive Science*, 12(2), pp.257-285.

Sweller, J., van Merriënboer, J. J. G., and Paas, F. G. W. C. (1998). "Cognitive architecture and instructional design," *Educational Psychology Review*, 10(3), pp.251-296.

Venkatesh, V., Thong, J. Y. L., and Xu, X. (2012). "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology," *MIS Quarterly*, 36(1), pp.157-178.

Wu, S., Liu, Y., Ruan, M., Chen, S., and Xie, X.-Y. (2025). "Human-generative AI collaboration enhances task performance but undermines human intrinsic motivation," *Scientific Reports*, 15(1), 15105.

Zhang, S., Mehta, N., Singh, P. V., and Srinivasan, K. (2021). "Frontiers: Can an Artificial Intelligence Algorithm Mitigate Racial Economic Inequality? An Analysis in the Context of Airbnb," *Marketing Science*, 40(5), pp.813-820.

Zheng, S., Tong, S., Kwon, H. E., Burtch, G., and Li, X. (2023). "Recommending what to search: Sales volume and consumption diversity effects of a query recommender system," Forthcoming at *Marketing Science*.

Zhou, E. and Lee, D. (2024). "Generative artificial intelligence, human creativity, and art," *PNAS Nexus*, 3(3), pgae052.

• The author Haeun Kim is a Ph.D. candidate at SNU Business School, Seoul National University. Her main research areas include digital marketing, retailing, advertising, and generative AI.

• The author Sungho Park is a tenured professor at SNU Business School, Seoul National University. His primary research areas include retail, AI and digital innovation.

• The author Juwon Hong is a Ph.D. candidate at SNU Business School, Seoul National University. Her main research areas include generative AI and AI agents.

• The author Dahye Lee graduated from the Master's program at SNU Business School, Seoul National University. Her primary research area is generative AI.