Do Open-Weight LLM Releases Affect Web Traffic to Proprietary AI Chatbots?

Evidence from Meta's Llama 3.1

Jikhan Jeong[1]

Version: October 16, 2026

Abstract

Despite the widespread use and market importance of proprietary AI chatbots, they remain largely closed systems: users cannot directly access the underlying pre-trained large language models (LLMs), the pre-training data, or the model weights. In contrast, open-weight LLMs make their model weights publicly available, allowing users to download and run the models, fine-tune them for specific tasks, and distill them into smaller, more efficient models in local computing environments, potentially mitigating cost and privacy concerns. However, few studies have examined, within a quasi-experimental framework, how releasing a frontier open-weight text-only LLM affects global online consumer attention to proprietary AI chatbots. To fill this gap, this paper estimates the impact of the release of Meta's open-weight frontier LLM, Llama 3.1, on global web traffic to proprietary AI chatbots using weekly web traffic data and synthetic difference-in-differences (SDID). The results show that the release increases web traffic to proprietary AI chatbots with a delay. In particular, the positive effects on web traffic to chatgpt.com, the dominant domain among proprietary AI chatbots, are larger in magnitude than those for the overall proprietary AI chatbot group and appear immediately. Further platform-specific analyses show that the positive, immediate effects on web traffic to perplexity.ai, an LLM-based search engine, and huggingface.co, an open-source development platform for AI models and datasets, also exceed those for the overall proprietary AI chatbot group. In addition, the release immediately increases web traffic to Meta's corporate website in the first few weeks, with positive effects gradually decreasing over time. Overall, proprietary AI chatbots and AI-related platforms show heterogeneous responses to the Llama 3.1 release in both magnitude and timing, depending on user type, service type, the openness of model weights, and the availability of Llama 3.1 at the application level. These findings align with the conjecture that releasing an open-weight LLM may increase overall online attention to open-source AI models worldwide, which may then spill over to proprietary AI chatbots rather than simply reducing web traffic to them through substitution. The approach and framework developed in this paper can be applied, extended, and scaled to other AI models to better understand user responses and market dynamics in emerging generative AI (Gen AI) markets.

KEYWORDS

Gen AI Markets, Open-weight LLMs, Proprietary AI Chatbots, Web Traffic, Synthetic Difference-in-Differences

JEL
L86, O33, C23

[1] Jaggi School of Business, Missouri University of Science and Technology, Rolla, MO 65409, USA
E-Mail: jikhan.jeong@mst.edu

Acknowledgements

The paper will be presented at the University of Missouri-St. Louis Information Systems and Technology Research Symposium: Technology in the Age of AI without proceedings. English editing of this article was completed using editorworld.com. This manuscript has not been published and is not under consideration for publication elsewhere.

Author statement

The author is solely responsible for all aspects of the work.

Conflict of interest

The author declares no conflict of interest.

Funding information

Partial funding for the Similarweb subscription was provided by the Missouri S&T OVCRI Seed Funding for Arts, Humanities, Social, and Behavioral Science Research.

Data availability

Some data were obtained from Similarweb and cannot be shared as they are paid data. Sample code and sample output for replication, excluding the proprietary data, are available upon request.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

OpenAI's ChatGPT, Microsoft's Copilot, and Grammarly were used to polish and correct grammar in the manually written draft to improve the quality of the manuscript. OpenAI's ChatGPT and Codex were used in some cases to assist with coding.

1. Introduction

In recent years, generative artificial intelligence (Gen AI) has been rapidly adopted in daily life and at work (Bick et al. 2026) and across a wide range of fields, including economics (Korinek 2023). Among Gen AI applications, proprietary AI chatbots have grown substantially in market size, user adoption, and technical capabilities since the launch of OpenAI's ChatGPT in November 2022. Within this market, ChatGPT has remained a dominant player despite the entry of several competing proprietary AI chatbots, including Microsoft Copilot, Google Gemini, Anthropic Claude, and others (Liu et al. 2025). These proprietary chatbots are built on pre-trained large language models (LLMs) and delivered to end users via web- or mobile-based chatbot interfaces. Despite their widespread use and market importance, proprietary AI chatbots remain largely closed systems: users cannot directly access the architecture of the underlying pre-trained LLMs, the pre-training data, or the model weights.

In contrast, open-weight LLMs make their model weights publicly available, letting users download and run the models, fine-tune them for specific tasks, and distill them into smaller, more efficient models on local hardware, potentially reducing cost and privacy concerns. A well-known example is Meta's Llama family of open-weight models (Touvron et al. 2023). Researchers, developers, and other users have widely adopted Llama and fine-tuned these models on their own data for specific purposes. As a result, open-weight LLMs can be adapted to user-specific tasks and provide a foundation for developing new LLMs and applications. Indeed, many derivative LLMs, such as Alpaca (Taori et al. 2023) and Vicuna (Chiang et al. 2023), have been built on or adapted from Llama across various fields and tasks. Therefore, releasing frontier open-weight LLMs can reduce resources and pre-training data constraints, lower costs for researchers studying

LLMs and for new firms entering the market, broaden the scope of LLM applications, and stimulate open innovation.

However, relatively few studies have examined how the release of frontier open-weight LLMs affects consumer attention to and demand for proprietary AI chatbots. This is partly because external users, including researchers, have limited access to internal user-level data across AI firms. In addition, some leading AI firms, such as OpenAI and Anthropic, remain privately held as of October 2026 and are therefore not subject to the same public-company disclosure requirements as publicly traded companies.

To address these challenges, prior studies have relied on alternative approaches. First, some AI chatbot firms and AI service providers examine user behavior using proprietary internal data (Yang et al. 2025; Aubakirova et al. 2026). Second, AI firms and researchers collaborate to conduct digital experiments or analyze internal data (Chatterji et al. 2025). Third, some researchers conduct digital, laboratory, or survey experiments with recruited participants (Agarwal and Sen 2026; Bick et al. 2026; Gholami et al. 2026; Yotzov et al. 2026). Fourth, other researchers collect observational data for associational analyses or quasi-experimental studies (Demirer et al. 2025; Feyzollahi and Habibi 2025; Nagle and Yue 2025; Rafizadeh 2025; Yeverechyahu et al. 2024; Zhao 2026). Finally, other studies use web traffic data to examine demand in Gen AI markets and assess Gen AI's impact on users, businesses, industries, fields, and society (Burtch et al. 2024; Liu et al. 2025; Padilla et al. 2025; Zhao and Berman 2025).

However, to the best of the author's knowledge, no prior study has empirically examined, using web traffic data and quasi-experimental methods, how frontier open-weight LLM releases affect global web traffic to proprietary AI chatbots and AI-related platforms.

Addressing this research question is important because it may improve our understanding of (1) how global online consumer attention to proprietary AI chatbots and AI-related platforms shifts after the release of open-weight LLMs and (2) how market conditions shift after such releases.

More broadly, the release of frontier open-weight LLMs may have either a positive or negative effect on the online presence of proprietary AI chatbots. On the one hand, frontier open-weight LLMs may increase overall online consumer attention to open-source AI models; some of this increased attention may spill over to proprietary AI chatbots. On the other hand, frontier open-weight LLMs may create low-cost outside options for proprietary AI chatbot users, thereby reducing global web traffic to proprietary AI chatbots. In this case, increased competitive pressure on proprietary AI chatbot providers may motivate them to respond strategically through pricing, product differentiation, marketing, and innovation. Accordingly, a quasi-experimental analysis using observational, real-world web traffic data may help clarify the direction and magnitude of the overall effect of a frontier open-weight LLM release on web traffic to proprietary AI chatbots.

To fill this research gap, this paper empirically examines the impact of the release of Meta's Large Language Model Meta AI (Llama) 3.1, a frontier open-weight LLM, on global web traffic to proprietary AI chatbots using web traffic data from Similarweb and two empirical approaches: synthetic difference-in-differences (SDID; Arkhangelsky et al. 2021; Clarke et al. 2024) and event-study SDID (Ciccia et al. 2024; Ciccia 2024). Specifically, Arkhangelsky et al. (2021) developed the SDID estimator, which uses unit and time weights with fixed effects to improve the pre-treatment fit between the treated and control groups. Under appropriate regularity conditions, the SDID estimator is consistent and asymptotically normal (Arkhangelsky et al. 2021). Thus, SDID offers a flexible quasi-experimental approach that can improve pre-treatment fit when the parallel trends assumption required for conventional difference-in-differences (DiD) may be less plausible.

Accordingly, this study uses SDID and event-study SDID as the primary empirical methods, while DiD and event-study DiD serve as robustness checks.

Meta's Llama 3.1 is a frontier open-weight multilingual LLM family released on July 23, 2024. Compared with its predecessor, Llama 3.1 features a longer context window, broader language supports, and stronger reasoning capabilities, with model sizes ranging from 8 billion to 405 billion parameters. In particular, Llama 3.1's 405B model, its largest, outperformed several foundation LLMs for proprietary AI chatbots, including OpenAI's GPT-4 and GPT-4o and Anthropic's Claude 3.5 Sonnet, on some benchmark tasks (Grattafiori et al. 2024). Taken together, Meta's Llama 3.1 provides an appropriate setting for estimating the impact of the release of a text-only frontier open-weight LLM on web traffic to proprietary AI chatbots and other AI-related applications.

The main results extend prior research by showing that the release of Llama 3.1 increases web traffic to proprietary AI chatbots that rely on closed-weight LLMs with a delay, rather than simply reducing web traffic to them through substitution. Specifically, the release's positive effects on web traffic to proprietary AI chatbots gradually increase in the weeks following the release. The release may increase online consumer attention to open-source AI models, and some of this increased attention may subsequently spill over to proprietary AI chatbots. Relatedly, Demirer et al. (2025) found no clear evidence that the release of an open-weight reasoning LLM, DeepSeek-R1, reduced demand for leading LLMs in the LLM application programming interface (API) market, while Azoulay et al. (2025) stated that Meta's Llama does not hinder the adoption of leading proprietary LLMs.

Subsequent platform-specific analysis shows that the positive effects of the Llama 3.1 release on web traffic to chatgpt.com, the dominant proprietary AI chatbot, are more than twice as large as those for the overall proprietary AI chatbot group. The Llama 3.1 release increases web traffic to

chatgpt.com starting in the release week, with the positive effects gradually increasing thereafter. By comparison, platform.openai.com, OpenAI's development platform, shows an immediate positive response in web traffic during the release week, followed by a decline. The estimated response then turns positive again and gradually increases over time. Users interact with ChatGPT through the chatbot interface on chatgpt.com, whereas developers, advanced users, and firms access closed-weight LLMs for ChatGPT through the API on platform.openai.com. Accordingly, the magnitude and timing of users' responses to a frontier open-weight LLM release may differ by user types and by how they access a proprietary AI chatbot service.

Perplexity.ai, a proprietary LLM-based search engine with a chatbot interface, shows a larger and more immediate positive response to the Llama 3.1 release than the overall proprietary AI chatbot group. The results for perplexity.ai may partly reflect the availability of Llama 3.1 on the platform. Specifically, Llama 3.1 was available to paid subscribers as one of the model options on perplexity.ai at the time of its release. Therefore, the release of a frontier open-weight LLM may increase online user attention to proprietary AI-based search engines that allow users to select and use an open-weight LLM, such as Llama 3.1, as the underlying LLM.

The Llama 3.1 release shows a large positive effect on web traffic to meta.com, Meta's corporate website, during the release week. However, the positive effect declines over time and turns negative four weeks after the release. These results suggest that releasing a frontier open-weight LLM may increase a firm's online presence in the short term, with the effect concentrated in the first few weeks rather than persisting longer. Therefore, a firm developing a frontier open-weight LLM may have an incentive to release the model to attract immediate global online attention and increase its online visibility.

Llama 3.1 was made available on huggingngface.co, an open-source platform for AI models and datasets, on its release date. Accordingly, global web traffic to huggingface.co may attract technically advanced users interested in open-source AI models, including open-weight LLMs. The findings show that the Llama 3.1 release increases web traffic to huggingface.co beginning in the release week, with the magnitude of the positive effects generally increasing over time despite some fluctuations. These results suggest that the Llama 3.1 release may increase global user attention to open-source AI models, particularly open-weight LLMs.

Taken together, this paper provides empirical evidence that the release of Meta's frontier open-weight LLM, Llama 3.1, increases global web traffic to proprietary AI chatbots, an AI-based search engine, Meta's corporate website, and an open-source development platform for AI models and datasets. The magnitude and timing of these responses vary by user type, service type, the openness of model weights, and the availability of Llama 3.1 at the application level.

This paper makes three contributions. First, it contributes to the literature on AI markets and adoption, particularly the literature on LLMs (Agarwal and Sen 2026; Aubakirova et al. 2026; Azoulay et al. 2025; Bick et al. 2026; Burtch et al. 2024; Chatterji et al. 2025; Demirer et al. 2025; Feyzollahi and Rafizadeh 2025; Gholami et al. 2026; Haag 2025; Korinek and Vipra 2025; Liu et al. 2025; Padilla et al. 2025; Yotzov et al. 2026; Zhao and Berman 2025). Second, the paper also contributes to the literature on the markets, impacts, and adoption of open-weight LLMs (Azoulay et al. 2025; Choi and Chang 2025; Demirer et al. 2025; Demirer et al. 2026; Habibi 2025; Nagle and Yue 2025; Yeverechyahu et al. 2024; Zhao 2026). Third, the paper contributes to the literature on applications of synthetic difference-in-differences (SDID) and event-study SDID (Arkhangelsky et al. 2021; Arkhangelsky and Samkov 2024; Ciccia et al. 2024; Ciccia 2024; Cobreros et al. 2026; Clarke et al. 2024; Grigoryan et al. 2026; Hu 2026; Hu and Winters 2025;

Ladu and Masi 2025; Muñiz-Fernández and Boto-García 2026; Saffer et al. 2025). In particular, this paper applies SDID to estimate the average treatment effect on the treated (ATT) of a frontier open-weight LLM release on global online consumer attention to proprietary AI chatbots and AI-related platforms, using weekly web traffic data. Moreover, this study uses an event-study SDID to estimate the dynamic treatment effects.

Overall, these findings can help consumers, developers, firms, and policymakers better understand how the release of frontier open-weight LLMs affects the online presence of proprietary AI chatbots and AI-related platforms, as well as the structure, dynamics, and competition in LLM markets. They also inform the behavior of for-profit firms that develop and deploy such models. Moreover, this paper's approach applies to and is extensible to both open- and closed-weight LLMs.

The remainder of this paper is organized as follows. Section 2 reviews the prior literature. Section 3 describes the data. Section 4 presents the empirical methods and results. Section 5 concludes and discusses topics for future research.

2. Literature Review

This paper contributes to the literature on LLM markets and adoption. Given the rapid adoption and widespread use of Gen AI, several scholars have examined LLM market structure and competition, as well as LLM adoption. For example, Korinek and Vipra (2025) investigated the market structure and competitive dynamics of LLMs, focusing specifically on foundation models. Azoulay et al. (2025) analyzed market structure and competition among foundation models in Gen AI from an innovation-economics perspective. Haag (2025) also examined international competition in AI.

Regarding firm-level LLM adoption, Demirer et al. (2025) examined the market structure and competitive dynamics of the LLM API market for business use, using data from OpenRouter and Microsoft Azure. Specifically, OpenRouter allows users to access various LLMs from different providers through a single unified API key. Therefore, OpenRouter can be regarded as an LLM API market. Aubakirova et al. (2026) also conducted an in-depth, data-driven analysis of both open-weight and closed-weight LLMs in the LLM API market using OpenRouter data. Furthermore, Yotzov et al. (2026) examined firms' AI adoption across countries using survey data.

Regarding consumer-side adoption of LLMs via chatbot interfaces, Liu et al. (2025) analyzed global Gen AI adoption using web traffic data. Bick et al. (2026) further examined Gen AI adoption among U.S. users at home and in the workplace using survey data. Chatterji et al. (2025) investigated OpenAI's ChatGPT usage behavior among free and paid users using internal data. Relatedly, Padilla et al. (2025) empirically estimated the effect of ChatGPT adoption on consumer search behavior using commercial clickstream data. Moreover, Gholami et al. (2026) examined the impact of LLM adoption on search activities, online behavior, and web navigation.

Furthermore, recent literature has begun to examine ChatGPT's impact on specific industries, platforms, and academic fields. For instance, Zhao and Berman (2025) documented the effect of ChatGPT's release on the online news industry using web traffic data. Similarly, Burtch et al. (2024) estimated ChatGPT's impact on web traffic to knowledge-sharing platforms. Moreover, Feyzollahi and Rafizadeh (2025) empirically demonstrated that ChatGPT's release increased LLM adoption in economics research.

This paper also contributes to the literature on the markets, adoption, and impacts of open-weight LLMs. In contrast to prior studies focusing on proprietary AI chatbots, a growing body of recent research has examined open-weight LLMs. For example, Demirer et al. (2025) showed that open-

source LLMs are a major driver of the rapid growth of the LLM API market and that their per-token API prices are significantly lower than those of closed-source LLMs with similar intelligence. Demirer et al. (2026) further highlighted the role of open-weight LLMs in LLM API markets, particularly in lowering prices and barriers to entry in downstream markets and stimulating experimentation. Nagle and Yue (2025) also showed that frontier open-source LLMs in the LLM API market deliver performance comparable to frontier closed-source LLMs at lower prices. However, frontier closed-source LLMs accounted for a dominant share of API usage on OpenRouter.

In addition, Lerner and Tirole (2005) noted that key participants in open-source production (e.g., open-source software development) include both individual contributors and for-profit firms. Because LLM development requires substantial resources, open-weight LLM development is more likely to be led by large firms, such as Meta, than by individual contributors. In particular, Habibi (2025) examined the rationale for a for-profit firm's decision to release an open-source LLM and empirically demonstrated that Meta's Llama 1, its initial open-weight LLM released on February 24, 2023, stimulated LLM-related research activity on GitHub, a commercial software development platform. Similarly, Yeverechyahu et al. (2024) documented the positive impact of GitHub Copilot, an LLM-based coding assistance tool, on open-source contributions to GitHub.

Another recent example of a frontier open-weight LLM is DeepSeek-R1 (Guo et al. 2025), which was released on January 20, 2025. Its release attracted significant attention from market participants because it demonstrated performance comparable to that of OpenAI's reasoning models while making its model weights publicly available. Similarly, Choi and Chang (2025) compared DeepSeek's advantages and limitations with those of OpenAI's ChatGPT. Beyond these comparisons, Zhao (2026) examined the effect of DeepSeek-R1's release on the stock values of

U.S. and Chinese firms. Importantly, Demirer et al. (2025) found no evidence that the release of DeepSeek-R1 affected demand for other leading LLMs on OpenRouter. Azoulay et al. (2025) also stated that Meta's Llama stimulated open innovation in open-source developer communities without slowing the rapid adoption of proprietary LLMs developed by leading firms.

Furthermore, this paper contributes to the literature on empirical applications of SDID and event-study SDID (Arkhangelsky et al. 2021; Arkhangelsky and Samkov 2024; Ciccia et al. 2024; Ciccia 2024; Clarke et al. 2024). SDID has been applied across a wide range of quasi-experimental studies in various empirical settings. In prior research on LLMs and digital platforms, Habibi (2025) estimated the effect of Meta's LLaMA 1 on LLM-related research activity on GitHub using SDID as a robustness check. Similarly, Yeverechyahu et al. (2024) estimated the effect of GitHub Copilot on open innovation on GitHub using SDID as an alternative specification. Beyond LLM-related applications, Grigoryan et al. (2026) estimated the impact of a war on neutral states' exports using SDID, while Cobreros et al. (2026) applied an SDID design to estimate the effect of gender-neutral financial incentives on the share of males in general practice. Saffer et al. (2025) examined the effects of flavored e-cigarette regulations on e-cigarette use and smoking using both DiD and SDID, and Ladu and Masi (2025) used SDID to estimate the economic effects of earthquakes. Muñiz-Fernández and Boto-García (2026) also estimated the effect of political instability on the housing market using SDID, while Hu (2026) examined how closures of food manufacturing plants affect local economies using SDID. Hu and Winters (2025) also estimated the effects of designating Econometrics and Quantitative Economics (EQE) programs as Science, Technology, Engineering, and Mathematics (STEM) programs on international economics degrees in the U.S. using SDID.

While prior studies have examined LLM markets and adoption, empirical studies assessing the impact of a frontier open-weight LLM on online consumer attention to proprietary AI chatbots and

AI-related platforms within a quasi-experimental framework remain relatively limited. To the best of the author's knowledge, this paper is the first to estimate the impact of the release of Llama 3.1, a frontier open-weight LLM, on global web traffic to proprietary AI chatbots and other AI-related platforms and to examine the dynamic effects of the release using SDID and event-study SDID.

3. Data

The author manually collects web traffic data through the Similarweb web interface. Daily visits and unique visitors for each website worldwide are downloaded and merged. The daily web traffic data are then aggregated to the weekly level because daily web traffic is highly volatile. This aggregation approach is consistent with a prior study using Similarweb data (Congiu et al. 2022).

Llama is an open-weight family of pre-trained foundation language models developed by Meta. Users can download Llama model weights to run and modify the models, fine-tune them on their own data for specific tasks, distill them into smaller, more efficient LLMs, develop new LLMs, and create new applications. In particular, Llama 3.1 is a family of multilingual, text-only foundation LLMs with model sizes ranging from 8 billion to 405 billion parameters. Compared with Llama 3, Llama 3.1 offers a longer context window, a larger maximum model size, and improved reasoning and multilingual capabilities. Additionally, Llama 3.1 405B, the flagship model of the Llama 3.1 family, performed competitively with, or outperformed, leading proprietary closed-weight LLMs, such as OpenAI's GPT-4 and GPT-4o, and Anthropic's Claude 3.5 Sonnet, on several benchmark tasks (Grattafiori et al. 2024). Accordingly, this paper aims to estimate the effect of the release of Llama 3.1, a frontier open-weight, text-only LLM, rather than the release of a closed-weight or multimodal LLM (e.g., a vision-language model), on web traffic to proprietary AI chatbots and AI-related platforms.

Llama 3, 3.1, and 3.2 were released on April 18, 2024, July 23, 2024, and September 25, 2024, respectively (Table 1). Releases within the Llama family may have affected web traffic to proprietary AI chatbots either immediately or with a delay. Because the treatment event examined is the release of Llama 3.1, the sample period is restricted to the period ending before the release of Llama 3.2. In addition, the previous domain for ChatGPT was chat.openai.com, which was replaced by chatgpt.com, the current official domain for ChatGPT, in May 2024. Therefore, the paper uses weekly data from June 18, 2024, to September 23, 2024, to mitigate potential confounding from the release of Llama 3, the subsequent release of Llama 3.2, and the ChatGPT domain change. Specifically, the daily data are aggregated to the weekly level using Tuesday-to-Monday periods to align the weekly aggregation window with the Llama 3.1 release on Tuesday, July 23, 2024.

Table 1 Model characteristics of Llama 3, Llama 3.1, and Llama 3.2

Model	Release date	Input	Output	Context length	Parameters	Reasoning
Llama 3	April 18, 2024	Text only	Text and code	8k	8, 70B	Yes (text)
Llama 3.1	July 23, 2024	Multilingual text	Multilingual text and code	128k	8B, 70B, 405B	Yes (text)
Llama 3.2	September 25, 2024	Multilingual text, image	Multilingual text, code, image	128k	Vision (11B, 90B) Text only (1B, 3B)	Yes (text and image)

* Sources: Meta (2024a, 2024b, 2024c)

The treated group in the main analysis comprises four leading proprietary AI chatbot domains that offer web-based AI chatbot interfaces: chatgpt.com for OpenAI's ChatGPT, claude.ai for Anthropic's Claude, gemini.google.com for Google's Gemini, and copilot.microsoft.com for Microsoft's Copilot. In particular, chatgpt.com accounts for 86.031% of visits and 87.596% of unique visitors to proprietary AI chatbots in the treated group (Table 2).

The control group comprises fourteen major digital platform domains that were expected to be less directly affected by the release of Llama 3.1. Moreover, Figures 1 and 2 show similar pre-treatment trends in average log visits and log unique visitors between the treated (i.e., proprietary

AI chatbots) and control groups, with a level difference between the groups. Summary statistics for visits and unique visitors for the treated and control groups are presented in Tables 2 and 3.

Table 2 Summary statistics for weekly visits from June 18, 2024, to September 23, 2024

Group	Domain	Sum	Mean	se(mean)	Max	Min
Proprietary AI chatbots	chatgpt.com	8,435,958,419	301,284,229	21,962,822	507,247,828	162,418,893
	claude.ai	217,995,300	7,785,546	904,250	14,125,710	2,746,399
	copilot.microsoft.com	100,448,845	3,587,459	327,050	6,530,284	1,640,443
	gemini.google.com	876,162,196	31,291,507	947,336	39,366,492	24,770,168
Control	airbnb.com	339,014,873	12,107,674	265,396	14,787,493	9,596,719
	amazon.com	8,736,526,620	312,018,808	3,771,166	385,231,092	287,463,020
	ebay.com	1,984,716,654	70,882,738	477,498	74,877,028	6,725,6173
	facebook.com	40,577,609,679	1,449,200,346	15,056,997	1,544,166,501	1,337,959,524
	indeed.com	1,381,703,175	49,346,542	386,503	52,998,560	45,576,637
	instagram.com	18,734,590,896	669,092,532	18,264,663	776,032,413	565,354,359
	linkedin.com	5,554,460,599	198,373,593	14,192,240	279,198,267	120,634,307
	lyft.com	34,342,992	1,226,535	178,212	2,403,952	286,492
	reddit.com	11,259,305,709	402,118,061	2,780,860	428,254,254	379,554,492
	spotify.com	1,550,611,582	55,378,985	3,371,224	74,957,844	35,049,770
	temu.com	2,090,824,345	74,672,298	4,344,308	106,062,569	48,669,915
	uber.com	238,806,894	8,528,818	985,973	14,251,220	3,369,283
	walmart.com	1,439,122,872	51,397,245	3,513,342	75,827,150	30,965,748
	youtube.com	93,829,366,903	3,351,048,818	357,616,126	5,371,423,945	1,464,835,729
		197,381,568,552	391,630,096	40,586,447	5,371,423,945	286,492

* Note: values are rounded to the nearest integer.

Table 3 Summary statistics for weekly unique visitors from June 18, 2024, to September 23, 2024

Group	Domain	Sum	Mean	se(mean)	Max	Min
Proprietary AI chatbots	chatgpt.com	4,650,652,699	166,094,739	12,631,866	279,846,185	89,992,822
	claude.ai	118,134,630	4,219,094	459,537	7,506,509	1,614,037
	copilot.microsoft.com	72,080,557	2,574,306	241,940	4,705,750	1,209,018
	gemini.google.com	564,892,343	20,174,727	738,232	26,627,192	15,226,345
Control	airbnb.com	250,041,979	8,930,071	237,865	11,222,217	6,805,542
	amazon.com	6,028,948,263	215,319,581	4,896,256	273,719,869	181,597,532
	ebay.com	1,346,963,226	48,105,830	950,916	53,918,692	42,247,416
	facebook.com	19,892,843,049	710,458,680	25,835,998	855,144,656	565,975,411
	indeed.com	979,137,184	34,969,185	490,847	39,051,478	31,351,423
	instagram.com	10,694,141,783	381,933,635	19,745,304	491,289,115	275,968,099
	linkedin.com	330,208,4491	117,931,589	5,377,156	148,107,370	87,190,843
	lyft.com	26,183,267	935,117	131,862	1,812,259	235,187
	reddit.com	6,855,399,511	244,835,697	8,839,059	299,566,846	196,967,763
	spotify.com	1,052,331,951	37,583,284	1,196,220	45,604,521	29,219,157
	temu.com	1,603,857,399	57,280,621	4,062,312	85,407,164	34,017,257
	uber.com	196,127,145	7,004,541	793,773	11,640,762	2,833,739
	walmart.com	1,165,077,772	41,609,920	3,398,205	64,572,219	22,477,147
	youtube.com	37,483,780,062	1,338,706,431	72,110,862	1,742,179,268	949,028,699
		90,876,917,082	231,828,870	19,085,665	1,742,179,268	235,187

* Note: values are rounded to the nearest integer.

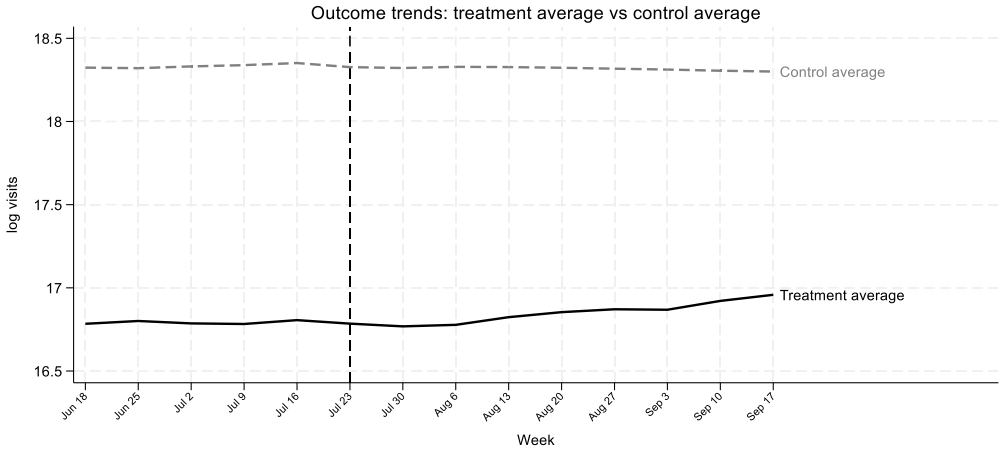


Figure 1 outcome trends in log visits

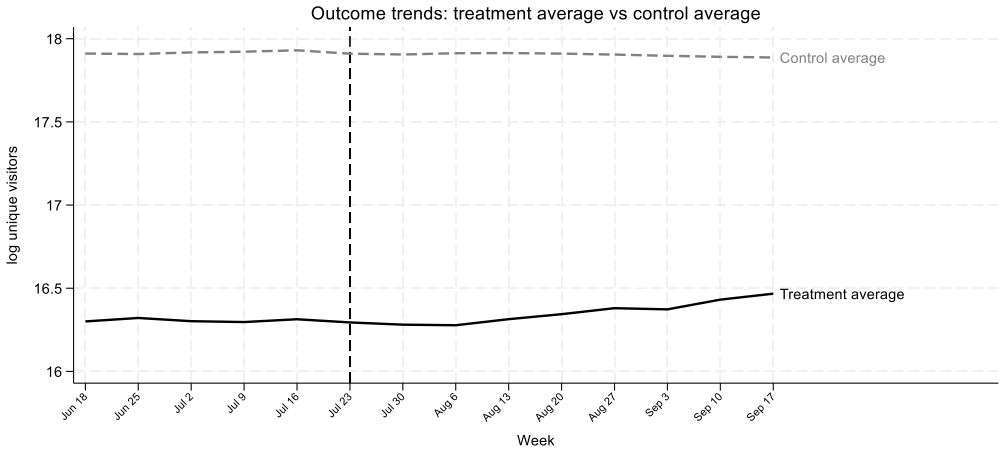


Figure 2 outcome trends in log unique visitors

The dependent variables are the natural logarithms of weekly visits (i.e., log visits) and weekly unique visitors (i.e., log unique visitors) for each domain worldwide. Weekly visits count the total number of visits to a domain in a given week, whereas weekly unique visitors count the number of distinct individuals who visit a domain during that week. The econometric specification includes week fixed effects and unit fixed effects for each domain-device pair to account for unobserved

week-specific shocks affecting all units and unobserved time-invariant characteristics of each domain-device pair, respectively.

4. Method

DiD and the synthetic control method (SCM; Abadie and Gardeazabal 2003; Abadie et al. 2010) are widely used quasi-experimental methods when randomized controlled trials (RCTs) are not feasible. In the digital economics literature, many prior studies have applied DiD (Burtch et al. 2024; Congiu et al. 2022; He et al. 2022; Padilla et al. 2025; Zhao and Berman 2025), and Burtch et al. (2024) also applied SCM.

DiD is a widely used quasi-experimental method in settings with many treated units when the parallel trends assumption holds. The parallel trends assumption implies that, without treatment, the treated group would have followed the same trend as the control group (Clarke et al. 2024). In contrast, SCM is a flexible approach for settings with a single or a small number of treated units because it reweights control units to ensure that a synthetic control group closely matches the treated group during the pre-treatment period (Clarke et al. 2024).

SDID is a recent quasi-experimental method for balanced panel data that combines features of DiD and SCM while addressing some limitations of both approaches (Arkhangelsky et al. 2021). Arkhangelsky et al. (2021) developed the estimator and provided a detailed explanation along with several empirical applications. Accordingly, this paper slightly modifies their notation to fit the research question examined here.

This paper applies the SDID method to balanced panel data and estimates the effect of the Llama 3.1 release on global web traffic to proprietary AI chatbots using the following two-way fixed-effects (TWFE) specification:

$$Y_{it} = \mu + \alpha_i + \gamma_t + \beta Release_{it} + \varepsilon_{it} \quad (1)$$

In the specification above, *i* indexes a unit, where each unit is defined as a domain-device pair, and *t* indexes a week. Y_{it} denotes the dependent variable for unit *i* in week *t*. Specifically, the dependent variable is either log visits or log unique visitors for unit *i* in week *t*. α_i denotes domain-by-device fixed effects, which account for unobserved time-invariant factors specific to each domain-device pair. Similarly, γ_t denotes week fixed effects, which capture unobserved time-varying common shocks affecting all domain-device pairs. $Release_{it}$ is a binary treatment indicator equal to one for the treated units after the Llama 3.1 release and zero otherwise. ε_{it} is an unobserved error term. Thus, this research design corresponds to a single-adoption setting rather than a staggered-treatment setting. The coefficient β represents the ATT.

The above specification can also be estimated using conventional TWFE DiD with balanced panel data. In addition, event-study DiD can decompose the overall DiD estimate into dynamic treatment effects using leads and lags relative to the Llama 3.1 release. Event-study DiD can also be used to assess potential violations of the parallel trends assumption underlying the DiD estimator by testing whether the pre-treatment lead coefficients are statistically different from zero using the following specification with a single adoption period:

$$Y_{it} = \mu + \alpha_i + \gamma_t + \sum_{j=2}^{J} \beta_j \, (lead\ j)_{\,it} + \sum_{\mathrm{k}=0}^{\mathrm{K}} \delta_{\mathrm{k}} \, (lag\ k)_{\,it} + \varepsilon_{it} \qquad (2)$$

where *j* indicates that a lead term corresponds to *j* weeks before the Llama 3.1 release, and *J* denotes the total number of lead terms. Likewise, *k* indicates that a lag term corresponds to *k* weeks after the release, and *K* denotes the total number of lag terms. Each lead or lag term is a binary indicator equal to one if a treated unit is observed *j* weeks before or *k* weeks after the release, respectively, and zero otherwise. Accordingly, β_j denotes the pre-treatment lead estimate j weeks before the release, while δ_k denotes the post-treatment estimate k weeks after the release. The one-week lead before the release (i.e., $(lead\ j = 1)_{\,it}$) is omitted from Equation (2) as the reference

period. Therefore, all lead and lag coefficients are interpreted relative to the reference period. In addition, this paper applies an event-study DiD analysis as a robustness check using the Stata package *'eventdd'* (Clarke and Tapia-Schythe 2021).

In particular, the statistical significance of the pre-treatment lead coefficients is commonly used to assess differential pre-trends, which may signal potential violations of the parallel trends assumption underlying DiD. If the average pre-treatment trends differ between the treated and control groups, the parallel trends assumption becomes less credible.

When the parallel trends assumption for conventional DiD is less plausible, SDID can estimate the ATT by applying unit and time weights to improve pre-treatment fit between the treated and control groups while incorporating unit and time fixed effects.

The SDID estimator of the ATT, $\hat{\beta}^{sdid}$, is defined as the difference between the average adjusted outcome for the treated units and the weighted average adjusted outcome for the control units, as follows:

$$\hat{\beta}^{sdid} = \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} \widehat{\theta_{\mathrm{i}}} - \sum_{i=1}^{N_{co}} \widehat{w}_i \widehat{\theta_{\mathrm{i}}} \tag{3}$$

where *N,* N_{co}*,* and N_{tr} $(= N - N_{co})$ denote the total number of units, the number of control units, and the number of treated units, respectively. $\widehat{\theta_{\mathrm{i}}}$ denotes the adjusted outcome for unit *i*, and $\widehat{w}_i$ denotes the unit weight.

The adjusted outcome for SDID $\widehat{\theta_{\mathrm{i}}}^{sdid}$ can be expressed as follows:

$$\widehat{\theta_{\mathrm{i}}} = \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it} - \sum_{t=1}^{T_{pre}} \hat{\lambda}_t Y_{it} \tag{4}$$

Where *T,* T_{pre}*,* and T_{post} $(= T - T_{pre})$ denote the total number of periods, the number of pre-treatment periods, and the number of post-treatment periods, respectively. Y_{it} is the dependent

variable for unit *i* in period *t*, and $\hat{\lambda}_t$ denotes the estimated time weight for a pre-treatment period *t*. The SDID estimates for the ATT from the TWFE model in Equation (1) are estimated as follows:

$$(\hat{\beta}^{SDID}, \hat{\mu}, \hat{\alpha}, \hat{\gamma}) = \underset{\beta,\mu,\alpha,\gamma}{\operatorname{argmin}} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \mu - \alpha_i - \gamma_t - \beta release_{it})^2 \hat{\lambda}_t^{sdid} \hat{w}_i^{sdid} \right\} \quad (5)$$

where μ denotes a constant term, while α_i, and γ_t denote unit and time fixed effects, respectively.

Unit weights are selected to improve the fit between the pre-treatment trends of the weighted average of control units and those of the average of treated units by solving the following regularized minimization problem:

$$(\widehat{w}_0, \widehat{w}) = \underset{w_0 \in \mathbb{R}, w \in \Omega}{\operatorname{argmin}} \sum_{t=1}^{T_{pre}} \left\{ w_0 + \sum_{i=1}^{N_{co}} w_i Y_{it} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} Y_{it} \right\}^2 + \Lambda^2 T_{pre} || w||_2^2 \quad (6)$$

where

$$\Omega = \left\{ w \in \mathbb{R}_+^N : \sum_{i=1}^{N_{co}} w_i = 1, w_i = \frac{1}{N_{tr}} \text{ for all treated unit} \right\}$$

where $\mathbb{R}_+^N$ denotes the set of *N*-dimensional vectors with nonnegative real numbers. w_0 is a constant term that allows for a constant level difference in pre-treatment outcomes between the treated and control units, rather than requiring an exact match in their pre-treatment outcome level. w_i denotes the unit weight for control unit *i*. Λ is a regularization parameter that penalizes the concentration of unit weights on a small number of control units.

Time weights are selected without regularization to assign greater weight to pre-treatment periods that are more similar to post-treatment periods by minimizing the difference between the weighted average of pre-treatment outcomes and the average post-treatment outcome for the control units (Clarke et al. 2024), as follows:

$$(\hat{\lambda}_0, \hat{\lambda}) = \underset{\lambda_0 \in \mathbb{R}, \lambda \in \Psi}{\operatorname{argmin}} \sum_{i=1}^{N_{co}} \left\{ \lambda_0 + \sum_{t=1}^{T_{pre}} \lambda_t Y_{it} - \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it} \right\}^2 \quad (7)$$

where

$$\Psi = \left\{ \lambda \in \mathbb{R}_+^T : \sum_{t=1}^{T_{pre}} \lambda_t = 1, \lambda_t = \frac{1}{T_{post}} \text{ for all post} - \text{treatment periods} \right\}$$

where $\mathbb{R}_+^T$ denotes the set of *T*-dimensional vectors with nonnegative real numbers. λ_0 is a constant term that allows a constant difference between the weighted average of pre-treatment outcomes and the average post-treatment outcome for the control units. λ_t denotes the time weight for period *t*.

Under appropriate regularity conditions, the SDID estimator is consistent and asymptotically normal. Its confidence intervals can be estimated using bootstrap, jackknife, or placebo inference. In particular, placebo inference is appropriate when there is a single treated unit or a small number of treated units. Therefore, this paper uses placebo inference because the treated group consists of only eight domain-device pairs. Additionally, the SDID analysis is implemented using the Stata package *'sdid'* (Clarke et al. 2024).

The SDID estimator can be extended to an event-study SDID framework under either single or staggered treatment adoption (Arkhangelsky et al. 2021; Ciccia 2024; Porreca 2022). In particular, Ciccia (2024) provided functional forms for estimating disaggregated dynamic treatment effects with the event-study SDID estimator. This paper slightly modifies Ciccia's (2024) notation to present the event-study SDID estimator at event time *l* in a single-adoption setting as follows:

$$\hat{\beta}_l^{SDID} = \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} Y_{i,T_{pre}+1+l} - \sum_{i=1}^{N_{co}} w_i Y_{i,T_{pre}+1+l} - \sum_{t=1}^{T_{pre}} \left(\frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} \lambda_t Y_{it} - \sum_{i=1}^{N_{co}} w_i \lambda_t Y_{it} \right) \quad (8)$$

where $l \in \{0, \ldots, T_{post}\}$.

In addition, this paper uses the Stata package *'sdid_event'* (Ciccia 2024) to implement the event-study SDID analysis.

5. Results

This section presents the main empirical results for proprietary AI chatbots using the SDID and event-study SDID approaches, followed by additional platform-specific analyses. It then presents robustness checks for the main results.

5.1 Main results

5.1.1. SDID results for proprietary AI chatbots

As shown in Table 4, the estimated ATTs for the Llama 3.1 release are statistically significant at the 5% level for both dependent variables: log visits and log unique visitors to proprietary AI chatbots. Specifically, the Llama 3.1 release increases visits and unique visitors by approximately 8.437% and 7.144%, respectively. Because this paper uses a log-linear model, the percentage effects are calculated as $100\times(\exp(\hat{\beta}^{SDID})-1)\%$.

Importantly, the release of Llama 3.1, a frontier open-weight LLM, increases visits and unique visitors to proprietary AI chatbots built on closed-weight LLMs, rather than reducing them through substitution. One plausible explanation is that the release may increase overall online consumer attention to open-weight LLMs. Some of this increased attention may then spill over to proprietary AI chatbots because of their potential advantages, such as user-friendly chatbot interfaces and strong brand recognition.

In addition, bootstrap and jackknife inference are less appropriate than placebo-based inference for both SDID and event-study SDID in settings with a single or a small number of treated units (Arkhangelsky et al. 2021). Accordingly, this paper uses placebo-based inference for the SDID and event-study SDID analyses because there are only eight treated domain-device pairs.

Table 4 SDID results for proprietary AI chatbots

Model	(1) SDID	(2) SDID
Dependent variable	Log visits	Log unique visitors
ATT	0.081**	0.069**
	(0.033)	(0.031)
Domain × device FE	Yes	Yes
Weekly FE	Yes	Yes
Observations	504	504

* Note: Placebo-based standard errors with 500 replications are reported in parentheses for the SDID estimates. *** p<0.01, **p<0.05, * p<0.1.

5.1.2. Event-study SDID results for proprietary AI chatbots

The SDID approach can be extended to an event-study framework to estimate dynamic treatment effects in a single-adoption setting (Ciccia 2024). Accordingly, this subsection estimates the dynamic treatment effects using event-study SDID.

Table 5 presents event-study SDID estimates of the dynamic treatment effects of the Llama 3.1 release. Event-study SDID estimates show that while the releases' immediate impact remains modest in the first two weeks, positive effects on visits and unique visitors begin to scale in the third week and become statistically significant by the fourth and fifth weeks, respectively (Figures 3 and 4).

Among estimates that are statistically significant at the 5% or 1% level, the release increases visits to proprietary AI chatbots by approximately 8.329%, 11.405%, 11.182%, 17.939%, and 22.753% four, five, six, seven, and eight weeks after the release, respectively. Similarly, the release increases unique visitors to proprietary AI chatbots by approximately 10.517%, 9.966%, 17.234%, and 21.774% five, six, seven, and eight weeks after the release, respectively (Table 5). Overall, these results suggest a delayed positive response in web traffic to proprietary AI chatbots after the release.

Additionally, the pre-treatment placebo effects for both log visits and log unique visitors are close to zero and statistically insignificant, except for the estimates four weeks before the release, which

are statistically significant at the 10% and 5% levels, respectively, suggesting limited evidence of differential pre-trends. Nevertheless, the Wald tests for both log visits and log unique visitors fail to reject the joint null hypothesis that all pre-treatment placebo effects are equal to zero (Table 5).

Table 5 Event-study SDID results for proprietary AI chatbots

Model	(1) Event-study SDID	(2) Event-study SDID
Dependent variable	Log visits	Log unique visitors
Placebo effect (t-5)	0.005 (0.016)	0.005 (0.015)
Placebo effect (t-4)	0.024* (0.013)	0.026** (0.012)
Placebo effect (t-3)	0.007 (0.011)	0.007 (0.009)
Placebo effect (t-2)	-0.004 (0.008)	-0.003 (0.006)
Placebo effect (t-1)	0.010 (0.013)	0.010 (0.013)
ATT (t)	0.010 (0.027)	0.006 (0.022)
ATT (t+1)	-0.002 (0.023)	-0.004 (0.020)
ATT (t+2)	-0.001 (0.021)	-0.017 (0.018)
ATT (t+3)	0.050 (0.031)	0.024 (0.028)
ATT (t+4)	0.080** (0.036)	0.053 (0.034)
ATT (t+5)	0.108*** (0.041)	0.100** (0.039)
ATT (t+6)	0.106** (0.048)	0.095** (0.047)
ATT (t+7)	0.165*** (0.048)	0.159*** (0.047)
ATT (t+8)	0.205*** (0.049)	0.197*** (0.049)
Domain × device FE	Yes	Yes
Weekly FE	Yes	Yes
Wald test: placebo effects jointly equal zero	0.4254	0.2887
Observations	504	504

* Note: Placebo-based standard errors with 500 replications are reported in parentheses for the SDID estimates. The Wald test presents the p-value for the joint null hypothesis that all pre-treatment placebo effects are equal to zero.
*** p<0.01, **p<0.05, * p<0.1.

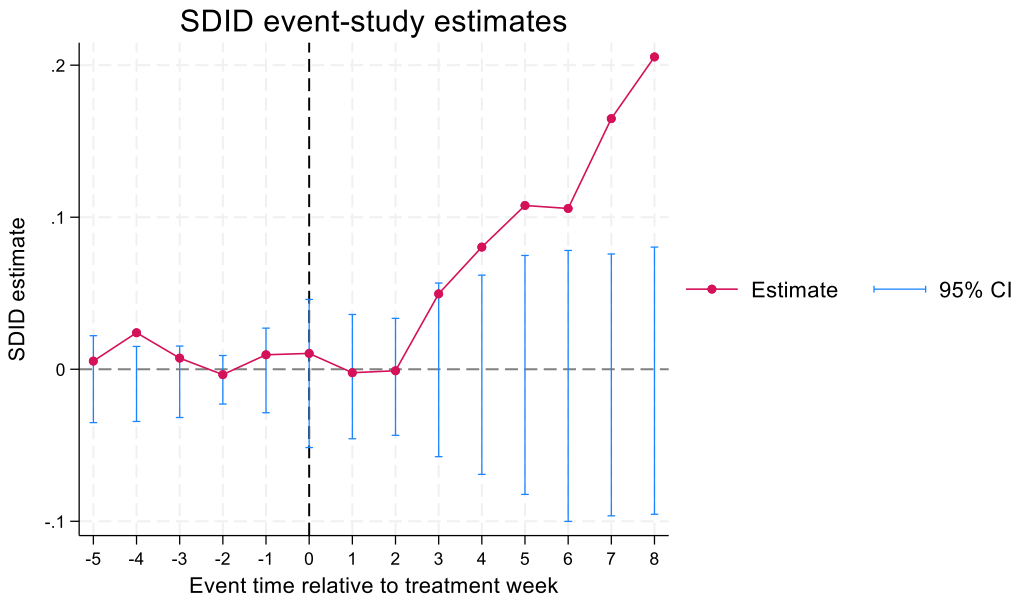


Figure 3 Event-study SDID estimates for log visits to proprietary AI chatbots

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

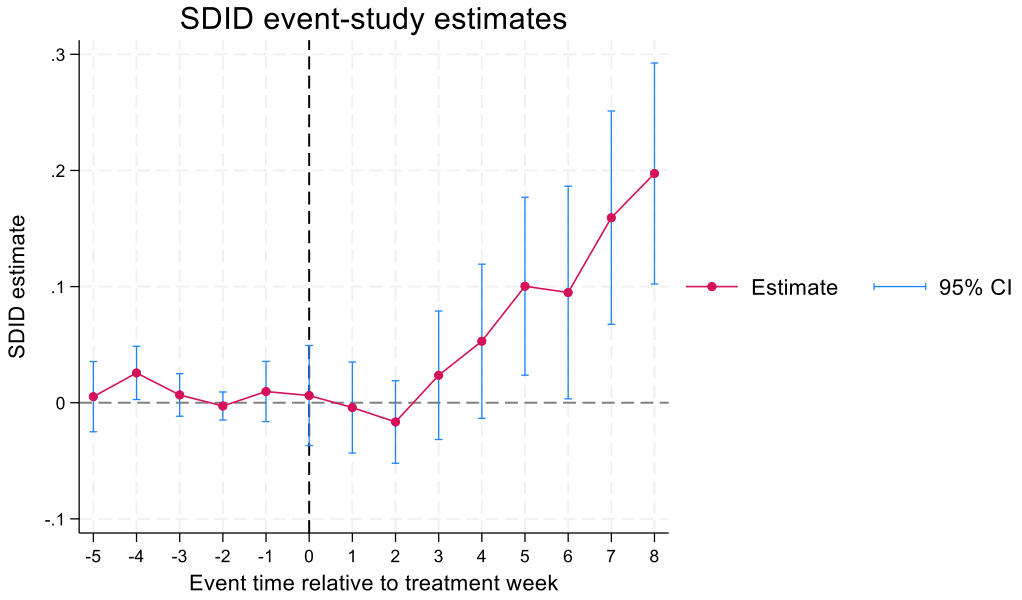


Figure 4 Event-study SDID estimates for log unique visitors to proprietary AI chatbots

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

5.2 Additional platform-specific analyses

This subsection presents platform-specific analyses of individual LLM-related platforms using SDID and event-study SDID. These methods suit settings with a single treated unit, whereas conventional DiD is less suitable.

Each platform-specific analysis focuses on one of the following domains: chatgpt.com for OpenAI's ChatGPT, platform.openai.com for OpenAI's developer platform, perplexity.ai for Perplexity's LLM-based search engine, google.com for Google's dominant global search engine, llama.meta.com for Meta's AI development platform, meta.com for Meta's corporate website, huggingface.co for an open-source AI model and dataset platform, github.com for a proprietary software development platform, and stackoverflow.com for a knowledge-sharing platform focused on computer programming. In addition, Tables 6 and 7 provide summary statistics on global visits and unique visitors to these domains during the sample period.

Table 6 Summary statistics for weekly visits during the sample period

Group	Domain	Sum	Mean	se(mean)	Max	Min
Proprietary AI chatbots	chatgpt.com	8,435,958,419	301,284,229	21,962,822	507,247,828	162,418,893
	claude.ai	217,995,300	7,785,546	904,250	14,125,710	2,746,399
	copilot.microsoft.com	100,448,845	3,587,459	327,050	6,530,284	1,640,443
	gemini.google.com	876,162,196	31,291,507	947,336	39,366,492	24,770,168
	platform.openai.com	89,317,765	3,189,920	205,933	5,099,009	1,790,155
	perplexity.ai	183,858,942	6,566,391	729,922	13,284,345	2,452,239
Specific domains for additional analyses	google.com	265,768,356,762	9,491,727,027	387,183,768	11,581,875,831	7,365,288,628
	llama.meta.com	4,856,299	173,439	27,922	754,733	4,053
	meta.com	61,495,020	2,196,251	80,629	2,974,699	1,576,623
	huggingface.co	55,625,190	1,986,614	186,151	3,367,214	946,942
	github.com	1,581,607,585	56,485,985	6,110,321	90,514,488	23,678,490
	stackoverflow.com	494,899,828	17,674,994	2,671,918	33,555,922	3,556,106

Table 7 Summary statistics for weekly unique visitors during the sample period

Group	Domain	Sum	Mean	se(mean)	Max	Min
Proprietary AI chatbots	chatgpt.com	4,650,652,699	166,094,739	12,631,866	279,846,185	89,992,822
	claude.ai	118,134,630	4,219,094	459,537	7,506,509	1,614,037
	copilot.microsoft.com	72,080,557	2,574,306	241,940	4,705,750	1,209,018
	gemini.google.com	564,892,343	20,174,727	738,232	26,627,192	15,226,345
	platform.openai.com	72,639,980	2,594,285	152,353	4,140,394	1,572,955
	perplexity.ai	108,065,238	3,859,473	419,628	7,777,869	1,504,525
Specific domains for additional analyses	google.com	122,506,285,057	4,375,224,466	333,808,319	6,147,977,318	2,627,070,894
	llama.meta.com	4,087,619	145,986	24,021	666,392	3,450
	meta.com	52,456,550	1,873,448	70,341	2,606,333	1,350,165
	huggingface.co	37,579,886	1,342,139	109,989	2,156,184	697,935
	github.com	1,004,903,868	35,889,424	3,004,579	53,109,313	19,391,186
	stackoverflow.com	315,512,821	11,268,315	1,577,820	20,449,043	2,901,984

5.2.1 chatgpt.com and platform.openai.com

The dominant domain for proprietary AI chatbots is chatgpt.com. Individual users can access ChatGPT on chatgpt.com via desktop or mobile devices and interact with it through a conversational AI chatbot interface. In contrast, advanced users, such as developers, researchers, and firms, access OpenAI's developer platform at platform.openai.com. Users visit platform.openai.com primarily to obtain and manage their API keys and to access specific Gen AI models via OpenAI's APIs, rather than to interact with ChatGPT through a chatbot interface. Accordingly, platform.openai.com is not classified as a proprietary AI chatbot domain.

Given these differences, this subsection applies SDID and event-study SDID separately to chatgpt.com and platform.openai.com. These analyses allow the estimated effects of the release on the proprietary AI chatbot group to be compared with the corresponding platform-specific effects on each site.

In addition, results for the other individual proprietary AI chatbots (i.e., claude.ai, copilot.microsoft.com, and gemini.google.com) are provided in the Appendix. However, the results for all three chatbot domains show evidence of differential pre-trends, so these estimates should be interpreted cautiously.

5.2.1.1 SDID results for chatgpt.com and platform.openai.com

The main SDID results for proprietary AI chatbots are shown in columns (1) and (2) of Table 8. The corresponding results for chatgpt.com are shown in columns (3) and (4), while those for platform.openai.com are shown in columns (5) and (6).

The release increases visits and unique visitors to chatgpt.com by approximately 21.531% and 18.294%, respectively, with both estimates statistically significant at the 1% level. The magnitudes of the effects of the release on both outcomes for chatgpt.com are larger than those for the overall

proprietary AI chatbot group. Additionally, there is no clear evidence of differential pre-trends (Table 9). In contrast, the SDID estimates for platform.openai.com are statistically insignificant for both outcomes.

Table 8 SDID results for chatgpt.com and platform.openai.com

Model	(1) SDID	(2) SDID	(3) SDID	(4) SDID	(5) SDID	(6) SDID
Treated group	Proprietary AI chatbots		chatgpt.com		platform.openai.com	
Dependent variable	Log visits	Log unique visitors	Log visits	Log unique visitors	Log visits	Log unique visitors
ATT	0.081**	0.069**	0.195***	0.168***	0.102	0.093
	(0.033)	(0.03131)	(0.065)	(0.061)	(0.065)	(0.061)
Domain × device FE	Yes	Yes	Yes	Yes	Yes	Yes
Weekly FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	504	504	420	420	420	420

* Note: Placebo-based standard errors with 500 replications are reported in parentheses in columns. *** p<0.01, **p<0.05, * p<0.1.

5.2.1.2 Event-study SDID results for chatgpt.com and platform.openai.com

Table 9 presents the event-study SDID estimates. The results for chatgpt.com generally mirror those for the overall proprietary AI chatbot group. However, chatgpt.com shows an immediate positive response in web traffic beginning in the release week, with the magnitude of the positive response gradually increasing over time (Figures 5 and 6). In contrast, platform.openai.com shows a positive response in web traffic during the release week, followed by a negative response one week after the release. The estimated effects are close to zero two and three weeks after the release. Subsequently, the positive response steadily increases over time (Figures 7 and 8).

Specifically, focusing on statistically significant estimates, the release increases visits to chatgpt.com by approximately 13.542%, 20.563%, 26.744%, 34.582%, 42.618%, and 48.438% three, four, five, six, seven, and eight weeks after the release, respectively, with the estimates statistically significant at the 10%, 5%, or 1% level. Similarly, the release increases unique visitors to chatgpt.com by approximately 10.407%, 16.766%, 22.630%, 29.175%, 36.889%, and 42.904%

three, four, five, six, seven, and eight weeks after the release, respectively, with the estimates statistically significant at the 10%, 5%, or 1% level.

For platform.openai.com, the release increases visits by approximately 30.735% and 36.479% seven and eight weeks after the release, respectively, and both estimates are statistically significant at the 1% level. Similarly, the release increases unique visitors by approximately 28.917% and 37.026% seven and eight weeks after the release, respectively, and both estimates are statistically significant at the 1% level.

Additionally, the pre-treatment placebo effects on both outcomes for chatgpt.com are statistically insignificant, and the Wald tests fail to reject the joint null hypothesis that all pre-treatment placebo effects are zero. These results provide no evidence of differential pre-trends.

By contrast, the pre-treatment placebo effects for log visits and log unique visitors to platform.openai.com are statistically insignificant, except for the placebo effects five weeks before the release, which are statistically significant at the 5% and 10% levels, respectively. Moreover, the Wald test for log unique visitors rejects the joint null hypothesis for platform.openai.com. Therefore, the event-study SDID results for platform.openai.com should be interpreted cautiously.

Overall, chatgpt.com shows larger and more immediate positive responses in both visits and unique visitors than the overall proprietary AI chatbot group. platform.openai.com also shows immediate positive responses, followed by negative and near-zero responses, and subsequently a gradual increase in the magnitude of the positive responses over time. However, the estimates for platform.openai.com are statistically significant only seven and eight weeks after the release.

Table 9 Event-study SDID results for chatgpt.com and platform.openai.com

Model	(1)	(2)	(3)	(4)	(5)	(6)
Treated group	Proprietary AI chatbots		chatgpt.com		platform.openai.com	
Dependent variable	Log visits	Log unique visitors	Log visits	Log unique visitors	Log visits	Log unique visitors
Placebo effect (t-5)	0.005	0.005	0.029	0.023	0.023	0.028
	(0.016)	(0.015)	(0.025)	(0.023)	(0.025)	(0.023)
Placebo effect (t-4)	0.024*	0.026**	0.008	0.003	0.042**	0.053***
	(0.013)	(0.012)	(0.021)	(0.017)	(0.021)	(0.017)
Placebo effect (t-3)	0.007	0.007	0.005	-0.002	0.023	0.028
	(0.011)	(0.009)	(0.022)	(0.017)	(0.022)	(0.017)
Placebo effect (t-2)	-0.004	-0.003	-0.001	-0.000	-0.006	-0.006
	(0.008)	(0.006)	(0.020)	(0.013)	(0.020)	(0.013)
Placebo effect (t-1)	0.010	0.010	0.008	0.003	0.022	0.022
	(0.013)	(0.013)	(0.019)	(0.023)	(0.019)	(0.023)
ATT (t)	0.010	0.006	0.042	0.027	0.072	0.069
	(0.027)	(0.022)	(0.057)	(0.043)	(0.057)	(0.043)
ATT (t+1)	-0.002	-0.004	0.060	0.041	-0.070	-0.067
	(0.023)	(0.020)	(0.051)	(0.041)	(0.051)	(0.041)
ATT (t+2)	-0.001	-0.017	0.073	0.058	0.027	0.013
	(0.021)	(0.018)	(0.048)	(0.038)	(0.048)	(0.038)
ATT (t+3)	0.050	0.024	0.127*	0.099*	0.002	-0.012
	(0.031)	(0.028)	(0.068)	(0.058)	(0.068)	(0.058)
ATT (t+4)	0.080**	0.053	0.187**	0.155**	0.081	0.080
	(0.036)	(0.034)	(0.076)	(0.068)	(0.076)	(0.068)
ATT (t+5)	0.108***	0.100**	0.237***	0.204***	0.101	0.106
	(0.041)	(0.039)	(0.086)	(0.078)	(0.085)	(0.078)
ATT (t+6)	0.106**	0.095**	0.297***	0.256***	0.136	0.130
	(0.048)	(0.047)	(0.099)	(0.093)	(0.099)	(0.093)
ATT (t+7)	0.165***	0.159***	0.355***	0.314***	0.268***	0.254***
	(0.048)	(0.047)	(0.097)	(0.093)	(0.097)	(0.093)
ATT (t+8)	0.205***	0.197***	0.395***	0.357***	0.311***	0.315***
	(0.049)	(0.049)	(0.099)	(0.096)	(0.098)	(0.096)
Domain × device FE	Yes	Yes	Yes	Yes	Yes	Yes
Weekly FE	Yes	Yes	Yes	Yes	Yes	Yes
Wald test: placebo effects jointly equal zero	0.4254	0.2887	0.8864	0.9591	0.1969	0.0090***
Observations	504	504	420	420	420	420

* Note: Placebo-based standard errors with 500 replications are reported in parentheses in columns. The Wald test presents the p-value for the joint null hypothesis that all pre-treatment placebo effects are equal to zero. *** p<0.01, **p<0.05, * p<0.1.

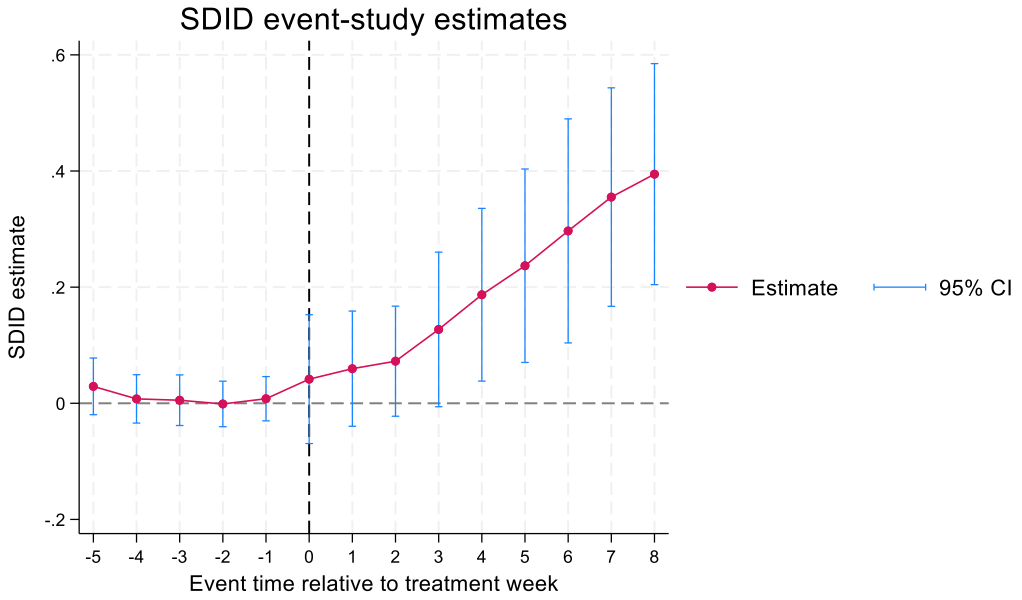


Figure 5 Event-study SDID estimates for log visits to chatgpt.com

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

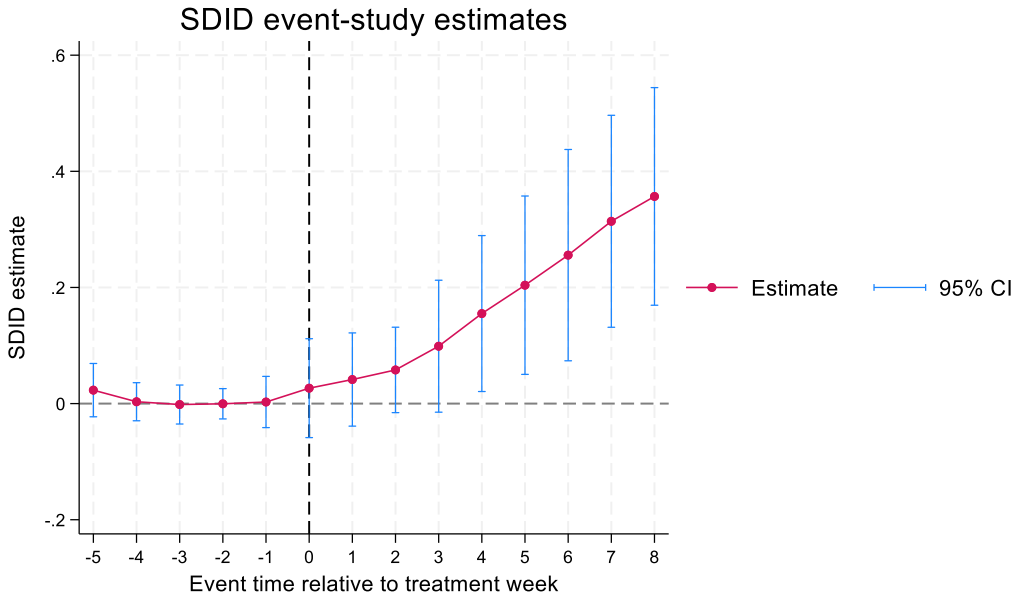


Figure 6 Event-study SDID estimates for log unique visitors to chatgpt.com

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

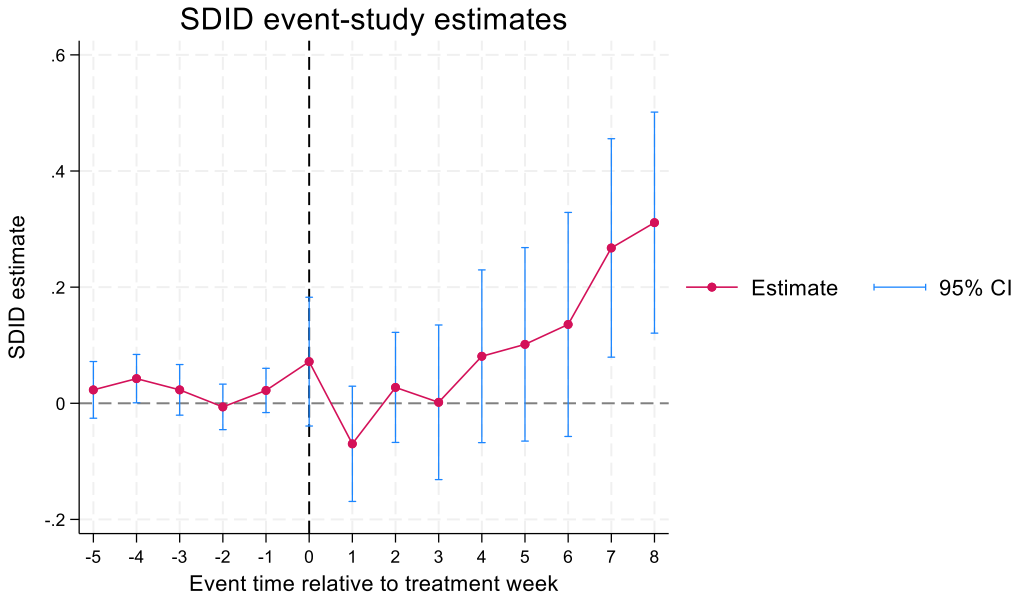


Figure 7 Event-study SDID estimates for log visits to platform.openai.com

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

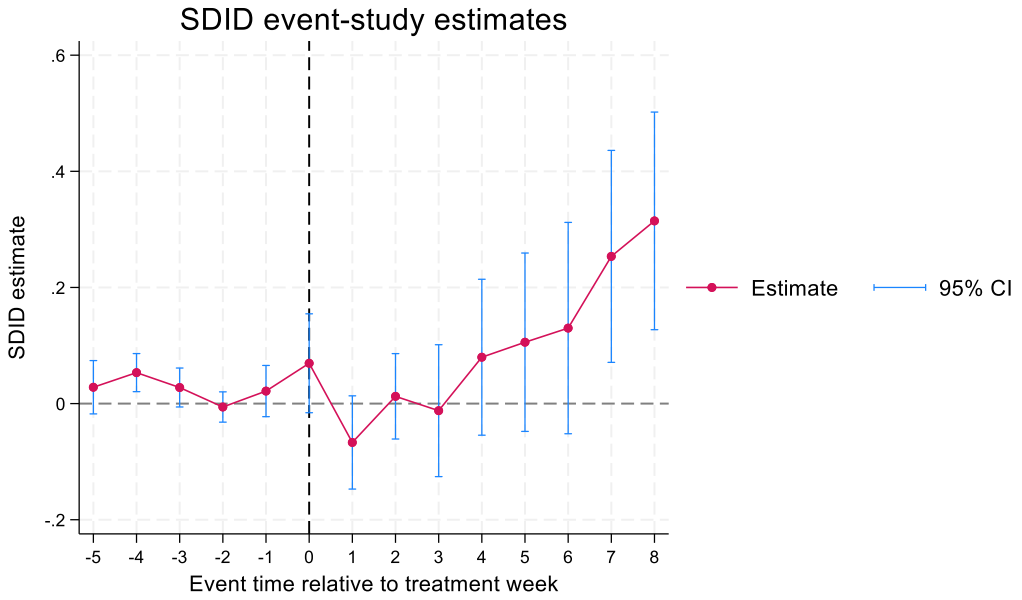


Figure 8 Event-study SDID estimates for log unique visitors to platform.openai.com

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

5.2.2 perplexity.ai and google.com

perplexity.ai is an LLM-based search engine with an AI chatbot interface that searches the web in real time and provides answers with citations and links that let users verify sources (Yang et al. 2025). However, despite sharing some characteristics with AI chatbots, its primary purpose is to serve as a search engine that provides verifiable information. Thus, it is not classified as a proprietary AI chatbot domain.

In addition, perplexity.ai allows users to select among different LLMs, and Llama 3.1 was made available to Perplexity Pro subscribers on July 23, 2024, the same day it was released. Therefore, the release may have directly affected user behavior on perplexity.ai.

Google is the dominant search engine worldwide. Therefore, google.com provides a useful comparison with perplexity.ai for assessing whether web traffic to google.com is affected by the release. However, Google also began providing Gen AI-based responses to search queries in the U.S., called AI Overviews (AIOs), on May 14, 2024 (Google 2024); therefore, google.com can also be broadly regarded as an LLM-based search engine. Relatedly, Agarwal and Sen (2026) examined the impact of Google's AIOs on users' online search behavior and web traffic to downstream websites using a field experiment. Moreover, Padilla et al. (2025) showed that adopting ChatGPT decreases consumer search behavior on google.com. However, to the best of the author's knowledge, no prior study has examined the effect of a frontier open-weight LLM on web traffic to google.com.

5.2.2.1 SDID results for perplexity.ai and google.com

The release increases visits and unique visitors to perplexity.ai by approximately 27.252% and 27.507%, respectively, with both estimates statistically significant at the 1% level (Table 10). The estimated effects of the release for perplexity.ai are approximately three times as large as those for

the overall proprietary AI chatbot group. In contrast, the SDID estimates for google.com are statistically insignificant for both outcomes.

Table 10 SDID results for perplexity.ai and google.com

Model	(1) SDID	(2) SDID	(3) SDID	(4) SDID	(5) SDID	(6) SDID
Treated group	Proprietary AI chatbots		perplexity.ai		google.com	
Dependent variable	Log visits	Log unique visitors	Log visits	Log unique visitors	Log visits	Log unique visitors
ATT	0.081**	0.069**	0.241***	0.243***	0.019	0.019
	(0.033)	(0.03131)	(0.065)	(0.061)	(0.065)	(0.061)
Domain × device FE	Yes	Yes	Yes	Yes	Yes	Yes
Weekly FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	504	504	420	420	420	420

* Note: Placebo-based standard errors with 500 replications are in parentheses. *** p<0.01, **p<0.05, * p<0.1.

5.2.2.2 Event-study SDID results for perplexity.ai and google.com

The event-study SDID estimates for perplexity.ai show larger, more immediate responses than those for the overall proprietary AI chatbot group, and the estimated effects on both outcomes for perplexity.ai gradually increase over time (Table 11; Figures 9 and 10). In contrast, those for google.com show smaller, more delayed responses than those for the overall proprietary AI chatbot group, and none of the estimates are statistically significant (Table 11; Figures 11 and 12).

Specifically, the release increases visits to perplexity.ai by approximately 9.308%, 13.542%, 22.630%, 30.213%, 44.196%, 56.674%, and 62.580% two, three, four, five, six, seven, and eight weeks after the release, respectively, with the estimates statistically significant at either the 10% or 1% level. Similarly, the release increases unique visitors to perplexity.ai by approximately 8.329%, 7.681%, 8.654%, 12.637%, 26.364%, 29.823%, 50.682%, 59.999%, and 60.801% during the release week and one, two, three, four, five, six, seven, and eight weeks after the release, respectively, with the estimates statistically significant at the 10%, 5%, or 1% level.

In addition, the pre-treatment placebo effects for log visits and log unique visitors to perplexity.ai are statistically insignificant, except for the placebo effect for log unique visitors four weeks before

the release, which is statistically significant at the 5% level. However, the Wald tests for both outcomes fail to reject the joint null hypothesis that all pre-treatment placebo effects are zero.

Overall, perplexity.ai, an LLM-based search engine that provides verifiable answers, shows larger and more immediate positive responses across both outcomes than the overall proprietary AI chatbot group. In contrast, all estimates for google.com are statistically insignificant.

Table 11 Event-study SDID results for perplexity.ai and google.com

Model	(1)	(2)	(3)	(4)	(5)	(6)
Treated group	Proprietary AI chatbots		perplexity.ai		google.com	
Dependent variable	Log visits	Log unique visitors	Log visits	Log unique visitors	Log visits	Log unique visitors
Placebo effect (t-5)	0.005	0.005	-0.013	0.016	0.005	0.005
	(0.016)	(0.015)	(0.025)	(0.023)	(0.025)	(0.023)
Placebo effect (t-4)	0.024*	0.026**	0.020	0.039**	0.008	0.006
	(0.013)	(0.012)	(0.021)	(0.017)	(0.021)	(0.017)
Placebo effect (t-3)	0.007	0.007	0.002	0.017	0.005	0.004
	(0.011)	(0.009)	(0.022)	(0.017)	(0.022)	(0.017)
Placebo effect (t-2)	-0.004	-0.003	-0.003	-0.004	-0.001	-0.001
	(0.008)	(0.006)	(0.020)	(0.013)	(0.020)	(0.013)
Placebo effect (t-1)	0.010	0.010	0.012	0.032	-0.002	-0.001
	(0.013)	(0.013)	(0.019)	(0.023)	(0.019)	(0.023)
ATT (t)	0.010	0.006	0.083	0.080*	0.010	0.010
	(0.027)	(0.022)	(0.057)	(0.043)	0.057	(0.043)
ATT (t+1)	-0.002	-0.004	0.068	0.074*	0.010	0.013
	(0.023)	(0.020)	(0.051)	(0.041)	(0.051)	(0.041)
ATT (t+2)	-0.001	-0.017	0.089*	0.083**	-0.002	0.003
	(0.021)	(0.018)	(0.049)	(0.038)	(0.048)	(0.038)
ATT (t+3)	0.050	0.024	0.127*	0.119**	0.001	0.006
	(0.031)	(0.028)	(0.068)	(0.058)	(0.068)	(0.058)
ATT (t+4)	0.080**	0.053	0.204***	0.234***	0.012	0.012
	(0.036)	(0.034)	(0.076)	(0.068)	(0.076)	(0.068)
ATT (t+5)	0.108***	0.100**	0.264***	0.261***	0.030	0.027
	(0.041)	(0.039)	(0.086)	(0.078)	(0.085)	(0.078)
ATT (t+6)	0.106**	0.095**	0.366***	0.410***	0.035	0.034
	(0.048)	(0.047)	(0.100)	(0.093)	(0.098)	(0.093)
ATT (t+7)	0.165***	0.159***	0.449***	0.470***	0.038	0.037
	(0.048)	(0.047)	(0.098)	(0.093)	(0.096)	(0.093)
ATT (t+8)	0.205***	0.197***	0.486***	0.475***	0.039	0.039
	(0.049)	(0.049)	(0.099)	(0.096)	(0.097)	(0.096)
Domain × device FE	Yes	Yes	Yes	Yes	Yes	Yes
Weekly FE	Yes	Yes	Yes	Yes	Yes	Yes
Wald test: placebo effects jointly equal zero	0.4254	0.2887	0.9091	0.1147	0.9985	0.9989
Observations	504	504	420	420	420	420

* Note: Placebo-based standard errors with 500 replications are in parentheses. The Wald test presents the p-value for the joint null hypothesis that all pre-treatment placebo effects are equal to zero. *** p<0.01, **p<0.05, * p<0.1.

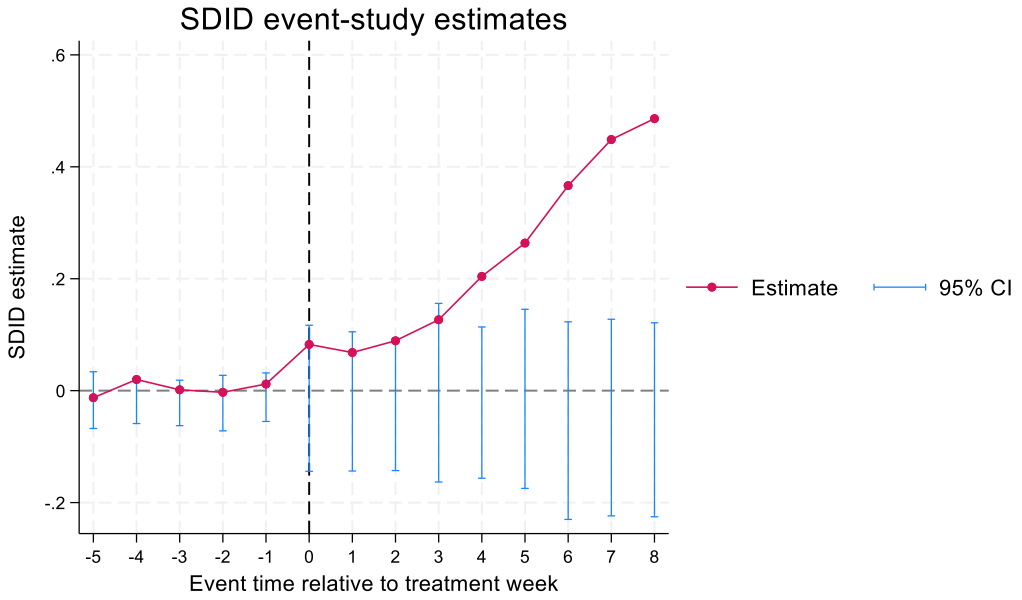


Figure 9 Event-study SDID estimates for log visits to perplexity.com

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

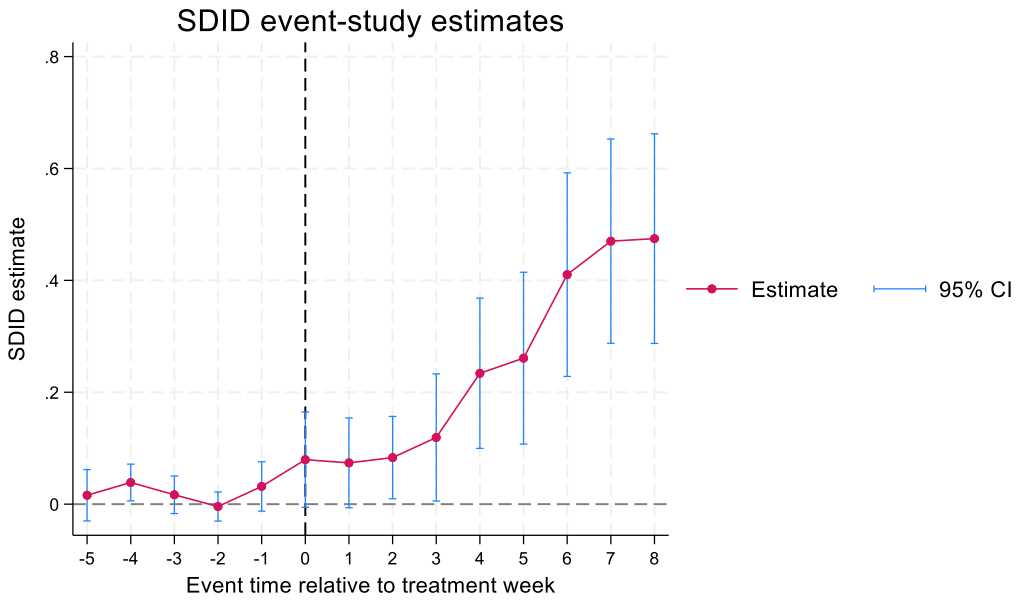


Figure 10 Event-study SDID estimates for log unique visitors to perplexity.com

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

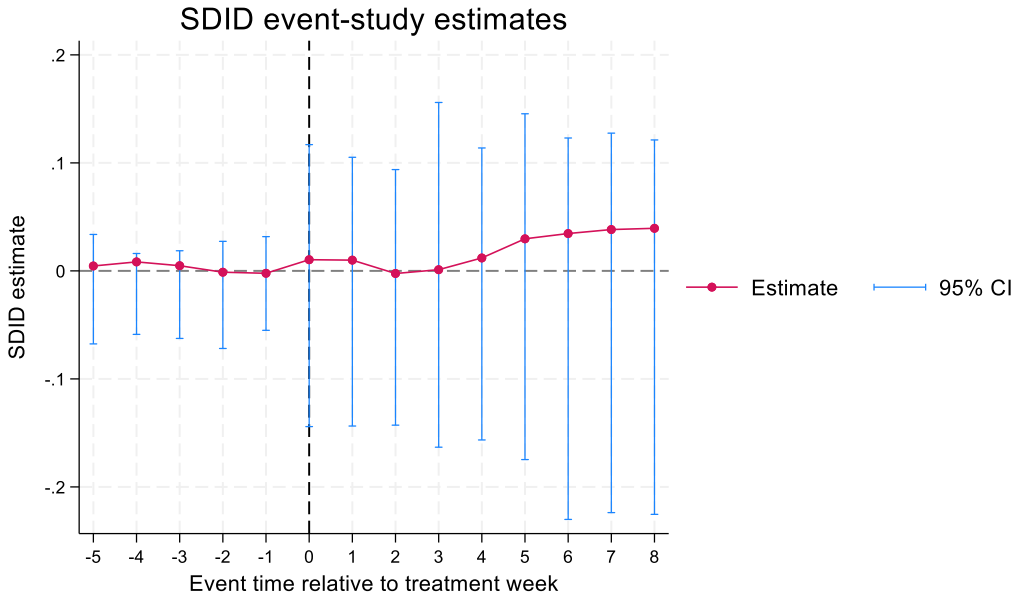


Figure 11 Event-study SDID estimates for log visits to google.com

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

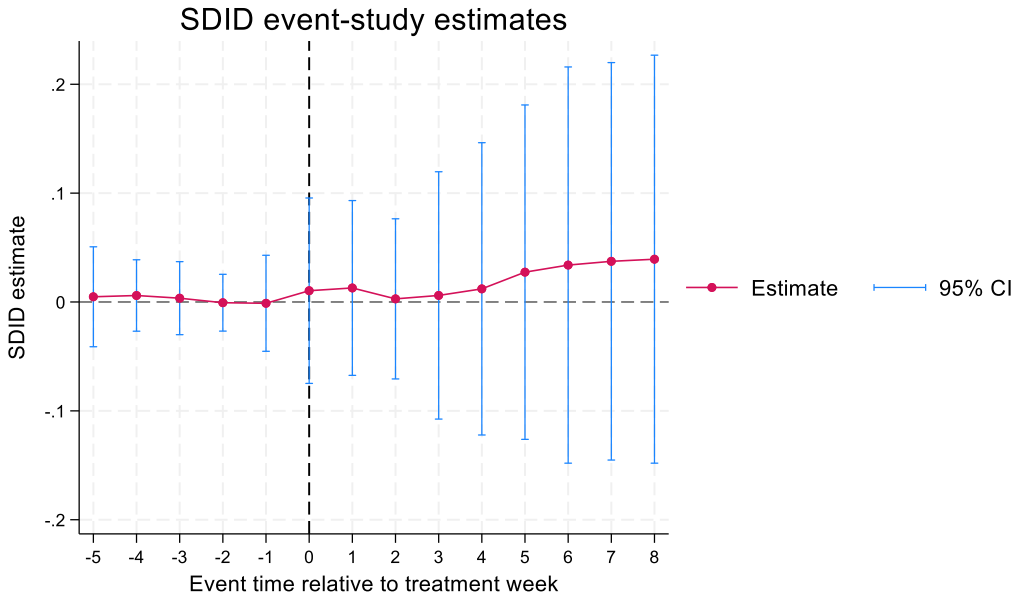


Figure 12 Event-study SDID estimates for log unique visitors to google.com

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

5.2.3 llama.meta.com and meta.com

llama.meta.com is not a proprietary AI chatbot interface but rather Meta's AI development platform. Importantly, Llama 3.1 was released on llama.meta.com and huggingface.co on July 23, 2024. Accordingly, the release of Llama 3.1 may have immediately affected user behavior on llama.meta.com, as users could download the model from either llama.meta.com or huggingface.co on the release date. Therefore, web traffic to llama.meta.com provides a useful measure of how the release of Meta's frontier open-weight model affects global online user attention to Meta's AI development platform.

In addition, meta.com is Meta's corporate website. Therefore, changes in web traffic to meta.com before and after the release of Llama 3.1 may provide evidence of how the release of a large technology firm's frontier open-weight LLM affects global web traffic to the firm's corporate website. Such changes may indicate whether the release attracts greater global online user attention to the firm and its brand.

5.2.3.1 SDID results for llama.meta.com and meta.com

As shown in Table 12, the effects of the Llama 3.1 release on both outcomes for llama.meta.com are statistically significant at the 1% level. However, the event-study SDID results in Table 13 indicate differential pre-trends between the treated and synthetic control groups. Therefore, the estimated ATTs from both the SDID and event-study SDID analyses should be interpreted cautiously as causal effects. In addition, the estimated effects of the Llama 3.1 release on both outcomes for meta.com are statistically insignificant. In contrast, the pre-treatment placebo effects in the event-study SDID analysis are statistically insignificant.

Table 12 SDID results for llama.meta.com and meta.com

Model	(1) SDID	(2) SDID	(3) SDID	(4) SDID	(5) SDID	(6) SDID
Treated group	Proprietary AI chatbots		llama.meta.com		meta.com	
Dependent variable	Log visits	Log unique visitors	Log visits	Log unique visitors	Log visits	Log unique visitors
ATT	0.081** (0.033)	0.069** (0.03131)	0.538*** (0.065)	0.519*** (0.061)	0.048 (0.065)	0.052 (0.061)
Domain × device FE	Yes	Yes	Yes	Yes	Yes	Yes
Weekly FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	504	504	420	420	420	420

* Note: Placebo-based standard errors with 500 replications are reported in parentheses in columns. *** p<0.01, **p<0.05, * p<0.1.

5.2.3.2 Event-study SDID results for llama.meta.com and meta.com

During the release week, web traffic to llama.meta.com and meta.com increases immediately. Subsequently, positive responses in web traffic to both domains gradually decline over time and turn negative eight weeks and four weeks after the release, respectively (Figures 13−16).

However, the statistical significance of the pre-treatment placebo effects and the Wald test results for llama.meta.com indicate differential pre-trends. Therefore, the estimated ATTs from the SDID and event-study SDID analyses should not be interpreted as causal effects (Table 13). In contrast, the event-study SDID results for meta.com show no evidence of differential pre-trends, as all pre-treatment placebo effects are statistically insignificant and the Wald tests fail to reject the joint null hypothesis that all pre-treatment placebo effects are equal to zero (Table 13).

Given these results, only the statistically significant estimates for meta.com are interpreted. Specifically, the release increases visits to meta.com by approximately 37.438% during the release week and by 26.744% one week after the release; both estimates are statistically significant at the 1% level. Likewise, the release increases unique visitors to meta.com by approximately 40.074%, 27.379%, and 6.503% during the release week and one and two weeks after the release, respectively, and the estimates are statistically significant at either the 1% or 10% level.

Overall, meta.com shows an immediate increase in both log visits and log unique visitors during the release week, though the positive effects gradually decline over time. Hence, the release of a large technology firm's frontier open-weight LLM may increase web traffic to its corporate website in the short term, with the positive effects gradually diminishing over time (Table 13).

Table 13 Event-study SDID results for llama.meta.com and meta.com

Model	(1)	(2)	(3)	(4)	(5)	(6)
Treated group	Proprietary AI chatbots		llama.meta.com		meta.com	
Dependent variable	Log visits	Log unique visitors	Log visits	Log unique visitors	Log visits	Log unique visitors
Placebo effect (t-5)	0.005	0.005	0.016	-0.013	-0.001	-0.010
	(0.016)	(0.015)	(0.025)	(0.023)	(0.025)	(0.023)
Placebo effect (t-4)	0.024*	0.026**	-0.030	-0.058***	-0.009	-0.002
	(0.013)	(0.012)	(0.021)	(0.017)	(0.021)	(0.017)
Placebo effect (t-3)	0.007	0.007	-0.196***	-0.169***	0.003	0.010
	(0.011)	(0.009)	(0.024)	(0.017)	(0.022)	(0.017)
Placebo effect (t-2)	-0.004	-0.003	0.004	0.006	0.001	0.000
	(0.008)	(0.006)	(0.020)	(0.013)	(0.020)	(0.013)
Placebo effect (t-1)	0.010	0.010	-0.153***	-0.162***	-0.019	-0.019
	(0.013)	(0.013)	(0.021)	(0.023)	(0.019)	(0.023)
ATT (t)	0.010	0.006	1.817***	1.812***	0.318***	0.337***
	(0.027)	(0.022)	(0.099)	(0.043)	(0.058)	(0.043)
ATT (t+1)	-0.002	-0.004	1.168***	1.125***	0.237***	0.242***
	(0.023)	(0.020)	(0.073)	(0.041)	(0.052)	(0.041)
ATT (t+2)	-0.001	-0.017	1.016***	0.964***	0.063	0.063*
	(0.021)	(0.018)	(0.067)	(0.038)	(0.048)	(0.038)
ATT (t+3)	0.050	0.024	0.830***	0.764***	0.060	0.058
	(0.031)	(0.028)	(0.078)	(0.058)	(0.068)	(0.058)
ATT (t+4)	0.080**	0.053	0.784***	0.731***	-0.010	-0.017
	(0.036)	(0.034)	(0.084)	(0.068)	(0.076)	(0.068)
ATT (t+5)	0.108***	0.100**	0.565***	0.552***	0.000	-0.008
	(0.041)	(0.039)	(0.089)	(0.078)	(0.085)	(0.078)
ATT (t+6)	0.106**	0.095**	0.356***	0.383***	-0.019	-0.031
	(0.048)	(0.047)	(0.100)	(0.093)	(0.098)	(0.093)
ATT (t+7)	0.165***	0.159***	0.014	0.051	-0.067	-0.053
	(0.048)	(0.047)	(0.096)	(0.093)	(0.096)	(0.093)
ATT (t+8)	0.205***	0.197***	-1.605***	-1.566***	-0.102	-0.091
	(0.049)	(0.049)	(0.120)	(0.096)	(0.097)	(0.096)
Domain × device FE	Yes	Yes	Yes	Yes	Yes	Yes
Weekly FE	Yes	Yes	Yes	Yes	Yes	Yes
Wald test: placebo effects jointly equal zero	0.4254	0.2887	0.0000	0.0000	0.9488	0.9427
Observations	504	504	420	420	420	420

* Note: Placebo-based standard errors with 500 replications are reported in parentheses in columns. The Wald test presents the p-value for the joint null hypothesis that all pre-treatment placebo effects are equal to zero. *** p<0.01, **p<0.05, * p<0.1.

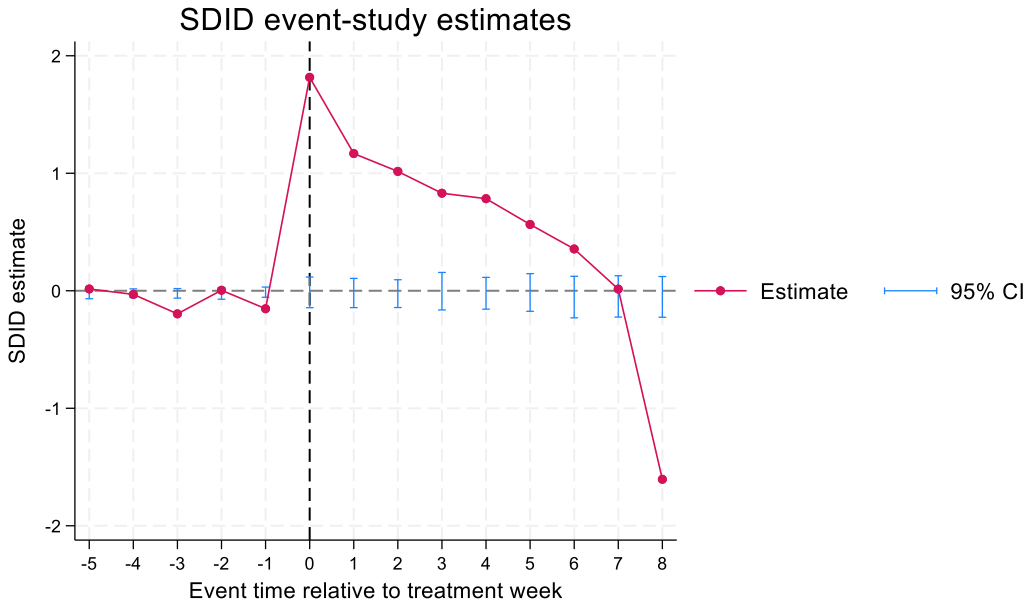


Figure 13 Time weights for log visits to llama.meta.com

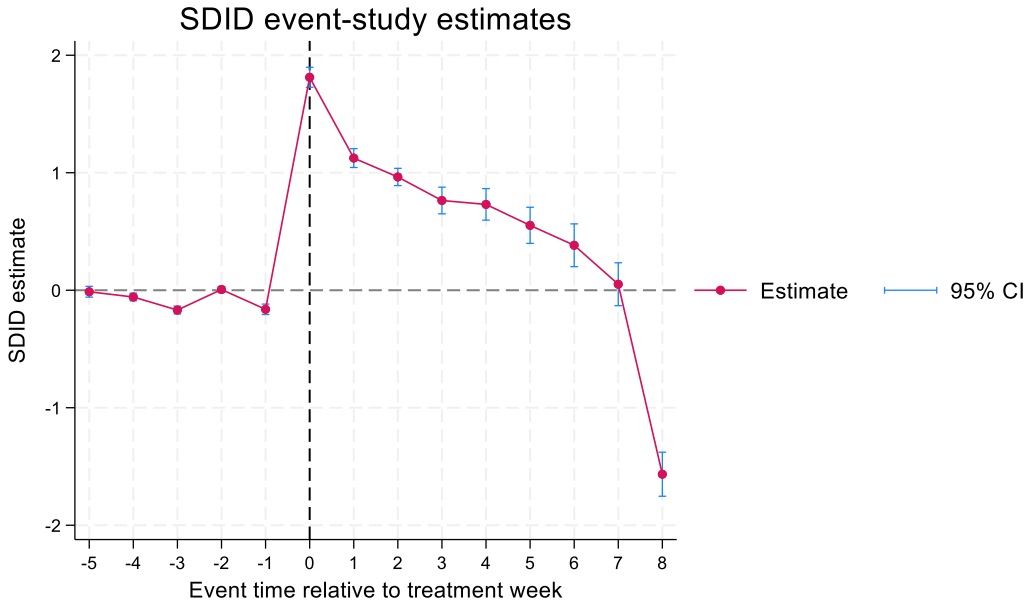


Figure 14 Time weights for log unique visitors to llama.meta.com

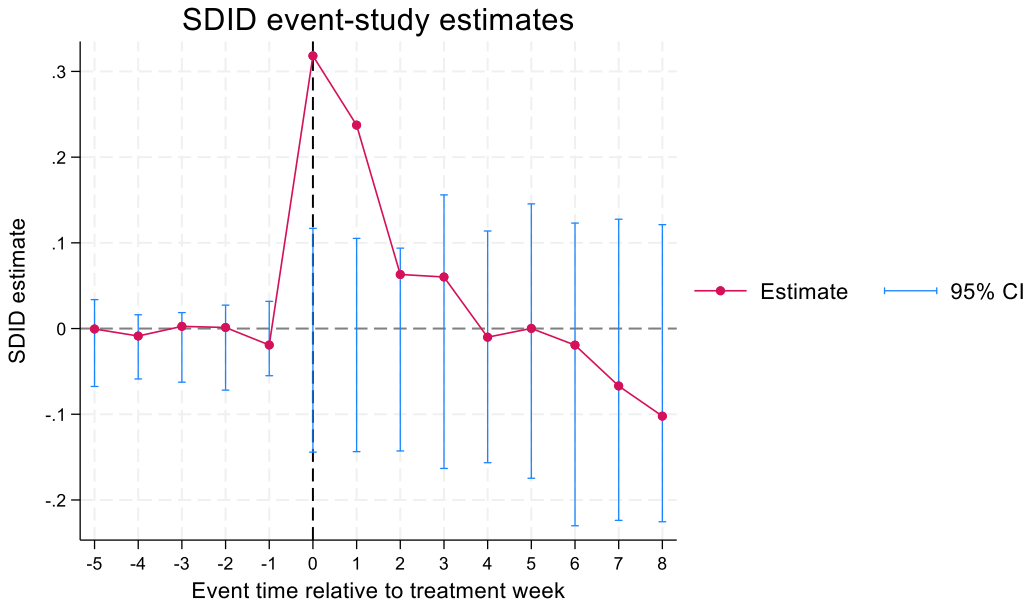


Figure 15 Time weights for log visits to meta.com

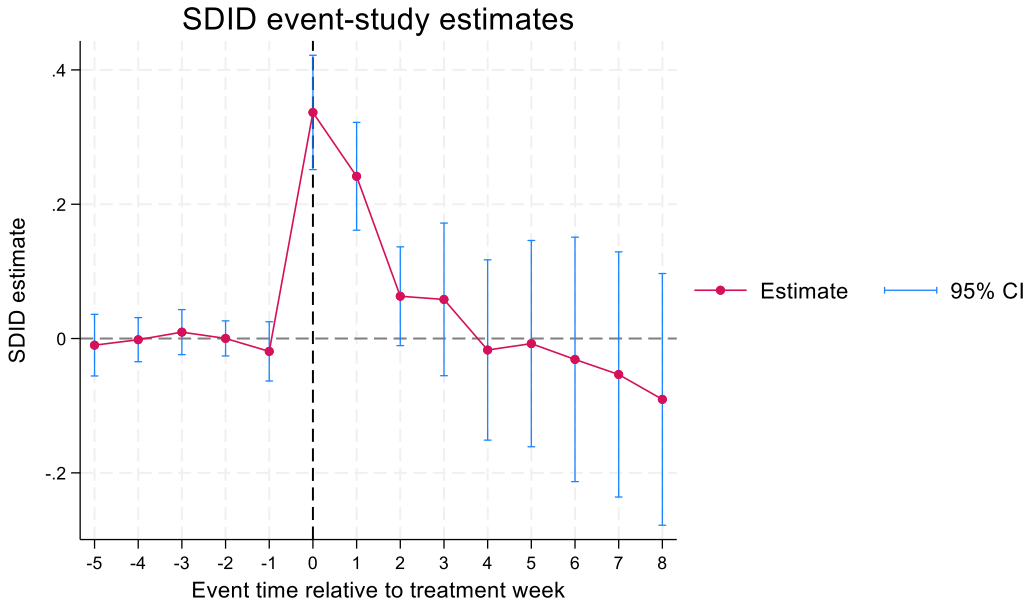


Figure 16 Time weights for log unique visitors to meta.com

5.2.4 huggingface.co and github.com

huggingface.co is a platform for open-source AI models and datasets that allows users to download, use, and upload open-weight or closed-weight AI models, as well as datasets. Therefore, users of huggingface.co may have more experience with LLMs than the general public. In addition, Llama 3.1 was made available on huggingface.co on its release date. Accordingly, the release of Llama 3.1, an open-weight LLM, may have directly affected user behavior on huggingface.co.

github.com is a proprietary development platform that enables users to store, manage, share, and collaboratively develop code. It is also widely used for open-source software development. Web traffic to github.com may reflect online attention to software development, while web traffic to huggingface.co may reflect online attention to open-source AI models and datasets.

5.2.4.1 SDID results for huggingface.co and github.com

The Llama 3.1 release increases visits and unique visitors to huggingface.co by approximately 11.963% and 13.769%, respectively, with the estimates statistically significant at the 10% and 5% levels, respectively (Table 14). The estimated effects of the release on both outcomes for huggingface.co are larger than those for the overall proprietary AI chatbot group. Additionally, the event-study SDID results show no clear evidence of differential pre-trends (Table 15). In contrast, the SDID estimates for github.com are statistically insignificant for both outcomes.

Table 14 SDID results for huggingface.co and github.com

Model	(1) SDID	(2) SDID	(3) SDID	(4) SDID	(5) SDID	(6) SDID
Treated group	Proprietary AI chatbots		huggingface.co		github.com	
Dependent variable	Log visits	Log unique visitors	Log visits	Log unique visitors	Log visits	Log unique visitors
ATT	0.081**	0.069**	0.113*	0.129**	0.045	0.045
	(0.033)	(0.03131)	(0.065)	(0.061)	(0.065)	(0.061)
Domain × device FE	Yes	Yes	Yes	Yes	Yes	Yes
Weekly FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	504	504	420	420	420	420

* Note: Placebo-based standard errors with 500 replications are in parentheses. *** p<0.01, **p<0.05, * p<0.1.

5.2.4.2 Event-study SDID results for huggingface.co and github.com

The release immediately increases web traffic to huggingface.co and github.com during the release week, and positive responses generally increase thereafter. However, the positive effects decrease in magnitude at five and four weeks after the release, respectively, and then gradually increase again (Figures 17-20).

Among the statistically significant estimates in Table 15, the event-study SDID estimates for huggingface.co show larger and more immediate positive responses than those for the overall proprietary AI chatbot group. Specifically, the release increases visits to huggingface.co by approximately 9.090%, 14.225%, 14.568%, 20.322%, and 20.804% two, three, four, seven, and eight weeks after the release, respectively. These estimates are statistically significant at the 10% level. Similarly, the release increases unique visitors to huggingface.co by approximately 9.636%, 10.186%, 15.258%, 14.683%, 22.385%, and 24.982% during the release week and two, three, four, seven, and eight weeks after the release, respectively. These estimates are statistically significant at the 5% level.

Additionally, the pre-treatment placebo effects for both outcomes on huggingface.co are statistically insignificant, and the Wald tests for both outcomes fail to reject the joint null hypothesis that all pre-treatment placebo effects are zero. Therefore, there is no evidence of differential pre-trends.

Overall, huggingface.co shows larger and more immediate positive responses across both outcomes than the overall proprietary AI chatbot group. Changes in web traffic to huggingface.co provide useful insights into how the release of a frontier open-weight LLM affects user behavior on an open-source AI platform. In this regard, the positive estimated effects for huggingface.co are consistent with the interpretation that the Llama 3.1 release may have increased overall attention

to open-source AI models, as huggingface.co primarily serves as a hub for open-source AI models. In contrast, all estimates for github.com are statistically insignificant.

Table 15 Event-study SDID results for huggingface.co and github.com

Model	(1)	(2)	(3)	(4)	(5)	(6)
Treated group	Proprietary AI chatbots		huggingface.co		github.com	
Dependent variable	Log visits	Log unique visitors	Log visits	Log unique visitors	Log visits	Log unique visitors
Placebo effect (t-5)	0.005	0.005	0.016	0.029	0.002	-0.004
	(0.016)	(0.015)	(0.025)	(0.023)	(0.025)	(0.023)
Placebo effect (t-4)	0.024*	0.026**	0.002	-0.001	-0.006	-0.010
	(0.013)	(0.012)	(0.021)	(0.017)	(0.021)	(0.017)
Placebo effect (t-3)	0.007	0.007	0.001	0.002	-0.004	-0.012
	(0.011)	(0.009)	(0.022)	(0.017)	(0.022)	(0.017)
Placebo effect (t-2)	-0.004	-0.003	-0.000	0.000	0.001	0.001
	(0.008)	(0.006)	(0.020)	(0.013)	(0.020)	(0.013)
Placebo effect (t-1)	0.010	0.010	0.007	-0.005	-0.010	-0.012
	(0.013)	(0.013)	(0.019)	(0.023)	(0.019)	(0.023)
ATT (t)	0.010	0.006	0.079	0.092**	0.019	0.021
	(0.027)	(0.022)	(0.057)	(0.043)	(0.057)	(0.043)
ATT (t+1)	-0.002	-0.004	0.029	0.043	0.037	0.044
	(0.023)	(0.020)	(0.051)	(0.041)	(0.051)	(0.041)
ATT (t+2)	-0.001	-0.017	0.087*	0.097**	0.048	0.053
	(0.021)	(0.018)	(0.049)	(0.038)	(0.048)	(0.038)
ATT (t+3)	0.050	0.024	0.133*	0.142**	0.035	0.036
	(0.031)	(0.028)	(0.068)	(0.058)	(0.068)	(0.058)
ATT (t+4)	0.080**	0.053	0.136*	0.137**	0.027	0.013
	(0.036)	(0.034)	(0.076)	(0.068)	(0.076)	(0.068)
ATT (t+5)	0.108***	0.100**	0.083	0.114	0.039	0.035
	(0.041)	(0.039)	(0.085)	(0.078)	(0.085)	(0.078)
ATT (t+6)	0.106**	0.095**	0.117	0.138	0.047	0.038
	(0.048)	(0.047)	(0.098)	(0.093)	(0.098)	(0.093)
ATT (t+7)	0.165***	0.159***	0.185*	0.202**	0.070	0.069
	(0.048)	(0.047)	(0.096)	(0.093)	(0.096)	(0.093)
ATT (t+8)	0.205***	0.197***	0.189*	0.223**	0.094	0.096
	(0.049)	(0.049)	(0.097)	(0.096)	(0.097)	(0.096)
Domain × device FE	Yes	Yes	Yes	Yes	Yes	Yes
Weekly FE	Yes	Yes	Yes	Yes	Yes	Yes
Wald test: placebo effects jointly equal zero	0.4254	0.2887	0.9894	0.9011	0.9956	0.9472
Observations	504	504	420	420	420	420

* Note: Placebo-based standard errors with 500 replications are in parentheses. The Wald test presents the p-value for the joint null hypothesis that all pre-treatment placebo effects are equal to zero. *** p<0.01, **p<0.05, * p<0.1.

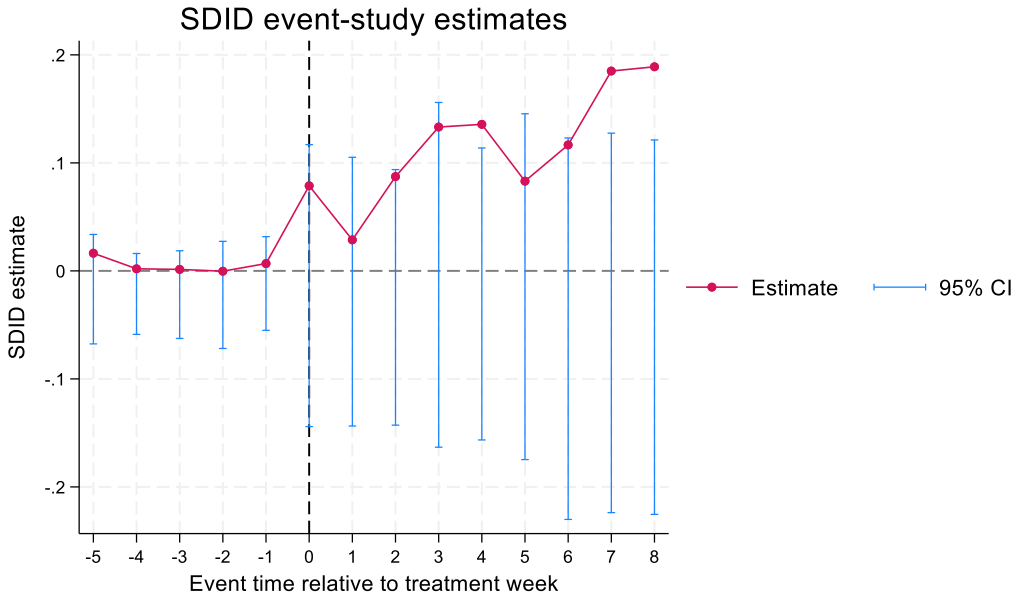


Figure 17 Event-study SDID estimates for log visits to huggingface.co

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

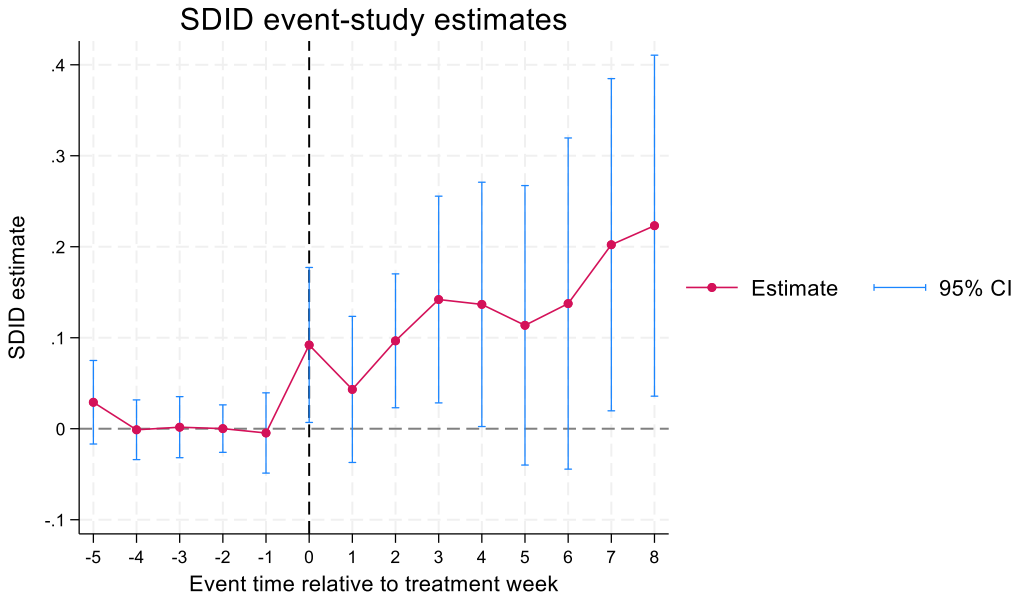


Figure 18 Event-study SDID estimates for log unique visitors to huggingface.co

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

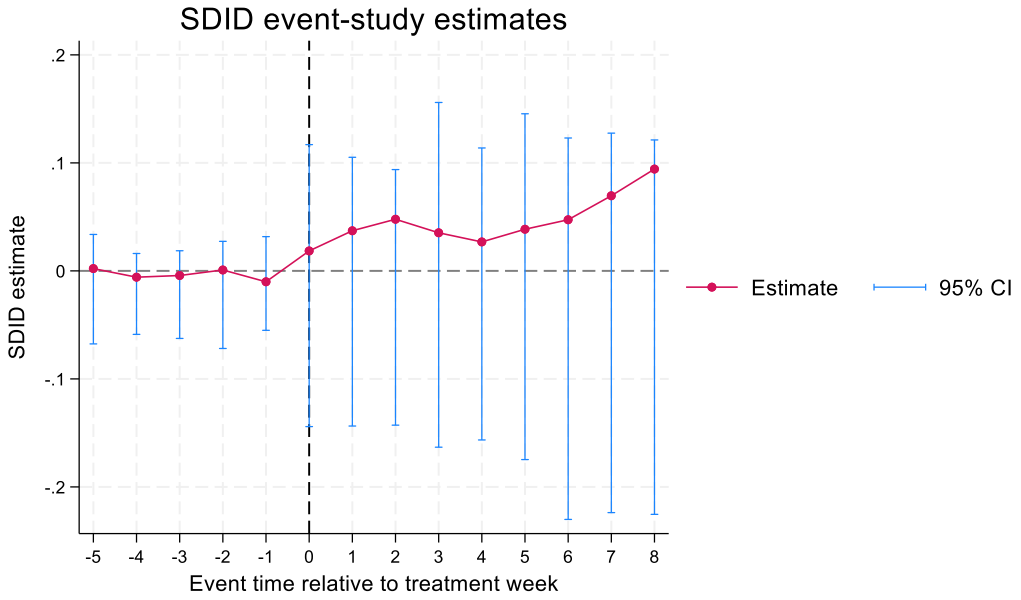


Figure 19 Event-study SDID estimates for log visits to github.com

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

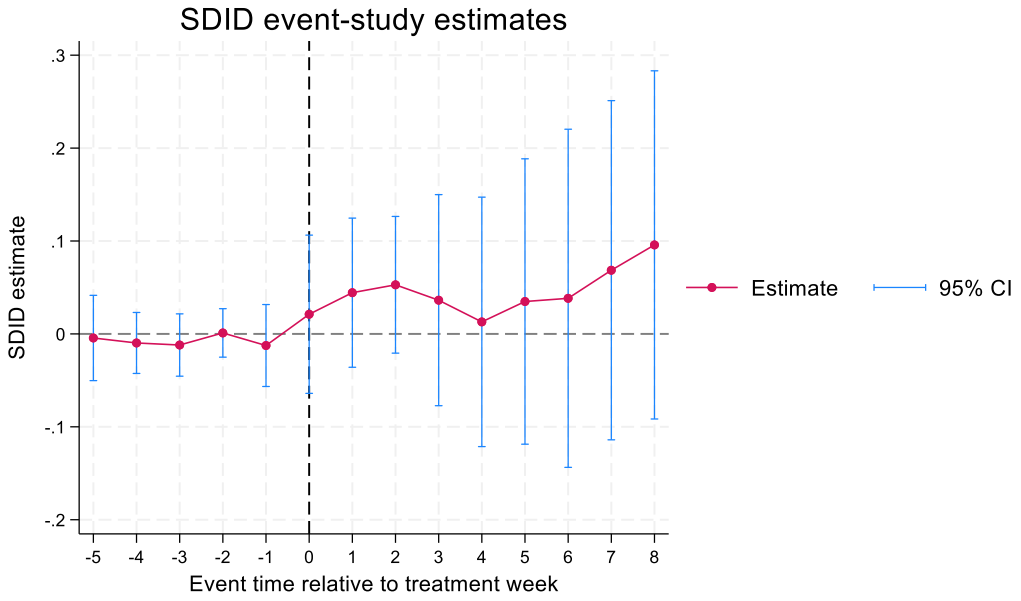


Figure 20 Event-study SDID estimates for log unique visitors to github.com

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

5.2.5 stackoverflow.com

5.2.5.1 SDID results for stackoverflow.com

stackoverflow.com is a knowledge-sharing platform where users ask and answer questions about computer programming. Recent studies have shown that proprietary closed-weight AI chatbots reduce web traffic to stackoverflow.com. For example, Burtch et al. (2024) showed that ChatGPT's release reduced web traffic to stackoverflow.com. Padilla et al. (2025) also provided evidence that ChatGPT's release negatively affected user visits to stackoverflow.com. However, to the best of the author's knowledge, no prior research has examined how the release of a frontier open-weight LLM affects web traffic to stackoverflow.com.

The SDID estimates of the effects of the release of Llama 3.1 on stackoverflow.com are directionally consistent with these prior studies, as the estimated effects are negative for both log visits and log unique visitors to stackoverflow.com. However, these estimates are statistically insignificant (Table 16). In addition, there is no evidence of differential pre-trends (Table 17).

Table 16 SDID results for stackoverflow.com

Model	(1) SDID	(2) SDID	(3) SDID	(4) SDID
Treated group	Proprietary AI chatbots		stackoverflow.com	
Dependent variable	Log visits	Log unique visitors	Log visits	Log unique visitors
ATT	0.081**	0.069**	-0.045	-0.037
	(0.033)	(0.03131)	(0.065)	(0.061)
Domain × device FE	Yes	Yes	Yes	Yes
Weekly FE	Yes	Yes	Yes	Yes
Observations	504	504	420	420

* Note: Placebo-based standard errors with 500 replications are in parentheses. *** p<0.01, **p<0.05, * p<0.1.

5.2.5.2 Event-study SDID results for stackoverflow.com

The release immediately reduces web traffic to stackoverflow.com, and the negative response gradually increases in magnitude for three weeks after the release. Thereafter, the negative effects gradually diminish over time (Figures 21 and 22).

However, the event-study SDID estimates are statistically insignificant, except for the ATT estimate for log unique visitors two weeks after the release. Specifically, the release of Llama 3.1 decreases unique visitors by approximately 6.012%, and this estimate is statistically significant at the 10% level.

In addition, the pre-treatment placebo effects for both outcomes are statistically insignificant, and the Wald tests for both outcomes fail to reject the joint null hypothesis that all pre-treatment placebo effects are equal to zero. Therefore, there is no clear evidence of differential pre-trends.

Table 17 Event-study SDID results for stackoverflow.com

Model	(1)	(2)	(3)	(4)
Treated group	Proprietary AI chatbots		stackoverflow.com	
Dependent variable	Log visits	Log unique visitors	Log visits	Log unique visitors
Placebo effect (t-5)	0.005	0.005	0.013	0.012
	(0.016)	(0.015)	(0.025)	(0.023)
Placebo effect (t-4)	0.024*	0.026**	-0.006	-0.001
	(0.013)	(0.012)	(0.021)	(0.017)
Placebo effect (t-3)	0.007	0.007	-0.015	-0.016
	(0.011)	(0.009)	(0.022)	(0.017)
Placebo effect (t-2)	-0.004	-0.003	0.001	0.000
	(0.008)	(0.006)	(0.020)	(0.013)
Placebo effect (t-1)	0.010	0.010	-0.007	-0.005
	(0.013)	(0.013)	(0.019)	(0.023)
ATT (t)	0.010	0.006	-0.019	-0.018
	(0.027)	(0.022)	(0.057)	(0.043)
ATT (t+1)	-0.002	-0.004	-0.047	-0.045
	(0.023)	(0.020)	(0.051)	(0.041)
ATT (t+2)	-0.001	-0.017	-0.064	-0.062*
	(0.021)	(0.018)	(0.048)	(0.038)
ATT (t+3)	0.050	0.024	-0.093	-0.081
	(0.031)	(0.028)	(0.068)	(0.058)
ATT (t+4)	0.080**	0.053	-0.067	-0.062
	(0.036)	(0.034)	(0.076)	(0.068)
ATT (t+5)	0.108***	0.100**	-0.066	-0.061
	(0.041)	(0.039)	(0.085)	(0.078)
ATT (t+6)	0.106**	0.095**	-0.032	-0.022
	(0.048)	(0.047)	(0.098)	(0.093)
ATT (t+7)	0.165***	0.159***	-0.011	0.006
	(0.048)	(0.047)	(0.096)	(0.093)
ATT (t+8)	0.205***	0.197***	-0.017	0.005
	(0.049)	(0.049)	(0.097)	(0.096)
Domain × device FE	Yes	Yes	Yes	Yes
Weekly FE	Yes	Yes	Yes	Yes
Wald test: placebo effects jointly equal zero	0.4254	0.2887	0.9675	0.9417
Observations	504	504	420	420

* Note: Placebo-based standard errors with 500 replications are in parentheses. The Wald test presents the p-value for the joint null hypothesis that all pre-treatment placebo effects are equal to zero. *** p<0.01, **p<0.05, * p<0.1.

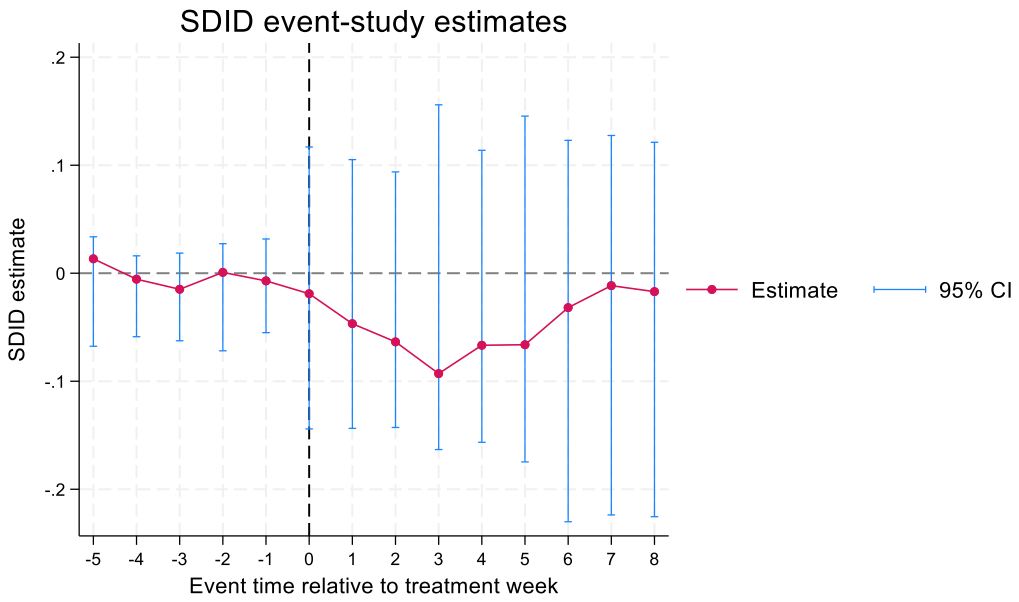


Figure 21 Event-study SDID estimates for log visits to stackoverflow.com

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

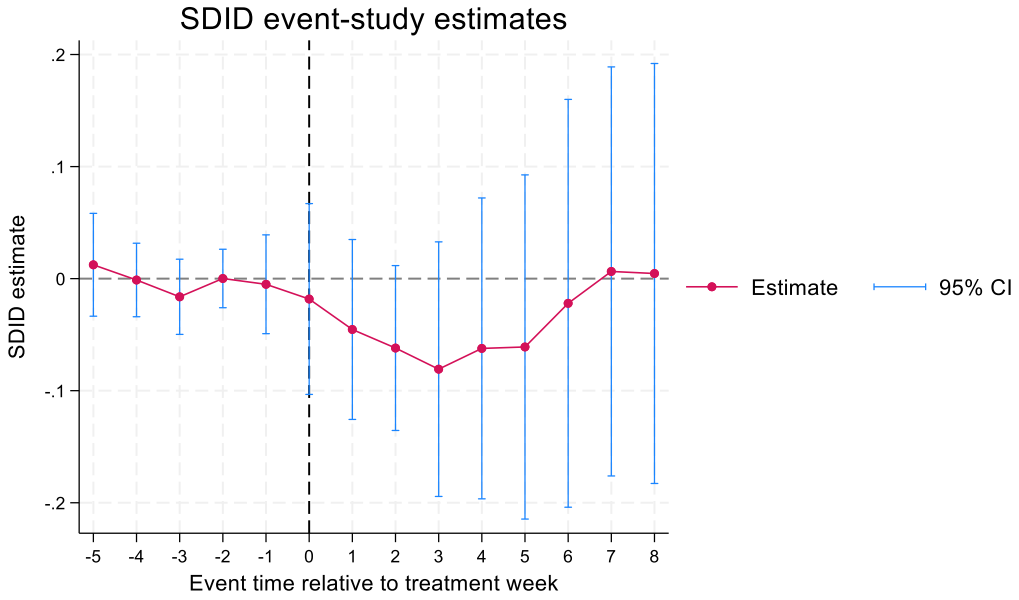


Figure 22 Event-study SDID estimates for log unique visitors to stackoverflow.com

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

5.3 Robustness Analysis

5.3.1 Alternative model: comparison with DiD

This subsection presents empirical results from DiD and event-study DiD for proprietary AI chatbots and compares them with the main SDID and event-study SDID results. Specifically, the DiD estimates of the ATT for the Llama 3.1 release have the same positive sign as the corresponding SDID estimates. However, they are slightly smaller in magnitude for both log visits and log unique visitors. These DiD estimates are statistically insignificant (Table 18), and the empirical evidence shows no differential pre-trends (Table 19).

Table 18 SDID and DiD results for proprietary AI chatbots

Model	(1) SDID	(2) SDID	(3) DiD	(4) DiD
Dependent variable	Log visits	Log unique visitors	Log visits	Log unique visitors
ATT	0.081**	0.069**	0.071	0.059
	(0.033)	(0.03131)	(0.044)	(0.035)
Domain × device FE	Yes	Yes	Yes	Yes
Weekly FE	Yes	Yes	Yes	Yes
Observations	504	504	504	504

* Note: Placebo-based standard errors with 500 replications are reported in parentheses for the SDID estimates. Clustered robust standard errors at the domain level are reported in parentheses for the DID estimates.

5.3.2 Alternative model: comparison with event-study DiD

The event-study DiD estimates are close to zero until two weeks after the release, then gradually rise for the overall proprietary AI chatbot group (Table 19; Figures 23 and 24). Therefore, the event-study DiD estimates indicate that the positive effects of the Llama 3.1 release on log visits and log unique visitors to proprietary AI chatbots gradually increase after a delay, consistent with the SDID and event-study SDID results.

Specifically, the release increases visits to proprietary AI chatbots by approximately 22.507% eight weeks after the release, and the event-study DiD estimate is statistically significant at the 5% level. Likewise, the release increases unique visitors to proprietary AI chatbots by approximately

17.000% and 21.774% seven and eight weeks after the release, respectively. These event-study DiD estimates are statistically significant at the 5% level.

In addition, the DiD analysis shows no evidence of a parallel-trends violation, as all pre-treatment estimates are statistically insignificant and the Wald tests fail to reject the joint null hypothesis that all pre-treatment estimates are \ zero.

Overall, the DiD and event-study DiD results are consistent with the SDID and event-study SDID results. However, the DiD estimates are statistically insignificant, and fewer event-study DiD estimates are statistically significant than their corresponding event-study SDID estimates.

Table 19 Event-study DiD results for proprietary AI chatbots

Model	(1) DiD	(2) DiD
Dependent variable	Log visits	Log unique visitors
Lead 5	0.006	0.006
	(0.115)	(0.097)
Lead 4	0.026	0.030
	(0.053)	(0.047)
Lead 3	0.001	0.001
	(0.019)	(0.014)
Lead 2	-0.011	-0.008
	(0.014)	(0.011)
Lag 0	0.004	0.002
	(0.018)	(0.016)
Lag 1	-0.007	-0.007
	(0.025)	(0.021)
Lag 2	-0.005	-0.018
	(0.038)	(0.030)
Lag 3	0.042	0.017
	(0.046)	(0.033)
Lag 4	0.076	0.051
	(0.083)	(0.062)
Lag 5	0.099	0.092
	(0.086)	(0.065)
Lag 6	0.102	0.093
	(0.092)	(0.070)
Lag 7	0.162	0.157**
	(0.095)	(0.072)
Lag 8	0.203**	0.197**
	(0.095)	(0.074)
Domain × device FE	Yes	Yes
Weekly FE	Yes	Yes
Wald test: pre-treatment effects jointly equal zero	0.1047	0.2985
Observations	504	504

* Note: Clustered robust standard errors at the domain level are reported in parentheses for the DiD estimates. The Wald test presents the p-value for the joint null hypothesis that all pre-treatment effects are equal to zero. *** p<0.01, **p<0.05, * p<0.1.

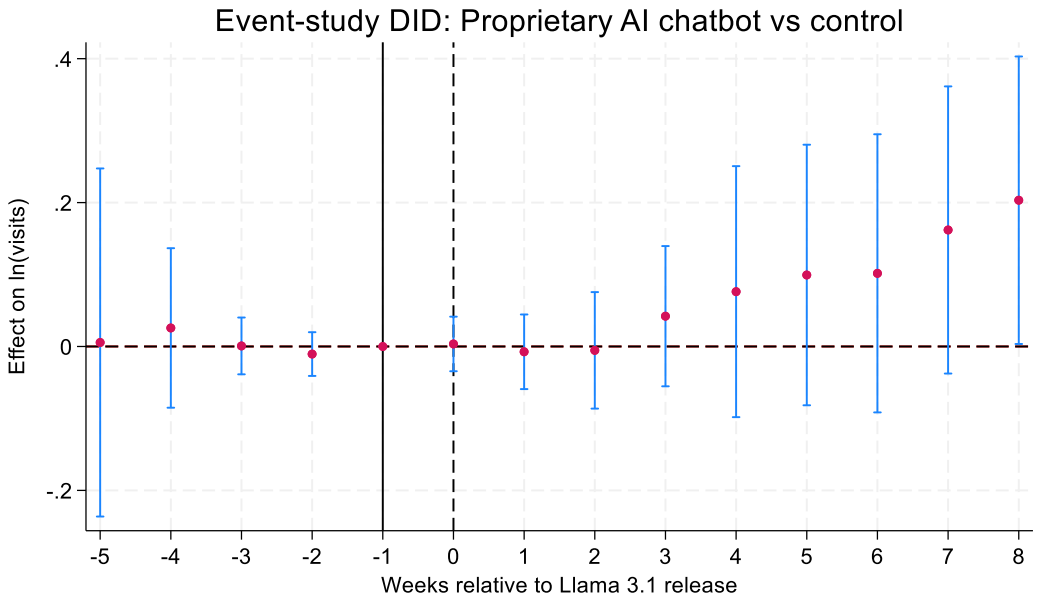


Figure 23 Event-study DiD estimates for log visits to proprietary AI chatbots

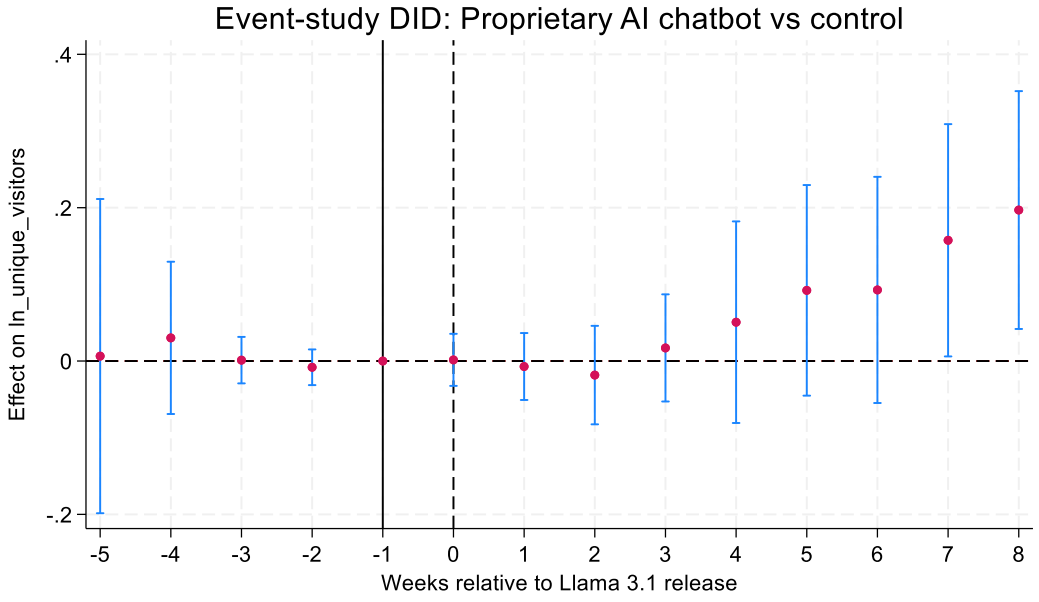


Figure 24 Event-study DiD estimates for log unique visitors to proprietary AI chatbots

5.3.3 Eight-week pre-treatment period

The main results and additional platform-specific analyses use a five-week pre-treatment period before the Llama 3.1 release to account for the Llama 3 release on April 18, 2024. Because web traffic to proprietary AI chatbots may have responded to the Llama 3 release with a delay, including earlier pre-treatment observations could raise concerns about potential confounding when estimating the effect of the Llama 3.1 release on web traffic to proprietary AI chatbots. Therefore, as a robustness check, this paper re-estimates the SDID and event-study SDID models for proprietary AI chatbots using a longer sample period with an eight-week pre-treatment period. The weekly data for the robustness analysis span from May 28, 2024, to September 23, 2024.

5.3.3.1 SDID results for proprietary AI chatbots: eight-week pre-treatment period

As a robustness check, Table 20 reproduces the main SDID results for log visits and log unique visitors to proprietary AI chatbots using a five-week pre-treatment period before the Llama 3.1 release in columns (1) and (2). It also presents the corresponding SDID and DiD results using an eight-week pre-treatment period before the release in columns (3)−(6).

Compared with the main results using a five-week pre-treatment period, the SDID estimates for both log visits and log unique visitors to proprietary AI chatbots using an eight-week pre-treatment period are slightly smaller in magnitude but remain positive. The estimated effects are statistically significant at the 5% and 10% levels, respectively. The DiD estimates for both log visits and log unique visitors using the eight-week pre-treatment period also remain positive. However, they are slightly smaller in magnitude than the corresponding SDID estimates. The DiD estimates are statistically insignificant (Table 20).

Table 20 DiD and SDID results for proprietary AI chatbots: eight-week pre-treatment period

Model	(1) SDID	(2) SDID	(3) SDID	(4) SDID	(3) DiD	(4) DiD
Dependent variable	Log visits	Log unique visitors	Log visits	Log unique visitors	Log visits	Log unique visitors
Pre-treatment periods	Five weeks (main results)		Eight weeks			
ATT	0.081** (0.033)	0.069** (0.03131)	0.078** (0.035)	0.063* (0.033)	0.065 (0.068)	0.057 (0.061)
Domain × device FE	Yes	Yes	Yes	Yes	Yes	Yes
Weekly FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	504	504	612	612	612	612

* Note: Placebo-based standard errors with 500 replications are reported in parentheses for the SDID estimates. Clustered robust standard errors at the domain level are reported in parentheses for the DID estimates.

5.3.3.2 Event-study SDID results for proprietary AI chatbots: eight-week pre-treatment period

Table 21 reproduces the event-study SDID estimates using a five-week pre-treatment period and presents the corresponding estimates using an eight-week pre-treatment period. The event-study SDID estimates based on the longer pre-treatment period are similar to the main event-study SDID results. However, the estimated positive effects based on the longer pre-treatment period are slightly smaller.

However, pre-treatment placebo effects for both log visits and log unique visitors at four, seven, and eight weeks before the Llama 3.1 release are statistically significant, and the Wald tests reject the joint null hypothesis that all pre-treatment placebo effects are zero. Accordingly, the evidence for differential pre-trends is clear. This evidence of differential pre-trends may reflect confounding from delayed responses in web traffic to proprietary AI chatbots following the release of Llama 3, the predecessor to Llama 3.1, on April 18, 2024. Taken together, these results suggest that using a longer pre-treatment period for the SDID and event-study SDID analyses may be less appropriate because of potential confounding from the earlier Llama 3 release.

Table 21 Event-study SDID result for proprietary AI chatbots: eight-week pre-treatment period

Model	(1) SDID	(2) SDID	(3) SDID	(4) SDID
Dependent variable	Log visits	Log unique visitors	Log visits	Log unique visitors
Pre-treatment periods	Five weeks (main results)		Eight weeks	
Placebo effect (t-8)			0.038***	0.025**
			(0.012)	(0.011)
Placebo effect (t-7)			0.029**	0.018*
			(0.011)	(0.011)
Placebo effect (t-6)			-0.021	-0.025*
			(0.014)	(0.013)
Placebo effect (t-5)	0.005	0.005	0.011	0.008
	(0.016)	(0.015)	(0.014)	(0.013)
Placebo effect (t-4)	0.024*	0.026**	0.025**	0.025**
	(0.013)	(0.012)	(0.011)	(0.010)
Placebo effect (t-3)	0.007	0.007	0.011	0.006
	(0.011)	(0.009)	(0.014)	(0.012)
Placebo effect (t-2)	-0.004	-0.003	-0.003	-0.006
	(0.008)	(0.006)	(0.011)	(0.009)
Placebo effect (t-1)	0.010	0.010	0.016	0.010
	(0.013)	(0.013)	(0.015)	(0.015)
ATT (t)	0.010	0.006	0.011	0.003
	(0.027)	(0.022)	(0.029)	(0.024)
ATT (t+1)	-0.002	-0.004	-0.001	-0.007
	(0.023)	(0.020)	(0.025)	(0.022)
ATT (t+2)	-0.001	-0.017	-0.002	-0.022
	(0.021)	(0.018)	(0.024)	(0.020)
ATT (t+3)	0.050	0.024	0.047	0.016
	(0.031)	(0.028)	(0.034)	(0.031)
ATT (t+4)	0.080**	0.053	0.070*	0.041
	(0.036)	(0.034)	(0.038)	(0.035)
ATT (t+5)	0.108***	0.100**	0.104**	0.090**
	(0.041)	(0.039)	(0.044)	(0.041)
ATT (t+6)	0.106**	0.095**	0.090*	0.080*
	(0.048)	(0.047)	(0.049)	(0.046)
ATT (t+7)	0.165***	0.159***	0.151***	0.146***
	(0.048)	(0.047)	(0.048)	(0.046)
ATT (t+8)	0.205***	0.197***	0.190***	0.184***
	(0.049)	(0.049)	(0.049)	(0.047)
Domain × device FE	Yes	Yes	Yes	Yes
Weekly FE	Yes	Yes	Yes	Yes
Wald test: placebo effects jointly equal zero	0.4254	0.2887	**0.0011****	**0.0116****
Observations	504	504	612	612

* Note: Placebo-based standard errors with 500 replications are in parentheses. The Wald test presents the p-value for the joint null hypothesis that all pre-treatment placebo effects are equal to zero. *** p<0.01, **p<0.05, * p<0.1.

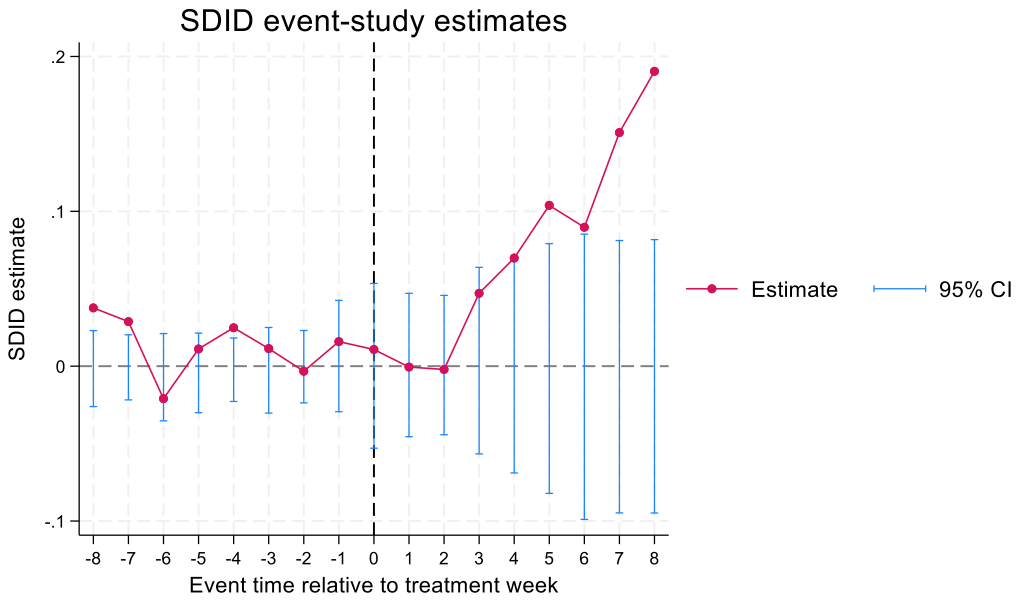


Figure 25 Event-study SDID estimates for log visits to proprietary AI chatbots: eight-week pre-treatment period

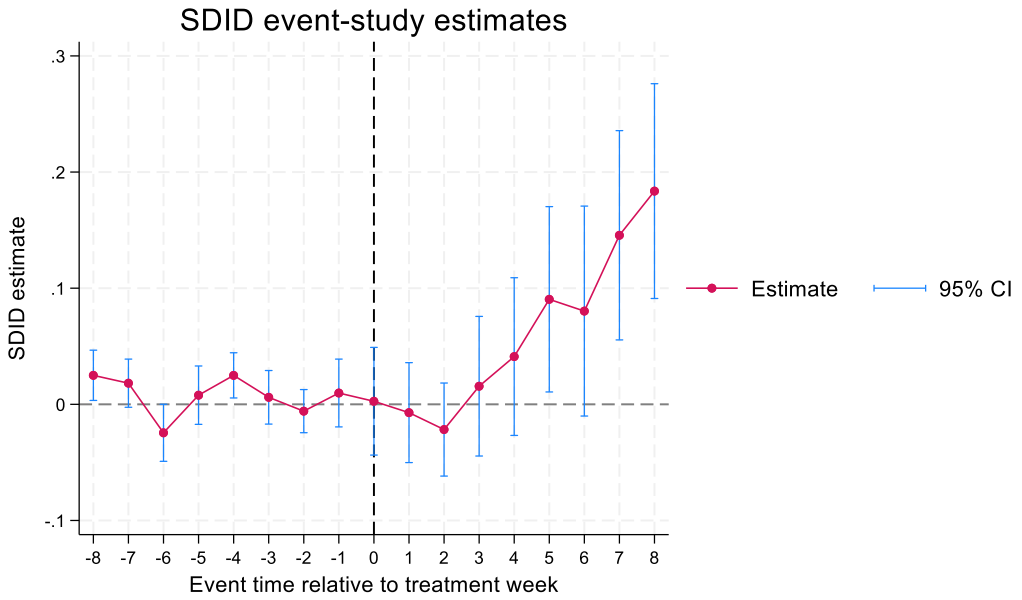


Figure 26 Event-study SDID estimates for log unique visitors to proprietary AI chatbots: eight-week pre-treatment period

5.3.3.3 Event-study DID results for proprietary AI chatbots: eight-week pre-treatment period

Table 22 presents event-study DiD estimates using either a five- or an eight-week pre-treatment period. The estimated dynamic effects are similar across the two specifications (Figures 27 and 28). However, the Wald tests for the eight-week pre-treatment period reject the joint null hypothesis that all pre-treatment estimates are equal to zero. Thus, the parallel trends assumption may be less plausible when using a longer pre-treatment period in the DiD analysis.

Table 22 Event-study DiD results for proprietary AI chatbots: eight-week pre-treatment period

Model	(1) DiD	(2) DiD	(3) DiD	(4) DiD
Dependent variable	Log visits	Log unique visitors	Log visits	Log unique visitors
Pre-treatment periods	Five weeks (main results)		Eight weeks	
Lead 8			0.046	0.033
			(0.198)	(0.174)
Lead 7			0.036	0.027
			(0.199)	(0.177)
Lead 6			-0.023	-0.024
			(0.186)	(0.165)
Lead 5	0.006	0.006	0.006	0.006
	(0.115)	(0.097)	(0.115)	(0.097)
Lead 4	0.026	0.030	0.026	0.030
	(0.053)	(0.047)	(0.053)	(0.047)
Lead 3	0.001	0.001	0.001	0.001
	(0.019)	(0.014)	(0.019)	(0.014)
Lead 2	-0.011	-0.008	-0.011	-0.008
	(0.014)	(0.011)	(0.014)	(0.011)
Lag 0	0.004	0.002	0.004	0.002
	(0.018)	(0.016)	(0.018)	(0.016)
Lag 1	-0.007	-0.007	-0.007	-0.007
	(0.025)	(0.021)	(0.025)	(0.021)
Lag 2	-0.005	-0.018	-0.005	-0.018
	(0.038)	(0.030)	(0.038)	(0.030)
Lag 3	0.042	0.017	0.042	0.017
	(0.046)	(0.033)	(0.046)	(0.033)
Lag 4	0.076	0.051	0.076	0.051
	(0.083)	(0.062)	(0.083)	(0.062)
Lag 5	0.099	0.092	0.099	0.092
	(0.086)	(0.065)	(0.086)	(0.065)
Lag 6	0.102	0.093	0.102	0.093
	(0.092)	(0.070)	(0.092)	(0.070)
Lag 7	0.162	0.157**	0.162	0.157**
	(0.095)	(0.072)	(0.095)	(0.072)
Lag 8	0.203**	0.197**	0.203**	0.197**
	(0.095)	(0.074)	(0.095)	(0.074)
Domain × device FE	Yes	Yes	Yes	Yes
Weekly FE	Yes	Yes	Yes	Yes
Wald test	0.1047	0.2985	0.0002***	0.0002***
Observations	504	504	612	612

* Note: Clustered robust standard errors at the domain level are reported in parentheses for the DiD estimates. The Wald test presents the p-value for the joint null hypothesis that all pre-treatment effects are equal to zero. *** p<0.01, **p<0.05, * p<0.1.

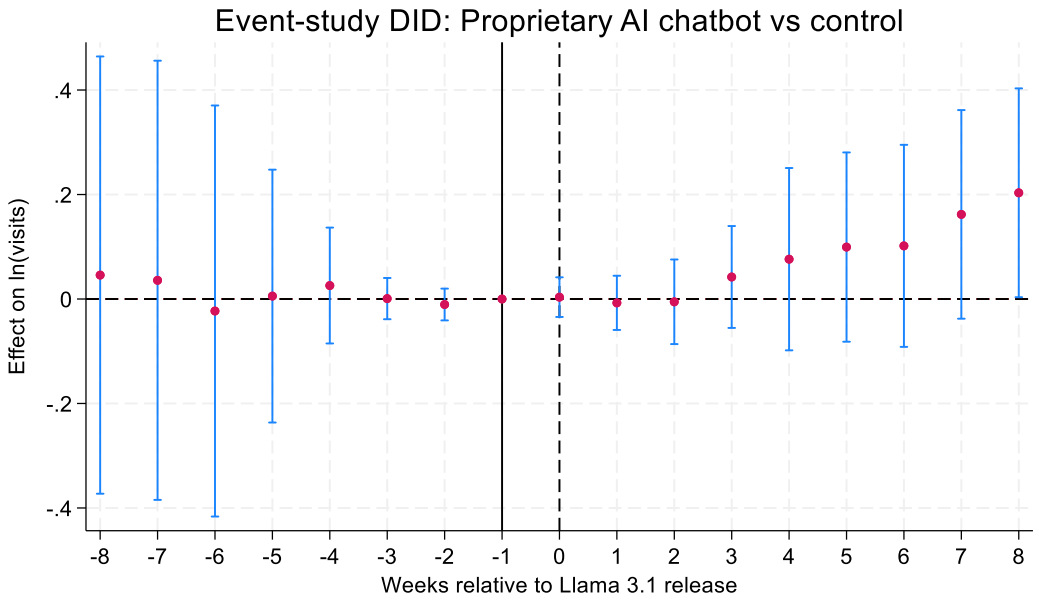


Figure 27 Event-study DiD estimates for log visits to proprietary AI chatbots (eight-week pre-treatment period)

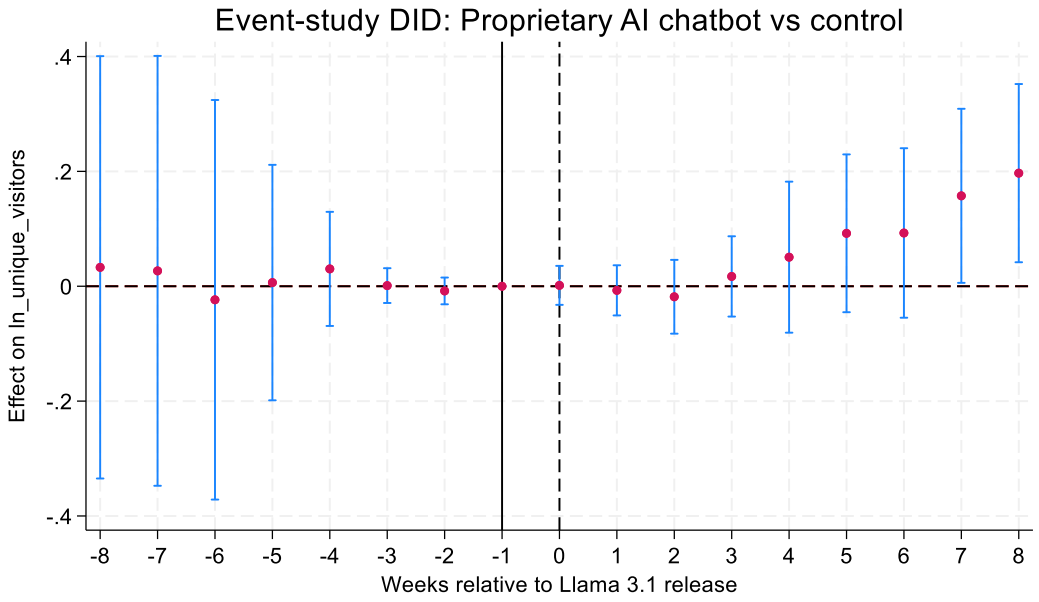


Figure 28 Event-study DiD estimates for log unique visitors to proprietary AI chatbots (eight-week pre-treatment period)

6. Conclusion

This paper examines the effect of the release of Meta's Llama 3.1, one of the most popular frontier open-weight text-only LLMs at the time of its release, on global web traffic to proprietary AI chatbots and AI-related platforms using weekly web traffic data and quasi-experimental methods.

The main empirical results show that the effects of the release on global web traffic to proprietary AI chatbots are small for the first two weeks after the release. However, the release's positive effects on global web traffic to the overall proprietary AI chatbot group gradually increase thereafter, rather than declining due to substitution.

While proprietary AI chatbots show delayed positive responses to the release, some platforms show immediate and larger positive responses. In particular, chatgpt.com shows larger and immediate positive responses to the release than the overall proprietary AI chatbot group. Likewise, platform.openai.com, OpenAI's development platform, shows positive web traffic responses during the release week. However, its response turns negative one week after the release, then near zero two and three weeks later, and then steadily increases over time.

Furthermore, perplexity.ai and huggingface.co show positive and immediate increases in web traffic following the release, which may reflect the availability of Llama 3.1 on both platforms on its release date. Therefore, increased online attention to open-weight LLMs may also increase overall attention to open-source AI models and applications that incorporate these open-weight LLMs. This increased attention may subsequently spill over to other LLM-based services, including proprietary AI chatbots.

Moreover, Meta's Llama 3.1 release increases web traffic to Meta's corporate website in the first few weeks after the release, but the positive effect gradually declines over time. Accordingly, the

release of Meta's frontier open-weight LLM may attract greater short-term online attention to Meta, the firm that develops and releases the model.

Overall, releasing a frontier open-weight, text-only LLM may increase global web traffic to proprietary AI chatbots, OpenAI's development platform, a proprietary LLM-based search engine, and an open-source AI model platform, rather than reduce web traffic to them through substitution. Importantly, proprietary AI chatbots and AI-related platforms show heterogeneous responses to the release of Llama 3.1 in both magnitude and timing, depending on user type (e.g., general versus advanced users), service type (e.g., chatbot versus API), the openness of model weights, and the availability of open-weight LLMs at the application level.

The empirical approach and framework used in this paper can be applied and extended to other Gen AI models to better understand user responses and market dynamics in emerging Gen AI markets. For example, future research could examine how releasing a low-cost, competitive open-weight LLM, such as DeepSeek-R1, affects consumer demand, firms' strategic behavior, and LLM markets across countries. Another open topic is the impact of AI agents on consumers, firms, industries, and markets, given the rapid innovation and deployment of AI agents across domains. For example, Yang et al. (2025) investigated users' adoption of Perplexity's AI agent and their behavior using internal behavioral data. Furthermore, Meta launched Muse, its personal agentic AI with free and paid subscription options, on September 8, 2026 (Meta 2026), attracting significant consumer attention and experiencing rapid adoption in Apple's U.S. App Store (Bobrowsky 2026). Muse's rapid adoption highlights the importance of understanding how free and paid AI agents affect consumer behavior, firms' strategic behavior, market structure and competition in Gen AI and digital markets, and broader economic outcomes.

Appendix

A.1 Individual proprietary AI chatbots: claude.ai, copilot.microsoft.com, gemini.google.com

The proprietary AI chatbot domains in this paper are OpenAI's chatgpt.com, Anthropic's claude.ai, Microsoft's copilot.microsoft.com, and Google's gemini.google.com. Because the empirical results for the overall proprietary AI chatbot group and chatgpt.com are already presented in the prior sections, this subsection presents the results for the other individual proprietary AI chatbots to examine their web traffic responses to the release of Llama 3.1.

Specifically, claude.ai relies on Anthropic's own pre-trained LLMs and performs well in coding and mathematics. Also, gemini.google.com relies on Google's own pre-trained models. In contrast, copilot.microsoft.com relies on OpenAI's GPT models because Microsoft partners with OpenAI and is a major investor.

A.1.1. Results for claude.ai, copilot.microsoft.com, gemini.google.com

The estimated SDID effects of the release on claude.ai, copilot.microsoft.com, and gemini.google.com are heterogeneous across the three chatbot domains (Table A1). Moreover, as shown in Table A2, some pre-treatment placebo effects four and/or five weeks before the release are statistically significant, and the Wald tests for both outcomes reject the joint null hypothesis for these chatbot domains. Accordingly, the evidence of differential pre-trends raises concerns about the causal interpretation of these estimates.

Table A1 SDID results for claude.ai, copilot.microsoft.com, gemini.google.com

Model	(1) SDID	(2) SDID	(3) SDID	(4) SDID	(5) SDID	(6) SDID
Treated group	claude.ai		copilot.microsoft.com		gemini.google.com	
Dependent variable	Log visits	Log unique visitors	Log visits	Log unique visitors	Log visits	Log unique visitors
ATT	-0.032	-0.019	0.164**	0.112*	0.023	0.044
	(0.065)	(0.061)	(0.065)	(0.061)	(0.065)	(0.061)
Domain × device FE	Yes	Yes	Yes	Yes	Yes	Yes
Weekly FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	420	420	420	420	420	420

* Note: Placebo-based standard errors with 500 replications are reported in parentheses in columns. *** p<0.01, **p<0.05, * p<0.1.

Table A2 Event-study SDID results for claude.ai, copilot.microsoft.com, gemini.google.com

Model	(1)	(2)	(3)	(4)	(5)	(6)
Treated group	claude.ai		copilot.microsoft.com		gemini.google.com	
Dependent variable	Log visits	Log unique visitors	Log visits	Log unique visitors	Log visits	Log unique visitors
Placebo effect (t-5)	-0.237***	-0.198***	0.112***	0.100***	0.098***	0.084***
	(0.025)	(0.023)	(0.025)	(0.023)	(0.025)	(0.023)
Placebo effect (t-4)	-0.030	-0.020	0.046**	0.041**	0.076***	0.085***
	(0.021)	(0.017)	(0.021)	(0.017)	(0.021)	(0.017)
Placebo effect (t-3)	0.018	0.017	-0.000	0.002	0.004	0.010
	(0.022)	(0.017)	(0.022)	(0.017)	(0.022)	(0.017)
Placebo effect (t-2)	0.004	0.002	-0.007	-0.005	-0.011	-0.009
	(0.020)	(0.013)	(0.020)	(0.013)	(0.020)	(0.013)
Placebo effect (t-1)	0.025	0.022	-0.005	0.011	0.009	0.011
	(0.019)	(0.023)	(0.019)	(0.023)	(0.019)	(0.023)
ATT (t)	0.031	0.030	-0.017	-0.020	-0.009	-0.002
	(0.057)	(0.043)	(0.057)	(0.043)	(0.057)	(0.043)
ATT (t+1)	-0.003	0.001	-0.029	-0.033	-0.037	-0.013
	(0.051)	(0.041)	(0.051)	(0.041)	(0.051)	(0.041)
ATT (t+2)	-0.067	-0.066*	0.018	-0.025	-0.060	-0.052
	(0.048)	(0.038)	(0.048)	(0.038)	(0.048)	(0.038)
ATT (t+3)	-0.060	-0.068	0.119*	0.035	-0.008	0.009
	(0.068)	(0.058)	(0.068)	(0.058)	(0.068)	(0.058)
ATT (t+4)	-0.137*	-0.124*	0.263***	0.159**	0.051	0.060
	(0.076)	(0.068)	(0.076)	(0.068)	(0.076)	(0.068)
ATT (t+5)	-0.131	-0.106	0.293***	0.217***	0.117	0.154*
	(0.085)	(0.078)	(0.085)	(0.078)	(0.085)	(0.078)
ATT (t+6)	-0.133	-0.088	0.281***	0.196**	0.099	0.124
	(0.098)	(0.093)	(0.098)	(0.093)	(0.098)	(0.093)
ATT (t+7)	-0.024	0.011	0.355***	0.280***	0.105	0.148
	(0.096)	(0.093)	(0.096)	(0.093)	(0.096)	(0.093)
ATT (t+8)	0.029	0.068	0.415***	0.343***	0.140	0.172*
	(0.097)	(0.096)	(0.097)	(0.096	(0.097)	(0.096)
Domain × device FE	Yes	Yes	Yes	Yes	Yes	Yes
Weekly FE	Yes	Yes	Yes	Yes	Yes	Yes
Wald test: placebo effects jointly equal zero	0.000***	0.000***	0.0001***	0.0002***	0.0000***	0.0000***
Observations	420	420	420	420	420	420

* Note: Placebo-based standard errors with 500 replications are reported in parentheses in columns. The Wald test presents the p-value for the joint null hypothesis that all pre-treatment placebo effects are equal to zero. *** p<0.01, **p<0.05, * p<0.1.

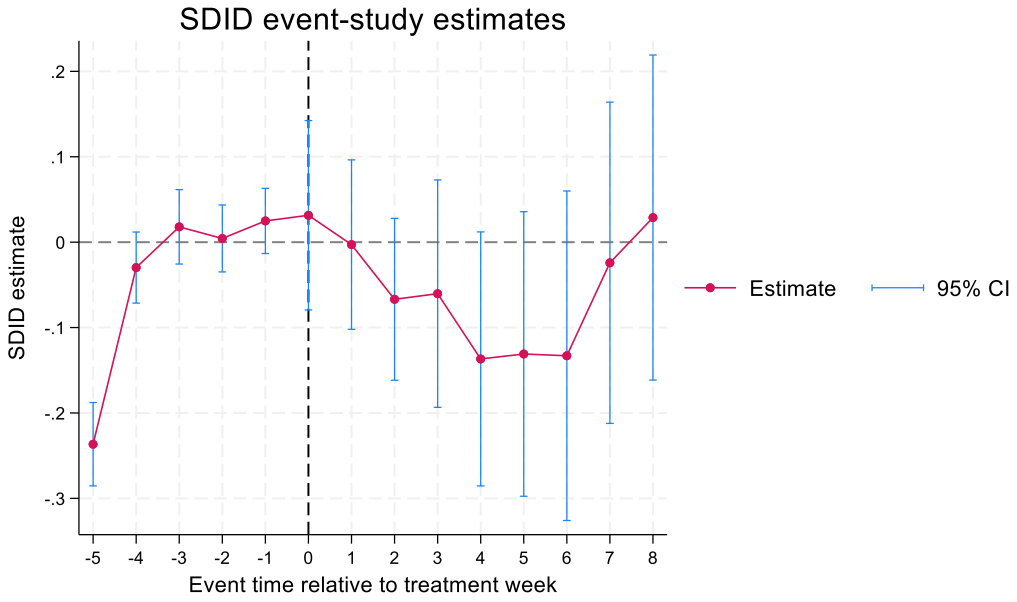


Figure A1 Event-study SDID estimates for log visits to claude.ai

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

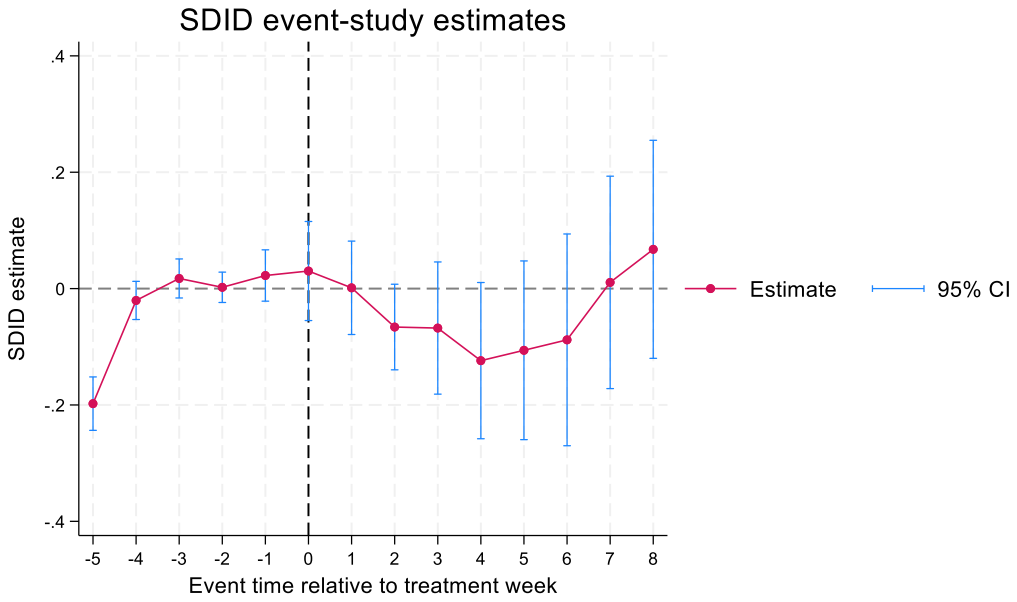


Figure A2 Event-study SDID estimates for log unique visitors to claude.ai

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

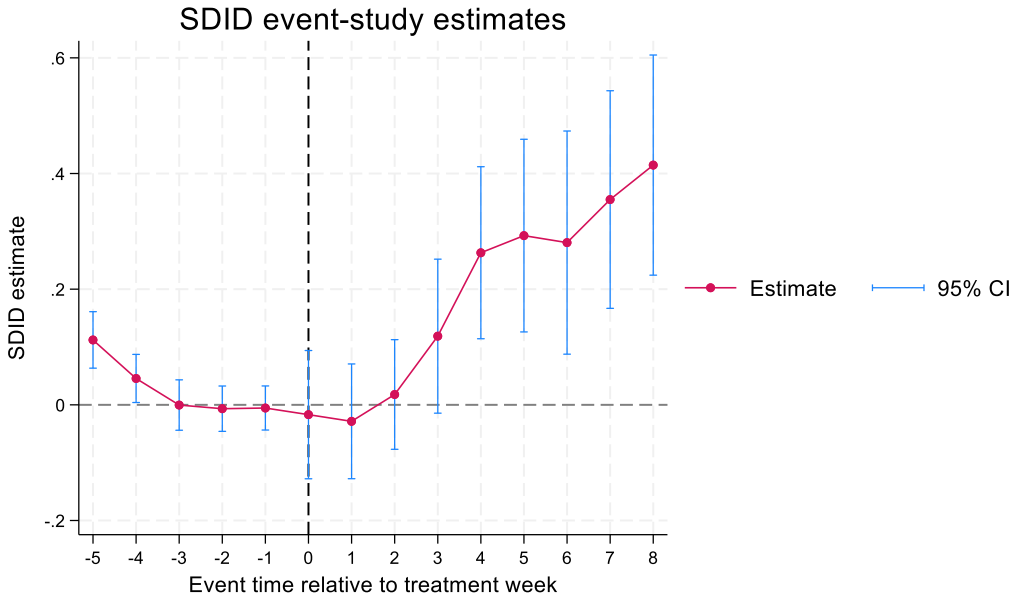


Figure A3 Event-study SDID estimates for log visits to copilot.microsoft.com

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

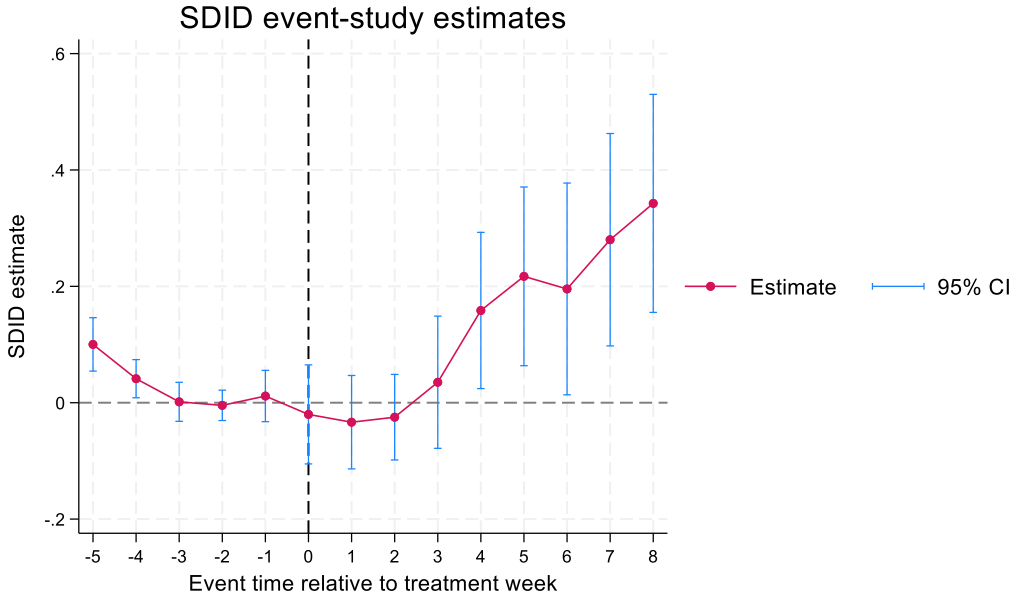


Figure A4 Event-study SDID estimates for log unique visitors to copilot.microsoft.com

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

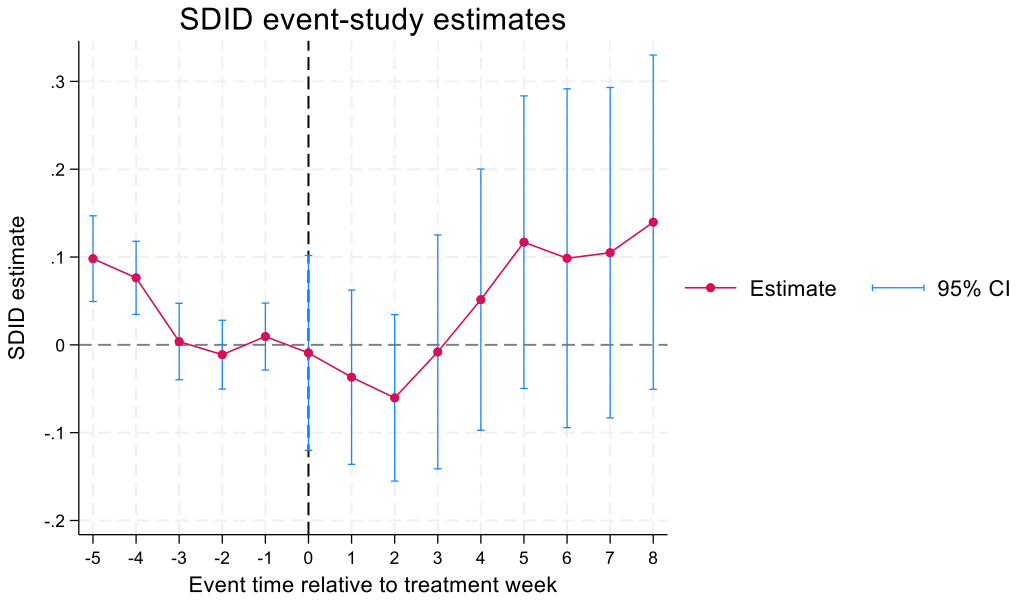


Figure A5 Event-study SDID estimates for log visits to gemini.google.com

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

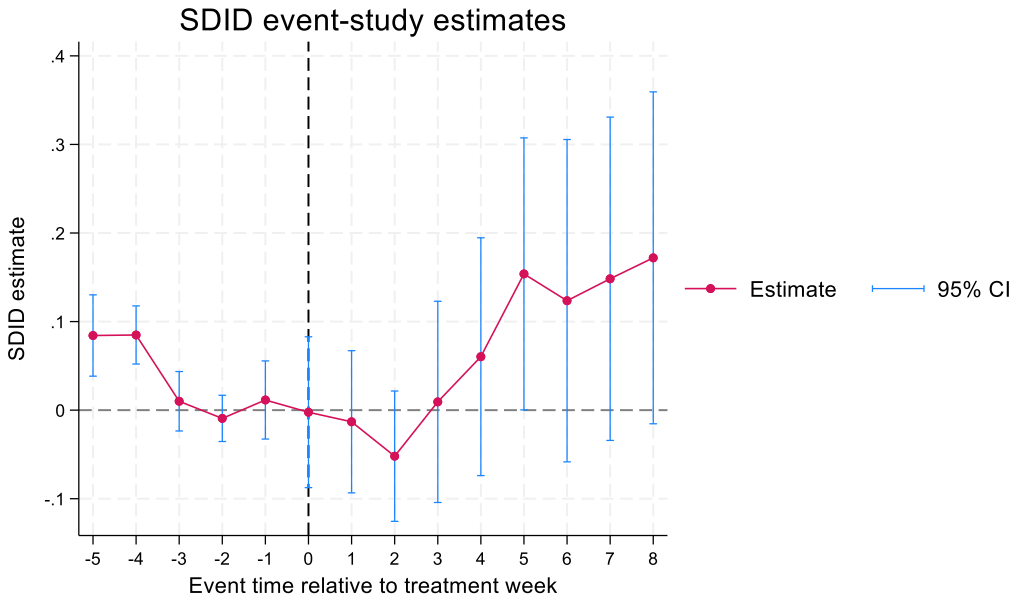


Figure A6 Event-study SDID estimates for log unique visitors to gemini.google.com

* Note: Each red point indicates the estimated coefficient for a given week. The blue vertical bars represent 95% placebo-based confidence intervals (CIs). The vertical dashed line indicates the release week.

Reference

Abadie, A., and J. Gardeazabal. 2003. The economic costs of conflict: A case study of the Basque Country. American Economic Review 93(1): 113–132.

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. Journal of the American Statistical Association, 105(490), 493-505.

Agarwal, S., & Sen, A. (2026). The impact of Google AI Overviews on publisher traffic and user experience: Evidence from a field experiment. SSRN Work. Pap, 6513059.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. American Economic Review, 111(12), 4088-4118.

Arkhangelsky, D., & Samkov, A. (2024). Sequential synthetic difference in differences. arXiv preprint arXiv:2404.00164.

Aubakirova, M., Atallah, A., Clark, C., Summerville, J., & Midha, A. (2026). State of AI: An empirical 100 trillion token study with OpenRouter. arXiv preprint arXiv:2601.10088.

Azoulay, P., Krieger, J., & Nagaraj, A. (2025). Old moats for new models: Openness, control, and competition in generative artificial intelligence. Entrepreneurship and Innovation Policy and the Economy, 4(1), 7-46.

Bick, A., Blandin, A., & Deming, D. J. (2026) The rapid adoption of generative AI. Management Science 0(0). https://doi.org/10.1287/mnsc.2025.02523

Burtch, G., Lee, D., & Chen, Z. (2024). The consequences of generative AI for online knowledge communities. Scientific Reports, 14(1), 10413.

Chatterji, A., Cunningham, T., Deming, D. J., Hitzig, Z., Ong, C., Shan, C. Y., & Wadman, K. (2025). How people use chatgpt (No. w34255). National Bureau of Economic Research.

Chiang, W. L., Li, Z., Lin, Z., Sheng, Y., Wu, Z., Zhang, H., ... & Xing, E. P. (2023). Vicuna: An open-source chatbot impressing gpt-4 with 90%* chatgpt quality. See https://vicuna. lmsys. org (accessed 14 April 2023), 2(3), 6.

Choi, W. C., & Chang, C. I. (2025). Advantages and limitations of open-source versus commercial large language models (LLMs): A comparative study of DeepSeek and OpenAI's ChatGPT.

Ciccia, D. (2024). A short note on event-study synthetic difference-in-differences estimators. arXiv preprint arXiv:2407.09565.

Ciccia, D., Clarke, D., & Pailañir, D. (2024). SDID_EVENT: Stata module providing synthetic difference-in-differences (SDID) event-study estimators. Statistical Software Components S459354, Boston College Department of Economics.

Clarke, D., Pailanir, D., Athey, S., & Imbens, G. (2024). On synthetic difference-in-differences and related estimation methods in Stata. The Stata Journal, 24(4), 557-598.

Clarke, D., & Tapia-Schythe, K. (2021). Implementing the panel event study. The Stata Journal, 21(4), 853-884.

Cobreros, L., Sunyer, C., & Castelló, J. V. (2026). Paying for parity: Incentivizing men in feminized occupations. Economics Letters, 113089.

Congiu, R., Sabatino, L., & Sapi, G. (2022). The Impact of privacy regulation on web traffic: Evidence from the GDPR. Information Economics and Policy, 61, 101003.

Demirer, M., Fradkin, A., Tadelis, N., & Peng, S. (2025). The emerging market for intelligence: Pricing, supply, and demand for LLMs (No. w34608). National Bureau of Economic Research.

Demirer, M., Fradkin, A., & Tadelis, N. (2026). The emerging market for intelligence: How firms buy and sell AI. Journal of Economic Perspectives, 40(3), 23-46.

Feyzollahi, M., & Rafizadeh, N. (2025). The adoption of large language models in economics research. Economics Letters, 250, 112265.

Gholami, S., Firullo, C., Cheyre, C., & Acquisti, A. (2026). Beyond search: LLM adoption and web traffic concentration. Available at SSRN 6238578.

Google (2024, May 14), Generative AI in search: Let Google do the searching for you, Available at Google blog post: https://blog.google/products-and-platforms/products/search/generative-ai-google-search-may-2024/

Grattafiori, A., Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., ... & Vasic, P. (2024). The llama 3 herd of models. arXiv preprint arXiv:2407.21783.

Guo, D., Yang, D., Zhang, H., Song, J., Wang, P., Zhu, Q., ... & He, Y. (2025). Deepseek-R1: Incentivizing reasoning capability in LLMs via reinforcement learning. arXiv preprint arXiv:2501.12948.

Haag, A. (2025, October 6). The state of AI competition in advanced economies, FEDS Notes. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/2380-7172.3930.

Habibi, M. (2025). Open sourcing GPTs: Economics of open sourcing advanced AI models. arXiv preprint arXiv:2501.11581.

Hu, S. (2026). Local economic impacts of food manufacturing plant closures in the Midwest. Department of Economics, Iowa State University. https://drive. google. com/file/d/1bPeXn0B0YeGBUNQ8bYL2pIxppPs AUwXj/view.

Hu, S., & Winters, J. V. (2025). Growing from the STEM? OPT classification and international students in Economics (No. 18336). IZA Discussion Papers.

Korinek, A. (2023). Generative AI for economic research: Use cases and implications for economists. Journal of Economic Literature, 61(4), 1281-1317.

Korinek, A., & Vipra, J. (2025). Concentrating intelligence: scaling and market structure in artificial intelligence. Economic Policy, 40(121), 225-256.

Ladu, M. G., & Masi, T. (2025). Decoding sectoral dynamics: Unveiling the heterogeneous effects of earthquakes. Italian Economic Journal, 1-24.

Lerner, J., & Tirole, J. (2005). The economics of technology sharing: Open source and beyond. Journal of Economic Perspectives, 19(2), 99-120.

Liu, Yan and Huang, Jingyun and Wang, He. (2025, September 08), Who on earth is using generative AI? Global trends and shifts in 2025. Available at SSRN: https://ssrn.com/abstract=5495020 or http://dx.doi.org/10.2139/ssrn.5495020

Meta (2024, April 18), Introducing Meta Llama 3: The most capable openly available LLM to date. Available at Meta blog post: https://ai.meta.com/blog/meta-llama-3/

Meta (2024, July 23), Introducing Llama 3.1: Our most capable models to date. Available at Meta blog post: https://ai.meta.com/blog/meta-llama-3-1/

Meta (2024, September 25), Llama 3.2: Revolutionizing edge AI and vision with open, customizable models. Available at Meta blog post: https://ai.meta.com/blog/llama-3-2-connect-2024-vision-edge-mobile-devices/

Meta (2026, September 9), Introducing Muse: The World's first personal ai agent built for everyone. Available at Meta blog post: https://about.fb.com/news/2026/09/introducing-muse-personal-ai-agent/s/

Muñiz-Fernández, Á., & Boto-García, D. (2026). Political instability and housing investment: evidence from the Catalan secessionist conflict. Economia Politica, 1-31.

Nagle, F., & Yue, D. (2025). The latent role of open models in the ai economy. Available at SSRN 5767103.

Saffer, H., Ozdogan, S., Grossman, M., Dench, D., & Dave, D. (2025). Comprehensive E-Cigarette flavor bans and tobacco use among youth and adults. Health economics, 34(12), 2238-2254.

Taori, R., Gulrajani, I., Zhang, T., Dubois, Y., Li, X., Guestrin, C., ... & Hashimoto, T. B. (2023, March). Stanford alpaca: An instruction-following llama model.

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). LLaMA: Open and efficient foundation language models. arXiv. https://doi.org/10.48550/arXiv.2302.13971

Padilla, N., Lam, H. T., Lambrecht, A., & Hollenbeck, B. (2025). The impact of LLM adoption on online user behavior. Available at SSRN.

Porreca, Z. (2022). Synthetic difference-in-differences estimation with staggered treatment timing. Economics Letters, 220, 110874.

Bobrowsky, M. (2026, September 22), Meta's new AI agent is an instant hit—and the backlash has already begun. The Wall Street Journal.

Yeverechyahu, D., Mayya, R., & Oestreicher-Singer, G. (2024). The impact of large language models on open-source innovation: Evidence from github copilot. arXiv preprint arXiv:2409.08379.

Yang, J., Yonack, N., Zyskowski, K., Yarats, D., Ho, J., & Ma, J. (2025). The adoption and usage of AI agents: Early evidence from perplexity. arXiv preprint arXiv:2512.07828.

Yotzov, I., Barrero, J. M., Bloom, N., Bunn, P., Davis, S. J., Foster, K. M., ... & Wang, B. Z. (2026). Firm data on AI (No. w34836). National Bureau of Economic Research.

Zhao, H., & Berman, R. (2025). The impact of LLMs on online news consumption and production. The Wharton School Research Paper.

Zhao, Yue. (2026) Open-source generative AI and firm value: evidence from the release of DeepSeek. 1-79. Available at: https://ink.library.smu.edu.sg/etd_coll/839