CES ifo Working Papers

13098
2026

October 2026

Integration and Competition in Generative AI: Evidence from Google's AI Mode Rollout

Michelangelo Rossi, Adèle Diehl, Louis-Daniel Pape

CES ifo

Imprint:

CESifo Working Papers

ISSN 2364-1428 (digital)

Publisher and distributor: Munich Society for the Promotion
of Economic Research - CESifo GmbH

Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740

Email office@cesifo.de
https://www.cesifo.org

Editor: Clemens Fuest

An electronic version of the paper may be downloaded free of charge
· from the CESifo website: www.ifo.de/en/cesifo/publications/cesifo-working-papers
· from the SSRN website: www.ssrn.com/index.cfm/en/cesifo/
· from the RePEc website: https://ideas.repec.org/s/ces/ceswps.html

Integration and Competition in Generative AI: Evidence from Google's AI Mode Rollout*

Michelangelo Rossi, Adèle Diehl, and Louis-Daniel Pape†

October 28, 2026

Abstract

Integrating generative AI into digital platforms may reshape competition among AI assistants. We study whether Google's introduction of AI Mode within Search reduces traffic to competing services. We exploit the August 2025 expansion of Google's AI Mode, which reached 25 sample countries, using website traffic estimates for ten AI assistants and a difference-in-differences design with 30 countries where AI Mode was unavailable. Making AI Mode available within Google Search reduced ChatGPT website visits by 5.9 percent over the first four weeks, implying 3.5 million fewer daily visits across treated countries. The decline is concentrated in direct visits, while organic-search referrals change little. Daily unique visitors decline by a similar proportion, with little change in visits per daily visitor or average visit duration, suggesting the reduction primarily reflects lower daily reach. Point estimates remain negative throughout the 40-day follow-up, after which comparison countries also gain access to AI Mode. We find no statistically significant average effect on Gemini or the eight smaller assistants combined over the first four weeks. These findings suggest integrating conversational AI into established search interfaces can reduce traffic to standalone competitors, highlighting distribution channels' importance for competition among AI assistants.

JEL codes: L13, L86, M31, O33.

Keywords: generative AI, search engines, platform integration, platform competition.

*This work benefited from Hi! PARIS and State funding managed by the French ANR under the France 2030 program, reference ANR-23-IACL-0005. All errors are ours.

†M. Rossi: HEC Paris and CESifo. E-mail: rossim@hec.fr. A. Diehl: Mines Paris - PSL. E-mail: adele.diehl@minesparis.psl.eu. L.-D. Pape: Télécom Paris, Center for Research in Economics and Statistics (CREST), Institut Polytechnique de Paris. E-mail: louis.pape@telecom-paris.fr.

1 Introduction

For more than three decades, search engines have played a central role in how consumers find information online. Whether comparing products, planning a trip, or learning about an unfamiliar topic, consumers have traditionally relied on search engines to navigate the web. Generative AI assistants are changing this pattern by allowing consumers to carry out these tasks directly through services such as ChatGPT, without necessarily visiting a search engine.

Yet incumbent digital platforms retain an important advantage: they can integrate AI assistants into interfaces that consumers already use. For example, Google and Bing have introduced conversational AI interfaces within their search engines, while Meta provides access to its AI assistant within its existing social and messaging services.[1] Such integration allows incumbent platforms to reach consumers without requiring them to visit a separate AI service, potentially diverting demand from standalone AI assistants. This matters for how AI providers attract and retain users, which intermediaries influence the information consumers encounter, and how traffic reaches brands and publishers. Competition in generative AI may therefore depend not only on the capabilities of different AI assistants, but also on how they are integrated into and distributed through established digital platforms.

Google's AI Mode provides a setting to examine how introducing an AI assistant within an established search engine affects traffic to standalone assistants. Unlike AI Overviews, which add generated summaries to conventional search results, AI Mode offers a dedicated conversational interface within Google Search, allowing consumers to ask questions, receive AI-generated answers, and engage in follow-up exchanges (Figure 1). Our analysis provides evidence on how integrating conversational AI into search affects demand for standalone AI assistants.

To identify this effect, we exploit a sharp expansion of AI Mode's international rollout. On August 21, 2025, Google extended public access to AI Mode in English to 25 countries in our sample, across five regions, while 30 other countries remained without access through September 30, 2025. Using daily website traffic estimates from Semrush for ten AI assistants, we compare changes in traffic before and after the rollout in the newly covered countries with contemporaneous changes in countries where AI Mode was not yet available. The analysis focuses on the first four weeks following launch and estimates the effect of making AI Mode publicly available within Google Search. Our data cover desktop and mobile website visits, excluding native-app use.

Making AI Mode available reduced ChatGPT website visits by 5.9 percent on average over the first four

[1] For official descriptions of these integrations, see Google (2025a) on AI Mode, Microsoft (2025) on Copilot Search, and Meta (2025) on Meta AI.

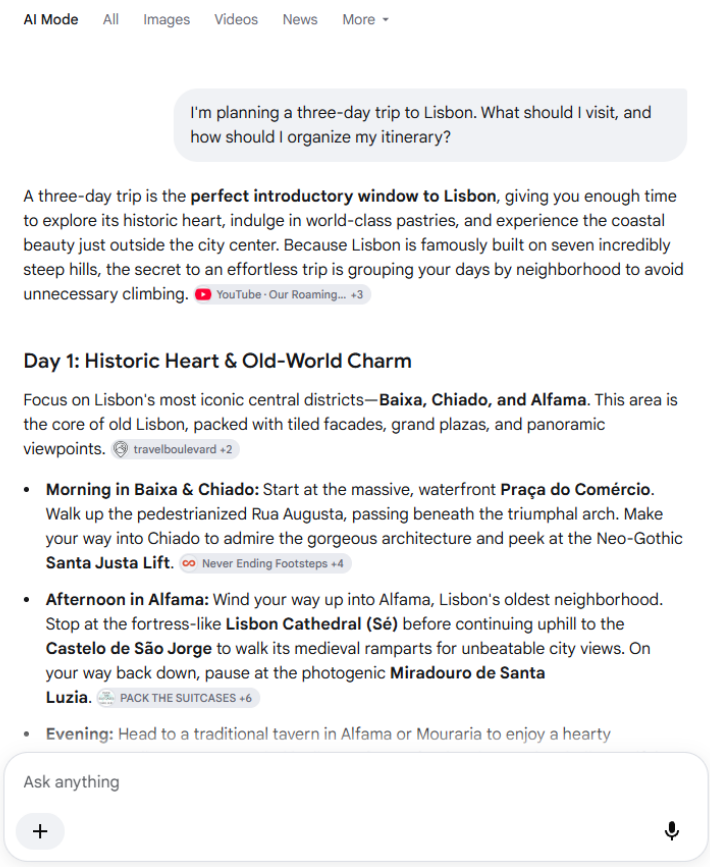


Figure 1. Google's AI Mode Interface

Notes: The figure illustrates Google's AI Mode within Google Search. Users can submit questions, receive AI-generated answers combining text, images, and links to external sources, and continue the interaction through follow-up questions. Screenshot taken in October 2026.

weeks. Assuming a common proportional effect across countries, this implies roughly 3.5 million fewer daily visits across the treated countries. The decline emerges immediately after launch, with negative point estimates throughout the four-week post-launch window. The estimated effect is robust to alternative specifications, weighting schemes, and estimators, while pre-rollout monthly placebo exercises do not produce similarly negative estimates. Most of the reduction is concentrated in direct visits, whereas organic-search referrals change little. Daily unique visitors decline by approximately the same proportion as visits, with little change in visits per daily visitor or average visit duration. These patterns suggest that the reduction in ChatGPT traffic primarily reflects lower daily reach, with little detectable change in average visit frequency or duration. For Gemini and the eight smaller AI assistants combined, we find no statistically significant average effect over the first four weeks.

These results point to the importance of distribution in competition among AI assistants. Introducing AI Mode within Google Search reduced traffic to ChatGPT, with the decline concentrated in direct visits

rather than organic-search referrals. This suggests that competition extends beyond visits originating from search results. Access to established distribution channels is therefore an important dimension of competition among AI assistants. For standalone providers, maintaining daily reach may become more challenging when competing assistants are embedded in platforms that consumers already use. More broadly, changes in the AI services consumers use may affect which intermediaries influence their information searches and how they discover brands and publishers.

1.1 Related Literature

First, our paper connects to research on how generative AI reshapes online attention. Experiments show that AI answers embedded in search reduce publisher referrals (Agarwal and Sen, 2026; Wang, Gleason, Bart, Wilson and Metaxa, 2026), while other studies document effects on Wikipedia traffic and Reddit engagement (Khosravi and Yoganarasimhan, 2026; Zhang, Cui and Zhang, 2026). Research on standalone assistants examines how LLM adoption changes traditional search and browsing (Padilla, Lam, Lambrecht and Hollenbeck, 2026; Gholami, Firullo, Cheyre and Acquisti, 2026), and how ChatGPT Search affects referrals and search use (Shi, Zhu and Gu, 2026). We complement these studies by examining how integration into search changes demand for competing AI intermediaries.

Second, we contribute to research on competition among AI providers. Using API data, Fradkin (2025) and Demirer, Fradkin, Tadelis and Peng (2025) document substitution, market expansion, and differentiation among models. Using consumer transaction data, Egan, Russel, Shi and Yao (2026) document substantial substitutability and limited differentiation among paid AI services. We complement this research by examining consumer demand for AI assistants, including free-to-use services, and providing causal evidence that integration into an established interface can reduce traffic to a leading standalone competitor. Our findings highlight access to consumers through existing platforms as an additional dimension of competition, alongside prices and model capabilities.

Finally, our setting relates to platforms' expansion into adjacent markets. Google's placement of Flight Search shifted clicks away from organic listings (Edelman and Lai, 2016), while Facebook's integration of Instagram increased its use and reduced demand for smaller rivals (Li and Agarwal, 2017). Related experiments identify the importance of defaults in search (Allcott, Castillo, Gentzkow, Musolff and Salz, 2025), and evidence from Google Maps shows how access points within Search shape navigation (Pape and Rossi, 2026). We extend this work to conversational AI: although Google's Gemini already existed, AI Mode introduces a new access point, allowing us to measure how integration within an established user interface affects demand for standalone AI assistants.

2 Context and Data

2.1 Google's AI Mode

AI Mode provides a conversational interface within Google Search, accessible through a dedicated tab. Users can submit questions through text, voice, or images and receive AI-generated answers with links to external websites and the possibility of asking follow-up questions (Figure 1). The service uses Google's Gemini models within its search infrastructure, issuing multiple related searches through *query fan-out*. Google promotes AI Mode for tasks such as comparing products, planning trips, and exploring complex questions (Google, 2025a,f).[2]

Unlike AI Overviews, which display summaries within conventional search results, AI Mode offers a dedicated conversational experience. It also differs from Gemini, Google's standalone assistant accessed separately from Search. AI Mode thus reaches consumers within an established interface, without requiring a visit to a separate assistant.

Our data do not directly measure AI Mode usage. However, U.S. desktop clickstream evidence indicates that AI Mode accounted for just over 1% of sampled Google Search sessions by early July 2025, involved fewer queries per session than conventional search, and generated no external website visit in approximately 93% of sessions (Harsel, 2025). By May 2026, Google reported more than one billion monthly AI Mode users worldwide (Reid, 2026).

2.2 International Rollout

Google launched public access to AI Mode in the United States on May 20, 2025, India on July 8, and the United Kingdom on July 28 (Google, 2025a,b,f). On August 21, it announced English-language access in more than 180 additional countries and territories, while many European countries remained excluded (Google, 2025c).[3] Availability expanded to more than 40 additional countries, including many across Europe, on October 7 (Google, 2025e). We assign public launch dates from these announcements.

Our main analysis exploits the August 21 expansion, comparing changes in assistant traffic in newly covered countries with those in countries that remained without access throughout the following month. Google Trends search interest in "AI Mode" rises around the respective expansions in treated countries and, subsequently, in the original control group, consistent with the announced rollout timing (Appendix Figure A1).

[2] Google described the transformation of Search through AI as follows: "AI is driving the most significant upgrade of the Google Search experience ever" (Reid, 2026).

[3] On September 8, Google added Hindi, Indonesian, Japanese, Korean, and Brazilian Portuguese (Google, 2025d).

Table 1. Mean daily visits in the four weeks before and after August 21, 2025: treated and comparison countries (group totals, in millions)

	Treated countries (25)			Comparison countries (30)			Difference
Assistant	Before	After	Change (%)	Before	After	Change (%)	(pp)
ChatGPT	53.17	55.40	4.2	43.18	47.74	10.5	−6.4
Gemini	3.78	4.83	27.8	1.44	1.68	17.3	10.5
Other AI	7.15	7.85	9.8	5.69	6.35	11.5	−1.7

Notes: Before and after cover days −28 to −1 and +1 to +28 relative to August 21, 2025. Treated countries are the 25 countries that gained access on that date; comparison countries are the 30 countries without access through September 18, 2025. Visits are mean daily visits summed across the countries of each group, in millions. Change is the percentage growth of summed visits; the difference is the growth in treated countries minus the growth in comparison countries, in percentage points. Growth rates and differences are calculated before rounding.

2.3 Data

We use daily website-traffic estimates from Semrush, an SEO and digital analytics provider, for ten AI assistants in 58 countries. The estimation sample covers June 2024 through September 2025; Appendix Figure A2 extends the descriptive series through September 2026. The data were downloaded in October 2026. These estimates are derived from clickstream panel data and cover desktop and mobile web browsing; native-app use is not included.

Daily observations allow us to examine short-run responses around the rollout before subsequent expansions leave very few countries without access to AI Mode. The websites are `chatgpt.com` (ChatGPT), `gemini.google.com` (Gemini), `claude.ai` (Claude), `copilot.microsoft.com` (Copilot), `deepseek.com` (DeepSeek), `grok.com` (Grok), `meta.ai` (Meta AI), `mistral.ai` (Mistral), `perplexity.ai` (Perplexity), and `qwen.ai` (Qwen).[4]

In the main analysis, we examine traffic to ChatGPT and Gemini separately and aggregate visits to the eight remaining assistants into a third category, Other AI. The appendix also reports estimates for each of these eight assistants individually.

The dataset covers 58 countries, comprising all 38 OECD members and 20 additional markets in Europe, Asia, Latin America, and Africa (Appendix Table A3). Our main analysis compares the 25 countries that received AI Mode in the August 21 expansion with the 30 countries where it remained unavailable through September 30, 2025. The treated countries span Europe, Asia, the Americas, Oceania, and Africa. The comparison group comprises the European Economic Area countries (all except Cyprus, Liechtenstein, and Malta), together with Switzerland, Türkiye, and Australia. The United States, India, and the United

[4] For DeepSeek, Mistral, and Qwen, the measured domain is the provider's main website, which also hosts pages other than its chat interface.

Kingdom received AI Mode earlier and are excluded from this main comparison.

ChatGPT accounts for most measured traffic: across the 28 countries that received AI Mode by September 2025, its share of visits to the ten websites declined from 84 percent in the week before each country's launch to 81 percent in the third week afterward (Appendix Tables A1 and A2). Comparing the four weeks before and after August 21, 2025, total ChatGPT visits grew by 4.2 percent across the 25 treated countries, compared with 10.5 percent across the 30 comparison countries (Table 1). Visits to Other AI grew at similar rates in the two groups (9.8 and 11.5 percent, respectively), while Gemini visits grew faster in treated countries (27.8 versus 17.3 percent).[5]

Roughly four fifths of ChatGPT visits are classified as direct, typically including users who type the address, use a bookmark, or access the site without following a link from another website. Organic search accounts for a further 10-12 percent (Appendix Table A4). Semrush classifies visits as direct whenever no referring website is recorded, so this category also includes visits whose referrer is not observed. Semrush also records referrals from Google AI Mode, which are negligible: in the third week after launch, they account for less than 0.01 percent of ChatGPT visits in treated countries. These referrals capture only direct click-throughs from AI Mode to ChatGPT. Our analysis instead examines whether access to AI Mode changes overall visits to AI assistants, including visits made without following a link from Google.

3 Empirical Strategy

We estimate the effect of AI Mode's public availability using a difference-in-differences (DiD) design, comparing the 25 countries included in the August 21, 2025 expansion with the 30 countries without access through September 18:

$$Y_{ct} = \alpha_{c,d(t)} + \omega_t + \delta\,\text{AI Mode}_c \times \text{Post}_t + \varepsilon_{ct}, \tag{1}$$

where Y_{ct} is log daily visits to ChatGPT, Gemini, or the combined Other AI category, AI Mode_c identifies treated countries, and Post_t identifies post-launch dates. Country-by-weekday fixed effects $\alpha_{c,d(t)}$ absorb recurring weekly patterns, while date fixed effects ω_t absorb common shocks. We compare days -28 to -1 with days 1 to 28, excluding the potentially partial launch day. The same comparison countries remain untreated through September 30, the endpoint of the extended daily event studies. Additional specifications adjust for regional linear trends, national holiday effects, or both, using days -84 to -29 to estimate these adjustments.

[5] These descriptive growth rates are calculated from aggregate visits across countries and need not coincide with the country-level regression estimates.

Identification requires parallel trends in log visits absent the launch, no anticipation, and no spillovers to comparison countries. A key concern is differential seasonality: comparison countries are mostly European, where the end of summer holidays could generate a stronger rebound in work- and study-related ChatGPT use. We assess this concern using a weekday-aligned comparison with 2024, in-time placebo launches, regional trends, and holiday adjustments (Section 4.1).

To examine pre-launch trends and the timing of the response, we estimate:

$$Y_{ct} = \alpha_{c,d(t)} + \omega_t + \sum_{k=-28}^{28} \delta_k \, \text{AI Mode}_c \times \mathbf{1}\{t = g + k\} + \varepsilon_{ct}, \tag{2}$$

where g is the launch date. The estimation window covers days -35 to 28, with days -35 to -29 forming the omitted reference week. Each δ_k measures the treated-comparison gap relative to the corresponding weekday in the omitted reference week.

Both equations use weighted least squares with fixed country weights proportional to mean daily visits to the relevant assistant over days -84 to -29, normalized within treated and comparison groups. Standard errors are based on 1,000 bootstrap replications that resample countries separately within each group.[6] Appendix B.5 reports weight concentration, leave-one-country-out estimates, and in-time placebos.

4 AI Mode and ChatGPT Traffic

Table 2 reports estimates of Equation (1) for ChatGPT visits. Our preferred specification, in Column (1), indicates a statistically significant 5.9 percent decline in visits relative to comparison countries, on average over the four weeks following launch. Columns (2)-(4) adjust for regional linear trends, national public holidays, or both. The estimates remain statistically significant across these specifications, with regional trends yielding a larger estimated decline. Column (5) assigns equal weight to each country and produces a negative, statistically significant estimate of similar magnitude.

Figure 2 presents the corresponding event study. It shows no clear differential pre-launch trend after accounting for country-specific weekly patterns. Visits decline relative to comparison countries over the first two weeks after launch; the gap subsequently narrows but remains negative through day 28.[7] Appendix Figure B1 reports an alternative specification with country fixed effects, using day -1 as the reference.

[6] Each replication draws countries with replacement, preserving group sizes and retaining each selected country's complete daily series. Baseline traffic weights are renormalized within each resampled group. Regional trend and holiday adjustments, where included, are re-estimated using days -84 to -29, before re-estimating the treatment effects. Standard errors are the standard deviations of the bootstrap estimates.

[7] The extended event study shows negative point estimates through day 40 (Appendix Figure B5).

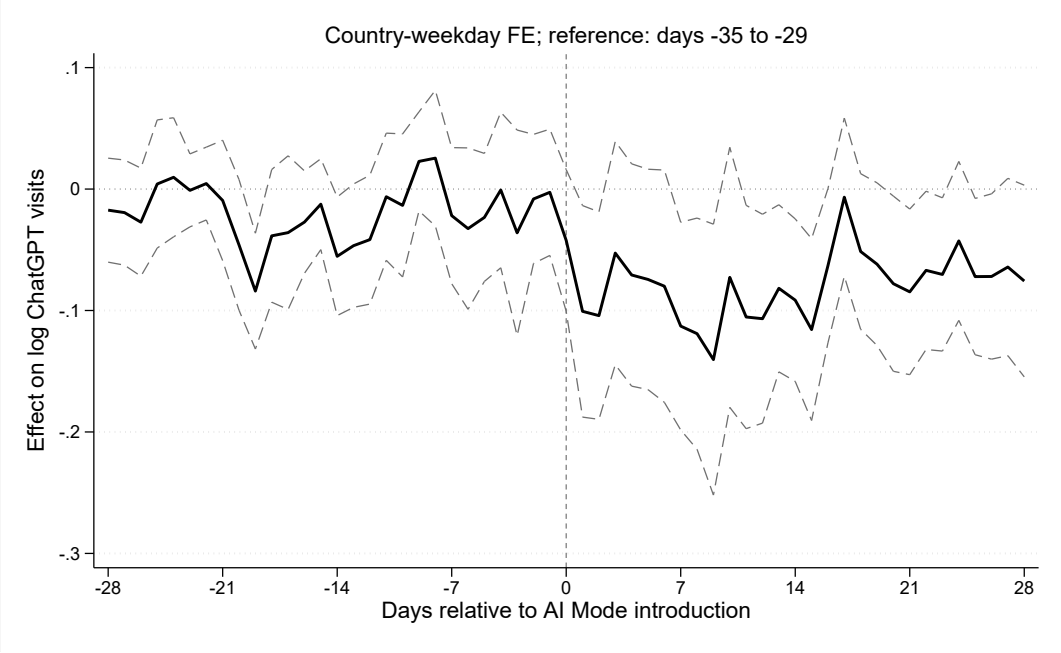


Figure 2. ChatGPT Traffic Around the Introduction of AI Mode

Notes: Traffic-weighted daily event-study estimates with country-by-weekday and date fixed effects. Each coefficient is relative to the same weekday in days −35 to −29; this reference week is omitted from the plot. Dashed lines show pointwise 95% normal-approximation confidence intervals using standard errors from 1,000 country-bootstrap replications, resampling treated and comparison countries separately. The vertical line marks August 21, 2025.

Table 2. AI Mode and ChatGPT visits

	(1)	(2)	(3)	(4)	(5)
AI Mode × Post	−0.061**	−0.076**	−0.058**	−0.077**	−0.055***
	(0.025)	(0.033)	(0.024)	(0.032)	(0.021)
Implied change (%)	−5.88	−7.36	−5.65	−7.43	−5.37
Country-by-weekday FE	Yes	Yes	Yes	Yes	Yes
Country weights	Traffic	Traffic	Traffic	Traffic	Equal
Regional linear trends	No	Yes	No	Yes	No
Holiday adjustment	No	No	Yes	Yes	No
N	3,080	3,080	3,080	3,080	3,080
R^2	0.993	0.992	0.993	0.993	0.996

Notes: Log visits; days 1–28 versus -28 to -1. Traffic weights use days -84 to -29. Regional and holiday components are fitted in that baseline period. Country-by-weekday and date FE. Overall weighted R^2 uses the adjusted outcome and includes FE; N counts country-days. Bootstrap SEs in parentheses (1000 draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

To illustrate the economic magnitude of the effect, the 25 treated countries recorded 55.4 million ChatGPT visits per day over the four post-launch weeks (Table 1). Assuming a common proportional effect across countries, the preferred estimate implies approximately 3.5 million fewer daily visits than in the absence of the launch.[8] For comparison, the standalone Gemini website received 4.8 million daily visits in the same countries over the same period.

Although Google's standalone assistant Gemini was already available, introducing a conversational interface within Google Search substantially reduced traffic to ChatGPT. This result highlights the importance of integration into established interfaces as a dimension of competition among AI assistants, at least in the initial weeks following launch.

4.1 Robustness Checks

The estimated decline in ChatGPT traffic is robust to alternative pre-launch reference periods, exclusion of the first post-launch week, regional trends, national holiday controls, alternative traffic weights, equal country weights, and synthetic difference-in-differences. Extended event studies consider up to 12 pre-launch weeks and a daily window spanning 60 days before to 40 days after launch (Appendix B).

To assess seasonality, we repeat the analysis around the corresponding weekday-aligned date in 2024. Differences between the 2025 and 2024 estimates are negative, similar in magnitude to the baseline effect, and statistically significant at the 10 percent level or better across both weighting schemes and reference periods. The sensitivity analysis of Rambachan and Roth (2023) also supports a negative effect when allowing differential linear trends extrapolated from 4, 8, or 12 pre-launch weeks, bounded nonlinear departures from those trends, and relative-magnitude restrictions with $\bar{M}$ up to 0.20-0.25 (Appendix B.4).

Estimates remain negative when excluding each country in turn. None of the seven monthly placebo launches from January to July 2025 produces an estimate as negative as the actual launch under either weighting scheme. Incorporating the earlier launches in the United States, India, and the United Kingdom also yields negative aggregate effects of similar magnitude. Each cohort is compared with countries that remain untreated throughout its estimation window (Callaway and Sant'Anna, 2021) (Appendix B.6).

Concurrent product releases are another potential concern (Appendix Table A5). Date fixed effects absorb common shocks, but not differential responses across treated and comparison countries. OpenAI released GPT-5 on August 7 (OpenAI, 2025); restricting the pre-launch period to the final week, entirely after that release, yields a similar decline of 6.0 percent (Appendix Table B1, column 3). Google released Gemini

[8] This difference is computed as $V^{\text{post}}[\exp(-\hat{\delta}) - 1]$, with a 95 percent confidence interval of 0.6 to 6.4 million obtained by transforming the bounds $\hat{\delta} \pm 1.96\,\text{SE}$.

2.5 Flash Image on August 26 (Fortin, Vernade, Kampf and Reshi, 2025), whereas the ChatGPT decline is already visible in the first days after AI Mode's launch (Figure 2).

4.2 Traffic Channels

We now examine which traffic channels account for the decline in ChatGPT visits following AI Mode's introduction. Appendix Table B8 applies our baseline difference-in-differences specification separately to direct and organic-search visits. Direct visits, which have no recorded referring website and make up about four fifths of ChatGPT traffic, decline by about the same rate as total visits, a reduction equivalent to 5.5 percentage points of baseline total visits. The estimated change in organic-search visits is small and imprecise. The lower limit of its 95 percent confidence interval corresponds to a decline of about 1.5 percentage points of baseline total visits, less than a quarter of the estimated total decline of 6.6 points. The remaining channels contribute a further decline of about 0.7 percentage points of baseline total visits. These contributions are estimated in visit levels normalized by baseline total traffic, so their sum need not equal the percentage effect obtained from the log specification. Direct visits therefore account for most of the reduction, while organic-search referrals contribute relatively little. The corresponding event studies show no pronounced differential pre-launch trends for either channel (Appendix Figure B10).

Google Trends data on search interest in "ChatGPT" provide complementary evidence (Appendix Table B9 and Figure B11). In the baseline traffic-weighted specification, search interest shows no significant change (-2.0 percent). With equal country weights, the estimate is negative and significant (-7.8 percent), while with regional trends the traffic-weighted estimates are about -9.5 percent but imprecise. Overall, the evidence is consistent with substitution occurring primarily among users who would otherwise have accessed ChatGPT directly. Although AI Mode may also divert organic-search referrals, their relatively small share of ChatGPT traffic limits the importance of this channel.

4.3 Visitors, Visit Frequency, and Duration

The decline in ChatGPT traffic primarily reflects fewer daily visitors, with little change in visit frequency or duration among daily visitors. Appendix Table B10 shows a statistically significant reduction in daily unique visitors that accounts for nearly all of the estimated decline in visits.[9] By contrast, visits per daily visitor and average visit duration change little: their 95 percent confidence intervals range from about -0.8 to $+0.4$ percent and from about -2.2 to $+0.9$ percent, respectively. Constructed total time on the site also

[9] Because the measure counts visitors per day, this is a reduction in visitor-days; it does not distinguish consumers who stopped using ChatGPT from those who used it on fewer days.

declines significantly, by a magnitude similar to that of visits.

The absence of a detectable change in average visit duration suggests that substitution need not be confined to brief information requests. ChatGPT supports a broad range of activities, including information seeking and practical guidance (Chatterji, Cunningham, Deming, Hitzig, Ong, Shan and Wadman, 2025), which overlap with potential uses of AI Mode, such as planning a trip or comparing products. These activities can involve both short exchanges and extended exploration, so their displacement need not substantially change average visit duration. Our aggregate data do not reveal which tasks shift between services, but the findings are consistent with AI Mode replacing some occasions of ChatGPT use across visits of different lengths.

5 Gemini and Other AI Assistants

We next examine whether the introduction of AI Mode affects visits to Gemini and Other AI, which aggregates traffic to the eight smaller AI assistants. Table 3 reports the five main specifications for Gemini (Panel A) and Other AI (Panel B). Neither outcome exhibits a statistically significant four-week average effect in any of the five specifications.

Appendix Figures C1 and C2 present the corresponding event studies using the same specification as Figure 2. For Other AI, the pre-treatment coefficients show no pronounced differential trend, while the post-treatment estimates show no clear, sustained departure from the pre-treatment pattern. This evidence is consistent with the absence of a statistically detectable average effect in Table 3. Gemini exhibits a more pronounced change: relatively stable pre-treatment coefficients are followed by an initial decline and a subsequent increase. The extended event studies in Appendix Figures C3 and C4 suggest that Gemini's late increase persists beyond the first four weeks, whereas Other AI continues to show no comparable shift.

Interpreting Gemini's response requires care, because Google operates both Gemini and AI Mode and several Gemini-specific events fall within the window (Appendix Table A5). Appendix Table C1 separates the first and second two weeks after the rollout, for all countries and excluding Indonesia, in each case with and without regional trends and holiday adjustment. In the full sample, Gemini visits decline by about 13 to 15 percent in weeks 1-2 and rise in weeks 3-4. The early decline is consistent with AI Mode cannibalizing some of Gemini's traffic, although this mechanism cannot be established directly.

The later increase coincides with Google's release of the Gemini 2.5 Flash Image model on August 26, whose popularity grew during September, particularly in parts of Asia (Fortin et al., 2025; Singh, 2025), and with the launch of the lower-priced AI Plus subscription in Indonesia in early September (Kansal, 2025). Ex-

Table 3. AI Mode and visits to Gemini and other AI assistants

Panel A. Gemini					
	(1)	(2)	(3)	(4)	(5)
AI Mode × Post	0.018	−0.003	0.021	−0.003	0.064
	(0.064)	(0.055)	(0.063)	(0.055)	(0.040)
Implied change (%)	1.81	−0.32	2.09	−0.25	6.63
Country-by-weekday FE	Yes	Yes	Yes	Yes	Yes
Country weights	Traffic	Traffic	Traffic	Traffic	Equal
Regional linear trends	No	Yes	No	Yes	No
Holiday adjustment	No	No	Yes	Yes	No
N	3,080	3,080	3,080	3,080	3,080
R^2	0.971	0.972	0.971	0.972	0.980

Panel B. Other AI					
	(1)	(2)	(3)	(4)	(5)
AI Mode × Post	−0.022	0.031	−0.020	0.031	0.003
	(0.040)	(0.049)	(0.040)	(0.048)	(0.033)
Implied change (%)	−2.17	3.18	−1.98	3.10	0.31
Country-by-weekday FE	Yes	Yes	Yes	Yes	Yes
Country weights	Traffic	Traffic	Traffic	Traffic	Equal
Regional linear trends	No	Yes	No	Yes	No
Holiday adjustment	No	No	Yes	Yes	No
N	3,080	3,080	3,080	3,080	3,080
R^2	0.984	0.984	0.985	0.984	0.990

Notes: Log visits; Other AI sums eight services. Days 1–28 versus −28 to −1. Own-service traffic weights and regional/holiday adjustments use days −84 to −29. Country-by-weekday and date FE. Overall weighted R^2 uses the adjusted outcome and includes FE; N counts country-days. Bootstrap SEs in parentheses (1,000 draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

cluding Indonesia reduces the estimated increase during weeks 3-4 and weakens the evidence for a positive late-period effect.[10] Date fixed effects absorb common shocks but not country-specific changes in adoption, so we do not attribute Gemini's late increase to AI Mode.[11]

5.1 Responses of Individual AI Assistants

Appendix Tables C3-C10 report exploratory estimates for each of the eight assistants included in Other AI. Most show no statistically significant effect in any specification. Qwen's estimates are negative and significant only when regional trends are included. Mistral's estimates are negative and significant across all five specifications, suggesting a possible decline in traffic following AI Mode's introduction. However, the traffic-weighted estimate roughly halves when France, Mistral's home market, is excluded. The equal-weight estimate changes little when France is excluded and remains statistically significant. The corre-

[10] By comparison, excluding Indonesia leaves the estimated average effect on ChatGPT negative.

[11] The comparison with 2024 reinforces this caution. Appendix Table C2 shows no statistically significant differences between the average effects estimated for 2024 and 2025. Figure C5 shows that the sharp rebound between weeks 2 and 3 in 2025 is not apparent at the corresponding point in 2024.

sponding window in 2024 also shows a decline, and PPML estimates in levels lose significance when regional trends are included (Appendix C.3). Overall, the evidence for Mistral is suggestive of a negative effect, but less robust than for ChatGPT.

6 Managerial and Policy Implications

Our findings identify distribution through established interfaces as an important dimension of competition among AI assistants. Following AI Mode's introduction, ChatGPT experienced fewer daily website visits and unique visitors, with little change in visits per daily visitor or average visit duration. For standalone providers, these results highlight daily reach as a competitive outcome: pressure from an integrated rival can emerge in direct traffic even when search referrals change little. Monitoring referrals alone may therefore miss an important competitive response. The evidence is strongest for ChatGPT; effects on most smaller assistants remain less conclusive.

For platforms with established consumer interfaces, integration offers a way to distribute AI services without requiring consumers to visit a standalone assistant. Our results suggest that this access can affect demand for standalone rivals, making distribution relevant alongside model capabilities and service quality. For standalone providers, the corresponding managerial challenge is to sustain consumers' reasons to open their service when conversational AI becomes available within products they already use.

Competition among assistants also matters for brands and publishers. Research documents how AI-generated answers affect referrals to external websites (Agarwal and Sen, 2026; Wang et al., 2026; Shi et al., 2026). Our findings highlight an additional consideration: integration can influence which assistants consumers use, potentially shaping where they encounter product information, brand recommendations, and publisher content. For marketers and publishers, changes in the relative use of assistants may therefore alter which services matter for visibility and traffic acquisition. Assessing these downstream effects requires understanding both how consumer demand is distributed across assistants and how each assistant selects, presents, and links to external content.

These findings are also relevant to competition-policy debates over AI integration and access to established distribution channels.[12] At the same time, reduced traffic to a rival does not necessarily imply con-

[12] These debates include the application of the Digital Markets Act's Article 6(5) prohibition on self-preferencing to AI Mode. Related interventions address different aspects of AI distribution: the September 2025 remedies decision in *United States v. Google* restricted exclusive distribution arrangements covering Google's Gemini app (US Department of Justice, 2025), while the Italian Competition Authority ordered Meta to suspend contractual terms excluding competing AI chatbots from WhatsApp (Italian Competition Authority, 2025). The UK Competition and Markets Authority's October 2025 strategic-market-status designation of Google's general search services includes AI Overviews and AI Mode but excludes the standalone Gemini assistant (Competition and Markets Authority, 2025).

sumer harm. Assessing the welfare implications of integration also requires considering its effects on convenience, service quality, consumer choice, and access to information.

7 Conclusion

This paper provides causal evidence on how integrating conversational AI into established digital platforms, particularly search engines, affects demand for standalone assistants. Exploiting the August 2025 expansion of Google's AI Mode, which reached 25 countries in our sample, we find that ChatGPT website visits declined by 5.9 percent on average over the first four weeks relative to countries without access. Direct visits account for most of the estimated decline, suggesting that traffic diversion extends beyond the displacement of search referrals. Daily unique visitors fall by a similar rate, while visits per daily visitor and average visit duration change little. Effects on other assistants are generally less conclusive. These findings highlight distribution as a dimension of competition in generative AI: integration into an established interface can reduce a leading standalone competitor's daily reach.

Our estimates capture the initial response to public availability. Future research can examine whether these effects persist, which consumers and information-seeking tasks shift between services, and how competition among assistants affects referrals to brands and publishers. Our identification strategy could also be extended to examine the effects of integrating AI assistants into established platforms beyond search engines. Linking individual assistant use to subsequent browsing and purchasing behavior would help establish the broader consequences for consumer decisions and the websites that serve them.

References

Agarwal, Saharsh, and Ananya Sen (2026) 'The Impact of Google AI Overviews on Publisher Traffic and User Experience: Evidence from a Field Experiment.' Working paper

Allcott, Hunt, Juan Camilo Castillo, Matthew Gentzkow, Leon Musolff, and Tobias Salz (2025) 'Sources of Market Power in Web Search: Evidence from a Field Experiment.' NBER Working Paper 33410, National Bureau of Economic Research

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager (2021) 'Synthetic Difference-in-Differences.' *American Economic Review* 111(12), 4088–4118

Budaraju, Hema (2025a) 'AI Overviews are now available in over 200 countries and territories, and more than 40 languages.' The Keyword (Google blog). May 20. https://blog.google/products-and-platforms/products/search/ai-overview-expansion-may-2025-update/

_ (2025b) 'We're bringing the helpfulness of AI Overviews to more countries in europe.' The Keyword (Google blog). March 25. https://blog.google/feed/were-bringing-the-helpfulness-of-ai-overviews-to-more-countries-in-europe/

Callaway, Brantly, and Pedro H. C. Sant'Anna (2021) 'Difference-in-Differences with Multiple Time Periods.' *Journal of Econometrics* 225(2), 200–230

Chatterji, Aaron, Thomas Cunningham, David J. Deming, Zoe Hitzig, Christopher Ong, Carl Yan Shan, and Kevin Wadman (2025) 'How people use ChatGPT.' NBER Working Paper 34255, National Bureau of Economic Research. September

Competition and Markets Authority (2025) 'CMA Confirms Google Has Strategic Market Status in Search Services.' Press release. October 10

Demirer, Mert, Andrey Fradkin, Nadav Tadelis, and Sida Peng (2025) 'The Emerging Market for Intelligence: Pricing, Supply, and Demand for LLMs.' NBER Working Paper 34608, National Bureau of Economic Research

Edelman, Benjamin, and Zhenyu Lai (2016) 'Design of Search Engine Services: Channel Interdependence in Search Engine Results.' *Journal of Marketing Research* 53(6), 881–900

Egan, Mark, Dominic Russel, Claire Shi, and Vincent Yao (2026) 'Competition in generative AI: Evidence from consumer demand.' Available at SSRN 7621763

Fortin, Alisa, Guillaume Vernade, Kat Kampf, and Ammaar Reshi (2025) 'Introducing Gemini 2.5 Flash Image, Our State-of-the-Art Image Model.' Google Developers Blog. August 26

Fradkin, Andrey (2025) 'Demand for LLMs: Descriptive Evidence on Substitution, Market Expansion, and Multihoming.' arXiv:2504.15440

Gholami, Samira, Cristiana Firullo, Cristobal Cheyre, and Alessandro Acquisti (2026) 'Beyond Search: LLM Adoption and Web Traffic Concentration.' Working paper, February 14

Google (2025a) 'AI in Search: Going beyond Information to Intelligence.' The Keyword (Google blog). May 20

__ (2025b) 'AI Mode in Google Search Rolling Out in India.' Google India blog. July 8

__ (2025c) 'AI Mode in Search Gets New Agentic Features and Expands Globally.' The Keyword (Google blog). August 21

__ (2025d) 'AI Mode Is Now Available in Five New Languages around the World.' The Keyword (Google blog). September 8

__ (2025e) 'AI Mode Is Now Available in More Languages and Locations around the World.' The Keyword (Google blog). October 7

__ (2025f) 'Google Search: Introducing AI Mode in the UK.' Google UK blog. July 28

Harsel, Luke (2025) 'Google AI Mode's Early Adoption and SEO Impact.' https://www.semrush.com/blog/google-ai-mode-seo-impact/. Accessed October 19, 2026

Italian Competition Authority (2025) 'Meta AI: The italian competition authority orders meta to suspend the terms excluding competing AI chatbots from WhatsApp.' Press release A576. December 24. https://en.agcm.it/en/media/press-releases/2025/12/A576

Kansal, Vikas (2025) 'Lakukan lebih banyak dengan AI: Pertama di dunia, Google AI Plus kini tersedia di Indonesia.' Google Blog. September 4, 2025

Khosravi, Mehrzad, and Hema Yoganarasimhan (2026) 'Impact of AI Search Summaries on Website Traffic: Evidence from Google AI Overviews and Wikipedia.' arXiv:2602.18455, version 6, September 2

Li, Zhuoxin, and Ashish Agarwal (2017) 'Platform Integration and Demand Spillovers in Complementary Markets: Evidence from Facebook's Integration of Instagram.' *Management Science* 63(10), 3438–3458

Meta (2025) 'Europe, meet your newest assistant: Meta AI.' Meta Newsroom. March. https://about.fb.com/news/2025/03/europe-meet-your-newest-assistant-meta-ai/

Microsoft (2025) 'Introducing Copilot Search in Bing.' Microsoft Bing Blog. April 4. https://blogs.bing.com/search/2025/4/Introducing-Copilot-Search-in-Bing/

Missoffe, Sébastien (2026) 'Lancement des aperçus IA et du mode IA dans la recherche Google en france.' Google France blog. July 22. https://blog.google/intl/fr-fr/nouveautes-produits/explorez-obtenez-des-r%C3%A9ponses/recherche-ia-apercus-mode/

Mistral AI (2025) 'Mistral AI Raises 1.7B€ to Accelerate Technological Progress with AI.' Mistral AI news. September 9

OpenAI (2025) 'Introducing GPT-5.' OpenAI blog. August 7

Padilla, Nicolas, H. Tai Lam, Anja Lambrecht, and Brett Hollenbeck (2026) 'The Impact of Early LLM Adoption on Online User Behavior.' *Management Science.* Forthcoming

Pape, Louis-Daniel, and Michelangelo Rossi (2026) 'Is Competition Only One Click Away? The Digital Markets Act's Impact on Google Maps.' *Marketing Science* 45(3), 596–613

Pichai, Sundar (2025) 'Bringing the Best of AI to College Students for Free.' The Keyword (Google blog). August 6

Rambachan, Ashesh, and Jonathan Roth (2023) 'A More Credible Approach to Parallel Trends.' *Review of Economic Studies* 90(5), 2555–2591

Reid, Elizabeth (2026) 'A new era for AI search.' The Keyword (Google blog). May 19. https://blog.google/products-and-platforms/products/search/search-io-2026/

Semrush (2026) 'Traffic Analytics Overview Report.' Semrush Knowledge Base. Accessed October 15, 2026

Shi, Qiaoni, Kai Zhu, and Kai Gu (2026) 'Answering Without Referring: How AI Search Rewrites the Web's Economic Bargain.' arXiv:2607.07652

Singh, Jagmeet (2025) 'India Leads the Way on Google's Nano Banana with a Local Creative Twist.' TechCrunch. September 17

US Department of Justice (2025) 'Department of Justice Wins Significant Remedies Against Google.' Press release. September 2

Wang, Stephanie T., Jeffrey Gleason, Yakov Bart, Christo Wilson, and Danaé Metaxa (2026) 'AI in Search Reduces Publisher Referrals Without Improving User Experience: Experimental Evidence.' arXiv:2608.18352

Zhang, Peibo, Ruomeng Cui, and Dennis J. Zhang (2026) 'The Impact of AI Search on the Online Content Ecosystem: Evidence from Google and Reddit.' arXiv:2605.16428, version 3, June 20

APPENDIX: Integration and Competition in Generative AI: Evidence from Google's AI Mode Rollout

A Descriptive Statistics

Table A1 reports mean daily visits to each assistant category in the 28 adopting countries in the week before launch and in the third week after it, and Table A2 decomposes Other AI into its eight services, with shares computed over all ten websites. Table 1 in the main text compares the four complete weeks before and after August 21, 2025, in treated and comparison countries; its growth rates of summed traffic are a descriptive counterpart of the traffic-weighted estimates, not the estimates themselves. Table A3 lists countries by launch group and region (Panel A) and the comparison countries available to each cohort (Panel B). Table A4 compares the August cohort with its 30 comparison countries on the same calendar days, adding unique visitors and traffic sources. Figure A2 shows daily visits for the same two groups from June 2024 through September 2026.

Table A4 reports more estimated unique visitors than visits for Gemini, implying visits-per-visitor ratios below one. We therefore exclude Gemini's visitor estimates from the analysis of visitors.[13] The displayed ChatGPT and Other AI aggregates do not exhibit this discrepancy, although that does not establish consistency at the country-day level. Other AI's visitor counts sum across services and are not deduplicated across assistants. Section 4.3 uses visitor estimates for ChatGPT only.

The separate modeling of visits and visitors also matters for interpretation: visits per visitor is a ratio of two estimated quantities, and its accounting relationship with visits and visitors holds by construction without independently validating either measure.

The Google AI Mode traffic-source label in Table A4 follows Semrush's classification. The reported shares round to 0.00 percent, except for Other AI in treated countries in the third week after launch, which rounds to 0.01 percent. These are category aggregates, not bounds on referrals to each individual assistant.

[13] Semrush reports visits and unique visitors as separate estimates, each with its own accuracy range (Semrush, 2026). According to Semrush (personal communication, 2026), the two measures are estimated with separate models based on clickstream panel data, without imposing consistency between them, so estimated unique visitors can exceed estimated visits.

Table A1. AI-assistant website traffic around AI Mode introduction

	Week before AI Mode (days −7 to −1)		Third week after AI Mode (days +15 to +21)		Change in
Assistant	Visits (000s)	Share (%)	Visits (000s)	Share (%)	visits (%)
ChatGPT	3,552.2	83.8	3,597.2	80.9	1.3
Gemini	196.0	4.6	276.1	6.2	40.9
Other AI	491.7	11.6	574.0	12.9	16.7
Total	4,239.9	100.0	4,447.3	100.0	4.9

Notes: Source: Semrush. Mean daily visits per country, in thousands, averaged with equal weights across the 28 adopting countries. Because every average uses the same countries, shares of these averages equal shares of summed traffic. Before covers days −7 to −1 and the third week after covers days +15 to +21, relative to each country's public launch of AI Mode. Desktop and mobile website visits combined; native-app usage is excluded. Shares use total visits to the ten tracked websites as their denominator; they are not shares of the entire AI market. Other AI pools eight services.

Table A2. Composition of Other AI traffic

	Week before AI Mode (days −7 to −1)		Third week after AI Mode (days +15 to +21)		Change in
Assistant	Visits (000s)	Share (%)	Visits (000s)	Share (%)	visits (%)
Claude	96.0	2.3	104.7	2.4	9.0
Copilot	59.3	1.4	62.7	1.4	5.7
DeepSeek	126.6	3.0	129.6	2.9	2.4
Grok	91.5	2.2	96.0	2.2	4.9
Meta AI	13.4	0.3	12.3	0.3	−7.8
Mistral	2.0	0.0	2.3	0.1	16.2
Perplexity	90.8	2.1	153.6	3.5	69.2
Qwen	12.1	0.3	12.8	0.3	6.4
Other AI subtotal	491.7	11.6	574.0	12.9	16.7

Notes: Equal-country averages across the 28 adopting countries, comparing the week before AI Mode's launch (days −7 to −1) with the third week after it (days +15 to +21). Visits are in thousands per day. Shares are percentages of traffic to all ten tracked websites, not shares within Other AI. The subtotal matches the Other AI row in Table A1.

Table A3. Countries by rollout cohort and comparison sets

Panel A. Countries by public launch date

Launch group	Countries	*N*
May 20, 2025	United States	1
July 8, 2025	India	1
July 28, 2025	United Kingdom	1
August 21, 2025	*Europe:* Albania, Bosnia & Herzegovina, Serbia, Ukraine; *Asia:* Armenia, Bangladesh, Georgia, Indonesia, Israel, Japan, Pakistan, Philippines, South Korea, Taiwan, Thailand, Vietnam; *Americas:* Brazil, Canada, Chile, Colombia, Costa Rica, Mexico, Peru; *Oceania:* New Zealand; *Africa:* South Africa	25
After Sept. 30, 2025	*Europe:* Austria, Belgium, Bulgaria, Croatia, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland; *Asia:* Turkiye; *Oceania:* Australia	30

Panel B. Comparison countries by treated cohort

Treated cohort	Comparison countries	*N*
United States	India, United Kingdom, August 21 group, and post-September group	57
India	August 21 group and post-September group	55
United Kingdom	Post-September group	30
August 21 group	Post-September group	30

Notes: Launch dates mark the start of public rollout. A country serves as a comparison for a cohort only if it remains untreated through day 28 of that cohort's window. The main analysis uses the August 21 group and the post-September group; the other cohorts appear in Table B7.

A.1 Search Interest Around the AI Mode Rollout

Figure A1 plots average weekly Google Trends search interest in "AI Mode" for the treated and control countries. Search interest rises in treated countries around the August 21 expansion, followed by an increase in the original control group around the October expansion. These patterns are consistent with the announced rollout timing.

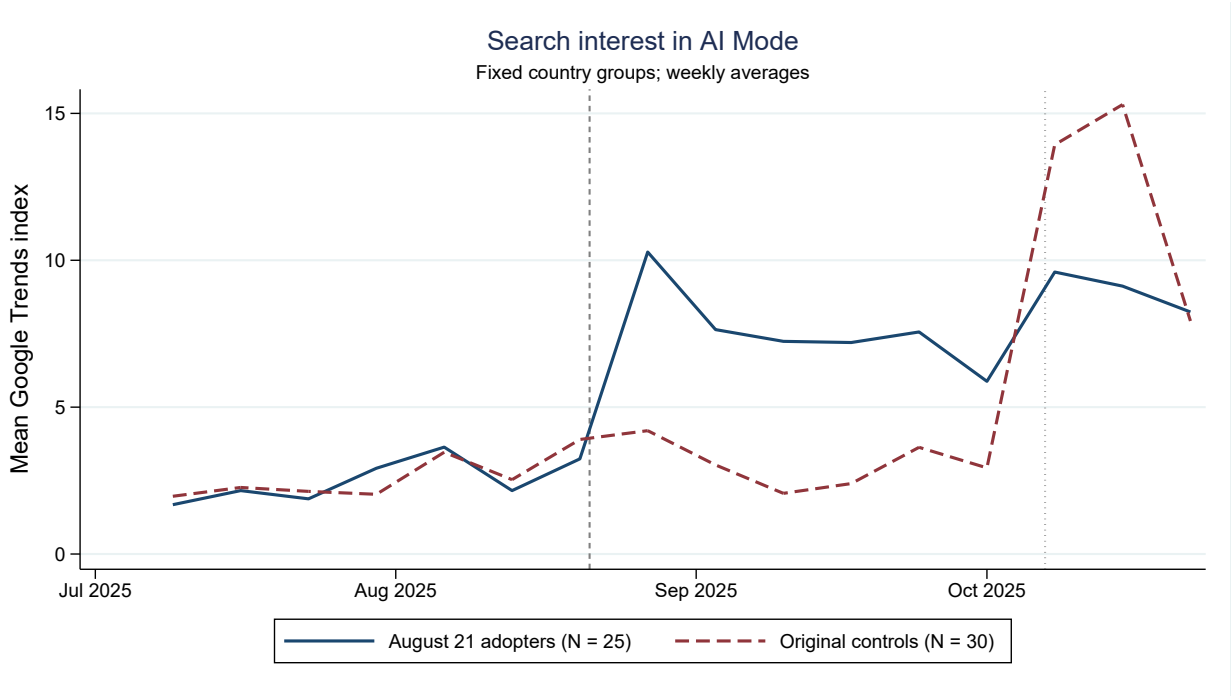


Figure A1. Search Interest Around the AI Mode Rollout

Notes: Lines show equal-weighted country averages of weekly Google Trends indices for "AI Mode" from July through October 2025. Country groups match the main analysis and remain fixed throughout. Vertical lines mark August 21 and October 7. The October expansion covered many, but not all, original control countries. Indices measure relative search interest rather than absolute search volumes.

Table A4. Visits, visitors, and traffic sources: August cohort and comparison countries

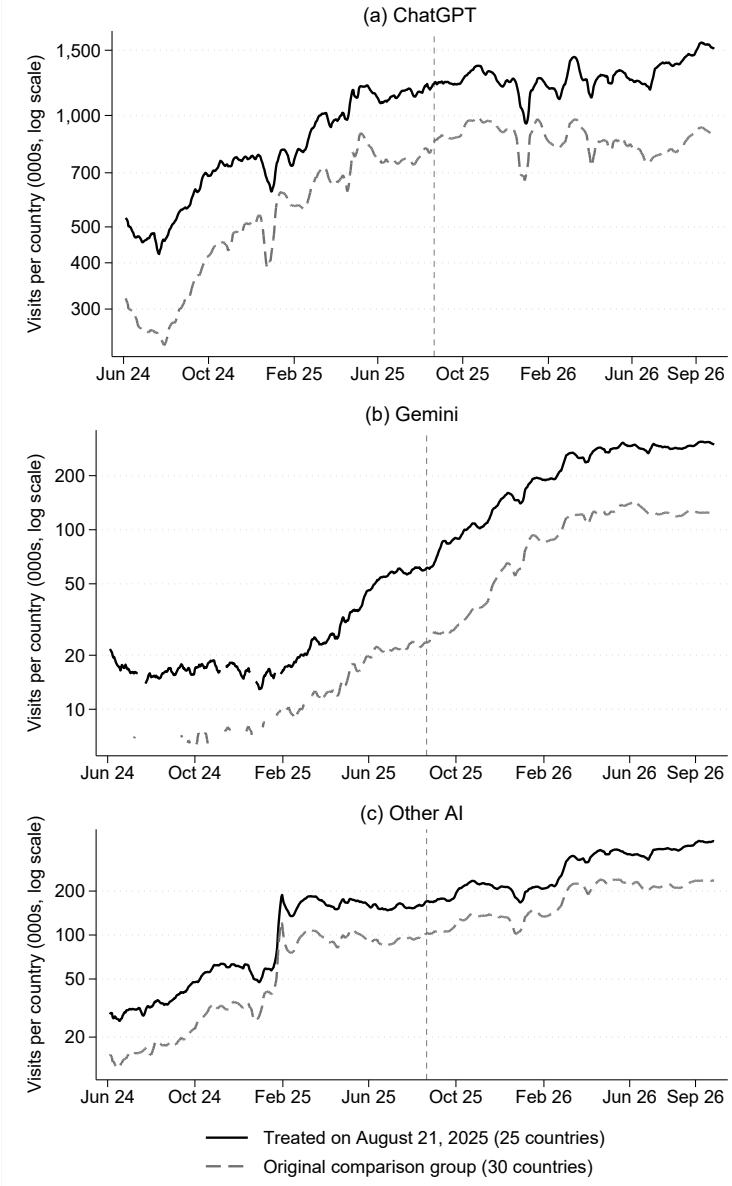
	August 21 cohort		Comparison countries	
	Week before	Third week after	Week before	Third week after
Panel A. ChatGPT				
Daily visits (000s)	2,152.6	2,235.6	1,456.8	1,592.8
Daily unique visitors (000s)	1,215.5	1,273.3	847.0	929.5
Visits per visitor	1.77	1.76	1.72	1.71
Direct (% of visits)	79.3	78.7	81.8	81.4
Organic search (% of visits)	10.6	12.2	11.2	11.6
Google AI Mode (% of visits)	0.00	0.00	0.00	0.00
Panel B. Gemini				
Daily visits (000s)	149.0	211.3	47.6	55.9
Daily unique visitors (000s)	190.4	268.0	69.7	87.1
Visits per visitor	0.78	0.79	0.68	0.64
Direct (% of visits)	92.1	91.8	93.4	93.5
Organic search (% of visits)	2.2	3.1	2.0	1.6
Google AI Mode (% of visits)	0.00	0.00	0.00	0.00
Panel C. Other AI				
Daily visits (000s)	293.6	324.2	191.4	216.5
Daily unique visitors (000s)	178.1	198.8	116.8	135.0
Visits per visitor	1.65	1.63	1.64	1.60
Direct (% of visits)	78.7	79.9	80.6	79.8
Organic search (% of visits)	12.2	11.5	12.3	12.7
Google AI Mode (% of visits)	0.00	0.01	0.00	0.00

Notes: Source: Semrush. The 25 countries that received AI Mode on August 21, 2025, and the 30 countries without access through September 18, 2025, on the same calendar days: the week before launch (August 14–20) and the third week after it (September 5–11). Values are equal-country averages of seven-day country means; ratios and shares are computed from these averages. Traffic sources are shares of website visits; the remaining visits come from other channels. Other AI sums eight services, so its unique visitors count a person once for each service visited. Visits exclude native-app usage.

A.2 Contemporaneous Events

Table A5 lists product launches and promotions around the August 21, 2025 rollout that could affect traffic to the assistants we study. Date fixed effects absorb events common to all countries, so an event matters for our estimates only to the extent that its effect differed between treated and comparison countries. Section 4.1 discusses GPT-5 and the Gemini image model, Section 5 discusses the Gemini-specific events, and Appendix C.3 discusses the Mistral funding announcement.

Figure A2. Daily visits to AI assistants, June 2024 to September 2026



Notes: Source: Semrush. Daily visits per country for the 25 countries that received AI Mode on August 21, 2025 (solid) and the 30 countries that received it after September 2025 (dashed), in thousands, on a log scale. Each line is the geometric mean across countries (the exponential of the equal-country mean of log visits), smoothed by averaging logs over a centered seven-day window before exponentiation. A group-day is omitted if any country has an undefined log, including a zero-visit observation; the smoother requires seven valid full-group days, and lines do not bridge gaps. Stable weekday cycles average out over complete weeks. Other AI sums eight services. Desktop and mobile website visits combined; native-app usage is excluded. The comparison countries gained access from October 2025 onward, so differences between the two groups after that date are descriptive only.

Table A5. Contemporaneous events around the August 21, 2025 rollout

Date (2025)	Day	Event	Countries in our sample
August 6	−15	Google offers its AI Pro plan free to university students (Pichai, 2025)	Brazil, Indonesia, Japan, South Korea (treated)
August 7	−14	OpenAI releases GPT-5 (OpenAI, 2025)	All
August 21	0	AI Mode expands to more than 180 countries and territories (Google, 2025c)	25 treated countries
August 26	5	Google releases Gemini 2.5 Flash Image ("Nano Banana"); adoption grows during September, particularly in parts of Asia (Fortin et al., 2025; Singh, 2025)	All
September 4	14	Google AI Plus subscription launches (Kansal, 2025)	Indonesia (treated)
September 8	18	AI Mode adds Hindi, Indonesian, Japanese, Korean, and Brazilian Portuguese (Google, 2025d)	Brazil, Indonesia, Japan, South Korea (treated)
September 9	19	Mistral announces a €1.7 billion funding round (Mistral AI, 2025)	All (French company)

Notes: Days are relative to August 21, 2025; the four-week estimation window covers days −28 to 28. School and university terms also resume in many countries during this period, at dates that differ across countries.

B The Effect of AI Mode on ChatGPT Traffic

This appendix reports the robustness checks summarized in Section 4.1. All estimates use the 25 countries treated on August 21, 2025, and the 30 comparison countries, unless stated otherwise.

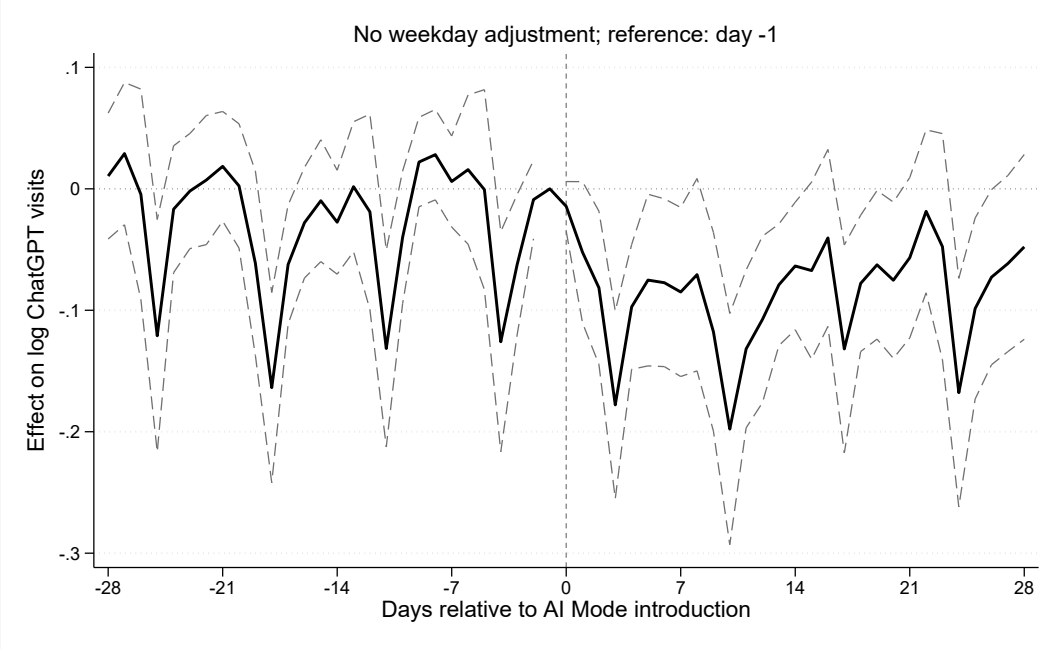


Figure B1. ChatGPT Traffic Around the Introduction of AI Mode: Without Weekday Adjustment

Notes: Traffic-weighted daily event-study estimates with country and date fixed effects, without country-specific weekday effects. The outcome is log daily ChatGPT visits. Country weights are proportional to mean daily visits on days −84 to −29, normalized within each group. The estimation period covers days −28 to 28, with day −1 (August 20, 2025) as the omitted reference. Dashed lines show pointwise 95% normal-approximation confidence intervals using standard errors from 1,000 country-bootstrap replications, resampling treated and comparison countries separately. The vertical line marks August 21, 2025.

B.1 Different Specifications

Reference period and post-launch window. Table B1 compares the preferred four-week reference period with two alternatives: days −35 to −29, the reference week of the main event study, and days −7 to −1, the final pre-launch week, which lies entirely after the release of GPT-5 on August 7. Column (4) excludes the first post-launch week and averages days 8-28. These specifications assess sensitivity to the reference period and to the initial rollout days.

Holidays. Column (3) of Table 2 adjusts for national public holidays.[14] Column (4) adjusts for both regional trends and holidays. Figure B2 shows the event studies for columns (2) to (4).

Regional trends. Column (2) of Table 2 allows linear trends to differ across Europe (the reference), Asia, the Americas, Oceania, and Africa. Slopes are estimated over days −84 to −29, conditional on country-by-weekday and calendar-date fixed effects, and the series is detrended before estimation. Thus, outcomes in the main event window do not enter the estimation of these trends. Because the comparison countries are mostly European, the non-European slopes come mainly from treated countries' pre-launch data, and those for Africa and Oceania are based on one country (South Africa) and two countries (Australia and New Zealand), respectively. Armenia, Georgia, Israel, and Türkiye are classified in Asia.

Weights and synthetic difference-in-differences. Table B2 considers common-period weights from March 25 to May 19, 2025, weights capped at 10 percent within each group, and equal weights. The cap threshold is re-estimated in each bootstrap replication and applies to each resampled country copy; repeated copies of one original country can together exceed 10 percent. Table B3 repeats the timing and regional-trend checks with equal weights, and Figure B3 shows the corresponding event study. Synthetic difference-in-differences (Arkhangelsky, Athey, Hirshberg, Imbens and Wager, 2021) uses unadjusted log visits over 56 pre-launch and 28 post-launch days, excluding day 0, and weights treated countries equally. Its standard error is based on 1,000 replications of the package's bootstrap.

[14] We use `python-holidays` (version 0.106) with observed holidays and no subnational or school calendars. A common holiday effect is estimated on days −84 to −29, conditional on country-by-weekday and date effects, and its predicted contribution is removed before estimation; the bootstrap repeats both steps. Routine Sunday holiday entries for Sweden are excluded because their recurring effects are absorbed by the country-by-weekday fixed effects.

Table B1. ChatGPT visits: reference period and post-launch window

	(1)	(2)	(3)	(4)
AI Mode × Post	−0.061**	−0.080***	−0.062***	−0.059**
	(0.025)	(0.028)	(0.022)	(0.026)
Implied change (%)	−5.88	−7.68	−6.01	−5.72
Reference period	Days −28 to −1	Days −35 to −29	Days −7 to −1	Days −28 to −1
Post-launch days	1–28	1–28	1–28	8–28
N	3,080	1,925	1,925	2,695
R^2	0.993	0.992	0.993	0.993

Notes: Traffic weights; reference and post periods are shown in the table. Country-by-weekday and date FE. Overall weighted R^2 uses the adjusted outcome and includes FE; N counts country-days. Bootstrap SEs in parentheses (1000 draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B2. ChatGPT visits: country weights and synthetic difference-in-differences

	(1)	(2)	(3)	(4)
AI Mode × Post	−0.062**	−0.061**	−0.055***	−0.063**
	(0.025)	(0.026)	(0.021)	(0.025)
Implied change (%)	−6.04	−5.90	−5.37	−6.12
Country weights	Common period	Capped	Equal	Equal (treated)
Estimator	DiD	DiD	DiD	Synthetic DiD
N	3,080	3,080	3,080	4,620
R^2	0.993	0.992	0.996	—

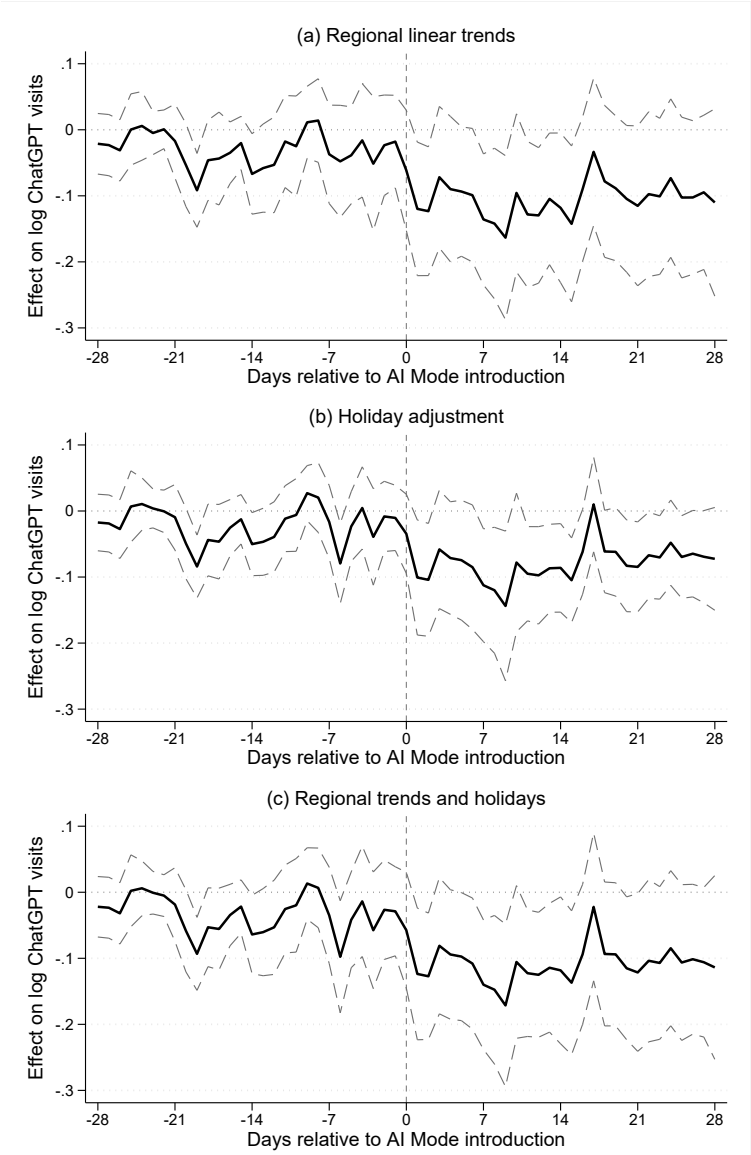
Notes: Common-period weights: March 25–May 19. Capped weights are recalculated in each draw. Synthetic DiD uses unadjusted log visits, 56 pre-days, 28 post-days and its package bootstrap; no separate weekday adjustment is applied and its R-squared is not applicable. Columns (1)–(3) include country-by-weekday and date fixed effects. Overall weighted R^2 uses the adjusted outcome and includes FE; N counts country-days. Bootstrap SEs in parentheses (1000 draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B3. ChatGPT visits: equal country weights

	(1)	(2)	(3)	(4)
AI Mode × Post	−0.055***	−0.082***	−0.099***	−0.046**
	(0.021)	(0.029)	(0.028)	(0.020)
Implied change (%)	−5.37	−7.86	−9.39	−4.47
Reference period	Days −28 to −1	Days −28 to −1	Days −35 to −29	Days −7 to −1
Regional linear trends	No	Yes	No	No
N	3,080	3,080	1,925	1,925
R^2	0.996	0.996	0.996	0.997

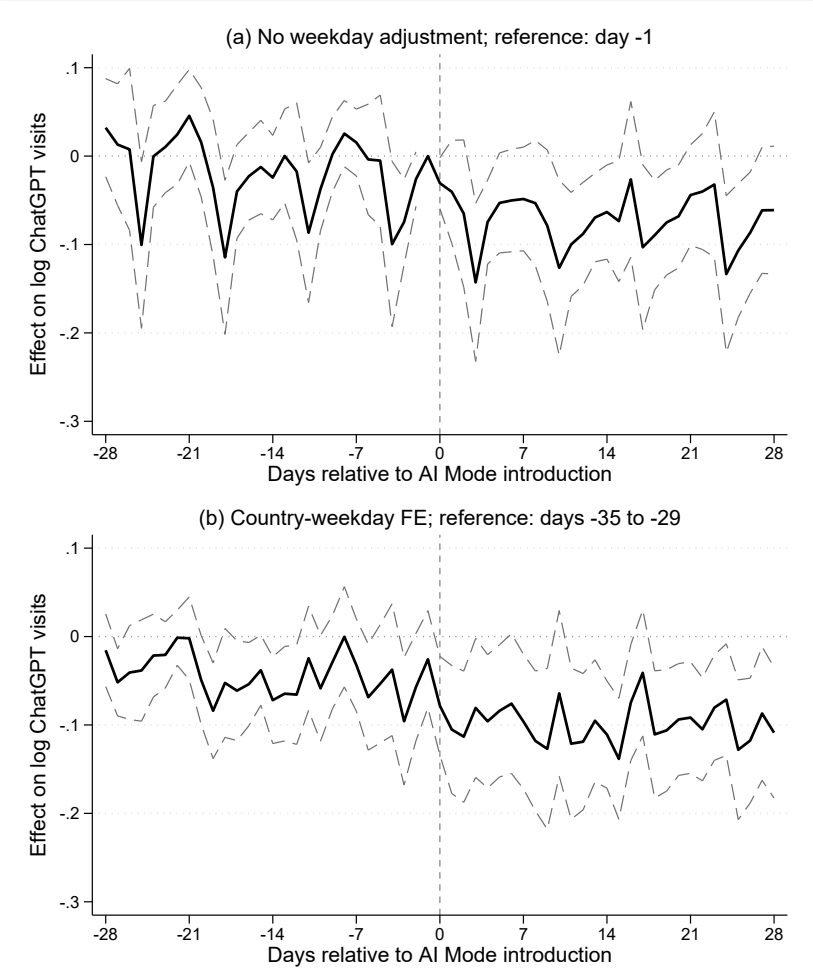
Notes: Equal country weights; reference periods and adjustments are shown in the table. Country-by-weekday and date FE. Overall weighted R^2 uses the adjusted outcome and includes FE; N counts country-days. Bootstrap SEs in parentheses (1000 draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure B2. ChatGPT traffic: regional trends and holidays



Notes: Daily event studies for columns (2) to (4) of Table 2: regional linear trends (a), holiday adjustment (b), and both (c); traffic weights. Each coefficient is relative to the same weekday in days −35 to −29, as in Figure 2. Dashed lines show pointwise 95% confidence limits from 1,000 country-bootstrap replications.

Figure B3. ChatGPT traffic: equal country weights



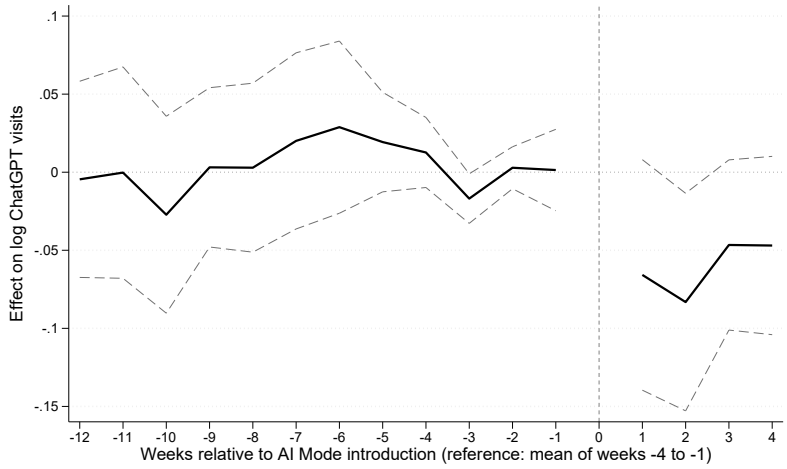
Notes: Daily event studies with equal country weights. Panel (a) uses country and date fixed effects without weekday adjustment, with day −1 as the reference. Panel (b) uses country-by-weekday and date fixed effects, with each coefficient relative to the same weekday in days −35 to −29. In both panels, the average coefficient over days 1 to 28 minus the average over days −28 to −1 equals the four-week estimate in column (1) of Table B3. Dashed lines show pointwise 95% confidence limits from 1,000 country-bootstrap replications.

B.2 Extended Event Windows

Weekly event study. Figure B4 replaces daily indicators with complete-week indicators, extending the pre-launch window to 12 weeks. Coefficients are normalized to average zero over weeks −4 to −1. The average coefficient over the four post-launch weeks therefore equals the preferred estimate.

Extended daily event study. We also extend the daily event study to 60 days before and 40 days after launch. The longer window provides additional evidence on the pre-launch path and on whether the estimated decline persists beyond the first four weeks. Coefficients use the same weekday-specific reference period as the main daily event study, allowing direct comparison over their common window.

Figure B4. ChatGPT traffic: complete-week event study



Notes: Preferred specification with complete weeks over the 12 weeks before launch; coefficients are normalized to average zero over weeks −4 to −1 (days −28 to −1), and the launch day is excluded. Weekly coefficients do not depend on the weekday adjustment. Dashed lines show pointwise 95% confidence limits from 1,000 country-bootstrap replications.



Figure B5. ChatGPT: extended event study

Notes: Preferred traffic-weighted August 21, 2025 specification on days -60 to +40 (June 22–September 30). The original sample of 25 treated and 30 comparison countries is unchanged; comparisons remain untreated throughout. Outcome transformation: log. Baseline traffic weights use days -84 to -29. Country-by-weekday and date fixed effects are included. Each day is relative to the same weekday in days -35 to -29; those reference dates are not plotted. Dashed lines show pointwise 95% normal-approximation intervals from 1000 country-bootstrap draws, stratified by treatment group, with group weights renormalized in each draw. The additional pre-launch coefficients are evaluated using the same fixed baseline weights and reference week. The main four-week estimate is unchanged.

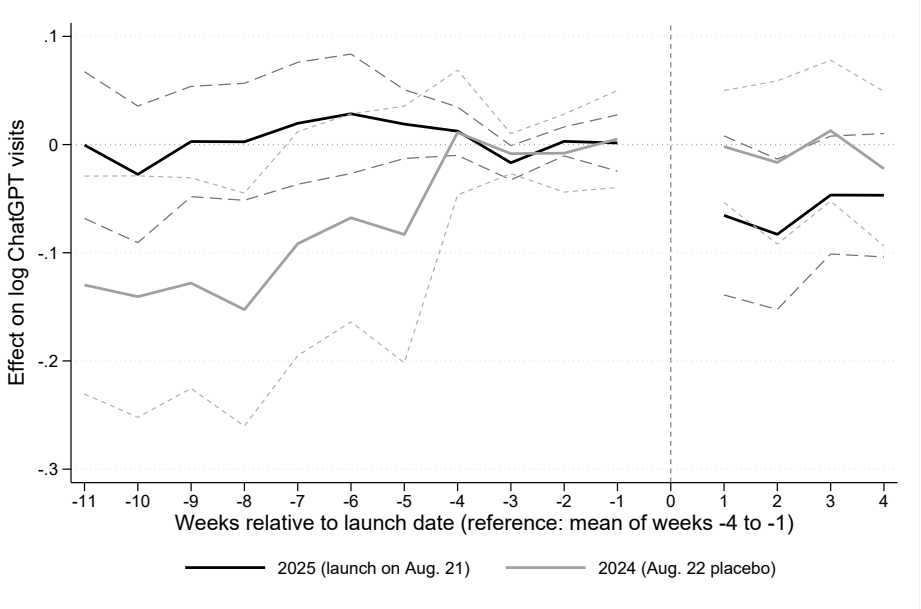
B.3 Seasonality: The Corresponding Window in 2024

Table B4 and Figure B6 repeat the analysis around August 22, 2024, exactly 52 weeks before the actual launch. This alignment preserves weekdays rather than calendar dates. Both years retain the four-week pre-launch and post-launch estimation windows. We use a slightly shorter training period in both years,[15] so the resulting 2025 estimates differ slightly from those in Table 2.

In Panel A, weights are each country's mean daily visits on days −77 to −29 of that year's window, so each year's comparison reflects that year's distribution of traffic. Panel B instead uses the same weights in both years, based on each country's mean daily visits from March 25 to May 19, 2025. This holds the weighting scheme fixed despite changes in countries' relative traffic between years. Both years use the same bootstrap draws of countries, yielding a joint bootstrap distribution for the difference between years.

Panels A and B report the estimates for each year and their paired differences under both reference periods. The differences are negative and similar in magnitude across weighting schemes. With the four-week reference period, they are statistically significant at the 10 percent level in both panels. With the last pre-launch week as the reference, they are significant at the 5 percent level with year-specific weights (Panel A) and at the 10 percent level with common weights (Panel B).

Figure B6. ChatGPT traffic: 2025 launch window and the corresponding window in 2024



Notes: Weekly event studies for the 25 treated countries relative to the 30 comparison countries around August 21, 2025 (black) and around Thursday, August 22, 2024 (gray), when no launch occurred, normalized to average zero over weeks −4 to −1. Countries are weighted by their mean daily ChatGPT visits on days −77 to −29 of each year's window. Dashed lines show pointwise 95% confidence limits for each year. See Table B4, Panel A.

[15] Because the daily data begin on June 1, 2024, we shorten the earlier training period to days −77 to −29 in both years.

Table B4. ChatGPT visits: the weekday-aligned window in 2024

	2025: AI Mode launch (1)	2024: placebo (2)	Difference (3)
Panel A. Matched countries, weights from each year's traffic			
Days 1–28 vs. days −28 to −1	−0.060**	−0.007	−0.054*
	(0.025)	(0.027)	(0.028)
N	3,080	3,080	6,160
R^2	0.993	0.989	—
Days 1–28 vs. days −7 to −1	−0.062***	−0.012	−0.050**
	(0.022)	(0.019)	(0.024)
N	1,925	1,925	3,850
R^2	0.993	0.992	—
Panel B. Matched countries, the same weights in both years			
Days 1–28 vs. days −28 to −1	−0.062**	−0.011	−0.051*
	(0.025)	(0.026)	(0.027)
N	3,080	3,080	6,160
R^2	0.993	0.990	—
Days 1–28 vs. days −7 to −1	−0.063***	−0.016	−0.047*
	(0.023)	(0.019)	(0.024)
N	1,925	1,925	3,850
R^2	0.993	0.992	—

Notes: Matched countries, August 21, 2025 versus August 22, 2024. Paired country-bootstrap SEs; the difference has no separate regression R-squared, and its N sums both year samples. Panel A uses year-specific weights; Panel B uses common-period weights. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

B.4 Sensitivity to Violations of Parallel Trends

Table B5 and Figure B7 assess the sensitivity of the average effect over weeks 1-4, relative to week −1, to departures from parallel trends, using Rambachan and Roth (2023). The main concern in our setting is a differential trend, for instance if ChatGPT adoption had been growing at different rates in treated and comparison countries before the launch. Allowing for a linear differential trend extrapolated from the pre-launch weeks leaves the effect negative and statistically significant, whether the trend is estimated from 12, 8, or 4 weeks. With 12 pre-launch weeks, the trend-adjusted 95 percent confidence interval is [−0.134, −0.055], further from zero than the conventional interval of [−0.106, −0.018].

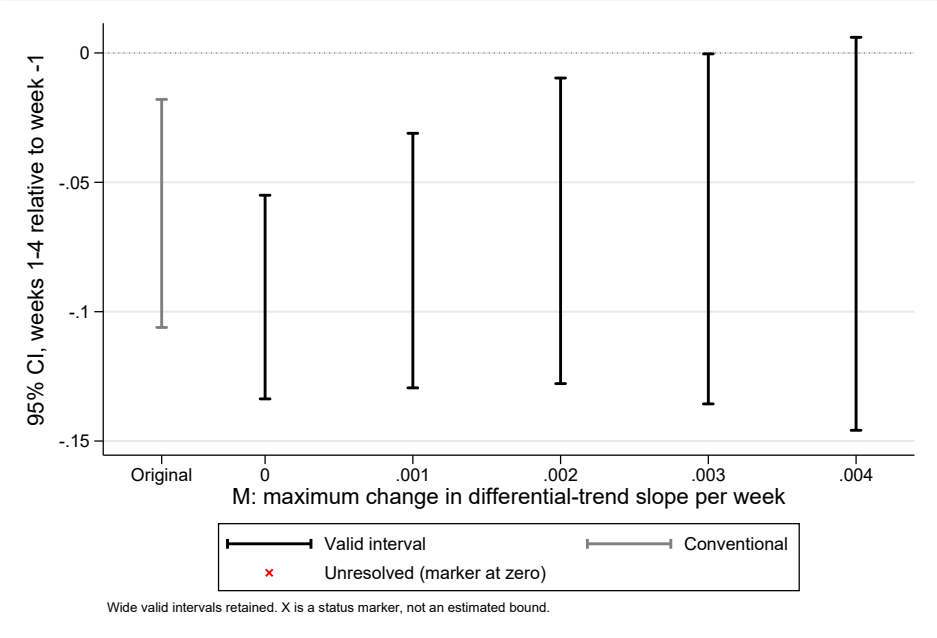
The result is also robust to nonlinear departures from a linear trend. Under the smoothness restriction, the effect remains significant when the slope of the differential trend may change by up to 0.003 log points between consecutive weeks with 12 or 8 pre-launch weeks, and by up to 0.001 with 4. Under the relative-magnitude restriction, which bounds week-to-week changes in the post-launch differential by a multiple $\bar{M}$ of the largest such change before launch, the effect remains significant at least through $\bar{M} = 0.20$ with 12 pre-launch weeks and $\bar{M} = 0.25$ with 8 or 4.

Table B5. Sensitivity to violations of parallel trends

Pre-launch weeks	Estimate (SE)	Conventional 95% CI	Linear trend CI, $M = 0$	Verified through M, smoothness	Verified through $\bar{M}$, rel. magnitude
12	−0.062 (0.022)	[−0.106, −0.018]	[−0.134, −0.055]	0.0030	0.20
8	−0.062 (0.022)	[−0.106, −0.018]	[−0.128, −0.033]	0.0030	0.25
4	−0.062 (0.022)	[−0.106, −0.018]	[−0.115, −0.003]	0.0010	0.25

Notes: HonestDiD sensitivity intervals for weeks 1–4 versus week -1, using country-bootstrap covariances. Verified-through values stop at zero inclusion or an unresolved grid point. These are sensitivity summaries, not separate regressions.

Figure B7. ChatGPT traffic: sensitivity to a nonlinear differential trend



Notes: Robust 95% confidence intervals (Rambachan and Roth, 2023) for the average effect over weeks 1-4 relative to week −1, using the preferred specification with 12 pre-launch weeks. The first interval is the conventional one; $M = 0$ allows a linear differential trend extrapolated from the pre-launch weeks, and $M > 0$ bounds the absolute change in its slope between consecutive weeks by M. Wide valid intervals are retained. See Table B5.

B.5 Weight Concentration and Diagnostic Benchmarks

Table B6 reports the concentration of the traffic weights and two diagnostic benchmarks that do not rely on the bootstrap. With traffic weights, the effective number of countries is 13.3 out of 25 in the treated group and 13.6 out of 30 in the comparison group. When we drop one country at a time from either group, the estimate ranges from −0.070 (dropping Mexico) to −0.045 log points (dropping Vietnam). The permutation benchmark reassigns treatment labels to 25 of the 55 countries at random 1,000 times, keeping each country's weight; the two-sided p-value is 0.045 with traffic weights and 0.016 with equal weights. Treatment was not randomly assigned, and the comparison group is mostly European, so this benchmark describes the dispersion of the estimator under random labels rather than providing design-based inference.

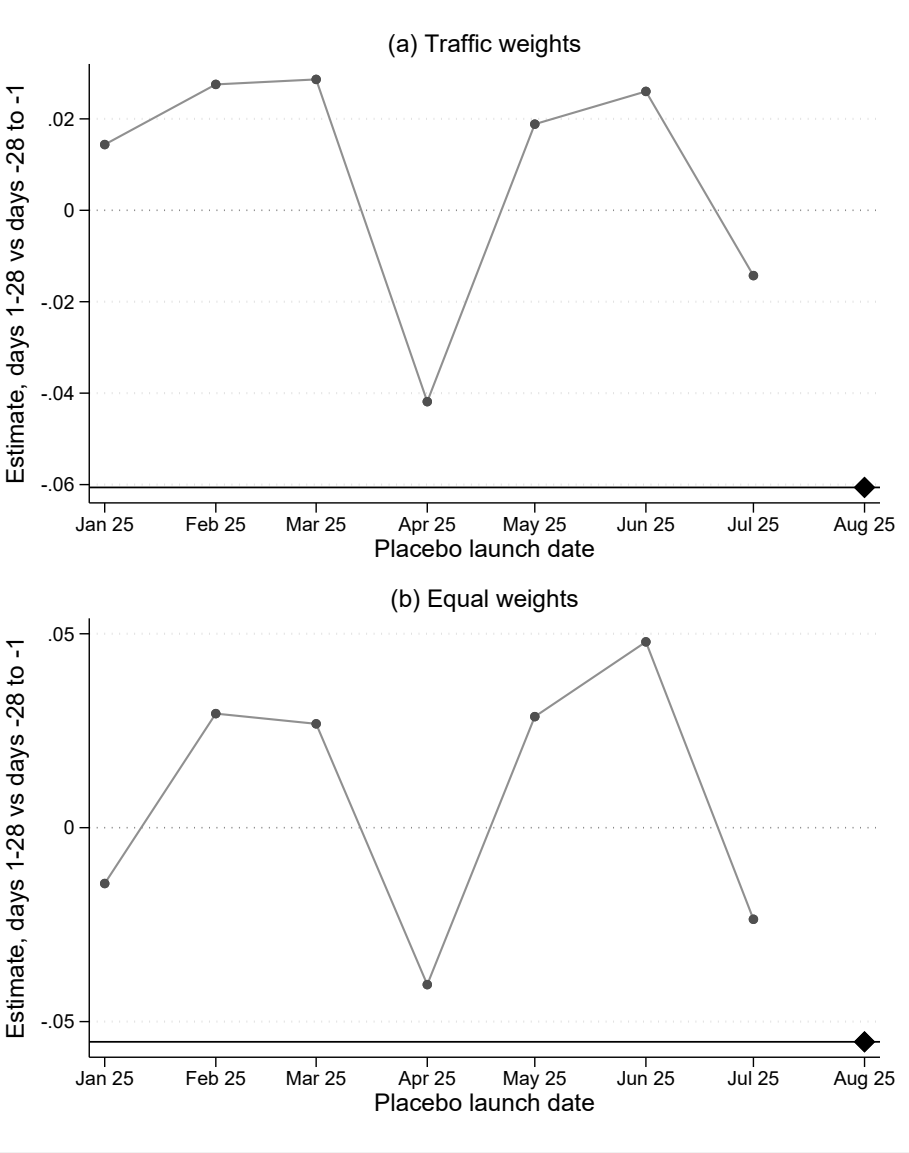
The in-time placebos in Table B6 and Figure B8 place a fake launch on the 21st of each month from January to July 2025, matching the day of the month of the actual launch. Each uses the full 84-day training and pre-launch period and a 28-day post-launch period ending before August 21, 2025; the July 21 window ends on August 18. None of these seven dates produces an estimate as negative as the actual one under either weighting scheme: the traffic-weighted estimates range from −0.042 (April 21) to 0.029, and the equal-weight estimates range from −0.040 to 0.048.

Table B6. ChatGPT visits: weight concentration and alternative inference

	Traffic weights (1)	Equal weights (2)
Effective treated countries (of 25)	13.3	25
Effective comparison countries (of 30)	13.6	30
Largest treated weight share	0.17 (Brazil)	0.04
Three largest treated weight shares	0.34	0.12
Largest comparison weight share	0.16 (Germany)	0.03
Three largest comparison weight shares	0.34	0.10
Leave-one-country-out estimates: range	[−0.070, −0.045]	[−0.062, −0.047]
dropping (lowest, highest)	Mexico, Vietnam	Luxembourg, Vietnam
Treatment-label permutations: two-sided p	0.045	0.016
In-time placebos at or below the estimate	0 of 7	0 of 7

Notes: Effective countries are inverse squared weight shares. LOO drops one country. Permutations are descriptive benchmarks, not randomized-treatment tests; monthly placebo windows overlap.

Figure B8. ChatGPT traffic: monthly in-time placebo estimates



Notes: Point estimates for fake launch dates on the 21st of each month from January to July 2025, for the 25 treated and 30 comparison countries, with the preferred estimator and traffic (a) or equal (b) weights. Each fake launch uses days −84 to −29 for weights and weekday means and compares days 1-28 with days −28 to −1; all post-launch days precede August 21, 2025. The black diamond (at August 21, 2025) and the horizontal line mark the actual estimate. Overlapping windows are not independent.

B.6 Staggered Rollout and the United States

Table B7 adds the three earlier single-country launches to the August expansion: the United States (May 20, 2025), India (July 8), and the United Kingdom (July 28). For each cohort, we estimate Equation (1) with its own launch date, using countries that remain untreated through day 28 of its window as comparisons. Already-treated countries therefore never serve as controls (Callaway and Sant'Anna, 2021), and the comparison sets differ across cohorts (Table A3, Panel B). Cohort estimates compare days 1-28 with days −28 to −1 and are aggregated using weights proportional to each treated cohort's baseline traffic (column 1) or its number of treated countries, (1, 1, 1, 25)/28 (column 2). In each bootstrap replication, countries are resampled jointly across cohort windows, with each country's resampling multiplicity applied to every window in which it appears. The three single-country cohorts are held fixed, so inference is conditional on them.

We do not use these cohorts in the main specification for three reasons. First, the US public launch coincided with a major international expansion of AI Overviews.[16] Because the May expansion changed access in some comparison countries, the US contrast may combine the response to AI Mode with contemporaneous changes in AI Overviews availability. Second, opt-in Labs access began on March 5 in the United States and on June 24 in India, so part of any response may precede the public launch. Third, each of these earlier cohorts contains only one country, leaving its estimate exposed to country-specific shocks whose uncertainty is not captured by the bootstrap.

The pooled estimate is −0.050 with traffic weights and −0.053 with equal weights. Figure B9 shows the pooled daily event studies.

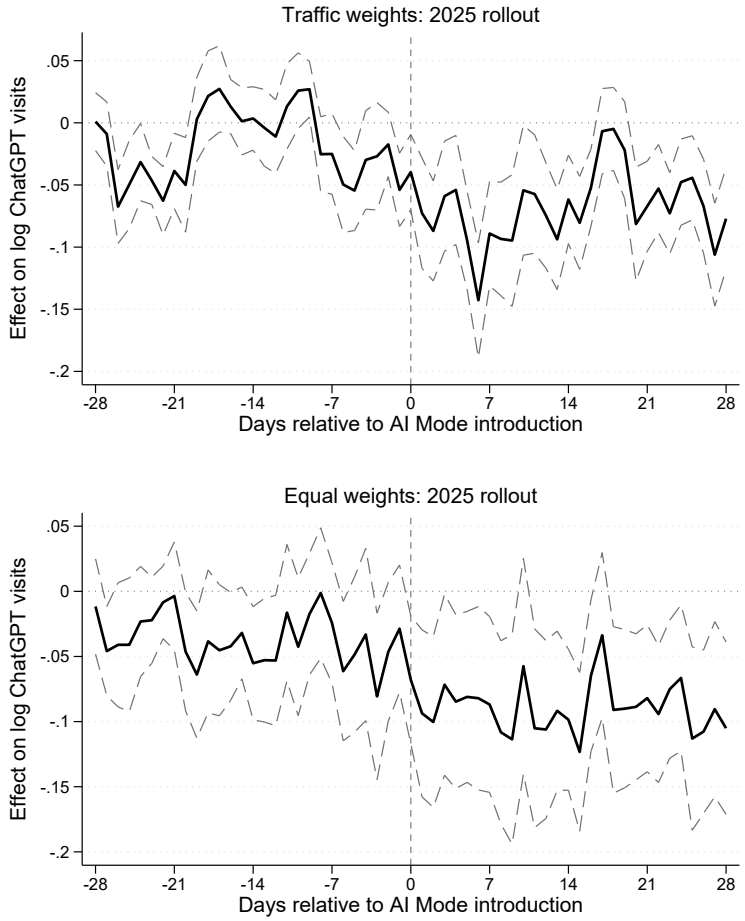
Table B7. ChatGPT visits: staggered rollout across all cohorts

	(1)	(2)
AI Mode × Post	−0.050***	−0.053***
	(0.014)	(0.019)
Implied change (%)	−4.86	−5.15
Cohort and country weights	Traffic	Equal
N	11,200	11,200
R^2	0.996	0.997

Notes: Cohort-specific comparisons use countries untreated through day 28. Bootstrap draws hold singleton launch cohorts fixed. N counts stacked country-days; R-squared is from the cohort-interacted regression. Country-by-weekday and date FE. Bootstrap SEs in parentheses (1000 draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

[16] On May 20, 2025, Google announced that AI Overviews were available in more than 200 countries and territories (Budaraju, 2025a). Availability differed across comparison countries: nine European countries had already received access in March 2025 (Budaraju, 2025b), whereas France received AI Overviews and AI Mode in July 2026 (Missoffe, 2026).

Figure B9. ChatGPT traffic: staggered rollout, 2025 cohorts



Notes: Daily event studies pooling the United States, India, United Kingdom, and August cohorts, with traffic weights and equal weights; see Table B7. Each cohort is compared with countries untreated through day 28 of its window. Each coefficient is relative to the same weekday in days −35 to −29 of the cohort's window. Dashed lines show pointwise 95% confidence limits from 1,000 country-bootstrap replications, holding the three single-country cohorts fixed. Weekly versions are available in the replication materials.

B.7 ChatGPT Traffic by Source

Table B8 examines which traffic sources account for the estimated change in ChatGPT visits.

Panel A applies the preferred specification to log direct visits and log organic-search visits and reports the implied percentage changes. Both sources have positive visits on every country-day in the required window. Direct visits decline significantly following the launch of AI Mode, whereas the estimated effect on organic-search visits is not statistically significant.

Panel B expresses changes in direct visits, organic-search visits, and all remaining channels combined relative to baseline total ChatGPT traffic, so that the three contributions are on a common scale and sum to the total. For each country c, let $\overline{V}_c$ denote mean total daily visits on days -84 to -29. For channel j, the outcome is $Y^j_{ct} = 100V^j_{ct}/\overline{V}_c$. Coefficients therefore measure changes in percentage points of baseline total visits.

Figure B10 presents the corresponding daily event studies. Neither channel exhibits a pronounced differential trend before launch. Direct visits show a clear decline immediately after AI Mode becomes available, while organic-search visits are noisier and display no comparable shift. Because direct visits make up about four fifths of ChatGPT traffic and decline in about the same proportion as total visits, they account for most of the overall decline.

Table B8. AI Mode and ChatGPT visits by traffic channel

Panel A. Log visits

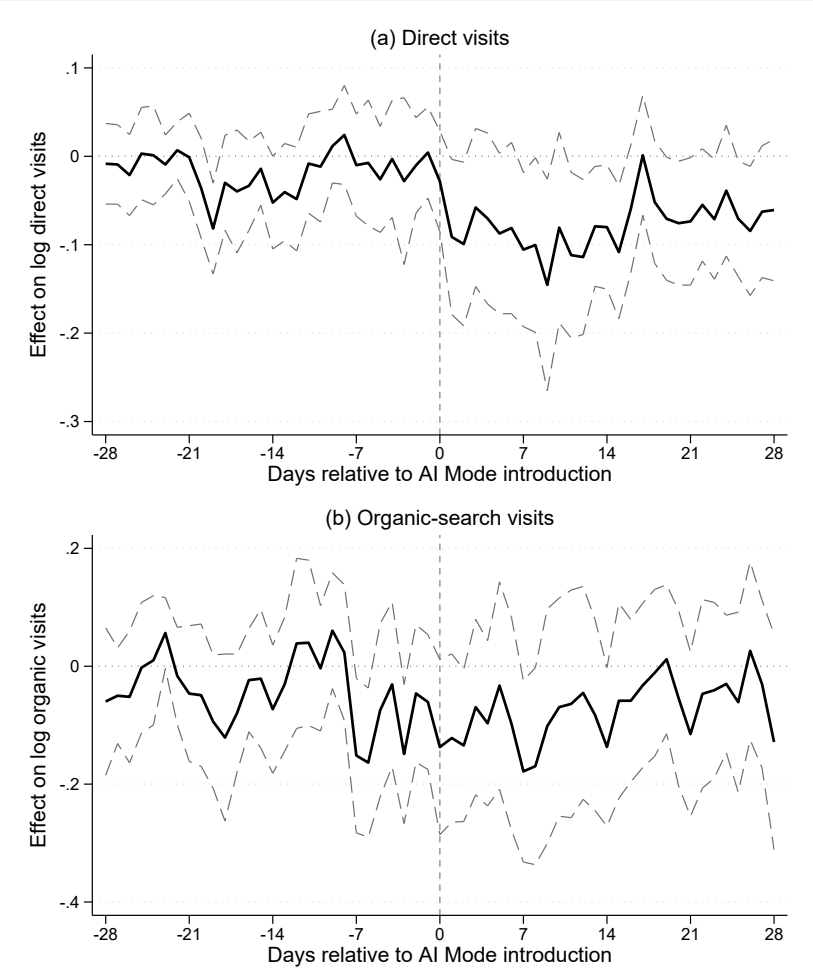
	Total	Direct	Organic search
	(1)	(2)	(3)
AI Mode × Post	−0.061**	−0.061**	−0.031
	(0.025)	(0.025)	(0.059)
Implied change (%)	−5.88	−5.91	−3.03
N	3,080	3,080	3,080
R^2	0.993	0.992	0.972

Panel B. Contributions, percentage points of baseline total visits

	Total	Direct	Organic search	Other channels
	(1)	(2)	(3)	(4)
AI Mode × Post	−6.57**	−5.46**	−0.44	−0.68***
	(2.61)	(2.13)	(0.56)	(0.24)
N	3,080	3,080	3,080	3,080
R^2	0.883	0.884	0.874	0.933

Notes: Panel B reports percentage points of baseline total visits; channel effects sum to the total. Other channels exclude direct and organic traffic. Panel A uses log visits for every channel and reports implied percentage changes. Country-by-weekday and date FE. Overall weighted R^2 uses the adjusted outcome and includes FE; N counts country-days. Bootstrap SEs in parentheses (1000 draws). $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure B10. ChatGPT traffic by channel



Notes: Daily event studies for log direct visits (a) and log organic-search visits (b), using the preferred specification; see Table B8. Each coefficient is relative to the same weekday in days −35 to −29. Dashed lines show pointwise 95% confidence limits from 1,000 country-bootstrap replications. Panels use separate vertical scales.

B.8 Google Search Interest in ChatGPT

We complement the organic-referral analysis with weekly Google Trends data for the search term "ChatGPT." These data measure relative search interest rather than visits referred to ChatGPT. The outcome is the log search-interest index. The series were downloaded separately for each calendar year, so each country's index has a consistent scale within 2025; our estimation window (May 25-September 20, 2025) lies within that year, which ensures that country fixed effects absorb country-specific normalization factors.

Table B9 follows the five specifications of the main ChatGPT analysis, using country and calendar-week fixed effects. We compare four complete weeks before the rollout (July 20-August 16, 2025) with four complete weeks after it (August 24-September 20), omitting the week containing the August 21 launch. Traffic weights use the original baseline Semrush ChatGPT visits. Regional trends and holiday adjustments are fitted over the preceding eight weeks and re-estimated in each bootstrap draw. All specifications use a common sample with positive, uncensored search-interest observations throughout the required weeks. In the baseline traffic-weighted specification, the estimate is small and statistically insignificant. With equal country weights, it is negative and significant; with regional trends, the traffic-weighted estimates are negative but imprecise. The two weighting schemes emphasize different markets.

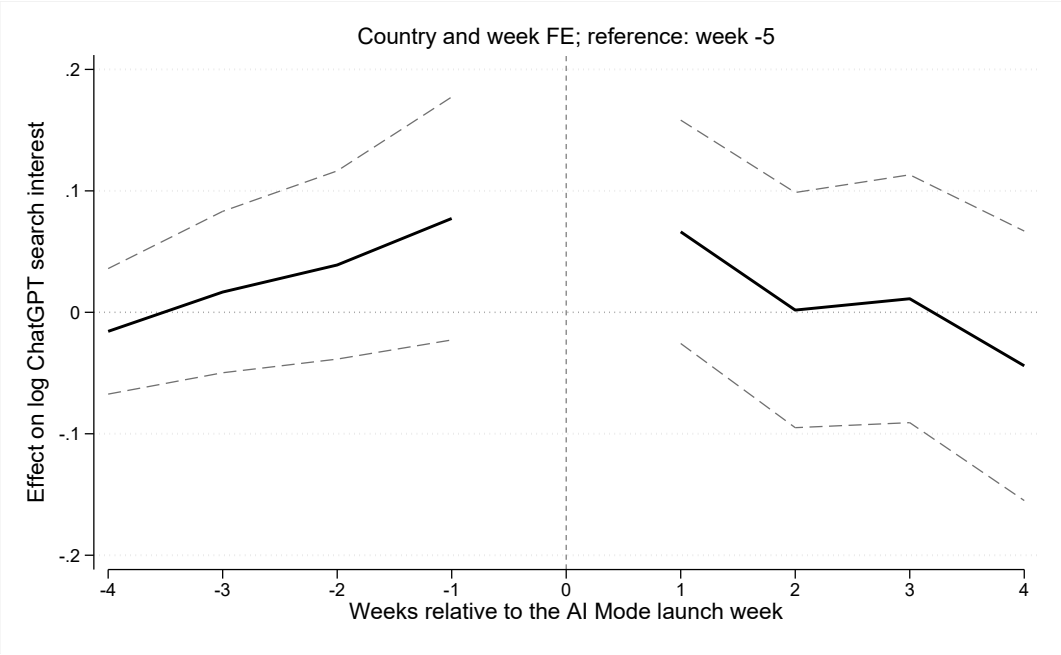
Figure B11 presents the corresponding traffic-weighted event study without additional adjustments. The point estimates rise over the four pre-launch weeks and subsequently decline, indicating that the small average DiD estimate masks changes over the event window. Together, the table and figure assess whether the rollout coincides with a change in search interest in ChatGPT relative to comparison countries, providing complementary evidence on the organic-search channel.

Table B9. AI Mode and Google search interest in ChatGPT

	(1)	(2)	(3)	(4)	(5)
AI Mode × Post	−0.020	−0.100	−0.015	−0.100	−0.081***
	(0.033)	(0.068)	(0.033)	(0.067)	(0.031)
Implied change (%)	−2.03	−9.52	−1.53	−9.53	−7.82
Country weights	Traffic	Traffic	Traffic	Traffic	Equal
Regional linear trends	No	Yes	No	Yes	No
Holiday adjustment	No	No	Yes	Yes	No
N	440	440	440	440	440
R^2	0.814	0.744	0.827	0.759	0.784

Notes: Log Google Trends interest for the search term ChatGPT. Four complete weeks before (July 20–August 16, 2025) versus after (August 24–September 20); the launch week is omitted. Country and calendar-week FE. Traffic weights use baseline Semrush ChatGPT visits on launch days −84 to −29. Adjustments are fitted on May 25–July 19 and refitted in each draw. Bootstrap SEs in parentheses (1,000 stratified country draws); overall R^2 uses the adjusted outcome, and N counts country-weeks. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure B11. AI Mode and Google search interest in ChatGPT



Notes: Weekly event-study estimates for log Google Trends search interest in ChatGPT, with country and calendar-week fixed effects and baseline Semrush ChatGPT traffic weights. Weeks run from Sunday to Saturday. Week 0 (August 17-23, 2025) contains the launch and is omitted. Coefficients are relative to week −5 (July 13-19). Dashed lines show pointwise 95% confidence intervals from 1,000 country-bootstrap replications, stratified by treatment group.

B.9 Visitors, Visit Frequency, and Duration

Table B10 examines how the change in ChatGPT traffic relates to the number of daily visitors, visit frequency, and visit duration. We apply the preferred specification to five outcomes, all measured in logs: daily visits, daily unique visitors, visits per unique visitor, average visit duration, and constructed total time on the site, defined as visits multiplied by average duration. Column (1) reproduces the main estimate.

Using the same sample and visit-based traffic weights across columns allows us to decompose the estimated change in visits into changes in daily unique visitors and visits per visitor. Similarly, the estimated change in constructed total time combines changes in visits and average visit duration.[17]

Figure B12 presents the corresponding daily event studies, showing the timing of each outcome's response and its pre-launch dynamics.

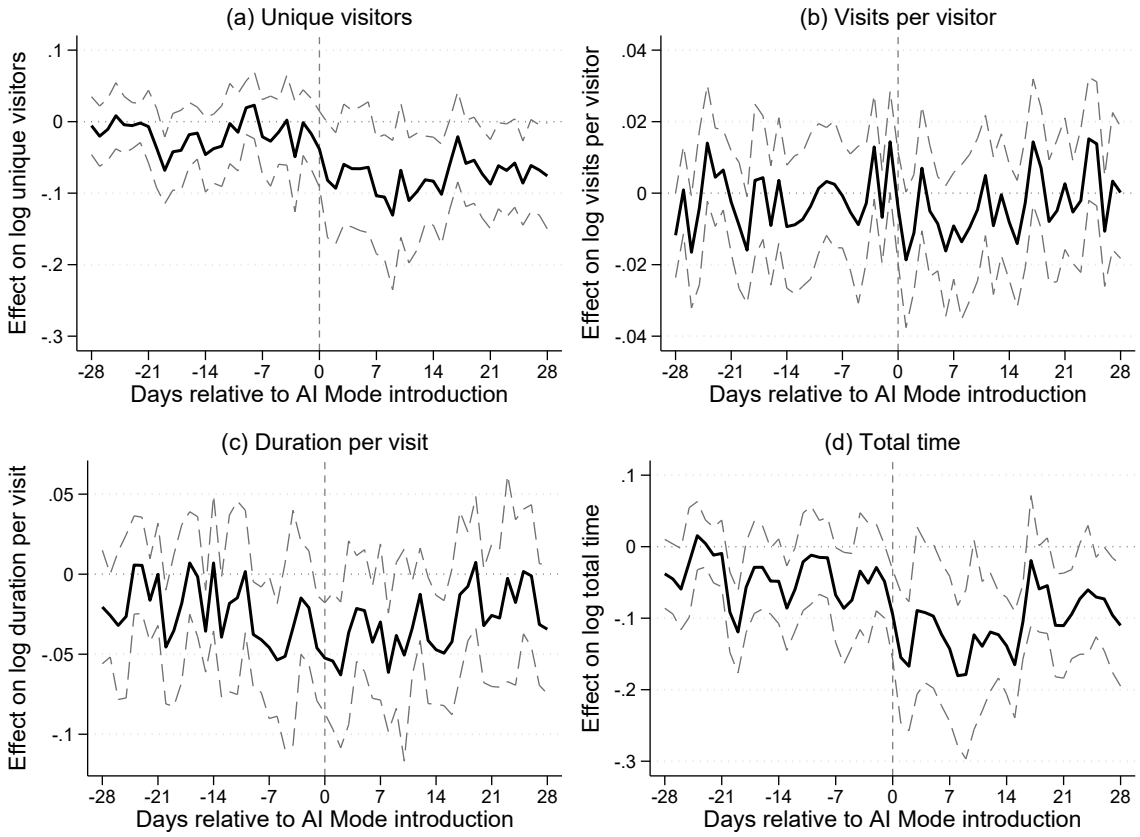
Table B10. AI Mode and ChatGPT: visitors, visit frequency, and visit duration

	Visits (1)	Visitors (2)	Visits per visitor (3)	Duration per visit (4)	Total time (5)
AI Mode × Post	−0.061**	−0.059**	−0.002	−0.007	−0.067**
	(0.025)	(0.023)	(0.003)	(0.008)	(0.028)
Implied change (%)	−5.88	−5.74	−0.15	−0.65	−6.50
N	3,080	3,080	3,080	3,080	3,080
R^2	0.993	0.993	0.862	0.953	0.991

Notes: Logged outcomes; baseline visit weights and common bootstrap draws. Visits and unique visitors are separately modeled by Semrush. No country is dropped. Country-by-weekday and date FE. Overall weighted R^2 uses the adjusted outcome and includes FE; N counts country-days. Bootstrap SEs in parentheses (1000 draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

[17] Writing visits as V, daily unique visitors as U, and average duration as D, the identities $\log V = \log U + \log(V/U)$ and $\log(VD) = \log V + \log D$ imply corresponding identities for the log-point coefficients. Before rounding, column (1) equals the sum of columns (2) and (3), and column (5) equals the sum of columns (1) and (4). All columns use the same bootstrap draws. These additive identities do not apply to percentage changes or standard errors.

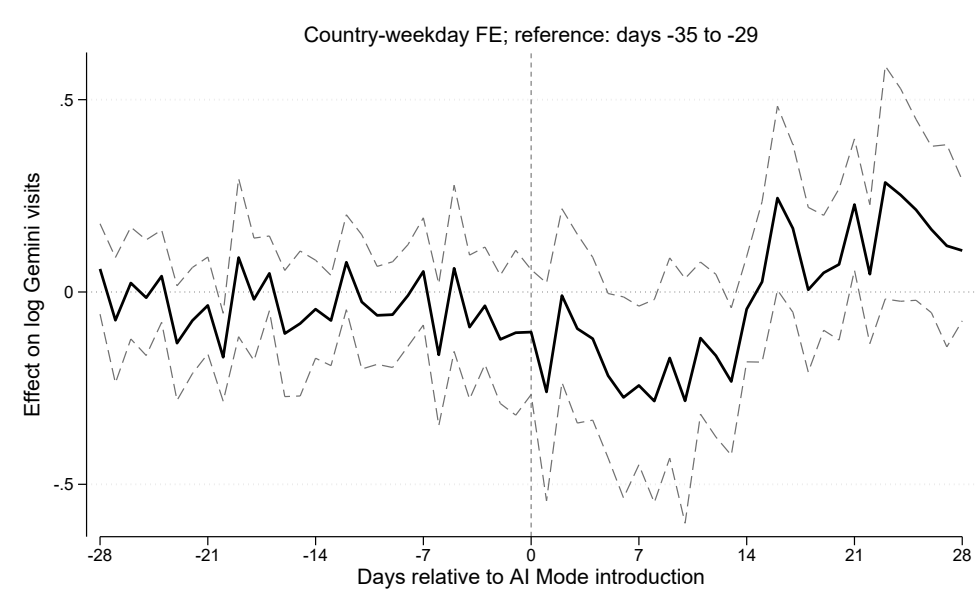
Figure B12. ChatGPT use: visitors, visit frequency, and duration



Notes: Daily event studies for columns (2)-(5) of Table B10: daily unique visitors (a), visits per visitor (b), average visit duration (c), and constructed total time (d). All outcomes are in logs and use the preferred specification. Each coefficient is relative to the same weekday in days −35 to −29. Dashed lines show pointwise 95% confidence limits. Panels use separate vertical scales.

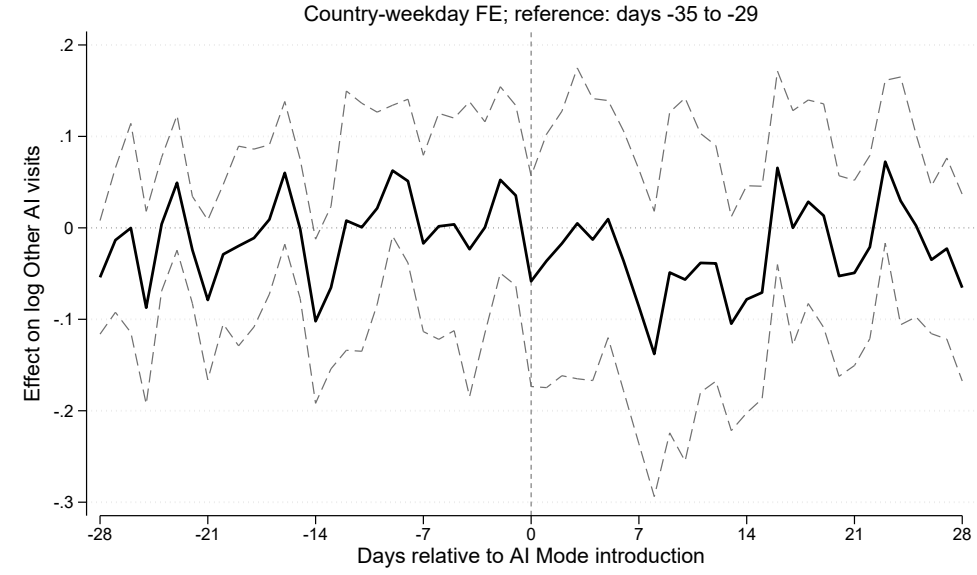
C Gemini and Other AI Assistants

Figure C1. AI Mode and visits to Gemini



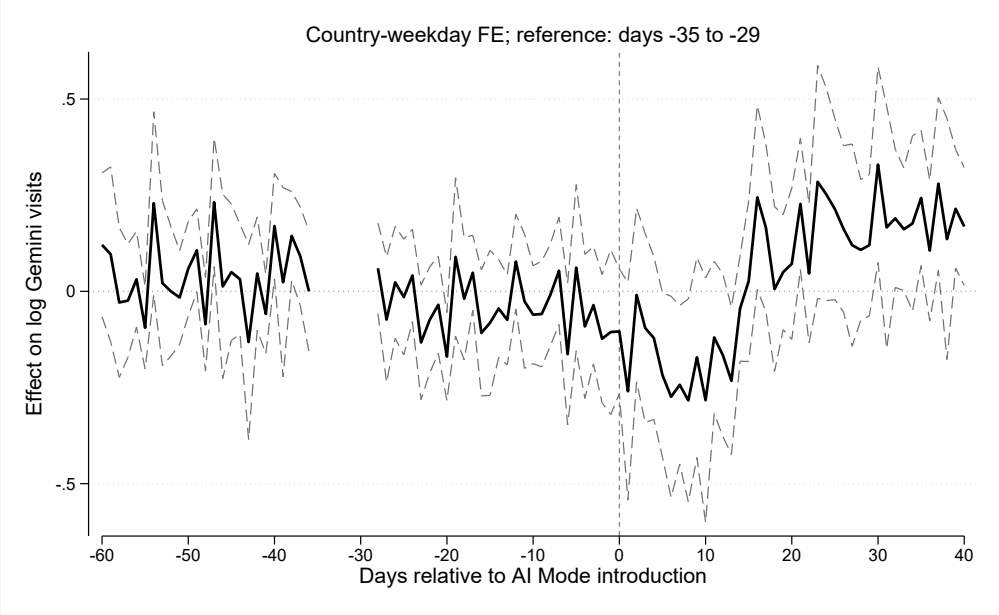
Notes: Daily event-study estimates for log visits to Gemini, with country-by-weekday and date fixed effects. Traffic weights use Gemini visits on days −84 to −29. Each coefficient is relative to the same weekday in days −35 to −29; day 0 is estimated separately. Dashed lines show pointwise 95% confidence intervals from 1,000 country-bootstrap replications.

Figure C2. AI Mode and visits to other AI assistants



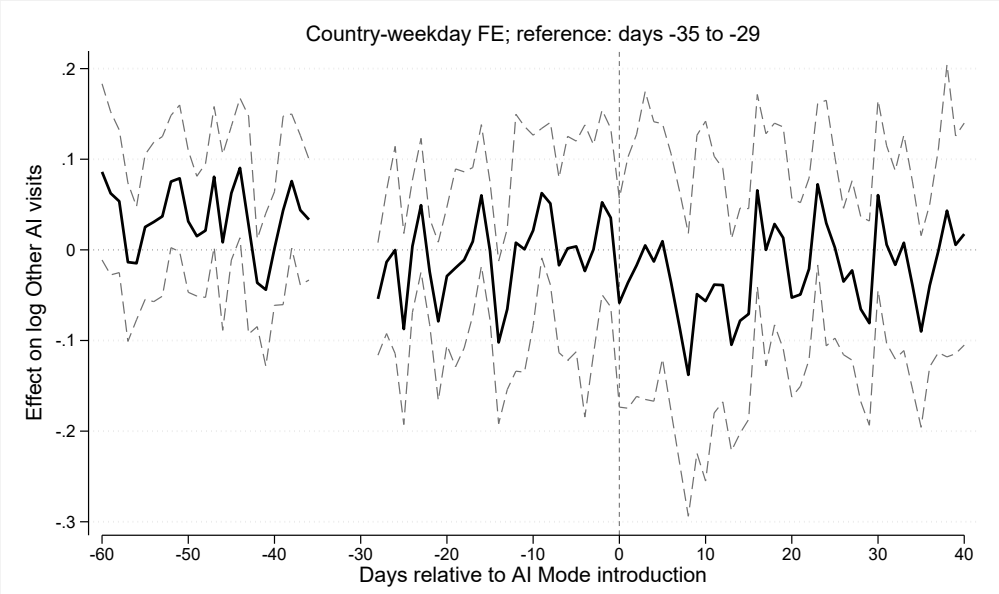
Notes: Daily event-study estimates for log visits to the eight smaller assistants combined, with country-by-weekday and date fixed effects. Traffic weights use combined visits on days −84 to −29. Each coefficient is relative to the same weekday in days −35 to −29; day 0 is estimated separately. Dashed lines show pointwise 95% confidence intervals from 1,000 country-bootstrap replications.

Figure C3. AI Mode and Gemini visits: extended event study



Notes: Daily event-study estimates for log Gemini visits over days −60 to 40, including Indonesia. The specification includes country-by-weekday and date fixed effects and uses Gemini traffic weights from days −84 to −29. Each coefficient is relative to the same weekday in days −35 to −29; this reference week is left blank. Comparison countries remain untreated throughout the window. Dashed lines show pointwise 95% confidence intervals from 1,000 country-bootstrap replications.

Figure C4. AI Mode and visits to other AI assistants: extended event study



Notes: The specification in Figure C2, extended to days −60 through 40. Traffic weights and the weekday-specific reference on days −35 to −29 are unchanged; the reference week is left blank. Comparison countries remain untreated throughout the window. Dashed lines show pointwise 95% confidence intervals from 1,000 country-bootstrap replications.

Table C1. AI Mode and Gemini visits: early and late effects

	All countries		Excluding Indonesia	
	(1)	(2)	(3)	(4)
AI Mode × Post (weeks 1–2)	−0.143*	−0.159***	−0.104	−0.116**
	(0.074)	(0.057)	(0.081)	(0.052)
AI Mode × Post (weeks 3–4)	0.179**	0.154*	0.126	0.105
	(0.087)	(0.093)	(0.080)	(0.092)
Country weights	Traffic	Traffic	Traffic	Traffic
Regional linear trends	No	Yes	No	Yes
Holiday adjustment	No	Yes	No	Yes
N	3,080	3,080	3,024	3,024
R^2	0.974	0.974	0.971	0.972

Notes: Log Gemini visits. Days 1–14 and 15–28 relative to days −28 to −1; day 0 excluded. Country-by-weekday and date FE. Traffic weights and regional/holiday adjustments use days −84 to −29. Overall R^2 includes FE and uses the adjusted outcome where applicable. Bootstrap SEs in parentheses (1,000 country draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

C.1 Gemini: Comparison with 2024

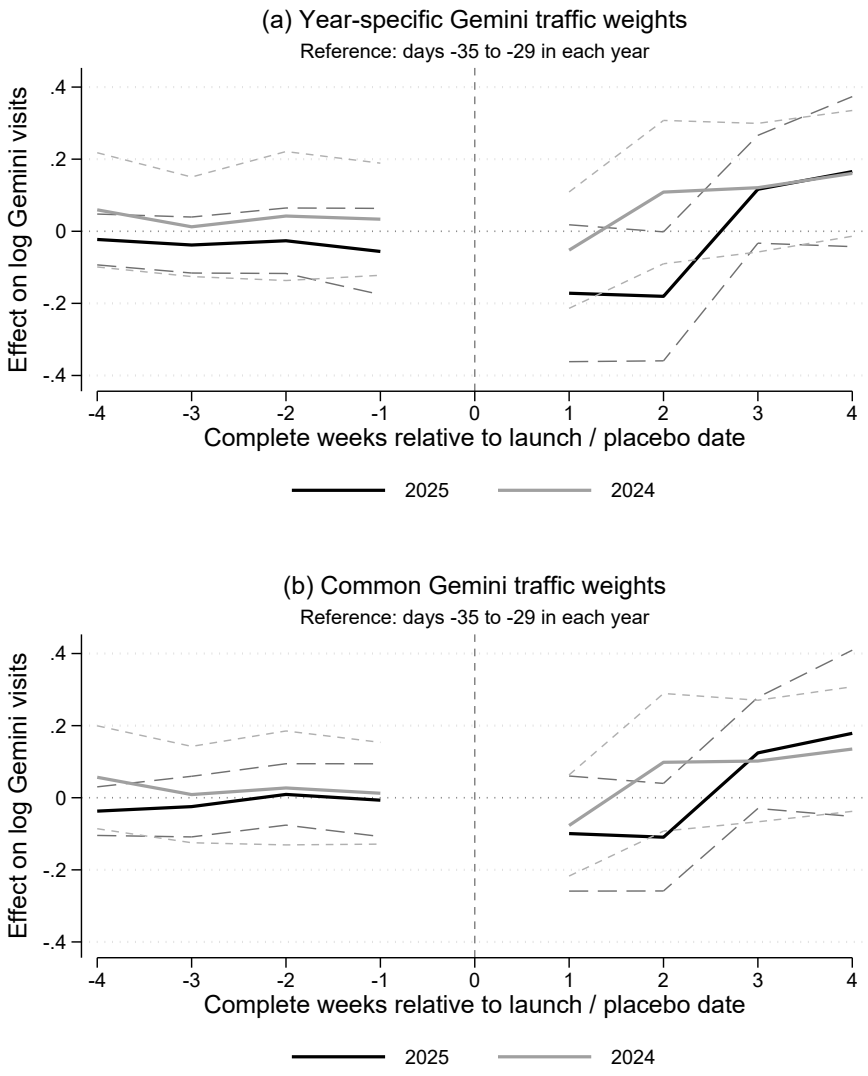
Table C2 and Figure C5 compare the evolution of Gemini visits around August 21, 2025 with the corresponding period in 2024. The placebo date is August 22, 2024, exactly 52 weeks earlier, preserving the alignment of weekdays. Both years use the same eligible countries and the treatment-group classification from 2025. Five countries (Albania, Bangladesh, Iceland, Luxembourg, and Slovenia) have at least one day with zero or missing Gemini visits in the 2024 window and are excluded, leaving 23 treated and 27 comparison countries. Panel A uses year-specific Gemini traffic weights; Panel B holds weights fixed across years. This comparison assesses whether the change in Gemini visits relative to comparison countries recurs in the previous year. Seasonal changes common to both groups are absorbed by date fixed effects.

Table C2. Gemini visits: the corresponding calendar window in 2024

	2025: AI Mode launch (1)	2024: placebo (2)	Difference (3)
Panel A. Matched countries, year-specific traffic weights			
Days 1–28 vs. days −28 to −1	0.018	0.048	−0.029
	(0.063)	(0.054)	(0.069)
N	2,800	2,800	5,600
R^2	0.970	0.924	—
Days 1–28 vs. days −7 to −1	0.038	0.051	−0.013
	(0.072)	(0.046)	(0.068)
N	1,750	1,750	3,500
R^2	0.969	0.938	—
Panel B. Matched countries, common traffic weights			
Days 1–28 vs. days −28 to −1	0.038	0.038	0.000
	(0.062)	(0.054)	(0.065)
N	2,800	2,800	5,600
R^2	0.969	0.924	—
Days 1–28 vs. days −7 to −1	0.030	0.052	−0.022
	(0.069)	(0.049)	(0.069)
N	1,750	1,750	3,500
R^2	0.968	0.939	—

Notes: Log Gemini visits; August 21, 2025 versus August 22, 2024 (52 weeks earlier). Same 23 treated and 27 comparison countries in both years. Country-by-weekday and date FE. Panel A weights use own-year days −77 to −29; Panel B uses March 25–May 19, 2025 in both years. Paired country-bootstrap SEs (1000 draws); difference is 2025 minus 2024, with N summed over years and no separate R^2. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure C5. Gemini visits: comparison of 2025 and 2024



Notes: Complete-week event-study estimates for log Gemini visits around August 21, 2025 (black) and August 22, 2024 (gray), using the same matched countries and country-by-weekday and date fixed effects. Each coefficient is relative to days −35 through −29 in its respective year; day 0 is excluded. Panel (a) uses Gemini traffic weights from days −77 through −29 in each year. Panel (b) uses weights from March 25-May 19, 2025 in both years. Dashed lines show pointwise 95% confidence intervals from 1,000 country-bootstrap replications, with the same country draws used across years. Table C2 reports the corresponding period contrasts and their cross-year differences.

C.2 Results for Individual AI Assistants

The following tables report exploratory estimates separately for the eight assistants included in Other AI. Each table presents the five specifications used in Table 3: traffic weights without additional adjustments, with regional trends, with holiday adjustment, with both adjustments, and equal country weights without additional adjustments. The outcome transformation is indicated in each table. For outcomes measured as $\log(1+V)$, coefficients are reported without percentage conversions.

Table C3. AI Mode and Claude visits

	(1)	(2)	(3)	(4)	(5)
AI Mode × Post	−0.050	−0.041	−0.049	−0.043	−0.012
	(0.038)	(0.045)	(0.038)	(0.045)	(0.048)
Country weights	Traffic	Traffic	Traffic	Traffic	Equal
Regional linear trends	No	Yes	No	Yes	No
Holiday adjustment	No	No	Yes	Yes	No
N	3,080	3,080	3,080	3,080	3,080
R^2	0.977	0.977	0.978	0.978	0.973

Notes: Log(1 + visits). Days 1–28 versus −28 to −1. Country-by-weekday and date FE. Own-service traffic weights use days −84 to −29. Regional/holiday adjustments use the baseline; overall R-squared refers to the adjusted outcome. Bootstrap SEs in parentheses (1,000 country draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table C4. AI Mode and Copilot visits

	(1)	(2)	(3)	(4)	(5)
AI Mode × Post	0.042	0.147	0.046	0.144	−0.009
	(0.077)	(0.089)	(0.077)	(0.090)	(0.066)
Country weights	Traffic	Traffic	Traffic	Traffic	Equal
Regional linear trends	No	Yes	No	Yes	No
Holiday adjustment	No	No	Yes	Yes	No
N	3,080	3,080	3,080	3,080	3,080
R^2	0.914	0.912	0.914	0.913	0.940

Notes: Log(1 + visits). Days 1–28 versus −28 to −1. Country-by-weekday and date FE. Own-service traffic weights use days −84 to −29. Regional/holiday adjustments use the baseline; overall R-squared refers to the adjusted outcome. Bootstrap SEs in parentheses (1,000 country draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table C5. AI Mode and DeepSeek visits

	(1)	(2)	(3)	(4)	(5)
AI Mode × Post	−0.005	−0.001	−0.004	−0.001	−0.055
	(0.035)	(0.054)	(0.035)	(0.054)	(0.038)
Implied change (%)	−0.55	−0.09	−0.43	−0.14	−5.39
Country weights	Traffic	Traffic	Traffic	Traffic	Equal
Regional linear trends	No	Yes	No	Yes	No
Holiday adjustment	No	No	Yes	Yes	No
N	3,080	3,080	3,080	3,080	3,080
R^2	0.978	0.978	0.978	0.978	0.979

Notes: Log visits. Days 1–28 versus −28 to −1. Country-by-weekday and date FE. Own-service traffic weights use days −84 to −29. Regional/holiday adjustments use the baseline; overall R-squared refers to the adjusted outcome. Bootstrap SEs in parentheses (1,000 country draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table C6. AI Mode and Grok visits

	(1)	(2)	(3)	(4)	(5)
AI Mode × Post	−0.088	0.021	−0.085	0.021	−0.001
	(0.081)	(0.062)	(0.080)	(0.061)	(0.047)
Country weights	Traffic	Traffic	Traffic	Traffic	Equal
Regional linear trends	No	Yes	No	Yes	No
Holiday adjustment	No	No	Yes	Yes	No
N	3,080	3,080	3,080	3,080	3,080
R^2	0.965	0.964	0.965	0.965	0.972

Notes: Log(1 + visits). Days 1–28 versus −28 to −1. Country-by-weekday and date FE. Own-service traffic weights use days −84 to −29. Regional/holiday adjustments use the baseline; overall R-squared refers to the adjusted outcome. Bootstrap SEs in parentheses (1,000 country draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table C7. AI Mode and Meta AI visits

	(1)	(2)	(3)	(4)	(5)
AI Mode × Post	−0.001	−0.033	−0.001	−0.033	0.009
	(0.081)	(0.139)	(0.082)	(0.139)	(0.093)
Country weights	Traffic	Traffic	Traffic	Traffic	Equal
Regional linear trends	No	Yes	No	Yes	No
Holiday adjustment	No	No	Yes	Yes	No
N	3,080	3,080	3,080	3,080	3,080
R^2	0.859	0.856	0.859	0.856	0.778

Notes: Log(1 + visits). Days 1–28 versus −28 to −1. Country-by-weekday and date FE. Own-service traffic weights use days −84 to −29. Regional/holiday adjustments use the baseline; overall R-squared refers to the adjusted outcome. Bootstrap SEs in parentheses (1,000 country draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table C8. AI Mode and Mistral visits

	(1)	(2)	(3)	(4)	(5)
AI Mode × Post	−0.270***	−0.398**	−0.265***	−0.400**	−0.233**
	(0.101)	(0.163)	(0.100)	(0.161)	(0.098)
Country weights	Traffic	Traffic	Traffic	Traffic	Equal
Regional linear trends	No	Yes	No	Yes	No
Holiday adjustment	No	No	Yes	Yes	No
N	3,080	3,080	3,080	3,080	3,080
R^2	0.941	0.940	0.940	0.939	0.741

Notes: Log(1 + visits). Days 1–28 versus −28 to −1. Country-by-weekday and date FE. Own-service traffic weights use days −84 to −29. Regional/holiday adjustments use the baseline; overall R-squared refers to the adjusted outcome. Bootstrap SEs in parentheses (1,000 country draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table C9. AI Mode and Perplexity visits

	(1)	(2)	(3)	(4)	(5)
AI Mode × Post	−0.024	0.056	−0.022	0.054	0.068
	(0.060)	(0.059)	(0.060)	(0.059)	(0.063)
Country weights	Traffic	Traffic	Traffic	Traffic	Equal
Regional linear trends	No	Yes	No	Yes	No
Holiday adjustment	No	No	Yes	Yes	No
N	3,080	3,080	3,080	3,080	3,080
R^2	0.967	0.967	0.967	0.967	0.970

Notes: Log(1 + visits). Days 1–28 versus −28 to −1. Country-by-weekday and date FE. Own-service traffic weights use days −84 to −29. Regional/holiday adjustments use the baseline; overall R-squared refers to the adjusted outcome. Bootstrap SEs in parentheses (1,000 country draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table C10. AI Mode and Qwen visits

	(1)	(2)	(3)	(4)	(5)
AI Mode × Post	−0.178	−0.293**	−0.177	−0.293**	0.021
	(0.112)	(0.146)	(0.112)	(0.146)	(0.111)
Country weights	Traffic	Traffic	Traffic	Traffic	Equal
Regional linear trends	No	Yes	No	Yes	No
Holiday adjustment	No	No	Yes	Yes	No
N	3,080	3,080	3,080	3,080	3,080
R^2	0.914	0.910	0.915	0.911	0.859

Notes: Log(1 + visits). Days 1–28 versus −28 to −1. Country-by-weekday and date FE. Own-service traffic weights use days −84 to −29. Regional/holiday adjustments use the baseline; overall R-squared refers to the adjusted outcome. Bootstrap SEs in parentheses (1,000 country draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

C.3 Robustness Checks for Mistral

We examine the sensitivity of Mistral's estimates to temporal aggregation, the inclusion of France, and the outcome specification. We also compare the 2025 event-study pattern with the corresponding calendar window in 2024. Two features of the setting call for caution. The measured domain, `mistral.ai`, also hosts the company's corporate pages, and Mistral announced a €1.7 billion funding round on September 9, 2025, day 19 of the window (Mistral AI, 2025). Either could affect measured traffic independently of AI Mode.

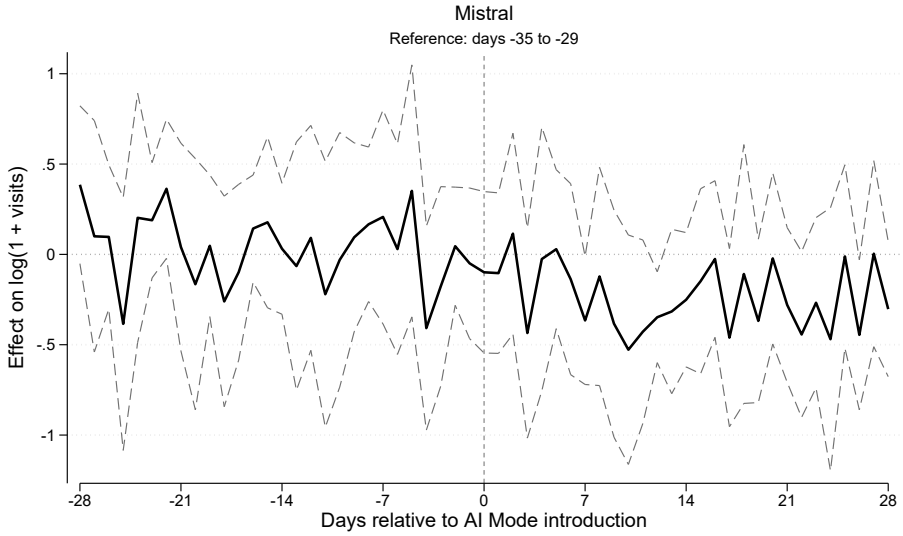
Table C11. Mistral: sensitivity to excluding France

	(1)	(2)	(3)	(4)
AI Mode × Post	−0.270***	−0.137*	−0.233**	−0.226**
	(0.101)	(0.071)	(0.098)	(0.106)
Sample	Full	No France	Full	No France
Country weights	Traffic	Traffic	Equal	Equal
N	3,080	3,024	3,080	3,024
R^2	0.941	0.854	0.741	0.714

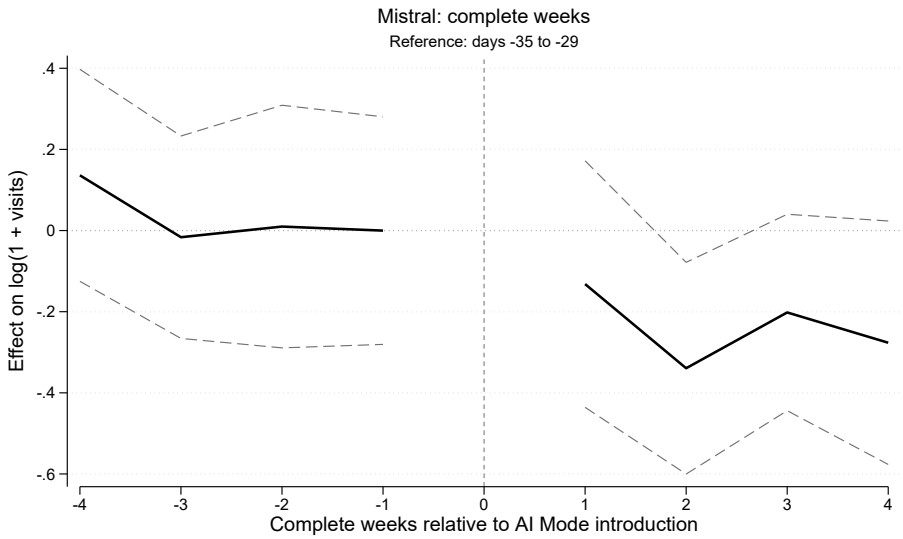
Notes: Log(1 + visits). Days 1–28 versus −28 to −1. Country-by-weekday and date FE. Own-service traffic weights use days −84 to −29. France is excluded before sample construction; weights are renormalized. Bootstrap SEs in parentheses (1,000 country draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure C6. Mistral: daily and complete-week event studies

(a) Daily

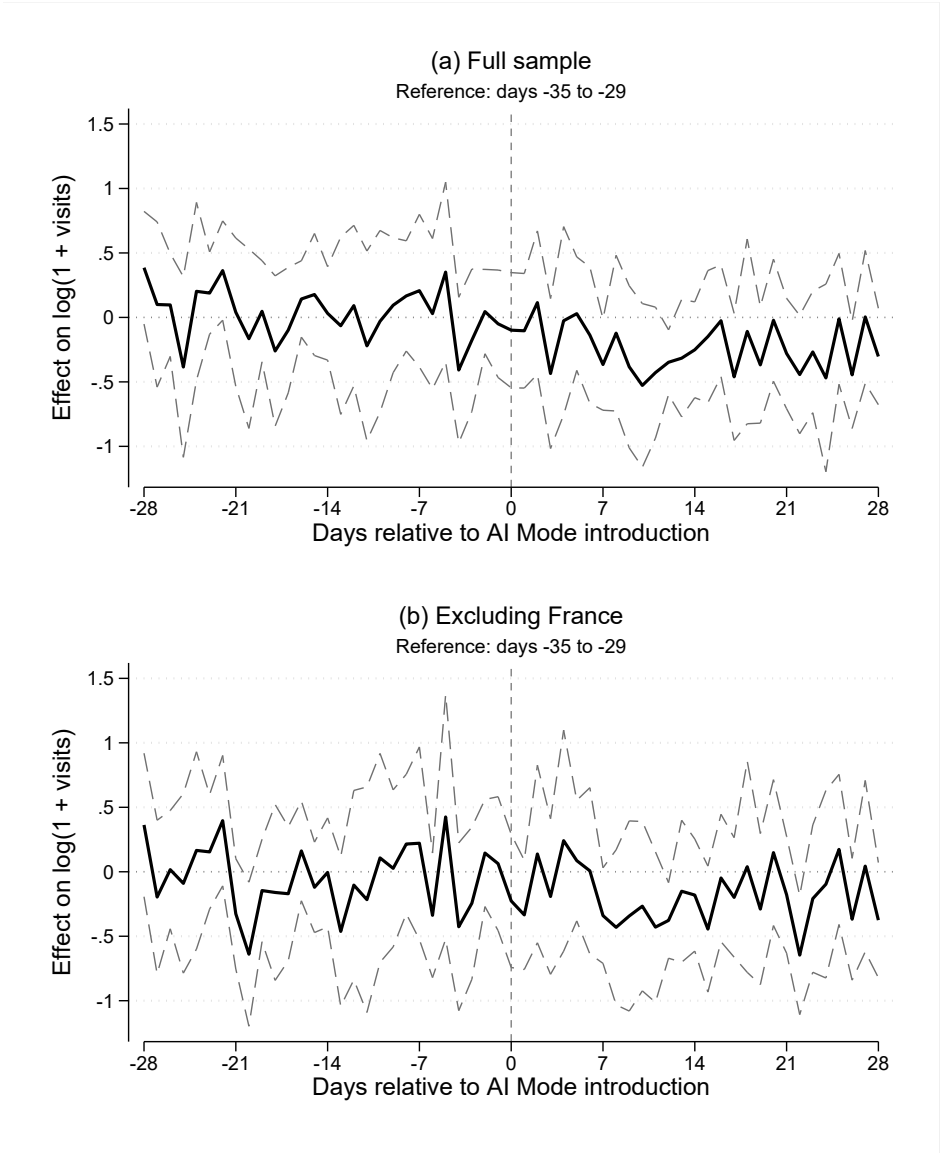


(b) Complete weeks



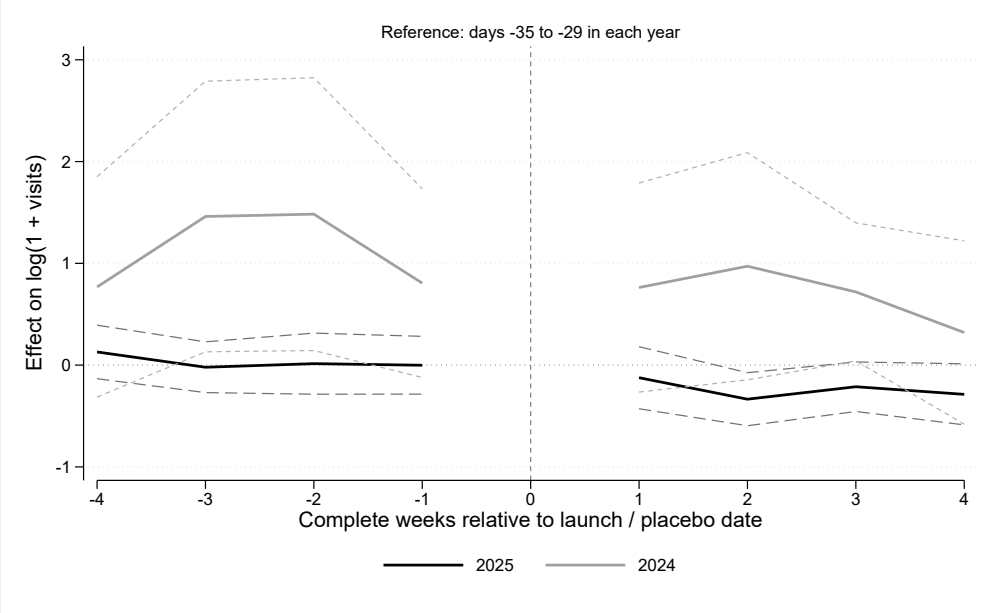
Notes: Outcome: $\log(1+V)$ for Mistral. The specification includes country-by-weekday and date fixed effects and uses Mistral traffic weights from days −84 to −29. Daily coefficients are relative to the same weekday in days −35 to −29; day 0 is estimated separately. Weekly coefficients average complete seven-day periods relative to days −35 to −29 (week −5); day 0 is excluded. Dashed lines show pointwise 95% confidence intervals from 1,000 country-bootstrap replications.

Figure C7. Mistral: sensitivity to excluding France



Notes: Daily event studies for $\log(1+V)$ in the full sample (a) and excluding France (b). Both panels include country-by-weekday and date fixed effects and use Mistral traffic weights from days -84 to -29, renormalized after exclusion. Coefficients are relative to the same weekday in days -35 to -29. Dashed lines show pointwise 95% confidence intervals from 1,000 country-bootstrap replications. Panels share the same vertical scale.

Figure C8. Mistral: the 2025 launch window and the corresponding window in 2024



Notes: Complete-week estimates for $\log(1+V)$ around August 21, 2025 (black) and August 22, 2024 (gray), both Thursdays. The comparison uses countries with complete observations and positive baseline traffic in both years, retaining their 2025 treatment classification. Year-specific Mistral traffic weights use days −77 to −29. Each series is relative to days −35 to −29 in its respective year; day 0 is excluded. Dashed lines show pointwise 95% confidence intervals from 1,000 country-bootstrap replications. The 2024 series is a descriptive placebo.

C.3.1 Poisson Pseudo-Maximum-Likelihood Estimates

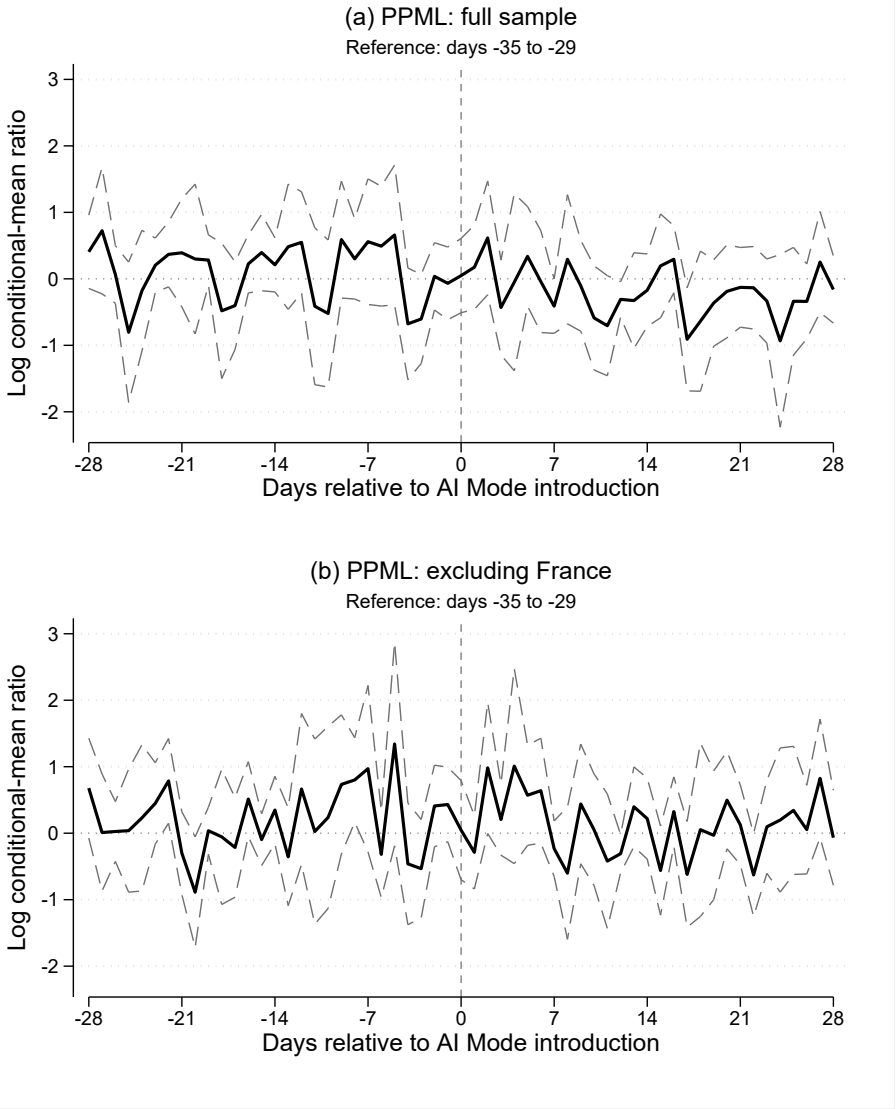
We also estimate a Poisson pseudo-maximum-likelihood model using visits in levels, including zero observations. The specification includes country-by-weekday and date fixed effects. Unlike the baseline adjustments used in the log specifications, regional trends and holiday controls enter the PPML conditional mean directly and are estimated jointly with the treatment coefficient. Table C12 reports the five specification variants; Figure C9 presents traffic-weighted daily event studies with and without France.

Table C12. Mistral: Poisson pseudo-maximum-likelihood estimates

	(1)	(2)	(3)	(4)	(5)
AI Mode × Post	−0.316**	−0.061	−0.315**	−0.058	−0.241**
	(0.148)	(0.129)	(0.148)	(0.128)	(0.112)
Implied change (%)	−27.08	−5.88	−27.02	−5.60	−21.43
Country weights	Traffic	Traffic	Traffic	Traffic	Equal
Regional linear trends	No	Yes	No	Yes	No
Holiday adjustment	No	No	Yes	Yes	No
N	3,080	3,080	3,080	3,080	3,080
Pseudo-R^2	0.985	0.985	0.985	0.985	0.954

Notes: Visits in levels, including zeros. Days 1–28 versus −28 to −1. Country-by-weekday and date FE. Own-service traffic weights use days −84 to −29. PPML in levels; trends and holidays enter the conditional mean directly. Bootstrap SEs in parentheses (1,000 country draws). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Figure C9. Mistral: PPML event studies



Notes: Daily PPML event studies using visits in levels, in the full sample (a) and excluding France (b). Both panels include country-by-weekday and date fixed effects and use Mistral traffic weights from days −84 to −29, renormalized after exclusion. No regional trends or holiday controls are included in these figures. Coefficients are log conditional-mean ratios relative to the same weekday in days −35 to −29; day 0 is estimated separately. Dashed lines show pointwise 95% confidence intervals from 1,000 country-bootstrap replications. Panels share the same vertical scale.