AI Overviews and Search Intent:

Google's Selective Deployment of Generative AI in Search

Robin Ng* Michael Wessel†

September 29, 2026

Abstract

In May 2024, Google introduced AI Overviews (AIO): AI-generated syntheses displayed directly on the Search Engine Results Page (SERP). An AIO can resolve a user's query on the SERP, speeding up the arrival of monetizable searches, but it is costly to generate and may displace advertising revenue. We model how a platform's incentive to deploy an AIO depends on the user's search intent: whether it is exploratory or targeted, and whether it is monetizable. We construct a novel dataset of SERPs for 1,575 search queries in two waves (November 2025 and August 2026). In line with our predictions, deployment is 32.3 percentage points higher for exploratory than targeted queries, and monetization reduces deployment by 13 percentage points. Among exploratory monetizable queries, deployment increases with advertisers' cost per click, suggesting that the risk of losing advertising revenue may not be a key consideration for AIO deployment. For exploratory searches, deployment is unrelated to how well organic search results resolve the user's query, rising when these results come from more authoritative sources. Of AIOs that cite sources, 98.1% cite a source from the top 10 organic results. Together, these findings suggest coopetition between Google and publishers that supply content for Overviews while competing with them for users' attention.

*Department of Economics and MaCCI, University of Mannheim; robin@robinng.com.

†Department of Digitalization, Copenhagen Business School and Business Information Systems, esp. E-Commerce and Digital Business, Friedrich Schiller University Jena; michael.wessel@uni-jena.de.

Funding: Robin Ng gratefully acknowledges support from the Deutsche Forschungsgemeinschaft (DFG) through the CRC TR 224, Project B05.

1 Introduction

Search engines organize information on the search engine results page (SERP) through a combination of organic listings, sponsored advertisements, and supplementary features such as knowledge panels and shopping carousels. Each of these elements represents a design choice about how to allocate user attention. Prior work has established that the arrangement of information on the SERP has first-order consequences for which links receive attention and traffic (L. Xu et al., 2012; Yang and Ghose, 2010). In May 2024, Google introduced AI Overviews (AIOs), which synthesize search results into AI-generated answers displayed on the SERP. An AIO is generated by a language model from several of the web results the platform has retrieved for the query and is typically displayed above them, with links to the pages it drew on (Google, 2025c). It differs in this respect from a featured snippet, which reproduces a passage from a single ranked page, and from a knowledge panel, which displays structured facts from a knowledge graph. The platform thus simultaneously controls the generation of the answer, the ranking of the sources that answer draws on, and the monetization of the surrounding advertising space.

Early research on the implications of AIOs suggests that this feature contributed to the growing prevalence of zero-click searches, in which users obtain what they are seeking directly on the SERP. Chapekis et al. (2025) find that users who encounter an AIO click a traditional result at roughly half the rate of users who do not, and Khosravi and Yoganarasimhan (2026) estimate that AIOs reduced daily traffic to English Wikipedia by approximately 15%. Using a randomized field experiment, S. Agarwal and Sen (2026) find that, conditional on an AI Overview appearing, outbound organic clicks fall by 39.8% and zero-click searches increase by 34.5%, with no detectable effect on sponsored clicks. Thus, because an AIO can absorb the content of several publishers into an object that competes with all of them, the feature contributes to growing concerns about declining publisher traffic and revenue.[1]

While this evidence documents the downstream consequences of AIOs, we lack insights into the equally consequential question of when and how AIOs are deployed in the first place. A natural

[1]Concerns about Google's market power are not new (e.g., Clemons and Madhani, 2010), and the platform has already been found to have engaged in illegal monopolization of the search market (*United States v. Google LLC*, 2024). Legal scholars have argued that the integration of AIOs into search may help extend this dominance into AI-mediated information access (Singh, 2025). Penske Media (publisher of *Rolling Stone*, *The Hollywood Reporter*, and other outlets) and Chegg have filed suit alleging that Google's AIOs caused substantial traffic and revenue losses. In the EU, a group of independent publishers have brought a competition complaint seeking interim measures.

interpretation is that AIOs are Google's answer to conversational AI systems such as ChatGPT, which reportedly triggered an internal "Code Red" (Singh, 2025). If that were their primary purpose, one would expect broad and uniform deployment across search queries, much like Google's AI Mode, which is explicitly designed to compete with conversational AI interfaces (Google, 2025c). Instead, AIOs appear selectively. Some queries consistently trigger an AIO, while others never or only sometimes do. Deployment is therefore not a product decision layered onto Google Search but an allocation decision integrated into the platform's broader optimization of the SERP.

An AIO that resolves a query on the SERP fulfills the user's need sooner, which brings forward the next, potentially monetizable, search. It also reduces click revenue on monetizable searches by reducing users' need to visit external websites. The platform therefore weighs what it gains by fulfilling the query against what it gives up by answering it, and the balance between the two is determined by the user's search intent. Not all search queries are alike: a user researching a topic creates very different monetization considerations than a user ready to complete a purchase.

We argue that search intent differs along two dimensions that determine when deploying an AIO is optimal. The first dimension is one of exploration. An AIO can help users resolve exploratory queries by synthesizing information from multiple sources. Conversely, for targeted queries where users primarily seek to reach a particular outcome, its added value is limited. Second, queries vary in their monetization capacity: some search intents are revenue generating, while others are not. If a user is exploring a non-monetizable domain (e.g., "Why is the sky blue?"), an AIO can resolve the query quickly while sacrificing little immediate revenue. If a targeted search query provides valuable monetization opportunities through sponsored advertisements (e.g., "Buy iPhone 17"), answering the query directly diverts attention away from paid results while adding little to fulfillment. AIOs are therefore best understood as an instrument through which the platform balances fulfillment against the monetization it forgoes, governed by the same economic tradeoffs that shape other aspects of SERP design.

This motivates our research question:

How does search intent affect a search platform's selective deployment of AI Overviews, and are deployment patterns consistent with a platform trading off the fulfillment an AI Overview provides against the monetization it forgoes?

To address this question, we build a theoretical framework in which a monopolist search plat-

form decides whether to deploy AIOs across queries that differ in search intent. Deploying an AIO incurs a cost per search but can improve intent fulfillment, allowing the user to acquire new intents sooner. We then extend the framework in three directions. First, we allow the value of deployment to depend on the quality of available source material, reflecting the retrieval-augmented nature of AIOs (Google, 2025c). Second, we study a platform that cannot perfectly infer a user's intent. Uncertainty can therefore make AIO deployment worthwhile even if it would be unprofitable if the user's intent were known. Third, we give users an outside option in the form of a competing generative AI system. This creates a deployment motive that is absent under monopoly: the platform deploys an AIO to retain a user that organic search alone would lose, even where deployment lowers the value of a user already committed to it. For each extension we characterize when the platform benefits from deployment and derive comparative statics that generate testable hypotheses.

To test the model's predictions, we construct a novel dataset of Google SERPs collected in two waves. Within each wave we collect multiple SERPs per query using a web scraper, alternating across device types, days, and times to capture within-query variation in deployment. The first wave was collected in November 2025 and the second in August 2026, on the same queries in the same locations, and together they cover 24,217 result pages nested within 1,575 query clusters. Because the two waves hold query composition fixed, the change in deployment between them likely reflects one of two market changes: changes to Google's AIO capabilities, or competitive pressure from other generative AI alternatives (Simon et al., 2025).

We find significant heterogeneity in AIO deployment. Exploratory queries receive an AI Overview on 52.0% of result pages against 19.8% for targeted queries. Monetization lowers deployment by 13.0 percentage points. This is driven largely by exploratory intents, where deployment would otherwise be the platform's default. Where a query's intent is ambiguous, targeted queries are 7.3 points more likely to receive an Overview and exploratory queries 7.5 points less likely. Organic fulfillment lowers deployment on targeted queries and has no significant effect on exploratory ones, the reverse of what the model predicts, and the advertising value of a query predicts more deployment rather than less. Deployment rises with source quality at low levels before flattening at higher levels, with no significant evidence of decline. Between the two waves, deployment rose by 9 percentage points, and the increase is concentrated on the queries from which the platform has the lowest monetization potential.

2 Related Literature

Our paper examines how an ad-supported search platform selectively deploys AI synthesis. It builds on research showing that intermediaries strategically design the search environments through which consumers discover and evaluate alternatives. Hagiu and Jullien (2011) examine incentives to divert search, while Dukes and L. Liu (2016) study how an intermediary chooses search costs that influence the breadth and depth of product evaluation. Hagiu and Wright (2024) show that facilitating discovery creates a tradeoff between attracting new buyers to sellers and exposing those sellers' existing customers to competing offers. Empirical work likewise demonstrates that platform design shapes matching and demand allocation on Airbnb and Amazon (Fradkin, 2017; Lee and Musolff, 2025). In this work the intermediary designs the environment in which users search among third-party offerings, but does not itself produce the answer. We study a platform that can, and ask when it chooses to fulfill a query on the SERP rather than route the user to a listing.

Research on search advertising shows that the monetization of search depends on the query and the information displayed alongside sponsored results. Ghose and Yang (2009) relate keyword characteristics and advertising position to clicks, conversions, and advertiser profitability. Yang and Ghose (2010) document positive interdependence between organic and sponsored search for a retailer, while L. Xu et al. (2012) model how organic listings affect sponsored-search bidding and advertising outcomes. A. Agarwal et al. (2015) show that competition from organic results can reduce sponsored click-through rates while increasing conversion rates conditional on a click. The tension between improving search and preserving advertising revenue is also examined in theoretical work. Taylor (2013) studies how better organic search attracts users while potentially cannibalizing sponsored clicks, and White (2013) analyzes the relationship between organic search quality and sponsored-search profitability. These works take the elements of the SERP as given and study how organic and sponsored results interact within it. We model the decision that adds a new element to the page, and show that it is governed by the same tension: an AI Overview accelerates fulfillment and brings forward subsequent searches, but displaces the click on which the current query is monetized.

A growing literature examines how generative AI changes the way consumers acquire information and make decisions. In a field experiment, Zhu et al. (2025) find that generative AI search

changes query formulation and browsing behavior and increases purchases. Spatharioti et al. (2025) show that search using LLMs can accelerate decision making, but that reliance on incorrect generated information can impair decisions. C. Kaiser et al. (2025) similarly document changes in task completion and external browsing when consumers use generative AI instead of a conventional search engine. This evidence establishes that generated answers can change the effort required to resolve a need and the extent to which consumers consult external websites. Related theoretical work examines the economic incentives surrounding generative search. X. Liu et al. (2025) analyze how AI Overviews affect websites' content investment and sponsored-search bidding, while Zhang et al. (2026) study the tradeoff between advertising revenue and subscription conversion in generative engines. Both strands take the presence of a generated answer as given and study responses to it. H. Xu et al. (2026) document variation in AIO activation and examine source quality and claim fidelity. We model how search intent affects the platform's decision to supply an AI Overview, and show that the queries selected for AIO deployment differ on the characteristics that predict the outcomes this literature measures.

Our paper also relates to research on coopetition on platforms, which arises when platforms intermediate third-party offerings while also competing with their providers. Calzada and Gil (2020) show that removing Google News reduced visits to news publishers, demonstrating that aggregation can generate valuable referrals. Related work examines how platforms allocate visibility to their own offerings (Zou and Zhou, 2025; Long and Amaldoss, 2024) and how platform entry affects providers' participation and innovation (Choi et al., 2025; Foerderer et al., 2018). In particular, Long and Amaldoss (2024) analyze the revenue tradeoff involved in allocating sponsored placement to a platform's private label. For generative search, Khosravi and Yoganarasimhan (2026) estimate that the introduction of AI Overviews reduced search referrals to Wikipedia. The logic in the coopetition literature is that the platform favors a product of its own over third-party content it intermediates. An AI Overview is different in that it is composed from the third-party results it displaces. This way, the same content supports organic fulfillment and the platform's substitute for it. We examine how the quality of that content shapes the decision to compose.

The competitive setting matters because the value of fulfilling a query depends on whether doing so affects the platform's broader relationship with the user. Telang et al. (2004) show how users' willingness to consult another search engine after an unsuccessful attempt can sustain competing

providers. Allcott et al. (2025) provide experimental evidence that switching frictions, beliefs about quality, and inattention influence search-engine choice. M. Kaiser and Schulze (2026) document how ChatGPT referrals generate lower conversion rates than traditional search engines, but perform better for complex product categories. This suggests that generative AI may compete with traditional search, particularly for complex products. This work studies whether users move between providers. We study how the incumbent responds to the possibility that they might, and show that the response differs according to whether users choose a provider for each query or commit to one for all of them.

3 Theoretical Framework

3.1 Search Intent Taxonomy

A search engine receives a query as a short string of words and returns a set of results. Which results are appropriate depends on what the user is trying to accomplish, and the literature on web search refers to this underlying purpose as search intent (Broder, 2002; Jansen et al., 2008). Classic information retrieval assumes that users seek information, whereas web queries may also aim to reach a particular site or carry out an activity (Broder, 2002). Search intent therefore determines what the results should accomplish: a user seeking a particular site needs results that direct them there, whereas a user seeking to understand a topic needs results that inform them. Results useful in one instance may serve poorly in another. Yet intent is not directly observed. The search engine must infer the user's purpose from a few words before selecting results, and this inference cannot be made with certainty (Jansen et al., 2008).

Broder (2002) distinguishes navigational queries seeking a specific website, informational queries seeking information, and transactional queries seeking to perform a web-mediated activity such as shopping, downloading a file, or consulting a database. Jansen et al. (2008) operationalize these definitions for automated classification. Dai et al. (2006) identify commercial intention as a separate dimension that cuts across these categories. Within purchase-related searches, they distinguish a research phase, in which users compare options, from a commitment phase, in which they seek to complete a transaction. Google's own guidelines likewise classify queries by user intent and recognize the possibility that a query might cover multiple intents (Google, 2025d).

Borrowing from these frameworks, we organize search intents along two dimensions: *exploration*

Table 1: Search intent classification by exploration and monetization capacity.

	Exploratory	**Targeted**
Non-monetizable	Informational	Navigational
Monetizable	Commercial	Transactional

and *monetization capacity*. These dimensions distinguish what an AIO can contribute to fulfillment from the monetization opportunity associated with the query. Exploratory searches involve acquiring information or comparing options, whereas targeted searches seek a specific destination or action. For exploratory searches, fulfillment requires users to process the information returned, and an AI-generated synthesis on the SERP can perform some of this work. For targeted searches, results fulfill the query by routing users to the desired destination or action.

Monetization capacity distinguishes searches along a path toward a purchase from those that are not (Dai et al., 2006). We classify the former as monetizable and the latter as non-monetizable. Because the user is on a path toward purchase, monetizable queries are valuable to advertisers and therefore offer the platform direct revenue from sponsored clicks. Crossing the two dimensions yields the four intent types in Table 1, which are also used by practitioners.[2]

3.2 Motivating Search Model

We first introduce a model of *organic search* in which the platform does not deploy AI Overviews.

Consider an infinite-horizon, discrete-time model with a monopolist search platform and a unit mass of users. Users may have one of the following search intents according to the exploration–monetization taxonomy in Table 1. Let $h \in \mathbb{H} = \{I, N, C, T\}$, where I, N, C, and T denote informational, navigational, commercial, and transactional intents, respectively. Let $x_h = 1$ indicate that intent h is exploratory, otherwise $x_h = 0$. Likewise, $m_h = 1$ indicates a monetizable intent, otherwise $m_h = 0$.

Intents are drawn with probability $P_h > 0$, with $\sum_{h \in \mathbb{H}} P_h = 1$, and are organically fulfilled with probability $f_h \in (0, 1)$ in each period.[3] If a search intent is unfulfilled, the user continues to hold the same intent. And if the search intent is resolved the user exits the market with probability ρ_h, otherwise she draws a new search intent. Exit is permanent and has value zero.

[2]See Ahrefs (`https://ahrefs.com/seo/glossary/search-intent`) and DataForSEO (`https://dataforseo.com/blog/keyword-search-intent-api-for-google-seo`), accessed September 2026.

[3]In practice search intents might be correlated. We make the independence assumption for tractability.

The platform discounts future payoffs by $\delta \in (0,1)$. Let $a_h > 0$ be its expected per-period click payoff for monetizable queries. The monetizable classification concerns whether the user's search lies on a path toward a purchase. To connect this classification to platform revenue, we make the simplifying assumption that only monetizable intents generate a positive current click payoff, so $m_h a_h = 0$ for non-monetizable intents.

Let $V_0 = \sum_{h \in \mathbb{H}} P_h V_h$ denote the expected value of a new intent under organic search, and the platform's value in intent state h is

$$V_h = m_h a_h + \delta\big[(1 - f_h)V_h + (1 - \rho_h) f_h V_0\big]. \tag{1}$$

Proposition 1 (Baseline equilibrium)**.** *Let $\lambda_h = P_h/[1 - \delta(1 - f_h)]$. The organic-search model has a unique equilibrium, with $V_h \geq 0$ for every h and*

$$V_0 = \frac{\sum_{h \in \mathbb{H}} m_h a_h \lambda_h}{1 - \delta \sum_{h \in \mathbb{H}} (1 - \rho_h) f_h \lambda_h} \geq 0.$$

3.3 AI Overview Deployment

The platform can deploy an AI Overview.[4] Deploying this feature incurs the cost $d > 0$ in each period of deployment.[5] It forgoes the current click payoff $m_h a_h$, giving current expected payoff $-d$.[6] It fulfills intent h with probability $f_h^{AI} \in (0,1]$, where $f_h^{AI} > f_h$ when $x_h = 1$ and $f_h^{AI} = f_h$ when $x_h = 0$. An unsuccessful attempt leaves the same intent unresolved next period under either method.[7] Whenever the platform is indifferent between deploying and not, we assume that it deploys the AIO.

For each intent h, the gain from deploying the AIO before deployment cost is

$$\Delta_h \equiv \underbrace{\delta\big[(f_h^{AI} - f_h)((1 - \rho_h)V_0 - V_h)\big]}_{\text{change in continuation value}} - \underbrace{m_h a_h}_{\text{forgone current payoff}}. \tag{2}$$

[4] AIOs have the potential to include other non-synthesis based features, but for simplicity in our model, we suppose they consist only of synthesis.

[5] AI-generated syntheses are created on the fly and therefore incur a cost per search, whereas snippets and other excerpts incur costs primarily when indexed.

[6] If the platform can embed advertisements, $b_h \geq 0$ per period, into the AIO, this lowers the effective deployment cost to $d - b_h$ which weakly increases the likelihood of deploying AIOs.

[7] Our results are qualitatively robust to sufficiently small positive AI fulfillment gains when $x_h = 0$.

Proposition 2 (Deployment condition)**.** *Deployment for intent h strictly raises the platform's equilibrium value if and only if $\Delta_h > d$.*

For an otherwise comparable query, exploration increases AI's fulfillment advantage over organic search. Faster AI fulfillment shortens the sequence of costly AI attempts and brings forward subsequent search when the user continues. A targeted AIO changes neither fulfillment nor exit and therefore only adds deployment costs and sacrifices any click payoff. Targeted deployment is consequently never profitable when intent is known. Corollary 1 in Appendix A.2 establishes the exploration ordering with the organic primitives fixed, motivating the following hypothesis.

Hypothesis 1 (Exploratory queries)**.** *AI Overviews are weakly more likely for exploratory than for targeted intents.*

An AIO also sacrifices clicks. For a monetizable intent, those clicks generate revenue both now and on later attempts if the query remains unresolved. Making an intent monetizable raises its value under organic search, while a policy that always deploys on that intent collects none of its click revenue. Deployment therefore becomes less attractive even after continuation values adjust. Corollary 2 in Appendix A.2 formalizes this comparison, motivating the following hypothesis.

Hypothesis 2 (Monetizable queries)**.** *AI Overviews are weakly less likely for monetizable than for non-monetizable intents.*

Monetization can reduce AIO deployment only among exploratory queries. Targeted queries receive no AIO even when they are non-monetizable, so monetization cannot reduce their deployment further. Proposition 3 in Appendix A.2 establishes this interaction and a larger reduction in the exploratory deployment gain when its expected per-period click payoff is at least as high. Repeated advertising opportunities may motivate that payoff ordering relative to an infrequent transaction commission. The deployment comparison motivates the following hypothesis.

Hypothesis 3 (Monetization and exploration)**.** *Monetization weakly reduces AI Overview deployment for exploratory queries and has no effect for targeted queries.*

When an exploratory AIO is profitable, faster fulfillment brings forward future search opportunities whose value outweighs the forgone click revenue and deployment cost. Better organic fulfill-

ment delivers more of this benefit while preserving the current click payoff, reducing AI's advantage. Holding AI fulfillment fixed, exploratory deployment therefore weakly decreases with organic fulfillment. Targeted deployment remains unprofitable. Proposition 4 in Appendix A.3 establishes these comparisons, motivating the following hypothesis.

Hypothesis 4 (Organic fulfillment)**.** *Higher organic fulfillment weakly reduces AI Overview deployment for exploratory queries and has no effect for targeted queries.*

A higher payoff from the focal monetizable intent strengthens the organic alternative. A policy that always deploys on that intent forgoes its revenue on every visit, so the increase in continuation value cannot overturn the direct revenue loss. Higher payoffs from other intents make subsequent search more valuable without adding to the current query's forgone revenue. This can encourage exploratory deployment when AI improves fulfillment and $\rho_h < 1$, both for non-monetizable queries and for monetizable queries whose own payoffs are unchanged. A non-monetizable query's own a_h has no effect because $m_h = 0$. Targeted deployment remains zero. Proposition 5 in Appendix A.4 establishes both revenue effects with equilibrium continuation values, motivating the following hypothesis.

Hypothesis 5 (Cost per click)**.** *For exploratory monetizable queries, AI Overviews are weakly less likely as the query's cost per click increases.*

Production and source quality. Organic search and AI fulfillment depend on the same underlying sources. Poor sources limit AI's ability to fulfill the query. At the highest source quality, organic search already fulfills it, so AI's fulfillment advantage disappears. Deployment is therefore unprofitable at both extremes. When AI fulfillment is strictly concave in underlying source quality and deployment is viable, Proposition 6 in Appendix A.5 establishes three source-quality regions: *data-void* at low quality, *deployment* at intermediate quality, and *redundancy* at high quality. Since AIOs do not improve fulfillment for targeted queries, they remain unused regardless of source quality. This motivates the following hypothesis.

Hypothesis 6 (Source quality)**.** *For exploratory queries, AI Overviews are more likely when underlying sources are of intermediate quality than when they are of low or high quality.*

Intent ambiguity. A known targeted intent never warrants an AIO in the baseline model. When the underlying intent is ambiguous, deployment for a query classified as targeted can sometimes be profitable if it actually turns out to be an exploratory intent. Hence, intent ambiguity can weakly increase the profitability of deployment on queries classified as targeted. For queries classified as exploratory, greater ambiguity weakly discourages deployment when the known exploratory intent would warrant an AIO. Proposition 7 in Appendix A.6 establishes this conditional comparison. These comparisons motivate the following hypothesis about average deployment patterns for targeted and exploratory queries.

Hypothesis 7 (Intent ambiguity)**.** *Greater intent ambiguity weakly increases AI Overview deployment for queries classified as targeted and weakly decreases it for queries classified as exploratory.*

The role of competing search methods. To examine the role of competing search methods, we extend the setting to one where the search platform competes with a rival AI provider.[8] The rival fulfills each intent with its own probability, and we allow the platform's AIOs to improve both exploratory and targeted fulfillment. When an intent is fulfilled, users exit with probability ρ_h and otherwise draw a new intent, regardless of which provider fulfilled it. Users observe the platform's AIO policy before choosing a provider.

For each search attempt, users choose the provider with higher fulfillment probability. Consider queries for which the platform initially does not deploy AIOs because doing so would lower profit conditional on serving them. When the rival performs better than organic search, the platform can deploy AIOs to retain the query. For non-monetizable queries, the platform gains nothing from paying to match the rival. Therefore, it chooses to free-ride on the rival's ability to speed up the arrival of monetizable queries.

Conversely, for monetizable queries, AIO deployment that was previously unprofitable may become worthwhile to retain the query. When the gain in fulfillment from deploying AIO is smaller for targeted than for exploratory queries, the platform must increase AIO deployment more for targeted queries to match the same improvement in the rival's fulfillment. Proposition 8 in Appendix A.7 establishes these comparisons.

Suppose instead that users commit to a provider for all future intents at the start of their search

[8]Simon et al. (2025) document growing use of generative AI for information seeking.

process. With commitment, securing the user allows the platform to capture her total lifetime value, so any AIO that improves fulfillment can be worthwhile, even if it yields no direct revenue. This can support deployment on targeted non-monetizable queries even when their fulfillment gains are small. Proposition 9 in Appendix A.7 establishes conditions for this response to stronger competition.

These arguments motivate the following hypotheses for competition without and with user commitment, respectively. The hypotheses focus on the parameter space where AIO deployment is feasible and profitable.

Hypothesis 8A (Competition for queries)**.** *For monetizable queries, stronger competing AI search alternatives weakly increase AI Overview deployment, and more so for targeted than exploratory queries.*

Hypothesis 8B (User commitment)**.** *When users commit to a search method, stronger competing AI alternatives increase AIO deployment for non-monetizable queries.*

4 Empirical Analysis

4.1 Data Collection

To study how search engine behavior depends on our intent classification, we require a dataset that covers the intent space and provides sufficient variation. Using existing search query data, such as top Google search phrases, would not help us address our research question because they exhibit selection bias toward popular (largely targeted) queries, and hence lack sufficient variation. We therefore construct a novel dataset combining search query characteristics with the corresponding SERPs collected from Google in two waves, in November 2025 and August 2026. We outline our data generation process here, and include details in Appendix B.

We generated candidate queries using an LLM seeded with high-frequency keywords from a commercial search analytics API covering the United States and a representative EU country (Ireland).[9] Queries varied across search intent, YMYL status, local relevance, question format, and length. We checked candidates against the API and required at least 10 monthly searches.[10] After

[9]We selected Ireland as the representative country for the EU sample as it facilitated a comparison between two English speaking populations. Including an EU country allows us to compare deployment across regulatory environments, such as compliance with the GDPR.

[10]We use the commercial provider DataForSEO, see `https://dataforseo.com/apis`.

deduplication and filtering, 1,575 unique queries remained. We also collected average cost per click and average monthly search volume over the preceding 12 months. We separately used an LLM to classify each query along the exploration and monetization dimensions. The resulting sample comprised 548 informational, 512 commercial, 323 navigational, and 192 transactional queries. Our data generation process follows the recommendations and validation practices set out by Carlson and Burbano (2025).

For each query, we collected SERPs using a web scraper for Google search results. AIO deployment is not fixed at the query level. The same query may or may not trigger an AIO across repeated searches depending on Google's real time assessment of helpfulness and response quality (Google, 2025c). We therefore collected multiple SERPs per query in each wave, alternating between mobile and desktop device contexts as well as day of the week and time of day. For local queries, we specified geographic coordinates through Google's UULE parameter.[11]

The analysis sample comprises 11,648 observations from the 2025 wave and 12,569 from the 2026 wave, nested within 1,575 query clusters.[12] For each SERP, the scraper returns the full HTML page as well as structured data including organic results (URLs, titles, snippets, and positions), paid results, related queries, and, in the 2026 wave, its text and cited sources.

4.2 Measures

Our primary dependent variable is *AIO deployment*, a binary indicator for whether Google displayed an AIO on the SERP. AIOs were detected through HTML parsing of Google's distinctive CSS markers. Detection accuracy was verified through manual inspection.

We measure *organic fulfillment* using Google's Search Quality Rater Guidelines (Google, 2025d), whose five-point Needs Met scale assesses how well a result satisfies the user's intent, from Fails to Meet (1) to Fully Meets (5). An LLM applies these criteria to the top ten organic results, given the query and user location. This assessment excludes the AIO where available. We aggregate the ratings into a SERP-level score that gives greater weight to higher-ranked results, capturing both fulfillment and prominence. We validated these ratings against Google's examples.

The guidelines impose different rating ceilings across intents, and the dispersion of the resulting

[11] We used Los Angeles and Dublin as the representative cities for United States and Irish queries respectively.
[12] Due to technical issues, we were unable to collect some SERPs in the 2025 wave.

scores differs across waves. We therefore standardize the measure within wave and exploration category, so that a one-unit increase represents one standard deviation in organic fulfillment among comparable queries in the same wave. Appendix C details the rating, aggregation, and validation procedures.

We measure *source quality* by the average domain authority across the top ten organic results. Higher scores indicate more authoritative and widely referenced sources.[13] Google-owned properties are excluded from this measure. Their presence among the results is plausibly a response to the quality of the available third-party sources, since the platform may rank its own properties precisely where external sources are weak, so including them would confound the availability of source material with the platform's response to it. That response is captured separately by the Google ecosystem score described below. Source quality is therefore undefined for the 30 SERPs whose top ten organic results consist entirely of Google-owned domains, mostly navigational searches for other Google products, and these SERPs are excluded from the analysis sample.

Cost per click (CPC) is the realized average price in USD per click. We use CPC as a proxy for the model's expected per-period click payoff.[14] In the model the click payoff enters only for monetizable intents. In practice, advertisers do bid on non-monetizable queries, typically to defend a brand or intercept a competitor's, so a CPC is observed there as well. A distinguishing feature of monetizable queries is that the click sits on the user's path to a purchase, not that a price exists for it. CPC is unavailable for 35.9% of observations, which we interpret as the absence of an advertising market for the query rather than missing data.

Using the intent definitions in Google's guidelines, we define *intent ambiguity* as an indicator equal to one when more than one intent plausibly applies to a query. For example, a search for *Netflix* may concern logging in or signing up. The measure captures the query's range of plausible interpretations rather than Google's own assessment of intent, which we do not observe. Under this definition, 47.8% of sample queries are ambiguous.

We use the collection waves as a proxy for competition from generative AI, and check for patterns consistent with our model's prediction. Both waves observe the same queries in the same locations, holding query composition fixed. The comparison does not isolate competition, however, because

[13] We use a commercial SEO tool which computes domain authority using backlink data, see `https://moz.com/learn/seo/domain-authority`.

[14] Expected revenue per search also depends on how often users click paid results, which we are unable to observe.

improvements in Google's own generative capabilities and reductions in AIO production costs could also increase deployment. The two competition extensions predict different patterns of deployment across the monetization dimension. When users do not commit, defensive deployment requires click revenue to cover its cost and is therefore confined to monetizable queries. With user commitment, deployment can also extend to non-monetizable queries because retaining the user preserves revenue from future searches. We therefore compare the wave effects across the monetization dimension to assess which regime is more consistent with the observed changes (Hypotheses 8A and 8B).

We control for mobile device (versus desktop), YMYL classification, local query status, question format, EU sample (Ireland versus United States), and the logarithm of average monthly search volume. YMYL (Your Money or Your Life) identifies queries where incorrect answers could affect users' health, financial stability, or safety. Google applies a higher standard for these queries (Google, 2025c). To capture the presence of Google's own properties in the organic results, we compute a weighted ecosystem score as the sum of the inverse ranks of Google owned domains appearing in the top ten results, representing an alternative fulfillment channel that may reduce the marginal value of deploying an AIO.

4.3 Empirical Model

We estimate a single pooled specification in which most hypotheses correspond to one marginal effect or one contrast of marginal effects, rather than separate regressions by intent type. Pooling the analysis provides a common specification for estimating and testing differences in deployment probabilities across intent types.

For SERP t of query i in wave w, we estimate

$$\begin{aligned}\Pr(\text{AIO}_{itw}=1) = \Lambda\big(\ & \beta_1 E_i + \beta_2 M_i + \beta_3 E_i M_i + \beta_4 W_w + \beta_5 E_i W_w + \beta_6 M_i W_w + \beta_7 E_i M_i W_w \\ & + \beta_8 f_{itw} + \beta_9 E_i f_{itw} \\ & + \beta_{10} D_{itw} + \beta_{11} D_{itw}^2 + \beta_{12} E_i D_{itw} + \beta_{13} E_i D_{itw}^2 \\ & + \beta_{14} A_i + \beta_{15} E_i A_i + X_{itw}'\gamma\ \big)\end{aligned} \tag{3}$$

where $\Lambda(\cdot)$ is the logistic function, E_i indicates an exploratory query, M_i a monetizable query,

W_w the second wave, f_{itw} organic fulfillment, D_{itw} source quality, A_i an ambiguous query, and X_{itw} the vector of controls. The three-way interaction among E_i, M_i, and W_w is included in full because Hypotheses 8A and 8B concern how the change between waves is distributed across both dimensions at once.

Each comparative static in the model holds the remaining primitives fixed, which motivates conditioning on their empirical proxies. We interpret f_{itw} primarily as query-specific fit and D_{itw} as source authority. Both can contribute to the model's organic fulfillment probability f_h. Including both lets us compare query fit at a given source authority and source authority at a given query fit, while entering either alone could mix these two associations. We nevertheless report equation (3) alongside a restricted specification containing only the intent dimensions, the wave, their interactions, and the controls, estimated on the same sample. We use the full specification for the main hypothesis tests. The restricted one shows that the results on the dimensions themselves do not depend on the moderator block.

Hypothesis 5 requires a separate estimation for two reasons. In the model the click payoff enters only for monetizable intents, and the hypothesis concerns exploratory monetizable queries specifically, so the relevant quantity is the effect of the click payoff in that single cell. In addition, CPC is observed only where advertisers bid on the query, and its absence is informative rather than a gap in the data, so replacing it with zero would identify the hypothesis partly off queries for which the click payoff is undefined. Because that cell is the commercial intent type, the test is a regression within it, on the SERPs where a cost per click is observed. The intent dimensions are constant there, so the specification carries the wave, the two page-level moderators that vary across a query's SERPs, and the controls.

Table 2 maps each hypothesis to the quantity that tests it and the sign the model predicts.

We report logistic estimates as our main results. A linear probability model need not keep predicted probabilities within the unit interval, and its estimates are biased when it does not (Horrace and Oaxaca, 2006). In our data, 7.3% of its predicted probabilities fall outside that interval. The within-query check in Section 5.7 uses a linear probability model instead, because a logit with a fixed effect for each query–location pair is subject to the incidental parameters problem.

Because most of our hypotheses are stated as a difference between the two sides of a dimension, we test them on average marginal effects and contrasts of marginal effects rather than on the coef-

Table 2: Hypotheses, test quantities, and predicted signs.

Hypothesis	Test quantity	Prediction
1 Exploratory queries	Marginal effect of E	Positive
2 Monetizable queries	Marginal effect of M	Negative
3 Monetization and exploration	Marginal effect of M at $E = 1$ versus at $E = 0$	Negative versus zero
4 Organic fulfillment	Marginal effect of f at $E = 1$ versus at $E = 0$	Negative versus zero
5 Cost per click	Marginal effect of $\ln \mathrm{CPC}$ within the commercial cell, on SERPs with an observed CPC	Negative
6 Source quality	Slopes at the lower and upper bounds of D and location of the turning point, for $E = 1$	Positive slope at the lower bound, negative slope at the upper bound, and an interior turning point
7 Intent ambiguity	Marginal effect of A at $E = 0$ versus at $E = 1$	Positive versus negative
8A Competition for queries	Marginal effect of W at $M = 1$, and contrast between $E = 0$ and $E = 1$ within $M = 1$	Positive, larger at $E = 0$
8B User commitment	Marginal effect of W at $M = 0$	≈ 0 under per-query choice, and positive under commitment

ficients themselves. In a logistic model the coefficient on a product term measures the interaction on the log-odds scale. The interaction effect on the probability of deployment depends on all covariates and can differ from the coefficient in sign and significance (Ai and Norton, 2003). Marginal effects avoid this and are directly interpretable as changes in the probability of deployment. We therefore report each test as an effect in percentage points with its confidence interval. Several of our hypotheses predict that a determinant has no effect on one side of a dimension, and an effect that is not distinguishable from zero is not evidence that the effect is absent. For these we report the interval and state the magnitudes it excludes, so that the claim rests on the precision of the estimate rather than on a failure to reject.

The same query is observed several times within each wave and in both waves, so we cluster standard errors on the query and location. Each query was generated for a single market, so a query–location cluster contains all SERPs collected for one query across both waves. The wave terms are therefore identified from repeated observation of the same queries rather than from differences in which queries are observed.

Hypothesis 6 predicts that, for exploratory queries, AI Overview deployment is more likely at

intermediate levels of source quality than at low or high levels. A significant quadratic coefficient alone is not sufficient to establish an interior maximum (Haans et al., 2016). We therefore apply the Lind and Mehlum (2010) procedure to test whether the slope is significantly positive at the lower bound of source quality, significantly negative at the upper bound, and whether the estimated turning point lies within the observed range. We regard the inverted U to be supported only when all three conditions are met, and report when they are not met. The hypothesis makes no claim about the targeted side, where the baseline model assigns no fulfillment gain, so we report that side descriptively.

5 Results

Table 3 reports summary statistics for the full sample and by intent cell. AI Overviews appear on 41.5% of SERPs overall, but deployment varies sharply along both intent dimensions: exploratory queries receive an AI Overview more than twice as often as targeted ones (52.0% against 19.8%), and non-monetizable queries receive one on 50.5% of SERPs against 30.3% for monetizable ones. The two dimensions combine so that deployment is highest for exploratory non-monetizable queries (69.8%) and lowest for targeted non-monetizable queries (17.7%), with the exploratory monetizable cell at 33.0% and the targeted monetizable cell at 23.2%. This pattern is consistent with the theoretical framework's central prediction: the default for an exploratory query is to deploy an AIO and the default for a targeted query is to withhold one, and monetization restrains deployment where it would otherwise be the default.

Several control variables merit a brief discussion before turning to the hypothesis tests. Mobile device context is associated with higher AIO deployment across all intent cells. This matters given the more constrained attention environment on mobile, where limited screen real estate amplifies the importance of the first element displayed (Ahn et al., 2018; Ghose et al., 2013): a deployed AIO can occupy half or more of the visible area even before the user expands it, displacing organic results below the fold and raising the likelihood of zero-click engagement. Local queries are associated with a strongly reduced deployment, at 17.1% against 48.9% for non-local ones, reflecting the difficulty of generating reliable syntheses for geographically specific needs. The EU sample is associated with lower deployment, possibly reflecting regulatory caution or differences in the availability of source

Table 3: Summary statistics by intent cell.

		Exploratory		Targeted	
	All	**Non-mon.**	**Monetizable**	**Non-mon.**	**Monetizable**
AIO deployment	0.415	0.698	0.330	0.177	0.232
	(0.493)	(0.459)	(0.470)	(0.382)	(0.422)
Organic fulfillment (0–1)	0.380	0.350	0.339	0.507	0.366
	(0.148)	(0.089)	(0.107)	(0.203)	(0.150)
Source quality (0–100)	70.03	72.55	62.10	78.86	69.19
	(15.93)	(12.26)	(16.13)	(15.73)	(15.01)
CPC (USD)	4.88	4.30	5.56	3.55	5.97
	(12.71)	(12.27)	(12.81)	(6.50)	(18.32)
Intent ambiguity (1=Yes)	0.481	0.468	0.592	0.462	0.256
	(0.500)	(0.499)	(0.492)	(0.499)	(0.436)
Search volume (thousands)	370.3	22.6	17.5	1,741.3	16.2
	(7,171.5)	(118.2)	(81.5)	(15,813.6)	(122.1)
Mobile device (1=Yes)	0.489	0.492	0.490	0.483	0.488
	(0.500)	(0.500)	(0.500)	(0.500)	(0.500)
YMYL (1=Yes)	0.386	0.438	0.317	0.323	0.527
	(0.487)	(0.496)	(0.465)	(0.468)	(0.499)
Local query (1=Yes)	0.233	0.113	0.425	0.143	0.213
	(0.423)	(0.317)	(0.494)	(0.350)	(0.410)
Question query (1=Yes)	0.267	0.491	0.178	0.056	0.215
	(0.442)	(0.500)	(0.383)	(0.229)	(0.411)
EU sample (1=Yes)	0.461	0.463	0.396	0.530	0.512
	(0.498)	(0.499)	(0.489)	(0.499)	(0.500)
Google ecosystem score	0.071	0.062	0.030	0.185	0.017
	(0.249)	(0.202)	(0.121)	(0.434)	(0.091)
SERPs	24,217	8,435	7,869	4,934	2,979
Queries	1,575	548	512	323	192

Notes: Means with standard deviations in parentheses, pooling the two collection waves (11,648 SERPs in 2025 and 12,569 in 2026). The four cells correspond to informational, commercial, navigational, and transactional queries respectively. CPC is observed for 15,531 SERPs and its mean is computed on those. Organic fulfillment is the position-discounted Needs Met score described in Section 4.2, reported here on its original 0–1 scale rather than standardized.

material. Deployment is higher for YMYL queries and for queries phrased as questions, and it declines with search volume, so the platform deploys less on the largest queries once intent and result quality are held fixed.

Table 4: Determinants of AIO deployment.

	Probability of an AI Overview		
	(1)	(2)	(3)
Exploratory	3.031*** (0.193)	3.111* (1.226)	
Monetizable	0.619** (0.230)	0.523* (0.266)	
Exploratory × Monetizable	-1.524*** (0.260)	-1.228*** (0.294)	
Wave 2026	1.885*** (0.170)	1.880*** (0.173)	0.149 (0.096)
Exploratory × Wave 2026	-1.324*** (0.188)	-1.273*** (0.192)	
Monetizable × Wave 2026	-1.243*** (0.229)	-1.040*** (0.236)	
Exploratory × Monetizable × Wave 2026	0.766** (0.257)	0.538* (0.264)	
Organic fulfillment		-0.360*** (0.091)	-0.014 (0.059)
Exploratory × organic fulfillment		0.367*** (0.100)	
Source quality		0.033 (0.032)	0.021*** (0.005)
Source quality squared		-1.44 (2.35)	
Exploratory × source quality		0.016 (0.038)	
Exploratory × source quality squared		-1.33 (2.88)	
Intent ambiguity		0.560*** (0.157)	
Exploratory × intent ambiguity		-0.971*** (0.192)	
ln(CPC)			0.472*** (0.111)
ln(search volume)	-0.107*** (0.017)	-0.104*** (0.017)	-0.061 (0.033)
Mobile device	0.335*** (0.035)	0.360*** (0.036)	0.487*** (0.071)
YMYL	0.931*** (0.083)	0.978*** (0.085)	0.257 (0.203)
Question query	0.723*** (0.102)	0.660*** (0.104)	0.391 (0.225)
EU sample	-1.009*** (0.093)	-0.942*** (0.094)	-0.927*** (0.180)
Google ecosystem score	-0.292 (0.193)	-0.228 (0.181)	0.143 (0.499)
Local query	-1.661*** (0.128)	-1.722*** (0.151)	-1.276*** (0.190)
Constant	-1.961*** (0.208)	-3.931*** (1.028)	-2.097*** (0.488)
SERPs	24,217	24,217	6,130
Query clusters	1,575	1,575	425
Pseudo R^2	0.276	0.288	0.170

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Logistic regression for the probability that an AI Overview is deployed. Standard errors clustered on query and location. Column (1) contains the intent dimensions, the wave, their interactions, and the controls. Column (2) adds the moderators. Column (3) tests Hypothesis 5 within the commercial cell, on the SERPs for which a cost per click is observed; the intent dimensions and their interactions are constant there and drop out, and the specification retains the wave and the two page-level moderators that vary across a query's SERPs. The reference category for intent ambiguity is the single-intent query. The squared source-quality terms are reported in units of 10^{-4}. Hypotheses are tested on the marginal effects in Table 5, except Hypothesis 5, whose marginal effect from column (3) is reported in Section 5.2.

Because our hypotheses concern differences between the two sides of an intent dimension, we test them using average marginal effects and contrasts between those effects, expressed in percentage points with 95% confidence intervals. Table 4 reports the logit coefficients and Table 5 the corresponding marginal effects from column (2), on which the hypothesis tests are based. Standard errors are clustered at the query–location level throughout.

Table 5: Average marginal effects on the probability of AIO deployment.

	Targeted		**Exploratory**		**Pooled / contrast**	
Adjusted deployment (%)	21.3		53.6			
Exploration (H1)					32.3	[29.4, 35.2]
Monetization, pooled (H2)					−13.0	[−16.1, −9.9]
Monetization, by side (H3)	−2.1	[−7.1, 2.9]	−18.7	[−22.9, −14.4]	−16.6	[−23.4, −9.8]
Organic fulfillment, per SD (H4)	−4.7	[−7.0, −2.4]	0.1	[−1.4, 1.6]	4.8	[2.1, 7.5]
Intent ambiguity (H7)	7.3	[3.2, 11.3]	−7.5	[−11.5, −3.5]	−14.8	[−20.4, −9.2]
Adjusted deployment by cell (%)						
non-monetizable	22.1		61.8			
monetizable	20.0		43.1			
Wave (H8A, H8B)						
non-monetizable	24.4	[20.7, 28.1]	11.3	[8.3, 14.3]	15.2	[12.9, 17.5]
monetizable	11.0	[6.7, 15.2]	2.0	[−1.1, 5.2]	4.8	[2.3, 7.2]
exploratory vs. targeted, monetizable					−8.9	[−14.2, −3.7]
monetizable vs. non-mon.					−10.4	[−13.8, −7.0]
SERPs	24,217					
Query clusters	1,575					

Notes: Average marginal effects in percentage points with 95% confidence intervals, computed from column (2) of Table 4 from 24,217 SERPs clustered across 1,575 queries. Adjusted deployment is the predicted probability on each side of the exploration dimension, and by cell, at the observed covariates. The final two columns carry effects that are not specific to one side of the exploration dimension: the exploration and pooled monetization effects, and the contrasts between the exploratory and targeted columns. In the wave block, the final columns carry the pooled wave effect on each side of the monetization dimension, the contrast between exploratory and targeted queries, and the contrast between monetizable and non-monetizable queries. The marginal effect of cost per click (Hypothesis 5) is computed from column (3) and reported in Section 5.2.

5.1 Intent Dimensions

Hypotheses 1 and 2 predict that deployment is higher for exploratory than for targeted queries and lower for monetizable than for non-monetizable queries. Table 5 reports the corresponding average marginal effects. Exploration increases the probability of deployment by 32.3 percentage points

(CI [29.4, 35.2]), and monetization decreases it by 13.0 points (CI [−16.1, −9.9]). Both effects are large relative to the overall deployment rate of 41.5%, and both are essentially unchanged when the moderators are excluded from the specification (column (1) of Table 4). These results support Hypotheses 1 and 2.

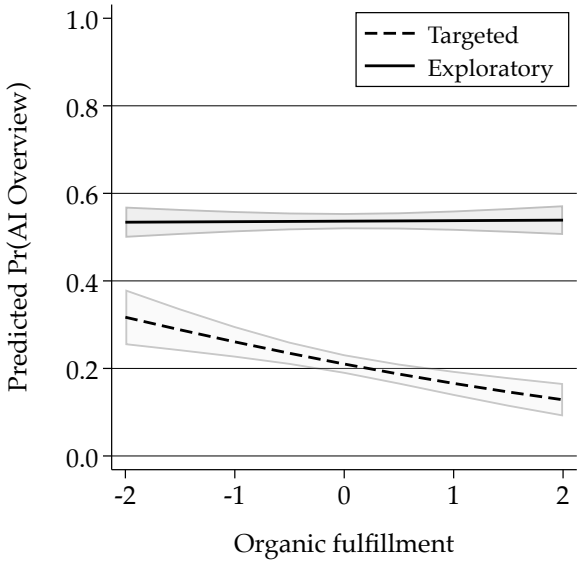
Hypothesis 3 predicts that the reduction in deployment associated with monetization is concentrated among exploratory queries, since only there is deployment the default and a click available to protect. Consistent with this prediction, the effect of monetization is −18.7 points among exploratory queries (CI [−22.9, −14.4]) and −2.1 points among targeted queries (CI [−7.1, 2.9]). The difference between the two, −16.6 points (CI [−23.4, −9.8]), is precisely estimated. In levels, the adjusted deployment probabilities are 61.8% for exploratory non-monetizable queries, 43.1% for exploratory monetizable queries, 22.1% for targeted non-monetizable queries, and 20.0% for targeted monetizable queries. These results support Hypothesis 3.

5.2 Organic Fulfillment and Cost per Click

Hypothesis 4 predicts that organic fulfillment reduces deployment for exploratory queries and has no effect for targeted queries. We find the opposite pattern. A one standard deviation increase in organic fulfillment is associated with a change of 0.1 points in deployment for exploratory queries (CI [−1.4, 1.6]), an estimate that rules out effects larger than two points in either direction, and with a decrease of 4.7 points for targeted queries (CI [−7.0, −2.4]). The difference between the two sides, 4.8 points (CI [2.1, 7.5]), has the opposite sign to the one predicted. Figure 1 illustrates this pattern: predicted deployment for exploratory queries is approximately flat across the observed range of fulfillment, whereas predicted deployment for targeted queries declines over the same range. Hypothesis 4 is therefore not supported. Organic results that address the user's query more closely may also improve the AI answer, whereas the prediction for exploratory queries holds AI quality fixed.

Hypothesis 5 predicts that, among exploratory monetizable queries, a higher cost per click reduces deployment. That cell is the commercial intent type, so the test is a within-cell comparison: among commercial queries for which an advertising market exists, does a higher cost per click lower deployment? Column (3) of Table 4 estimates this specification on the 6,130 such SERPs. A one-unit increase in log cost per click is associated with an *increase* of 8.3 points in deployment (CI

Figure 1: Predicted probability of AIO deployment by organic fulfillment.



Notes: Predicted probabilities from column (2) of Table 4. Organic fulfillment is standardized within wave and exploration category, so one unit on the horizontal axis is one standard deviation among comparable queries in the same wave. Shaded areas are 95% confidence intervals.

$[4.7, 12.0]$). Predicted deployment rises from 26.2% at the first quartile of cost per click to 29.2% at the median and 36.1% at the third quartile. The estimate is precise and has the opposite sign to the prediction, so Hypothesis 5 is not supported. Rather than protecting the queries whose clicks are worth most, the platform deploys most readily on exactly those queries. This test conditions on the existence of an advertiser market and therefore says nothing about whether its presence matters. Estimating the same specification on all 7,869 commercial SERPs, with an indicator for an observed cost per click in place of its level, we fidn that queries with an advertiser market are 6.1 points less likely to receive an Overview than those without one (CI $[-13.3, 1.1]$). The estimate has the sign the model predicts but is imprecise, and the interval includes zero.

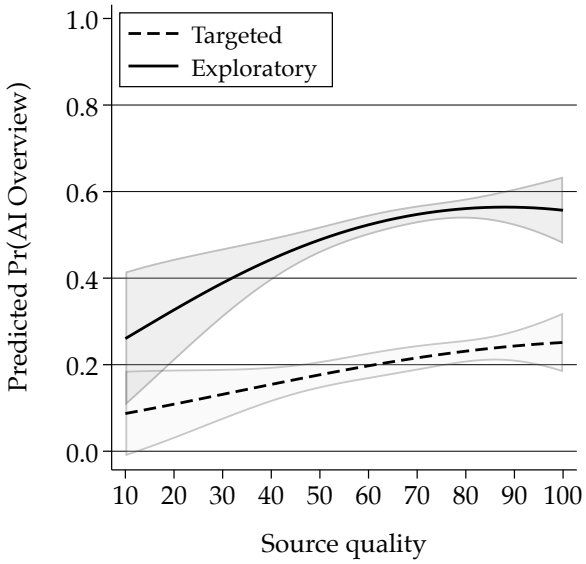
5.3 Source Quality

Hypothesis 6 predicts an inverted U-shaped relationship between source quality and deployment for exploratory queries, with deployment withheld where the available sources are too poor to support a reliable synthesis (the data-void region) and again where the organic results are good enough to make synthesis redundant (the redundancy region). This comparison assumes that, at a given f_{itw},

higher D_{itw} raises organic fulfillment f_h, with AI fulfillment changing according to $f_h^{AI} = \Psi_h(f_h)$. Predicted deployment for exploratory queries rises from 26.0% at a domain authority of 10 to 48.9% at 50 and 56.4% at 90. The Lind and Mehlum (2010) test formalizes this observation. The slope at the lower bound of the observed range is positive (0.047, CI $[0.007, 0.087]$) and the slope at the upper bound is not significantly different from zero (-0.006, CI $[-0.029, 0.017]$). The turning point falls inside the observed range, at 88.1 (CI $[57.9, 118.3]$) against an observed maximum of 99. The data are thus consistent with the data-void region, which also corresponds to Google's stated policy of withholding AI Overviews where reliable sources are lacking (Google, 2025c), but provide no evidence of a redundancy region. We regard Hypothesis 6 as partially supported.

Figure 2 plots the relationship for both sides of the exploration dimension. Deployment on targeted queries also rises with source quality, from 8.7% at a domain authority of 10 to 17.7% at 50 and 24.3% at 90, and a joint test of the two targeted source-quality terms rejects the null of no relationship ($\chi^2(2) = 8.26$, $p = 0.016$). The hypothesis makes no claim about the targeted side, since the baseline model assigns no fulfillment gain there, and we do not treat this as a test of it. We return to the targeted response in Section 5.5, where the change between waves gives it an interpretation.

Figure 2: Predicted probability of AIO deployment by source quality.



Notes: Predicted probabilities from column (2) of Table 4, pooling the two collection waves. Source quality is average domain authority across the top ten organic results, which ranges from 3.5 to 99 in the estimation sample. Shaded areas are 95% confidence intervals.

5.4 Intent Ambiguity

Hypothesis 7 predicts that intent ambiguity moves deployment away from the default on each side of the exploration dimension: ambiguity increases deployment for targeted queries, where the default is to withhold, and decreases it for exploratory queries, where the default is to deploy. The estimates in Table 5 are consistent with both parts of the prediction. Among targeted queries, those carrying more than one of the guidelines' user intents are 7.3 points more likely to receive an AIO than single-intent queries (CI $[3.2, 11.3]$). Among exploratory queries, they are 7.5 points less likely (CI $[-11.5, -3.5]$). The difference between the two effects is -14.8 points (CI $[-20.4, -9.2]$). In levels, predicted deployment for targeted queries increases from 17.6% for single-intent to 24.8% for multiple-intent queries, and for exploratory queries it decreases from 57.6% to 50.1%. These results support Hypothesis 7.

This result also helps reconcile the model with the observed level of deployment on targeted queries. The model predicts that a query whose users all hold a targeted intent should not receive an AIO, yet targeted queries receive one on roughly one fifth of SERPs. The ambiguity result indicates that this deployment is concentrated among queries that admit more than one plausible intent. In our sample, 93% of the ambiguous queries combine a targeted intent with an informational one, so ambiguity operates predominantly across the exploration dimension rather than within one side of it.

5.5 Competing AI Systems

Hypotheses 8A and 8B are derived from the same comparative static under two alternative assumptions about user behavior. In both cases, a stronger competing AI system increases deployment. The hypotheses differ in where the increase falls along the monetization dimension. Under per-query choice, the platform defends only those queries on which click revenue can cover the cost of deploying, so the increase is confined to monetizable queries (Hypothesis 8A). Under commitment, the platform accepts losses on non-monetizable queries to retain the user for the queries that pay, so the increase extends to non-monetizable queries (Hypothesis 8B). We test these predictions using the change in deployment between the two collection waves. Deployment rose from 36.6% of SERPs in 2025 to 46.0% in 2026, and the two hypotheses concern the distribution of that increase

across the monetization dimension.[15]

Table 5 reports the wave effects. Among monetizable queries, deployment increased by 4.8 points between waves (CI $[2.3, 7.2]$), and the increase was larger for targeted queries (11.0 points, CI $[6.7, 15.2]$) than for exploratory queries (2.0 points, CI $[-1.1, 5.2]$), a difference of -8.9 points (CI $[-14.2, -3.7]$). This ordering is consistent with the matching logic of the model: because an AIO improves fulfillment less for a targeted query, matching a given improvement in the competitor requires a larger increase in deployment there. These results are consistent with Hypothesis 8A.

Among non-monetizable queries, deployment increased by 15.2 points (CI $[12.9, 17.5]$). Under per-query choice this increase should be zero. Instead, it exceeds the increase among monetizable queries by 10.4 points (CI $[7.0, 13.8]$). The four intent cells order accordingly: the increase is 24.4 points for targeted non-monetizable, 11.3 points for exploratory non-monetizable, 11.0 points for targeted monetizable, and 2.0 points for exploratory monetizable queries. The expansion is therefore largest where the platform forgoes the least click revenue, and within each level of monetization it is larger on the targeted side. These results are consistent with Hypothesis 8B and indicate that the commitment regime describes the data better than per-query choice.

The wave comparison does not isolate competition, since the platform's own capacity to produce AIOs improved and its production cost fell over the same period. One feature of the data bears on that alternative, namely where the increase falls across the source-quality distribution. If the expansion reflected a general improvement in the platform's ability to generate a reliable synthesis, it could concentrate where the available material is weakest, since those are the queries a less capable system could not serve. Estimating the model separately on each wave, the increase on targeted queries instead grows with source quality: predicted deployment rises by 6.6 points between waves at a domain authority of 10 and by 20.2 points at 90. On exploratory queries the increase is roughly constant across the range, at 13.8 and 13.7 points respectively. The platform expanded onto targeted queries where the material to compose an AIO was strongest rather than where it was scarcest, which is consistent with matching a rival rather than general improvements. Separately, the first-wave estimates support the inverted-U relationship examined in Section 5.3. In 2025 the exploratory relationship satisfies all three conditions for an interior maximum, with a

[15] In 2025, AIOs were never generated for 42% of queries, sometimes generated for 46%, and generated in every instance for 12%. In 2026, the share never generated fell slightly to 40%, while the share always generated rose to 32%.

turning point at 72.2 (CI [63.9, 80.5]) inside the observed range, a positive slope at the lower bound (0.070, CI [0.031, 0.110]), and a negative slope at the upper bound (-0.027, CI $[-0.050, -0.005]$). In 2026, neither endpoint slope is statistically distinguishable from zero. A joint test of the four wave interactions does not reject equality across waves ($\chi^2(4) = 6.50$, $p = 0.16$), so we read the wave difference as suggestive rather than established.

5.6 The Intensive Margin: Placement, Sourcing, and Depth

The preceding sections examine whether an AIO is deployed. Conditional on deployment, the platform makes three further choices: where to position the AIO relative to the other elements of the SERP, which sources to compose it from, and how deep it should go in answering the query. Each choice extends the same tradeoff, since a prominently placed, long, and broadly sourced AIO resolves more of the query on the page and displaces more of the organic results than a low-placed restatement of the top result. In this section we examine whether the intent dimensions that govern deployment also govern these choices.

These additional analyses serve two purposes. Where deployment follows the model, finding that placement follows the same intent dimensions indicates that the platform's logic extends beyond the decision to deploy to how much of the answer appears on the page. Where deployment departs from the model, as it does for organic fulfillment on the targeted side, the composition of the deployed AIOs allows us to examine competing explanations for the departure.

An AI Overview can appear in one of four positions on the page: as the first element (79.5% of the 10,049 deployed AIOs), above the organic results but below a product carousel or featured snippet (10.7%), below paid results or the first organic result (8.7%), or embedded within the organic results (1.1%). Since four fifths of AI Overviews occupy the first position, the choice the platform makes is essentially whether the AIO leads the page or not. We therefore model placement as a binary outcome, equal to one if the AIO is the first element on the page, using the same specification as for deployment.[16] Table 6 reports the marginal effects for placement alongside those for deployment.

The scraper captured both the content of the AIO and the sources, i.e., the websites used by

[16] An ordered logistic regression over the four positions yields the same signs for every effect reported in Table 6. The proportional-odds assumption cannot be tested because the lowest position is empty in some cells of the three-way interaction. Separate estimates for desktop and mobile show the same pattern.

Table 6: Deployment and placement: average marginal effects.

	Deployment	**Placement**
	Pr(AI Overview appears)	Pr(AI Overview is first)
Effect of exploration	32.3 [29.4, 35.2]	31.0 [24.8, 37.2]
Effect of monetization, exploratory queries	−18.7 [−22.9, −14.4]	−10.4 [−14.4, −6.5]
targeted queries	−2.1 [−7.1, 2.9]	17.3 [6.4, 28.1]
contrast	−16.6 [−23.4, −9.8]	−27.7 [−38.9, −16.5]
Effect of organic fulfillment, exploratory queries	0.1 [−1.4, 1.6]	−0.4 [−1.7, 0.8]
targeted queries	−4.7 [−7.0, −2.4]	−10.6 [−15.7, −5.4]
contrast	4.8 [2.1, 7.5]	10.2 [4.9, 15.4]
Effect of intent ambiguity, targeted queries	7.3 [3.2, 11.3]	0.7 [−8.4, 9.9]
exploratory queries	−7.5 [−11.5, −3.5]	1.6 [−2.0, 5.1]
contrast	−14.8 [−20.4, −9.2]	0.8 [−9.0, 10.7]
Effect of the wave, targeted non-monetizable	24.4 [20.7, 28.1]	−49.4 [−60.4, −38.3]
targeted monetizable	11.0 [6.7, 15.2]	−16.8 [−28.2, −5.4]
exploratory non-monetizable	11.3 [8.3, 14.3]	4.6 [1.8, 7.4]
exploratory monetizable	2.0 [−1.1, 5.2]	8.6 [4.1, 13.1]
SERPs	24,217	10,049
Query clusters	1,575	1,141

Notes: Average marginal effects in percentage points with 95% confidence intervals. The deployment column is computed from column (2) of Table 4. The placement column is computed from a logistic regression with the same right-hand side for whether a deployed AI Overview is the first element on the page, estimated on the SERPs where an AIO appears. The second column conditions on the decision the first models and describes the placement of the AIOs the platform chose to deploy, not an effect on placement identified separately from that choice. Standard errors clustered on query and location.

Google's LLM for retrieval-augmented generation of the answer. These sources are displayed in a panel beside the AIO and as inline citations within its text, and because they are websites, we can match them to the organic results on the same SERP.[17] For each AIO we record how many distinct external domains it cites, what share of those domains also appear among the organic top ten, whether the top-ranked organic result is among them, and the average domain authority of the cited sources. We also record the length of the AIO in words.[18] Together these measures describe whether the AIO draws on the results it displaces and how much of the query it resolves on the page. Table 7 reports them by intent cell.

Exploration increases the probability that a deployed AIO is the first element on the page by 31.0 points (CI $[24.8, 37.2]$), close to its 32.3-point effect on deployment, and monetization decreases it by 10.4 points among exploratory queries (CI $[-14.4, -6.5]$). The pattern for organic fulfillment also carries over: fulfillment has no effect on placement among exploratory queries and reduces the probability of first placement by 10.6 points per standard deviation among targeted queries (CI $[-15.7, -5.4]$), a difference of 10.2 points (CI $[4.9, 15.4]$). Intent ambiguity is the exception. Its effect on placement is indistinguishable from zero on both sides of the dimension and the two do not differ from one another, so ambiguity governs whether an AIO appears but not where it appears. We make no prediction about placement. Instead, this result highlights that ambiguity is largely associated with the deployment decision, while the other three determinants carry through to the second decision as well.

Table 7 shows that the AI Overview is composed largely from the results it displaces. Of the 5,761 AI Overviews in the second wave that cite at least one external source, 98.1% cite a domain ranked in the organic top ten, 75.3% cite the top-ranked result, and on average 58.2% of the domains an AIO cites appear on the same page. The cited sources have a mean domain authority of 67.9, compared with 71.8 for the organic top ten on the same page (paired difference -3.9, $t = -24.1$), and among cited sources those that also rank on the page have a higher mean authority (72.9) than those that do not (61.4). The sources an AIO cites from outside the organic top ten therefore come from websites with lower domain authority than those on the SERP.

[17] Due to technical limitations, this data is only available for the 2026 wave.

[18] Sources are the union of the panel and inline citations, deduplicated by URL. Google-owned properties are excluded, as described in Section 4.2. Length is the word count of the AI Overview as first displayed, not of any expanded view.

Table 7: Sourcing and depth of the deployed AI Overview by intent cell, 2026.

		Cites top (%)	Source domains	On page (%)	Domain authority Cited	Domain authority Top 10	Words
Targeted	Non-monetizable	90.6	3.6	62.8	73.9	78.8	110
	Monetizable	81.5	6.1	65.7	66.4	67.0	129
Exploratory	Non-monetizable	71.3	6.2	56.9	68.0	72.6	165
	Monetizable	74.9	6.3	56.4	64.9	67.6	162
All deployed AI Overviews		75.3	5.9	58.2	67.9	71.8	155

Notes: Second wave, analysis sample, 5,761 AIOs citing at least one external source (721 targeted non-monetizable, 426 targeted monetizable, 3,245 exploratory non-monetizable, 1,369 exploratory monetizable). *Cites top* is the share of AIOs citing the domain ranked first among the organic results. *Source domains* is the mean number of distinct external domains cited. *On page* is the mean share of those domains that rank in the organic top ten. *Domain authority* reports the mean authority of the cited sources and of the organic top ten on the same page. *Words* is the median length of the AIO text, computed on AIOs with rendered text, which differs from the counts above by fewer than thirty in each cell. Cells correspond to navigational, transactional, informational, and commercial queries respectively.

AIOs on targeted queries are narrower and shorter than those on exploratory queries. They cite 4.6 domains on average compared with 6.3, cite the top-ranked result on 87.2% of pages compared with 72.4%, and in three quarters of cases draw on no result ranked below the first. Their median length is 119 words compared with 164. The difference in length is not explained by the difference in sourcing: holding the number of cited domains, organic fulfillment, source quality, and the page controls constant, an AIO on a targeted query is 40 words shorter (CI $[31, 49]$). The gap narrows as AIOs cite more domains, from 45 words at three cited domains to 35 at six and 25 at nine. These differences are insignificant.

The composition data allow us to examine two explanations for the fulfillment result in Section 5.2. If AIOs appear on targeted queries because the organic results fail to deliver the destination the user seeks, these AIOs should draw on results beyond the first. Instead, they cite the top-ranked result more often than AIOs on exploratory queries do. If they appear because the query is not in fact targeted, AIOs on ambiguous targeted queries should resemble those on exploratory queries. Instead, they are indistinguishable from AIOs on unambiguous targeted queries, citing the top result on 86.8% and 87.6% of pages and drawing on 4.7 and 4.4 domains, respectively.

Finally, the two margins together characterize the expansion between waves. Among targeted non-monetizable queries, deployment increased by 24.4 points while the predicted share of deployed AIOs placed first on the page fell from 77.3% to 27.9%. Among targeted monetizable queries, de-

ployment increased by 11.0 points and predicted first placement fell from 76.0% to 59.2%. Among exploratory queries, predicted first placement increased on both sides of the monetization dimension, to 90.6% and 81.9%. Since for a targeted query the correct result is itself the fulfillment, an AIO placed below that result and restating it is unlikely to interfere with the user's search journey. The platform thus expanded where it forgoes the least on the extensive margin and in the form that forgoes the least on the intensive margin.

5.7 Robustness Checks

We conduct three checks, each addressing one threat to the design. Appendix D reports the full estimates.

The fulfillment and source-quality effects in Table 4 are identified from variation both across and within queries. Because queries with strong organic results may differ from other queries in ways the controls do not capture, we re-estimate the model as a linear probability model with fixed effects for each query–location pair and each collection batch, so that the effects are identified only from variation across a query's own SERPs. The targeted-side fulfillment effect persists at -2.1 points per standard deviation (CI $[-4.1, -0.2]$), the exploratory-side effect remains indistinguishable from zero, and the contrast between them remains positive (2.8 points, CI $[0.6, 5.1]$). The ordering of the wave effects across intent cells is also preserved.

The positive slope of deployment in advertising value could reflect the platform monetizing the AIO itself. The model allows for this: if the platform can place ads within an AIO, the effective deployment cost falls to $d - m_h b_h$, and if that revenue rises with the query's advertising value, deployment would rise with CPC rather than fall. Google has placed ads within AIOs in the United States since 2024 but in no EU country, and excludes finance, health, and other sensitive verticals (Google, 2025a), so if that were the source of the slope it should be confined to United States queries outside YMYL, which covers the finance and health verticals in our sample. Interacting the CPC slope with the EU sample and with YMYL status shows the reverse ordering. The slope is 7.4 points per log point of CPC in the United States (CI $[3.4, 11.5]$) and 14.1 in the EU sample (CI $[8.8, 19.3]$), a difference of 6.6 points (CI $[1.0, 12.2]$). It is 12.7 points outside YMYL and 7.1 within, but the attenuation is as large in the EU sample as in the United States, so it does not reflect the advertising exclusion. In the cell with no in-AI Overview advertising under either rule,

YMYL queries in the EU sample, the slope is 13.3 points (CI $[7.7, 18.9]$). The reversal is a property of the deployment rule, not of the AIO's own advertising.

The wave comparison serves as our proxy for competition, but the platform's own capability improved over the same period: in January 2026, between our two collection waves, Google made Gemini 3 the default model for AI Overviews globally, having previously routed only the most difficult questions to it (Stein, 2026). That change, or a fall in the cost of producing an AIO, would raise deployment without any change in the outside option. A competing generative system may be less effective at serving a local query for want of a location signal and a business-listings feed, a limitation that benchmark evidence on local-service queries confirms is not closed by web access alone (He et al., 2026).

Under this interpretation, competition may contribute less to deployment increases for local queries. We find increases in both groups, with a smaller increase on local queries: deployment rose by 11.2 points on non-local queries (CI $[9.1, 13.2]$) and by 6.2 points on local ones (CI $[3.2, 9.2]$), a difference of -4.9 points (CI $[-8.7, -1.2]$). This pattern is consistent with the role of competition, but improvements in Google's capabilities and reductions in production costs may also affect the two groups differently. The difference therefore does not identify or bound competition's contribution. The incidence of the wave effect across the monetization dimension also holds within non-local queries, where the increase is 16.5 points for non-monetizable and 5.7 points for monetizable queries. Google's local pack and mapping product may also influence deployment on local queries. Anecdotally, we know competing LLMs' local capability was not zero by 2026, so these queries may also be exposed to competition. Finally, local queries are skewed towards commercial intents, which is why the estimate conditions on both intent dimensions.

5.8 Summary of Findings

Table 8 collects the tests. The two intent dimensions and the ambiguity prediction are supported, and the two competition predictions are consistent with the observed changes between waves. Two predictions are not supported. Organic fulfillment reduces deployment on the targeted side rather than the exploratory one, and cost per click raises deployment where the model predicts it should lower it. The source-quality prediction is partially supported: the relationship exhibits the data-void region the model implies but not the redundancy region.

Table 8: Summary of hypothesis tests.

Hypothesis	Prediction	Finding	Verdict
1 Exploratory queries	Positive	$+32.3$ pp $[29.4, 35.2]$	Supported
2 Monetizable queries	Negative	-13.0 pp $[-16.1, -9.9]$	Supported
3 Monetization and exploration	Negative at $E = 1$, zero at $E = 0$	-18.7 pp vs. -2.1 pp (n.s.)	Supported
4 Organic fulfillment	Negative at $E = 1$, zero at $E = 0$	$+0.1$ pp (n.s.) vs. -4.7 pp	Not supported
5 Cost per click	Negative	$+8.3$ pp $[4.7, 12.0]$	Not supported
6 Source quality	Positive slope at the lower bound, negative at the upper bound, interior turning point	Lower slope $+0.047$; upper slope ≈ 0	Partially supported
7 Intent ambiguity	Positive at $E = 0$, negative at $E = 1$	$+7.3$ pp vs. -7.5 pp	Supported
8A Competition for queries	Positive at $M = 1$, larger at $E = 0$	$+4.8$ pp; $+11.0$ pp vs. $+2.0$ pp	Consistent
8B User commitment	Positive at $M = 0$	$+15.2$ pp $[12.9, 17.5]$	Consistent

Notes: Marginal effects in percentage points from Table 5, with 95% confidence intervals, except for Hypothesis 5, whose marginal effect is computed from column (3) of Table 4 (Section 5.2). "Supported" and "not supported" apply to the hypotheses tested on a single marginal effect or contrast. Hypotheses 8A and 8B are stated in terms of the strength of competing systems, which we proxy by the collection wave; we therefore report whether the observed pattern is consistent with the prediction rather than treating the wave comparison as a direct test.

6 Discussion and Conclusion

We address the question of how search intent shapes a search platform's selective deployment of AI Overviews, and whether deployment is consistent with a platform trading off the fulfillment an AIO provides against the monetization it forgoes. Using two waves of Google search results, we find that deployment is consistent with such a tradeoff along the two intent dimensions we propose. Exploratory queries receive more AI Overviews, and the negative association between monetization and deployment is concentrated on the exploratory side. On other margins, deployment departs from this tradeoff. Organic fulfillment is unrelated to deployment for exploratory queries and negatively associated with it for targeted queries, and among exploratory monetizable queries deployment increases rather than decreases with cost per click. The source-quality prediction is supported only in part, since on exploratory queries deployment rises with the authority of the organic sources where these are weak but does not significantly decline where they are strong. Between the two waves, deployment expanded most on non-monetizable queries.

Our key contribution is to connect the selective provision of AI-generated syntheses to the economics of search intent. We organize established search intent categories along two dimensions: whether it is exploratory, and whether it is monetizable for the platform. Organizing search along these dimensions allows us to connect a user's search intent with a platform's incentive to provide a generative AI synthesis. In our model, an AIO can accelerate fulfillment and bring forward subsequent searches, but it sacrifices current click revenue and incurs a production cost for each search for which an AIO is generated. The platform can therefore find it worthwhile to deploy AIOs to answer a query that yields no immediate revenue while withholding an AIO on a monetizable query. Empirically, commercial queries receive fewer AIOs than informational queries, yet deployment rises with CPC within commercial queries. Lower average deployment on commercial searches therefore does not imply that the most expensive comparison keywords receive fewer AI answers.

Intent ambiguity helps explain departures from these broad patterns. A query classified as targeted with uncertainty can warrant an AIO when an exploratory interpretation makes deployment sufficiently attractive. Conversely, ambiguity can weaken the incentive to deploy on a query classified as exploratory. The opposite associations between ambiguity and deployment across the two categories are consistent with this mechanism. The framework thus connects the platform's choice

to both the economic value of resolving a search and the information available about what the user seeks.

The competition extension shows how the value of retaining users can encourage deployment for queries with no immediate advertising revenue. When users choose a provider for each search attempt, the platform can leave non-monetizable queries to a rival and still compete for subsequent monetizable searches. When users commit to a provider, serving those queries can help secure future revenue. The larger expansion on non-monetizable queries between November 2025 and August 2026 is consistent with this retention mechanism. Since we do not have clear measures of technological advancement and deployment costs, our results cannot be claimed to be causal. However, additional results on the placement of AIOs on a SERP lend support to the competition narrative: among targeted queries, deployment increased while the share of deployed AI Overviews appearing first on the page fell. In other words, AIOs were made more available but not necessarily made prominent. One possible interpretation is that the platform offers AIOs to match a feature its competitors provide so as to retain users, rather than to answer the query.

Two empirical findings merit further discussion. First, among exploratory monetizable queries, deployment increases with cost per click, whereas the model predicts that higher advertising value discourages deployment. Advertising placed within Overviews does not account for this, since the slope is at least as steep in the EU sample, where Google shows no such ads. The sign the model predicts appears, although imprecisely, on the other margin of advertising value: whether a commercial query carries an advertiser market at all is associated with fewer Overviews. One potential explanation for the reversal of CPC value is that an Overview helps users move from product comparison to a subsequent transactional search, and that where clicks on the comparison query are expensive, clicks on that subsequent search are valuable too. In the model, a higher payoff from subsequent searches encourages deployment, so the observed slope may combine the negative effect of a query's own click value with the positive effect of the searches it leads to.

Second, organic fulfillment does not significantly affect deployment for exploratory queries and has a negative effect on targeted queries. A potential explanation is that the platform's uncertainty about query intent and organic fulfillment are not independent. When Google is less certain about the user's intent, organic results may cover several interpretations and fit the user's query less well. Such uncertainty would then make AI Overviews more attractive for classified targeted intents and

less attractive for similarly classified exploratory intents. This interdependence may not be fully captured by our ambiguity indicator.

Our findings have implications for how stakeholders interact with search engines. For publishers, our findings point to coopetition with the platform. In the 2026 wave, 98.1% of AI Overviews that cite an external source cite at least one domain among the top ten organic results, and deployment on exploratory queries rises with the authority of those results. Credible content therefore supports AIO production but also competes with AIOs for users' attention. This has two implications. If users are clicking through to the AIOs' citations then publishers now have an incentive to produce content for AI citations. Conversely, if users are not viewing the underlying source, then AIOs may only serve as competition. For advertisers, our findings show that more valuable ad slots, with higher CPC, are more likely to compete with AIOs. This suggests that advertisers bidding on comparison queries may face more competition from AIOs than the lower deployment on monetizable queries overall would imply. AIO deployment and advertising prices may be endogenous. Advertising value influences the platform's deployment incentives, while deployment may affect CPC. For example, if AIOs reduced sponsored clicks and advertisers do not change spending targets, then CPC could rise.

Our analysis is also relevant for policy makers. We provide a first look at how search platforms deploy AI Overviews. Our results suggest that the platform's incentives in deploying AIOs need not align with those of the users, publishers, and advertisers who depend on it. Google states that AIOs are shown only where its systems determine that they add to classic Search, and that they therefore often do not appear (Google, 2025b). Monetization is not among the stated criteria, yet it is associated with lower deployment on exploratory queries, either because commercial value enters the decision or because Google judges AIOs less additive where users are comparing offers. Google also states that total organic clicks from Search to websites have been relatively stable year over year, without publishing supporting figures (Reid, 2025), while Chapekis et al. (2025) find that users who see an AI summary click a traditional result at roughly half the rate of those who do not. Because the first is an aggregate and the second a rate per search, the two can coexist if search volume rose, and whether the exposure we document is a loss for publishers depends on behavior our data do not capture. The European Commission opened a formal investigation in December 2025 into whether AIOs and AI Mode rely on publishers' content without appropriate

compensation and without the option to refuse such use without losing access to Google Search (European Commission, 2025). Our composition results indicate how closely AIOs draw on the results they appear alongside. In June 2026, the Competition and Markets Authority required Google to rank organic results in the United Kingdom, including those within AIOs, by objective and non-discriminatory criteria and to be more transparent about how rankings work, with six months to comply (Competition and Markets Authority, 2026b). Our findings concern the prior decision of whether an AIO is shown at all, and suggest that transparency should extend to that decision, including whether commercial value enters it. The Competition and Markets Authority has also required Google to let publishers in the United Kingdom opt out of their content being used in AI features such as AIOs while remaining in the organic results (Competition and Markets Authority, 2026a). Opt-outs change the pool of sources from which AIOs are composed, and how deployment responds to that change is a question for future research.

Our empirical analysis has a number of limitations. First, we constructed our sample to span the intent space rather than reproduce the actual distribution of search queries. We therefore cannot examine how a user's search journey or individual search patterns affect AIO deployment, or whether deployment differs across users. Second, we are unable to observe how users interact with AIO citations, limiting our ability to evaluate the value of being cited by AIOs. Third, we do not have clear measures of competition or AI quality improvements over time. Hence, our wave analysis should be taken as a crude proxy of how competition affects AIO deployment. By connecting search intent to deployment incentives, our framework explains where an ad-supported platform finds AI-generated answers worth providing. This provides a foundation for examining how generative search affects consumers and firms' opportunities to reach them.

References

Agarwal, Ashish, Kartik Hosanagar, and Michael D. Smith (2015). "Do Organic Results Help or Hurt Sponsored Search Performance?" *Information Systems Research* 26.4, pp. 695–713.

Agarwal, Saharsh and Ananya Sen (2026). "The impact of Google AI Overviews on publisher traffic and user experience: Evidence from a field experiment". *SSRN* 6513059.

Ahn, Jae-Hyeon, Yoon-Soo Bae, Jaehyeon Ju, and Wonseok Oh (2018). "Attention Adjustment, Renewal, and Equilibrium Seeking in Online Search: An Eye-Tracking Approach". *Journal of Management Information Systems* 35.4, pp. 1218–1250.

Ai, Chunrong and Edward C. Norton (2003). "Interaction terms in logit and probit models". *Economics Letters* 80.1, pp. 123–129.

Allcott, Hunt, Juan Camilo Castillo, Matthew Aaron Gentzkow, Leon Musolff, and Tobias Salz (2025). *Sources of market power in web search: Evidence from a field experiment*. Tech. rep. National Bureau of Economic Research.

Brand, James, Ayelet Israeli, and Donald Ngwe (2025). "Using LLMs for Market Research". Harvard Business School Marketing Unit Working Paper No. 23-062, revised October 2025.

Broder, Andrei (2002). "A Taxonomy of Web Search". *ACM SIGIR Forum*. Vol. 36. 2. ACM New York, NY, USA, pp. 3–10.

Calzada, Joan and Ricard Gil (2020). "What Do News Aggregators Do? Evidence from Google News in Spain and Germany". *Marketing Science* 39.1, pp. 134–167.

Carlson, Natalie and Vanessa Burbano (2025). "The Use of LLMs to Annotate Data in Management Research: Foundational Guidelines and Warnings". *Strategic Management Journal* 46.5, e70023. DOI: 10.1002/smj.70023.

Chapekis, Athena, Anna Lieb, Sono Shah, and Aaron Smith (2025). *What Web Browsing Data Tells Us About How AI Appears Online*. Tech. rep. Pew Research Center. URL: https://www.pewresearch.org/data-labs/2025/05/23/what-web-browsing-data-tells-us-about-how-ai-appears-online/.

Choi, Jay Pil, Doh-Shin Jeon, and Byung-Cheol Kim (2025). ""Sherlocking" and Platform Information Policy". *Management Science* 71.4, pp. 3073–3092. DOI: 10.1287/mnsc.2023.01187.

Clemons, Eric K and Nehal Madhani (2010). "Regulation of digital businesses with natural monopolies or third-party payment business models: Antitrust lessons from the analysis of Google". *Journal of Management Information systems* 27.3, pp. 43–80.

Competition and Markets Authority (June 3, 2026a). *CMA secures fairer deal for publishers and improves Google search services in UK*. URL: https://www.gov.uk/government/news/cma-secures-fairer-deal-for-publishers-and-improves-google-search-services-in-uk (visited on 09/29/2026).

— (June 17, 2026b). *Further CMA action to secure a fairer deal for businesses and improve Google search services in UK*. URL: https://www.gov.uk/government/news/further-cma-action-to-secure-a-fairer-deal-for-businesses-and-improve-google-search-services-in-uk (visited on 09/29/2026).

Dai, Honghua, Lingzhi Zhao, Zaiqing Nie, Ji-Rong Wen, Lee Wang, and Ying Li (2006). "Detecting Online Commercial Intention". *Proceedings of the 15th International Conference on World Wide Web*, pp. 829–837.

Dukes, Anthony and Lin Liu (2016). "Online Shopping Intermediaries: The Strategic Design of Search Environments". *Management Science* 62.4, pp. 1064–1077. DOI: 10.1287/mnsc.2015.2176.

European Commission (Dec. 9, 2025). *Commission opens investigation into possible anticompetitive conduct by Google in the use of online content for AI purposes*. URL: https://ec.europa.eu/commission/presscorner/detail/en/ip_25_2964 (visited on 09/29/2026).

Foerderer, Jens, Thomas Kude, Sunil Mithas, and Armin Heinzl (2018). "Does Platform Owner's Entry Crowd Out Innovation? Evidence from Google Photos". *Information Systems Research* 29.2, pp. 444–460. DOI: 10.1287/isre.2018.0787.

Fradkin, Andrey (2017). "Search, Matching, and the Role of Digital Marketplace Design in Enabling Trade: Evidence from Airbnb". *Working Paper*.

Ghose, Anindya, Avi Goldfarb, and Sang Pil Han (2013). "How Is the Mobile Internet Different? Search Costs and Local Activities". *Information Systems Research* 24.3, pp. 613–631.

Ghose, Anindya and Sha Yang (2009). "An Empirical Analysis of Search Engine Advertising: Sponsored Search in Electronic Markets". *Management Science* 55.10, pp. 1605–1622.

Google (2025a). *About ads and AI Overviews*. Undated help page. The 2025 version lists the United States only; the 2026 version adds eleven non-EU countries. Google Ads Help. URL: https://support.google.com/google-ads/answer/16297775?hl=en (visited on 09/29/2026).

— (Dec. 10, 2025b). *AI features and your website*. Google Search Central. URL: https://developers.google.com/search/docs/appearance/ai-features (visited on 09/29/2026).

— (2025c). *AI Overviews and AI Mode in Search*. Tech. rep. URL: https://search.google/pdf/google-about-AI-overviews-AI-Mode.pdf.

— (Sept. 11, 2025d). *General Guidelines*. Search Quality Rater Guidelines. No author stated; published and described by Google as its Search Quality Rater Guidelines. URL: https://static.googleusercontent.com/media/guidelines.raterhub.com/en//searchqualityevaluatorguidelines.pdf (visited on 09/29/2026).

Haans, Richard F. J., Constant Pieters, and Zi-Lin He (2016). "Thinking About U: Theorizing and Testing U- and Inverted U-Shaped Relationships in Strategy Research". *Strategic Management Journal* 37.7, pp. 1177–1195. DOI: 10.1002/smj.2399.

Hagiu, Andrei and Bruno Jullien (2011). "Why Do Intermediaries Divert Search?" *The RAND Journal of Economics* 42.2, pp. 337–362.

Hagiu, Andrei and Julian Wright (2024). "Optimal discoverability on platforms". *Management Science* 70.11, pp. 7770–7790.

He, Hang, Chuhuai Yue, Chengqi Dong, Mingxue Tian, Hao Chen, Zhenfeng Liu, Jiajun Chai, Xiaohan Wang, Yufei Zhang, Qun Liao, Guojun Yin, Wei Lin, Chengcheng Wan, Haiying Sun, and Ting Su (2026). "LocalSearchBench: Benchmarking Agentic Search in Real-World Local Life Services". *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. ACM. DOI: 10.1145/3770855.3817466.

Horrace, William C. and Ronald L. Oaxaca (2006). "Results on the bias and inconsistency of ordinary least squares for the linear probability model". *Economics Letters* 90.3, pp. 321–327.

Jansen, Bernard J., Danielle L. Booth, and Amanda Spink (2008). "Determining the Informational, Navigational, and Transactional Intent of Web Queries". *Information Processing & Management* 44.3, pp. 1251–1266.

Järvelin, Kalervo and Jaana Kekäläinen (2002). "Cumulated gain-based evaluation of IR techniques". *ACM Transactions on Information Systems (TOIS)* 20.4, pp. 422–446.

Kaiser, Carolin, Jakob Kaiser, Rene Schallner, and Sabrina Schneider (2025). "A New Era of Online Search? A Large-Scale Study of User Behavior and Personal Preferences during Practical Search Tasks with Generative AI versus Traditional Search Engines". *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, pp. 1–7.

Kaiser, Maximilian and Christian Schulze (2026). "Frontiers: ChatGPT Referrals to E-Commerce Websites: How Do LLMs Compare Against Traditional Channels?" *Marketing Science* 45.4, pp. 699–715.

Khosravi, Mehrzad and Hema Yoganarasimhan (2026). "Impact of AI Search Summaries on Website Traffic: Evidence from Google AI Overviews and Wikipedia". *Available at SSRN 6164926*.

Lee, Kwok Hao and Leon Musolff (2025). "Two-Sided Markets Shaped by Platform-Guided Search". *Working Paper*.

Lind, Jo Thori and Halvor Mehlum (2010). "With or Without U? The Appropriate Test for a U-Shaped Relationship". *Oxford Bulletin of Economics and Statistics* 72.1, pp. 109–118. DOI: 10.1111/j.1468-0084.2009.00569.x.

Liu, Xu, Zhe Wang, and Dengpan Liu (2025). "Generative Engine Optimization and Sponsored Search Bidding: Strategic Implications of AI Overviews for the Search Ecosystem". Working paper. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5365040.

Long, Fei and Wilfred Amaldoss (2024). "Self-preferencing in e-commerce marketplaces: The role of sponsored advertising and private labels". *Marketing Science* 43.5, pp. 925–952.

Reid, Liz (Aug. 6, 2025). *AI in Search is driving more queries and higher quality clicks*. Google. URL: https://blog.google/products-and-platforms/products/search/ai-search-driving-more-queries-higher-quality-clicks/ (visited on 09/29/2026).

Simon, Felix, Rasmus Kleis Nielsen, and Richard Fletcher (2025). *Generative AI and News Report 2025: How People Think about AI's Role in Journalism and Society*. Tech. rep. Reuters Institute for the Study of Journalism, University of Oxford. DOI: 10.60625/risj-5bjv-yt69.

Singh, Madhavi (2025). *The Antitrust Case Against AI Overviews*. Harvard Journal of Law and Technology Digest. October 6, 2025. URL: https://jolt.law.harvard.edu/digest/the-antitrust-case-against-ai-overviews.

Spatharioti, Sofia Eleni, David Rothschild, Daniel G. Goldstein, and Jake M. Hofman (2025). "Effects of LLM-based Search on Decision Making: Speed, Accuracy, and Overreliance". *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, pp. 1–15.

Stein, Robby (Jan. 2026). *AI Mode in Google Search and AI Overviews get Gemini upgrades*. Google blog, 27 January 2026. URL: https://blog.google/products-and-platforms/products/search/ai-mode-ai-overviews-updates/.

Taylor, Greg (2013). "Search Quality and Revenue Cannibalization by Competing Search Engines". *Journal of Economics & Management Strategy* 22.3, pp. 445–467. DOI: 10.1111/jems.12027.

Telang, Rahul, Uday Rajan, and Tridas Mukhopadhyay (2004). "The market structure for Internet search engines". *Journal of Management Information Systems* 21.2, pp. 137–160.

White, Alexander (2013). "Search Engines: Left Side Quality Versus Right Side Profits". *International Journal of Industrial Organization* 31.6, pp. 690–701. DOI: 10.1016/j.ijindorg.2013.04.003.

Xu, Haofei, Umar Iqbal, and Jacob M. Montgomery (2026). "Measuring Google AI Overviews: Activation, Source Quality, Claim Fidelity, and Publisher Impact". arXiv preprint 2605.14021. DOI: 10.48550/arXiv.2605.14021.

Xu, Lizhen, Jianqing Chen, and Andrew Whinston (2012). "Effects of the Presence of Organic Listing in Search Advertising". *Information Systems Research* 23.4, pp. 1284–1302.

Yang, Sha and Anindya Ghose (2010). "Analyzing the Relationship Between Organic and Sponsored Search Advertising: Positive, Negative, or Zero Interdependence?" *Marketing Science* 29.4, pp. 602–623.

Zhang, Luyang, Cathy Jiao, Beibei Li, and Chenyan Xiong (2026). "An Economic Framework for Generative Engines: Advertising or Subscription?" arXiv preprint 2603.29071. DOI: 10.48550/arXiv.2603.29071.

Zhu, Yuting, Shuang Zheng, Xin Ye, and Liang Shen (2025). "Generative Search: Evidence from a Large-Scale Field Experiment". *Available at SSRN 5297194*.

Zou, Tianxin and Bo Zhou (2025). "Self-preferencing and search neutrality in online retail platforms". *Management Science* 71.5, pp. 4087–4107.

Online Appendices

A Formal results and proofs

A.1 Baseline equilibrium and deployment

Proof of Proposition 1. Equation (1) gives $P_h V_h = \lambda_h[m_h a_h + \delta(1-\rho_h) f_h V_0]$. Substituting into the definition of V_0 yields the stated expression for V_0. Since $0 < \delta f_h \lambda_h < P_h$ and $0 \le \rho_h \le 1$, the aggregate continuation coefficient is nonnegative and satisfies $\delta \sum_h (1-\rho_h) f_h \lambda_h < \sum_h P_h = 1$. The aggregate value mapping is therefore an affine contraction, so the equilibrium is unique and the denominator is positive. The numerator is nonnegative, so $V_0 \ge 0$. Equation (1) then gives $V_h \ge 0$. □

Proof of Proposition 2. At common continuation values, the organic action yields $m_h a_h + \delta[(1 - f_h)V_h + (1-\rho_h) f_h V_0]$, whereas the AI action yields $-d + \delta[(1 - f_h^{AI})V_h + (1-\rho_h) f_h^{AI} V_0]$. Their difference is $\Delta_h - d$, with Δ_h defined in Equation (2).

Let $\lambda_k^{AI} = P_k / [1 - \delta(1 - f_k^{AI})]$ for every intent k. For a proposed new-intent value v, define the organic-search mapping $F(v) = \sum_k \lambda_k [m_k a_k + \delta(1-\rho_k) f_k v]$. Deploying repeatedly for intent h replaces its term with $\lambda_h^{AI}[-d + \delta(1-\rho_h) f_h^{AI} v]$. Let $F_h^{AI}(v)$ denote the resulting mapping. The inequalities $0 < \delta f_k \lambda_k < P_k$ and $0 < \delta f_h^{AI} \lambda_h^{AI} < P_h$, together with $0 \le \rho_k \le 1$, imply that both mappings are affine contractions with nonnegative slopes below $\sum_k P_k = 1$.

Let $V_0^{AI,h}$ be the fixed point of F_h^{AI}. Primes on these affine mappings denote their slopes with respect to v. By Equation (1), $P_h V_h = \lambda_h[m_h a_h + \delta(1-\rho_h) f_h V_0]$. Substitution gives $F_h^{AI}(V_0) - F(V_0) = \lambda_h^{AI}(\Delta_h - d)$. Since $F(V_0) = V_0$ and the new mapping is affine,

$$V_0^{AI,h} - V_0 = \frac{\lambda_h^{AI}(\Delta_h - d)}{1 - (F_h^{AI})'}. \tag{4}$$

The denominator and λ_h^{AI} are positive. Thus repeated deployment raises the equilibrium value exactly when $\Delta_h > d$, and the platform deploys at equality.

With expected AIO advertising revenue $m_h b_h$ per period, where $b_h \ge 0$, the same calculation replaces $\Delta_h - d$ in Equation (4) by $\Delta_h + m_h b_h - d$. The coefficient is positive and independent of b_h, proving the weak increase in deployment for $m_h = 1$.

For any fixed stationary deployment policy on the other intents, replace their terms in both mappings by the corresponding organic or AI terms. The mappings remain affine contractions, and the same comparison holds at the continuation values of the policy without deployment on h.

If the other intents are optimized under each choice for h, the Bellman equation for each $k \neq h$ instead gives the contribution

$$\max\left\{\lambda_k[m_k a_k + \delta(1-\rho_k) f_k v],\ \lambda_k^{AI}[-d + \delta(1-\rho_k) f_k^{AI} v]\right\}.$$

Both terms have slopes in $[0, \delta P_k]$, so the resulting aggregate mappings are nondecreasing contractions with Lipschitz constant at most δ. At the fixed point without deployment on h, the difference between the deployment and no-deployment mappings remains $\lambda_h^{AI}(\Delta_h - d)$. Each mapping minus its argument is strictly decreasing and vanishes at its unique fixed point, so this difference has the sign of the change in equilibrium value. Thus the threshold also holds when other deployment choices are reoptimized, including indifference at $\Delta_h = d$. □

A.2 Intent dimensions

Corollary 1 (Exploration dimension)**.** *Deployment is weakly increasing in f_h^{AI} and is zero at $f_h^{AI} = f_h$.*

Proof of Corollary 1. At a common baseline value V_0, the proof of Proposition 2 makes deployment equivalent to $(\lambda_h^{AI}/P_h)[-d + \delta(1-\rho_h) f_h^{AI} V_0] \geq V_h$. Holding monetization status, organic fulfillment, and the exit probability fixed leaves V_h unchanged. The left side is strictly increasing in f_h^{AI}, since its derivative is $\delta(\lambda_h^{AI}/P_h)^2[(1-\rho_h)(1-\delta)V_0 + d] > 0$. Deployment is therefore weakly increasing in AI fulfillment. At the targeted value $f_h^{AI} = f_h$, Equation (2) gives $\Delta_h = -m_h a_h < d$, ruling out targeted deployment.

Since $\Delta_h = -m_h a_h < d$ at $f_h^{AI} = f_h$, continuity preserves targeted non-deployment for sufficiently small positive AI fulfillment gains. Other intents' gains are unchanged because their comparisons leave the targeted intent organic.

□

Corollary 2 (Monetization dimension)**.** *Changing m_h from zero to one strictly reduces Δ_h and*

weakly reduces deployment.

Proof of Corollary 2. The organic and AI mapping slopes do not depend on m_h. Changing m_h from zero to one raises V_0 by $a_h\lambda_h/(1-F')$, while $V_0^{AI,h}$ is unchanged because that policy collects no revenue from intent h. Applying Equation (4) in the two organic environments gives

$$\Delta_h|_{m_h=0} - \Delta_h|_{m_h=1} = \frac{\lambda_h}{\lambda_h^{AI}} a_h \frac{1-(F_h^{AI})'}{1-F'} > 0. \tag{5}$$

Both slope denominators are positive by the contraction argument. Equation (5) therefore proves the ordering with all continuation-value effects included. □

For each monetizable intent, measure the deterrent from monetization by comparing its deployment gain with the counterfactual in which only its monetization indicator is set to zero. Recompute the organic equilibrium in each counterfactual. Consider an exploratory monetizable intent h and a targeted monetizable intent j with $P_h = P_j$, $f_h = f_j$, and $\rho_h = \rho_j$.

Proposition 3 (Monetization and exploration)**.** *Monetization weakly reduces deployment on h and leaves deployment on j at zero. If $a_h \geq a_j$, the reduction in the deployment gain is strictly larger for h than for j.*

Proof of Proposition 3. Equation (5) gives the reduction for each intent. The organic weights and F' do not depend on AI fulfillment or payoffs. For a fixed intent, the factor multiplying its payoff is positive and increases with AI fulfillment because

$$\frac{\partial}{\partial f_h^{AI}}\left[\frac{\lambda_h}{\lambda_h^{AI}}\left(1-(F_h^{AI})'\right)\right] = \frac{\delta\lambda_h}{P_h}\left[1-(1-\rho_h)P_h - \delta\sum_{k\neq h}(1-\rho_k)f_k\lambda_k\right] > 0.$$

The bracket is positive because $\delta\sum_{k\neq h}(1-\rho_k)f_k\lambda_k < 1-P_h$ and $(1-\rho_h)P_h \leq P_h$. At $f_h^{AI} = f_h$, the AI and organic continuation slopes coincide, so the factor in Equation (5) equals one. Raising AI fulfillment strictly increases this factor, while the targeted factor remains one. Thus $a_h \geq a_j > 0$ establishes the larger exploratory reduction in gain. Corollary 2 gives weakly lower exploratory deployment after monetization. For the targeted intent, $\Delta_j = -m_j a_j < d$ in both monetization states, so its deployment is unchanged at zero. This proves the deployment interaction without requiring a distribution of queries around the threshold. □

Corollary 3 (Deployment levels among monetizable intents)**.** *Suppose $f_C = f_T$, $\rho_C = \rho_T$, and $a_C > a_T$. Then $\Delta_T > \Delta_C$, and neither monetizable intent receives an AIO for any $d > 0$.*

Proof of Corollary 3. Proposition 1 gives $V_0 > 0$ because both monetizable intents have positive payoffs and arrival probabilities. For $k \in \{I, N\}$, Equation (1) gives $V_k = \delta(1-\rho_k)f_k(\lambda_k/P_k)V_0 < V_0$, since $\delta(1-\rho_k)f_k < 1 - \delta(1-f_k)$. Matching organic fulfillment and exit probabilities gives $V_C - V_T = (\lambda_C/P_C)(a_C - a_T) > 0$. If $V_C \leq V_0$, then V_T, V_I, and V_N would all be strictly below V_0, contradicting $V_0 = \sum_k P_k V_k$ with positive weights summing to one. Hence $V_C > V_0$.

Since $f_T^{AI} = f_T$ and $f_C^{AI} > f_C$, subtracting the gains in Equation (2) yields

$$\Delta_T - \Delta_C = (a_C - a_T) + \delta(f_C^{AI} - f_C)\big[V_C - (1-\rho_C)V_0\big] > 0.$$

The organic value equation also implies

$$\Delta_C = \delta f_C^{AI}\big[(1-\rho_C)V_0 - V_C\big] - (1-\delta)V_C < 0.$$

Thus $\Delta_C < d$ for every $d > 0$, so Proposition 2 rules out deployment for C. For T, the gain is $\Delta_T = -a_T < d$, so targeted deployment is also zero. The strict ordering of deployment gains therefore produces no difference in deployment levels under these assumptions. The comparison uses the common organic equilibrium and requires no equality of arrival probabilities. □

A.3 Organic fulfillment

Vary the organic fulfillment probability of one intent while holding its expected per-period payoff, exit probability, arrival probability, and all other primitives fixed. For an exploratory intent, hold f_h^{AI} fixed and vary f_h within $(0, f_h^{AI})$. For a targeted intent, maintain $f_h^{AI} = f_h$. In both comparisons, recompute all organic continuation values.

Proposition 4 (Organic fulfillment)**.** *For an exploratory intent, $\partial\Delta_h/\partial f_h < 0$ if and only if*

$$m_h a_h < \frac{(1-\rho_h)(1-\delta)\sum_{k\neq h} m_k a_k \lambda_k}{1-(1-\rho_h)P_h - \delta\sum_{k\neq h}(1-\rho_k)f_k\lambda_k}.$$

The derivative is zero at equality and positive when the inequality reverses. Exploratory deploy-

ment is weakly decreasing in f_h. For targeted intents, $\Delta_h = -m_h a_h$ and deployment is zero.

Proof of Proposition 4. Differentiating the organic fixed point, allowing all continuation values to adjust, gives

$$\frac{\partial V_0}{\partial f_h} = \frac{\delta \lambda_h^2}{P_h(1-F')}\big[(1-\rho_h)(1-\delta)V_0 - m_h a_h\big].$$

For an exploratory intent with AI fulfillment fixed, the mapping F_h^{AI} and its fixed point are independent of f_h, since this policy never uses organic search for intent h. Equation (4) gives $\partial \Delta_h / \partial f_h = -[1-(F_h^{AI})'](\partial V_0/\partial f_h)/\lambda_h^{AI}$, so the gain has the opposite derivative sign. Proposition 1 and $[1-\delta(1-f_h)]\lambda_h = P_h$ give

$$\begin{aligned}
&(1-F')\big[(1-\rho_h)(1-\delta)V_0 - m_h a_h\big] \\
&\quad = (1-\rho_h)(1-\delta)\sum_{k\neq h} m_k a_k \lambda_k \\
&\qquad - m_h a_h\Big[1-(1-\rho_h)P_h - \delta\sum_{k\neq h}(1-\rho_k) f_k \lambda_k\Big].
\end{aligned}$$

Here $1-F' > 0$, and $\delta\sum_{k\neq h}(1-\rho_k)f_k\lambda_k < 1-P_h$ makes the bracket multiplying $m_h a_h$ strictly positive. The displayed identity therefore gives the stated sign condition. Neither sum contains the focal f_h, so the sign holds throughout its comparison interval. Equality and reversal give the remaining gain comparisons.

The organic value equation and Equation (2) also give

$$\Delta_h = \frac{\delta(f_h^{AI} - f_h)\lambda_h}{P_h}\big[(1-\rho_h)(1-\delta)V_0 - m_h a_h\big] - m_h a_h.$$

If the displayed revenue condition fails, the bracket is nonpositive and $\Delta_h \leq -m_h a_h < d$ throughout the comparison interval. If the condition holds, the gain strictly decreases with f_h, so deployment can only be withdrawn as organic fulfillment rises. This proves weakly decreasing exploratory deployment without a revenue restriction. For a targeted intent, $f_h^{AI} = f_h$ gives $\Delta_h = -m_h a_h < d$ for every f_h, proving both invariance of its gain and zero deployment. □

A.4 Revenue and continuation value

Vary one monetizable intent's expected per-period payoff a_h, holding all other payoffs and primitives fixed. Recompute the organic equilibrium after each change.

Proposition 5 (Own revenue and continuation value)**.** *Increasing a_h for a monetizable intent h strictly reduces Δ_h and weakly reduces deployment on that intent. It weakly increases Δ_k and deployment on every other intent k. The latter gain increases strictly exactly when k is exploratory and $\rho_k < 1$. Targeted gains are unchanged. Changing a_k for a non-monetizable intent has no effect.*

Proof of Proposition 5. Proposition 1 gives $\partial V_0/\partial a_h = \lambda_h/(1-F') > 0$. The AI policy for h receives no payoff from that intent, so $V_0^{AI,h}$ is independent of a_h. The coefficient in Equation (4) is also independent of payoffs. Hence $\partial\Delta_h/\partial a_h = -(\lambda_h/\lambda_h^{AI})[1-(F_h^{AI})']/(1-F') < 0$. Equivalently, Equation (5) is linear in a_h, so the own-revenue effect strictly dominates the increase in continuation value.

For $k \neq h$, the focal payoff $m_k a_k$ is unchanged. Substituting its organic value into Equation (2) and differentiating gives

$$\frac{\partial\Delta_k}{\partial a_h} = \delta(1-\rho_k)(1-\delta)(f_k^{AI}-f_k)\frac{\lambda_k\lambda_h}{P_k(1-F')} \geq 0.$$

All factors other than $1-\rho_k$ and $f_k^{AI}-f_k$ are strictly positive. The derivative is therefore positive exactly for exploratory intents with $\rho_k < 1$, and zero for targeted intents or $\rho_k = 1$. Multiple increases in payoffs from monetizable intents also weakly increase deployment gains on other exploratory intents. Finally, if $m_k = 0$, its own a_k enters neither the organic values nor any deployment gain. □

A.5 Production and source quality

Fix an exploratory intent h. Index underlying source quality by $f_h \in [0,1]$ and let it determine AI fulfillment through $f_h^{AI} = \Psi_h(f_h)$. The function $\Psi_h : [0,1] \to [0,1]$ is continuous and strictly concave, with $\Psi_h(0) = 0$ and $\Psi_h(1) = 1$. The endpoint values extend the baseline by continuity. Hold all other primitives fixed. Let $V_0(f_h)$ and $V_h(f_h)$ be the corresponding organic equilibrium values, and evaluate $\Delta_h(f_h)$ using Equation (2) at those values.

Proposition 6 (Source quality)**.** *Deployment is absent if* $\max_{f_h \in [0,1]} \Delta_h(f_h) < d$*. Otherwise, there are unique cutoffs* $0 < \underline{f}_h \leq \bar{f}_h < 1$ *such that deployment occurs exactly for* $f_h \in [\underline{f}_h, \bar{f}_h]$*, with* $\underline{f}_h = \bar{f}_h$ *if and only if* $\max_{f_h \in [0,1]} \Delta_h(f_h) = d$*.*

Proof of Proposition 6. The weights $\lambda_h = P_h/[1 - \delta(1 - f_h)]$ are continuous on $[0, 1]$. For every intent, $0 \leq \delta(1 - \rho_h) f_h \lambda_h \leq \delta P_h$, so the denominator in Proposition 1 is at least $1 - \delta > 0$. Hence $V_0(f_h)$ and $V_h(f_h)$ are continuous, and continuity of Ψ_h makes $\Delta_h(f_h)$ continuous. This accounts for the change in all organic continuation values as source quality varies.

At $f_h = 0$ and $f_h = 1$, AI and organic fulfillment coincide, giving $\Delta_h(0) = \Delta_h(1) = -m_h a_h < d$. No additional endpoint condition is needed. Using the identity in the proof of Proposition 4, the organic value equations give

$$\begin{aligned}
&[1 - \delta(1 - f_h)](1 - F')[\Delta_h(f_h) + m_h a_h] \\
&\quad = \delta[\Psi_h(f_h) - f_h] \Bigg[(1 - \rho_h)(1 - \delta) \sum_{k \neq h} m_k a_k \lambda_k \\
&\qquad - m_h a_h \Big(1 - (1 - \rho_h) P_h - \delta \sum_{k \neq h} (1 - \rho_k) f_k \lambda_k \Big) \Bigg].
\end{aligned}$$

The multiplier on the left is positive and equals $[1-\delta(1-f_h)][1-\delta \sum_{k \neq h}(1-\rho_k) f_k \lambda_k] - \delta(1-\rho_h) f_h P_h$, which is affine in f_h. If the coefficient of $\Psi_h(f_h) - f_h$ is nonpositive, then $\Delta_h(f_h) \leq -m_h a_h < d$ everywhere. Otherwise, multiplying $\Delta_h(f_h) - d$ by this multiplier gives a strictly concave function that is negative at both endpoints. Its nonnegative set is therefore empty, a singleton, or a closed interval contained in $(0, 1)$. When the maximum gain equals d, strict concavity permits only a singleton. When the maximum gain exceeds d, this set is a nondegenerate interval with unique boundary points. Proposition 2 gives deployment exactly on this set, including its boundaries. A targeted intent has $\Delta_h(f_h) = -m_h a_h$ throughout and never receives an AIO. □

A.6 Intent ambiguity

When a new intent arrives, the platform observes a signal σ from a finite set and forms the Bayesian posterior $\mu_h(\sigma) = \Pr(h \mid \sigma)$. A policy conditions on the initial signal and is maintained until fulfillment. Compare the policy without AIOs with the stationary policy that deploys whenever one fixed signal σ occurs and otherwise uses organic search. The signal decision recurs on future new-

intent draws. Evaluate gains at the common organic continuation values. Consider only signals with positive probability.

Fix a focal classification h and the conditional distribution $(\nu_k)_{k\neq h}$ over alternative intents, with $\nu_k \geq 0$ and $\sum_{k\neq h} \nu_k = 1$. Classification confidence is the posterior probability $\mu_h(\sigma)$. Compare signals that differ only in this confidence, holding all intent primitives and conditional weights fixed. The organic continuation values are common to these comparisons.

Proposition 7 (Intent confidence)**.** *Greater confidence in classification h weakly increases deployment when $\Delta_h \geq d$ and weakly decreases it when $\Delta_h < d$. For queries classified as targeted, deployment weakly decreases with confidence and is zero when the intent is known with certainty.*

Proof of Proposition 7. Write $w_\tau = \Pr(\tau)$ for each possible signal. Bayes' rule gives $\sum_\tau w_\tau \mu_k(\tau) = P_k$. For a proposed new-intent value v, the organic intent value is $U_k(v) = (\lambda_k/P_k)[m_k a_k + \delta(1-\rho_k) f_k v]$, and the organic aggregate mapping is $F(v) = \sum_k P_k U_k(v)$. Deployment at signal σ replaces $U_k(v)$ by $(\lambda_k^{AI}/P_k)[-d + \delta(1-\rho_k) f_k^{AI} v]$ with weight $w_\sigma \mu_k(\sigma)$. Denote the resulting aggregate mapping by $F_\sigma(v)$.

The organic and AI intent-value slopes are $\delta(1-\rho_k) f_k \lambda_k / P_k$ and $\delta(1-\rho_k) f_k^{AI} \lambda_k^{AI}/P_k$, respectively. Both lie in $[0, \delta]$, so the aggregate mappings are affine contractions with nonnegative slopes at most $\delta < 1$. At $v = V_0$, $U_k(V_0) = V_k$. The proof of Proposition 2 gives the conditional intent-value difference $(\lambda_k^{AI}/P_k)(\Delta_k - d)$. The weight λ_k^{AI}/P_k counts discounted AI attempts until fulfillment. Let V_0^σ be the fixed point of F_σ. Affinity gives

$$V_0^\sigma - V_0 = \frac{w_\sigma \sum_k \mu_k(\sigma)(\lambda_k^{AI}/P_k)(\Delta_k - d)}{1 - F_\sigma'(V_0)}. \tag{6}$$

The denominator and w_σ are positive. Thus deployment is determined by the sign of the duration-weighted sum in Equation (6), with all subsequent intent arrivals included.

The sign of the posterior policy gain is the sign of the affine expression

$$\mu_h(\sigma)\frac{\lambda_h^{AI}}{P_h}(\Delta_h - d) + (1 - \mu_h(\sigma)) \sum_{k\neq h} \nu_k \frac{\lambda_k^{AI}}{P_k}(\Delta_k - d).$$

At certainty, the displayed expression has the sign of $\Delta_h - d$. If $\Delta_h \geq d$, deployment at any confidence implies deployment at every higher confidence, since the intervening values are convex

combinations of two nonnegative values. If $\Delta_h < d$, non-deployment at any confidence implies non-deployment at every higher confidence, since the expression is negative both there and at certainty. This proves both monotonicity results, with deployment at equality. For a query classified as targeted, $\Delta_h = -m_h a_h < d$, so deployment weakly decreases with confidence and is zero when the intent is known with certainty. A positive posterior gain requires some exploratory intent with $\Delta_k > d$. □

A.7 Competition

We study two competition extensions. Both extensions follow from the baseline setting with two modifications. First, we relax the restriction on targeted intents and allow $f_h^{AI} \geq f_h$ for all h. Second, we allow the platform to commit to an AIO deployment probability $z_h \in [0, 1]$ for each intent h before users choose a provider. This probability is independent of time and histories. AI fulfillment, organic fulfillment, payoffs, and exit probabilities are fixed as competition changes.

Competition without user commitment. Without user commitment, the user chooses a platform anew before every search attempt after observing its stationary AIO policy. Users choose the platform with the higher ex ante fulfillment probability, with ties favoring the platform. Following fulfillment on either platform, the user exits with probability ρ_h and otherwise draws a new intent next period. Following failure, the same intent persists and the user chooses again. Later choices depend on the current intent and the available policies, not on the platform used previously. Rival service gives the platform zero current payoff but preserves the opportunity to serve subsequent intents.

Fix a focal intent h with $f_h^{AI} > f_h$. All other intents continue to be served organically by the platform. The rival fulfills h with probability $q_h \in (0, 1]$. Initially, rival fulfillment is no greater than f_h, so the platform's organic values are those of Proposition 1. Evaluate Δ_h using Equation (2) with the permitted AI fulfillment probability, and assume $\Delta_h < d$. At profit indifference, choose AIO deployment.

Proposition 8 (Competition without user commitment)**.**

(i) For $f_h < q_h \leq f_h + (f_h^{AI} - f_h) m_h a_h / (m_h a_h + d)$, an optimal policy is

$$z_h^*(q_h) = \frac{q_h - f_h}{f_h^{AI} - f_h}.$$

Otherwise, $z_h^*(q_h) = 0$. *For* $m_h = 0$, *deployment is zero for every* q_h.

(ii) If a targeted intent j *and an exploratory intent* h *satisfy* $\Delta_j, \Delta_h < d$ *and* $0 < f_j^{AI} - f_j < f_h^{AI} - f_h$, *with* q_j *and* q_h *strictly between their respective bounds in part (i) and all other intents organic in each case, then* $dz_j^*(q_j)/dq_j > dz_h^*(q_h)/dq_h > 0$.

Proof of Proposition 8. Write $\mathbf{z} = (z_k)_{k \in \mathbb{H}} \in [0,1]^{\mathbb{H}}$ for the deployment policy. For an attempt served by the platform, fulfillment and the continuation weight are

$$f_k(z_k) = f_k + z_k(f_k^{AI} - f_k), \qquad \lambda_k(z_k) = \frac{P_k}{1 - \delta(1 - f_k(z_k))},$$

and the current expected payoff is $(1 - z_k)m_k a_k - z_k d$. Let $V_0(\mathbf{z})$ denote the platform's value of a new intent conditional on serving every attempt. The baseline value equations give

$$V_0(\mathbf{z}) = \frac{\sum_k \lambda_k(z_k)[(1 - z_k)m_k a_k - z_k d]}{1 - \delta \sum_k (1 - \rho_k) f_k(z_k) \lambda_k(z_k)}.$$

Since $0 \leq \delta(1 - \rho_k) f_k(z_k) \lambda_k(z_k) \leq \delta P_k$, the denominator is at least $1 - \delta > 0$. These values are therefore well defined and smooth. Use one-sided derivatives at boundary policies. Organic values satisfy $V_0(\mathbf{0}) = V_0$ and $\lambda_k(0) = \lambda_k$. Write $z\mathbf{e}_h$ for the policy with AIO probability z on intent h and zero on all other intents.

Multiplying the numerator and denominator of $V_0(z\mathbf{e}_h)$ by $1 - \delta(1 - f_h(z))$ makes both affine in z. The denominator is positive, so the derivative has constant sign. At zero, the organic value equations give

$$\left.\frac{dV_0(z\mathbf{e}_h)}{dz}\right|_{z=0} = \frac{\lambda_h(\Delta_h - d)}{1 - \delta \sum_k (1 - \rho_k) f_k \lambda_k} < 0.$$

Hence $V_0(z\mathbf{e}_h)$ strictly decreases on $[0,1]$.

If $q_h \leq f_h$, every policy retains the query and $z = 0$ is optimal. Otherwise, retention requires

$$z \geq \frac{q_h - f_h}{f_h^{AI} - f_h}.$$

If this bound exceeds one, retention is infeasible. If feasible, the best retaining policy sets z equal to the bound.

Let $\widehat{V}_0^h(q_h)$ denote the platform's value when the rival serves h. At the matching probability, the two policies induce identical fulfillment and continuation probabilities, so

$$V_0(z\mathbf{e}_h) - \widehat{V}_0^h(q_h) = \frac{\lambda_h(z)[(1-z)m_h a_h - zd]}{1 - \delta\sum_{k\neq h}(1-\rho_k)f_k\lambda_k - \delta(1-\rho_h)f_h(z)\lambda_h(z)}.$$

The platform therefore matches exactly when $z \leq m_h a_h/(m_h a_h + d)$, including equality. Combining this condition with the required matching probability proves (i). Within the positive-deployment interval,

$$\frac{dz_h^*(q_h)}{dq_h} = \frac{1}{f_h^{AI} - f_h} > 0.$$

The ordering of fulfillment gains gives (ii). □

Competition with user commitment. Users instead commit to one provider for all future intents after observing $\mathbf{z}$. A user receives utility one upon fulfillment, zero otherwise, discounts by δ, and receives zero continuation utility upon exit.

Let $\Phi_0(\mathbf{z})$ denote her utility from committing to the platform and write $\Phi_0 = \Phi_0(\mathbf{0})$. If the rival offers utility Φ_{LLM}, the user chooses the platform exactly when $\Phi_0(\mathbf{z}) \geq \Phi_{\text{LLM}}$. The platform earns $V_0(\mathbf{z})$ if chosen and zero otherwise.

Proposition 9 (Commitment and optimal deployment)**.** *(i) For $f_k^{AI} > f_k$, $\partial_{z_k}\Phi_0(\mathbf{z}) > 0$ at every policy. For $f_k^{AI} = f_k$, $\partial_{z_k}\Phi_0(\mathbf{z}) = 0$, $\partial_{z_k}V_0(\mathbf{z}) < 0$, and $z_k = 0$ in every profit-maximizing policy attracting users.*

(ii) Suppose $f_h^{AI} > f_h$, $\Delta_h < d$, $V_0 > 0$, and all other intents remain organic. If $\Phi_{\text{LLM}} \notin (\Phi_0, \Phi_0(\mathbf{e}_h)]$, zero deployment is optimal. Otherwise, z_h is uniquely defined by

$$\Phi_0(z_h\mathbf{e}_h) = \Phi_{\text{LLM}}, \qquad z_h \in (0,1].$$

Deploying with probability z_h is optimal if $V_0(z_h\mathbf{e}_h) \geq 0$. Otherwise, zero deployment is optimal. While z_h is optimal, $dz_h/d\Phi_{\text{LLM}} > 0$.

(iii) Under the assumptions of (ii), for $m_h = 0$, any positive optimal deployment with positive profit satisfies $0 < V_0(z_h \mathbf{e}_h) < V_0$, whereas $z_h^*(q_h) = 0$ without commitment.

(iv) For a targeted intent j and an exploratory intent h, suppose $V_0 > 0$, $\rho_j = \rho_h$, $f_k^{AI} > f_k$, and $\Delta_k < d$ for $k \in \{h, j\}$.

With no AIO deployment, targeted deployment has a strictly lower marginal profit cost per unit of additional user utility if and only if

$$\frac{[1 - \delta(1 - f_j^{AI})](m_j a_j + d)}{f_j^{AI} - f_j} < \frac{[1 - \delta(1 - f_h^{AI})](m_h a_h + d)}{f_h^{AI} - f_h}.$$

Proof of Proposition 9. The user's value equations give

$$\Phi_0(\mathbf{z}) = \frac{\sum_k f_k(z_k)\lambda_k(z_k)}{1 - \delta \sum_k (1 - \rho_k) f_k(z_k)\lambda_k(z_k)}.$$

Differentiating yields

$$\frac{\partial \Phi_0(\mathbf{z})}{\partial z_k} = \frac{\lambda_k(z_k)(f_k^{AI} - f_k)(1 - \delta)[1 + \delta(1 - \rho_k)\Phi_0(\mathbf{z})]}{\big[1 - \delta \sum_\ell (1 - \rho_\ell) f_\ell(z_\ell)\lambda_\ell(z_\ell)\big][1 - \delta(1 - f_k(z_k))]}.$$

All factors apart from $f_k^{AI} - f_k$ are strictly positive. If $f_j^{AI} = f_j$, fulfillment and continuation weights are independent of z_j, and

$$\frac{\partial V_0(\mathbf{z})}{\partial z_j} = -\frac{\lambda_j(m_j a_j + d)}{1 - \delta \sum_k (1 - \rho_k) f_k(z_k)\lambda_k(z_k)} < 0.$$

Reducing z_j therefore raises profit without changing user utility, proving (i).

Under the assumptions of (ii), the proof of Proposition 8 establishes that $V_0(z\mathbf{e}_h)$ strictly decreases in z, while the utility derivative above is strictly positive. Thus $z = 0$ is optimal when $\Phi_{\text{LLM}} \leq \Phi_0$. For $\Phi_0 < \Phi_{\text{LLM}} \leq \Phi_0(\mathbf{e}_h)$, continuity and strict monotonicity give a unique z_h satisfying the commitment constraint with equality. The retaining policies are exactly those with $z \geq z_h$, so z_h maximizes profit among them. It is optimal when its profit is nonnegative. Otherwise, or when $\Phi_{\text{LLM}} > \Phi_0(\mathbf{e}_h)$, zero deployment is optimal and the rival serves the user. While z_h is optimal,

$$\frac{dz_h}{d\Phi_{\text{LLM}}} = \frac{1}{\partial_{z_h} \Phi_0(z_h \mathbf{e}_h)} > 0.$$

For (iii), the strict decrease of $V_0(z\mathbf{e}_h)$ gives $0 < V_0(z_h\mathbf{e}_h) < V_0$ whenever optimal deployment is positive and profitable. Since $m_h = 0$, the focal intent generates no click revenue. The profit comes from other monetizable intents. Without user commitment, Proposition 8(i) gives zero deployment.

For (iv) and $k \in \{h, j\}$, let $\kappa_k = -\partial_{z_k} V_0(\mathbf{0})/\partial_{z_k}\Phi_0(\mathbf{0})$ be the marginal profit cost per unit of user utility. Both the profit cost and utility gain are positive under the stated assumptions. The value derivatives and Equation (2) give

$$\kappa_k = \frac{1}{(1-\delta)[1+\delta(1-\rho_k)\Phi_0]} \times \left[\frac{[1-\delta(1-f_k^{AI})](m_k a_k + d)}{f_k^{AI} - f_k} - \delta d - \delta(1-\rho_k)(1-\delta)V_0\right].$$

When $\rho_j = \rho_h$, the positive prefactor and the terms outside the fraction in brackets are common to both intents. Hence $\kappa_j < \kappa_h$ is equivalent to the stated inequality.

□

B Search Query Generation and Classification

Testing our theoretical predictions requires a query sample that spans the intent space, varies along dimensions that could affect AIO deployment (e.g., YMYL status, local relevance, question format, length), and reflects queries that real users actually search. Google does not publish query logs, and publicly available datasets such as those derived from search engine competitions would likely overrepresent popular head queries and underrepresent the long tail of specific, intent-rich queries that are central to our analysis. We therefore employed a structured generation approach that combines LLM-based query production with external validation against a commercial search analytics API, and then classified the resulting queries on the two intent dimensions with a separate, independently validated instrument.

A central concern with LLM-generated research data is that the generated items may reflect the model's training distribution rather than the target construct (Carlson and Burbano, 2025; Brand et al., 2025). Our design separates the two tasks the LLM performs. In generation, the model produces candidate query text only. Whether a query is actually searched, and at what advertising value, is determined externally by the API from Google's own signals. In classification, a different model assigns the intent labels used in the analysis, and is validated against author-coded labels before deployment. Neither step depends on the other's judgment. This appendix describes the generation protocol, the filtering criteria, and the classification instrument in detail.

B.1 Generation Protocol

We used Claude Haiku 4.5 (Anthropic) accessed through the Anthropic API.[19] Each API call requested 180 candidate queries in structured JSON format, with the model instructed to return each query's text, whether it was phrased as a question, and its approximate length category (short: 2–4 words, long: 5–10 words).

To ensure the generated queries resemble actual search behavior rather than generic LLM completions, we seeded each prompt with *anchor examples*: high-frequency keywords extracted from the DataForSEO search analytics API for the United States and Ireland. These anchors

[19] For reproducibility, we document the model identifier (`claude-haiku-4-5-20251001`), temperature (0.9), maximum output tokens (8,000), and system prompt ("You are a data generator that strictly adheres to constraints and outputs valid JSON."). Prompt templates, anchor keyword lists, and the complete generation and validation code are available upon request.

serve a similar function to few-shot examples in classification tasks (Carlson and Burbano, 2025), grounding the model's output in the vocabulary and phrasing conventions of real searches. For instance, anchor examples for commercial queries included brand names and retail chains (e.g., "old navy," "papa john's pizza"), while anchors for navigational queries included login and service pages (e.g., "facebook login").

Each generation call was parameterized along four dimensions that define the combinatorial design of our query sample:

1. **Search intent** (4 levels): informational, navigational, commercial, transactional. The prompt mapped each intent to Google's own query classification vocabulary, drawing on Google's Search Quality Rater Guidelines (Google, 2025d). For example, informational queries were described as "KNOW (User wants to know a fact, concept, or answer)," while transactional queries were described as "DO (User wants to buy, download, or complete an action NOW. Must use specific models/brands)."

2. **YMYL status** (2 levels): YMYL queries were constrained to health, financial, legal, and civic topics. Non-YMYL queries covered entertainment, hobbies, technology, and general knowledge.

3. **Local relevance** (2 levels): local queries were instructed to include geographic modifiers ("near me," "in my area," "nearby," or a specific city name), with approximately 75% using generic location terms and 25% using the city name. Non-local queries were explicitly prohibited from containing any location terms.

4. **Geographic market** (2 levels): United States (Los Angeles) and Ireland (Dublin), each with market-specific anchor examples and location parameters.

This yields $4 \times 2 \times 2 \times 2 = 32$ unique configuration cells. To generate sufficient candidate volume for post-validation filtering, we repeated each configuration 8 times, producing $32 \times 8 = 256$ batches of 180 queries each, for a target of approximately 46,000 raw candidate queries.[20]

Because the generation step uses the four intent types to structure the candidate pool, the prompt provided positive and negative examples for the categories hardest to separate. For instance,

[20]The actual yield was 33,082 candidates after accounting for occasional API failures and malformed responses.

it specified that "best italian restaurants in my area" is commercial (comparison and research), while "library in my area" is navigational. These examples govern generation only. The labels used in the analysis come from the classification instrument described in B.4.

B.2 External Validation

Each candidate query was submitted to the DataForSEO Keywords Data API, which returns Google's own signals:

- **Search volume**: the average monthly search volume over the preceding 12 months, indicating whether the query reflects actual user search behavior.
- **Cost per click**: the average CPC in USD, which we use as a proxy for the click payoff in our empirical specifications.

B.3 Filtering and Final Sample

We applied the following sequential filters to the raw candidate pool:

1. **Deduplication**: Exact-match deduplication on query text, retaining the first occurrence. Because the same configuration was repeated 8 times with high temperature, some queries appeared in multiple batches.
2. **Search volume**: We excluded queries with fewer than 10 average monthly searches, which ensures that each query in our sample corresponds to a real search phrase used by actual users, rather than a plausible but synthetic combination that no one actually searches.

B.4 Classification on the Two Intent Dimensions

The labels used in the analysis come from a separate instrument that classifies each query on the two dimensions of Table 1: whether the user is on a path toward a purchase, and whether the user has already selected a concrete destination, provider, or action. The instrument receives the query text and the country in which it was searched, judges the two dimensions independently, and returns a dominant label and a confidence score for each. The four intent types are derived from the two labels. No result, advertising, or deployment information enters the classification.

Instrument. The deployed configuration is GLM 5.2 (FP8) at temperature zero with a fixed seed. Each distinct query and country pair received one valid classification, and calls that failed were re-run. The prompt, output schema, few-shot examples, script, and input are recorded by SHA-256 hash in the output, together with the model identifier, and the full prompt text is available upon request.

Development. The prompt was developed against a stratified pilot of 100 queries drawn separately from the validation set described below. Each revision addressed a specific boundary case on which the pilot showed the definitions were not being applied as intended: whether commitment attaches to what is transacted or to where it is obtained, how to treat a search for a category of local provider as against a named one, and whether accessing a service already paid for is a monetizable search. Where rule text alone failed to transfer to queries outside the pilot, we added nine worked examples adjudicated by one author, each recorded with the case it illustrates.

Validation. We drew 200 queries from a 600-query random sample of the corpus that had not been used in the pilot, comprising 150 drawn at random and 50 drawn from the two smallest cells so that per-cell precision is estimable. One author reviewed the instrument's labels and corrected those judged incorrect, against agreement thresholds fixed before any label was inspected. The prompt at that stage agreed with the corrected labels on 95.5% of intent types and cleared every threshold. The review then informed further revisions of the prompt, in which 18 of the 200 queries were adopted as worked examples and 18 labels were re-adjudicated. Tables B1 and B2 therefore report agreement between the deployed instrument and the corrected labels, overall and by intent type, on the 169 queries whose text and label did not enter the instrument. Because the review informed the prompt and the instrument's labels were visible during it, these figures are an upper bound on accuracy, and we report the agreement rate rather than an inter-rater reliability statistic (Carlson and Burbano, 2025).

Six of the 169 queries are classified differently by the instrument and the reviewer. Five lie at the boundary of the exploration dimension: two searches for a category of nearby facility, where the reviewer judged the user to be gathering information and the instrument judged them to be selecting a destination, and three searches of the form "where to buy X", where the reviewer judged the purchase decision already made and the instrument judged the retailer still open. The sixth differs on the monetization dimension only.

Table B1: Agreement between the deployed classification and author-corrected labels.

	Agreement
Monetization dimension	99.4%
Exploration dimension	97.0%
Intent type	96.4%

Notes: $n = 169$. Excluded are the 29 queries adopted as worked examples in the prompt or whose label was re-adjudicated after the review, and two further queries that returned unparseable output. Because the instrument's labels were visible during review, agreement is an upper bound on accuracy.

Table B2: Classification precision and recall by intent type.

Intent type	n	**Precision**	**Recall**
Informational	28	100.0%	89.3%
Commercial	66	94.3%	100.0%
Navigational	38	95.0%	100.0%
Transactional	37	100.0%	91.9%

Notes: n is the number of queries assigned to each intent type by the author-corrected labels. Precision is the share of queries the instrument assigns to a type that the corrected labels also assign to it. Recall is the share of queries the corrected labels assign to a type that the instrument also assigns to it.

Table B3 reports the final distribution of queries across intent types, and Table B4 presents illustrative queries.

Table B3: Final query sample by intent type.

	Exploratory	**Targeted**
Non-monetizable	548	323
Monetizable	512	192

Notes: Cells correspond to informational, navigational, commercial, and transactional intents respectively. $N = 1{,}575$ queries in the analysis sample.

Table B4: Sample search queries by intent type.

Intent	Example query	YMYL	Local	Intent ambiguity	AIO rate
Informational	Best beaches near me	No	Yes	Yes	73%
	Best comedy tv shows	No	No	Yes	50%
	Capital gains tax rates	Yes	No	No	88%
	DNA structure	No	No	No	94%
	Fun things to do nearby	No	Yes	Yes	93%
	How to tie a tie	No	No	Yes	19%
	What are menopause symptoms	Yes	No	No	93%
	What is potential energy	No	No	No	94%
Commercial	Affordable homeowners insurance	Yes	No	Yes	100%
	Best coffee makers	No	No	Yes	75%
	Best dui attorney in my area	Yes	Yes	Yes	73%
	Best espresso machine for home use	No	No	Yes	88%
	Best video hosting platforms	No	No	Yes	94%
	Dental implant cost	Yes	No	Yes	100%
	What is the best credit card for rewards	Yes	No	Yes	100%
	What restaurants near me	No	Yes	No	13%
Navigational	Airbnb sign in	Yes	No	No	50%
	BBC weather	No	No	Yes	0%
	Grammarly official site	No	No	No	12%
	Indeed jobs	No	No	Yes	0%
	Library in my area	No	Yes	No	6%
	Netflix login	No	No	No	0%
	Paypal sign in	Yes	No	No	0%
	Trauma center nearby	Yes	Yes	No	81%
Transactional	Adidas shop nearby	No	Yes	No	12%
	Bose headphones buy	No	No	No	6%
	Buy instant pot duo	No	No	No	38%
	Pharmacy near me open late	Yes	Yes	No	31%
	Purchase vmware workstation	Yes	No	No	75%
	Where to buy davinci resolve studio	Yes	No	Yes	40%
	Where to buy iPhone 15 pro max	No	No	Yes	43%
	Whole foods market in my area	No	Yes	No	38%

Notes: Intent ambiguity indicates that the query plausibly carries more than one of the four user intents defined in the Search Quality Rater Guidelines (Google, 2025d). AIO rate is the share of SERP observations for the query, across both waves, in which an AI Overview was deployed. Queries are illustrative examples selected to span topic sensitivity, local relevance, intent ambiguity, and deployment rate within each intent type.

C Measuring Organic Fulfillment

Our theoretical model requires a measure of baseline organic fulfillment, capturing how well organic results satisfy the user's intent absent any AIO. We construct this measure following Google's Search Quality Rater Guidelines (Google, 2025d), which specify rating criteria, decision rules, and worked examples for each search intent type. This appendix describes the rating instrument, the aggregation to a SERP-level score, and the validation.

C.1 Design

Three choices were fixed before any deployment regression was estimated (Carlson and Burbano, 2025). The unit of rating is a single organic result block, evaluated on the evidence a user sees on the SERP: title, displayed URL, and snippet. The scale is the guidelines' own five-point Needs Met scale, from Fails to Meet (1) through Slightly, Moderately, and Highly Meets to Fully Meets (5). The aggregation to a SERP-level score is the position-discounted measure described in C.3. Fixing the primary measure in advance matters because several defensible aggregations of the same ratings are available, and selecting among them after seeing regression output would amount to specification search.

The rating covers only organic results. Where a SERP contains an AIO, a knowledge panel, a product carousel, or any other non-organic element, that element is removed before rating. Organic fulfillment as measured therefore describes the results a user would face in the absence of an AI Overview, which is the quantity the model requires and which cannot be affected mechanically by the deployment decision we study.

C.2 Rating Instrument

Each rating task presents the query, the user's location, and the top ten organic results in rank order. The prompt reproduces the guidelines' rating conditions for each point on the scale, and encodes as explicit rules the conditions under which the guidelines make the highest ratings available or unavailable:

- **Website intent.** Where the query's intent includes reaching a specific site and a result block is that site, the guidelines direct that it be rated Fully Meets.

- **Dual intent.** Where a query carries both a website and a visit-in-person intent, the guidelines cap at Highly Meets any result that satisfies only one of them.
- **Simple factual queries.** Where the query has a single unambiguous answer and the result displays it directly, the guidelines direct Fully Meets. Where the answer is present but not displayed, the rating is lower.
- **Queries that cannot Fully Meet.** The guidelines state that broad queries, ambiguous queries without a dominant interpretation, and queries with multiple reasonable interpretations cannot receive Fully Meets for any result.
- **Location mismatch.** For queries with a visit-in-person intent, results that do not fit the user's country and location are rated Fails to Meet.

The last of these makes the measure sensitive to the location supplied with the query, and we supply city-level rather than country-level locations for this reason.

Because several of these rules are conditioned on the query's intent in the guidelines' own sense, each rating task states that intent, classified in the separate pass described in Section 4.2. The rater applies the rules to the intent it is given and does not infer it. This dependency runs in one direction only: the intent classification does not observe the results, and the rater does not revise the intent.

We ran the instrument on Gemma 4 (31B) at temperature zero with a fixed seed, so that each result receives a single deterministic rating, and both waves were rated with the same model, prompt, and parameters, so that differences between waves reflect the results rather than the instrument. Model identifier, quantization, digest, prompt version, and seed are recorded in the output header, following the documentation practice recommended for open-weight models (Carlson and Burbano, 2025).

C.3 Aggregation

Ratings are aggregated to the SERP level by a discounted gain that weights results by their position and normalizes against a page on which every result fully meets the need. Writing g_p for the

numeric rating at position p, the score is

$$\frac{\sum_p \left(2^{\,g_p - 1} - 1\right) / \log_2(p+1)}{\sum_p \left(2^{\,5-1} - 1\right) / \log_2(p+1)}.$$

The exponential gain and logarithmic position discount follow the discounted-gain tradition (Järvelin and Kekäläinen, 2002) and match the guidelines' emphasis on prominence: a highly rated result at the top of the page contributes more than the same result further down. The normalization is deliberately against a perfect page rather than against the ideal reordering of the observed results. Standard normalized discounted cumulative gain divides out the level, so that a page of uniformly moderate results and a page of uniformly excellent results receive the same score. Since the model requires the level of organic fulfillment rather than the quality of its ordering, we normalize against the perfect page and retain the ordering-only variant as a diagnostic. We also store the maximum rating on the page, the mean of the top three, and the mean of all rated results.

C.4 Validation

The September 2025 edition of the guidelines contains 99 worked examples in which Google states the rating a given result should receive for a given query. We encode these as rating tasks and score the instrument against Google's own ratings. Examples cited within the prompt are excluded, so the validation set is held out. A gate was fixed before any model was scored: exact agreement of at least 60%, agreement within one rating category of at least 90%, and no rating category at zero recall. Five candidate models were scored in the production configuration, and the decision rule, also fixed in advance, was to select the model with the highest exact agreement among those clearing the gate.

The deployed instrument agrees exactly with Google's rating on 77.8% of the worked examples and within one category on 93.9%, with a mean absolute error of 0.29 on the five-point scale. Per-category recall is uneven: the instrument recovers Fails to Meet and Highly Meets reliably, Fully Meets in roughly two thirds of cases, and the two intermediate categories less often, reflecting that the guidelines themselves treat the boundary between Slightly and Moderately Meets as a matter of judgment.

D Robustness Checks

This appendix reports the full estimates for the three checks described in Section 5.7.

D.1 Within-query identification

Table D1 re-estimates the deployment model as a linear probability model absorbing a fixed effect for every query–location pair and for every collection batch within a wave. The intent dimensions, intent ambiguity, and the page-type controls are constant within a query–location pair and are absorbed by construction, so the table reports only the terms identified from variation across a query's own SERPs: the wave interactions, organic fulfillment, and source quality. Coefficients are in probability units, so a coefficient of -0.021 on organic fulfillment corresponds to -2.1 percentage points per standard deviation.

The targeted-side fulfillment slope is -2.1 points per standard deviation (CI $[-4.1, -0.2]$), the exploratory-side slope is 0.7 points and indistinguishable from zero (CI $[-0.4, 1.8]$), and the contrast between them is 2.8 points (CI $[0.6, 5.1]$). The wave increments retain the ordering reported in Table 5: relative to targeted non-monetizable queries, the increase is 12.0 points smaller for exploratory queries and 10.7 points smaller for monetizable queries, with the three-way term indistinguishable from zero.

D.2 In-AI Overview advertising and the CPC slope

Table D2 re-estimates column (3) of Table 4 with the CPC slope interacted with the EU sample and, separately, with YMYL status. Ads within AI Overviews were available in the United States but not in any EU country during both collection waves. Google excludes ads for topics such as finance, health, and other sensitive verticals, which our YMYL indicator approximates (Google, 2025a) (Section 5.7). The slope is 7.4 points per log point of CPC in the United States (CI $[3.4, 11.5]$) and 14.1 in Ireland (CI $[8.8, 19.3]$), a difference of 6.6 points (CI $[1.0, 12.2]$, $p = 0.020$). It is 12.7 points outside YMYL (CI $[7.1, 18.3]$) and 7.1 within (CI $[2.9, 11.3]$), a difference of -5.6 points (CI $[-12.4, 1.2]$, $p = 0.106$). A three-way interaction gives slopes of 10.9 (CI $[4.2, 17.6]$) for United States non-YMYL, 5.7 (CI $[0.9, 10.5]$) for United States YMYL, 19.9 (CI $[12.4, 27.4]$) for Irish non-YMYL, and 13.3 (CI $[7.7, 18.9]$) for Irish YMYL queries. The estimation sample is the

Table D1: Within-query identification: linear probability model with query–location and batch fixed effects.

	Pr(AI Overview)
Exploratory × Wave 2026	-0.120*** (0.024)
Monetizable × Wave 2026	-0.107*** (0.031)
Exploratory × Monetizable × Wave 2026	0.018 (0.037)
Organic fulfillment	-0.021* (0.010)
Exploratory × organic fulfillment	0.028* (0.011)
Source quality	0.005 (0.005)
Source quality squared	-0.02 (0.35)
Exploratory × source quality	-0.004 (0.005)
Exploratory × source quality squared	0.27 (0.44)
ln(search volume)	-0.014 (0.018)
Google ecosystem score	0.021 (0.030)
Constant	0.341** (0.123)
Query–location fixed effects	Yes
Batch × wave fixed effects	Yes
SERPs	24,217
Query clusters	1,575
Within R^2	0.022

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Linear probability model estimated by `reghdfe`, absorbing a fixed effect for each query–location pair and each collection batch within a wave. Standard errors clustered on query and location. The intent dimensions, their interaction, intent ambiguity and its interaction with exploration, and the page-type controls (YMYL, question format, EU sample, local status) are constant within a query–location pair and are absorbed. The wave main effect is absorbed by the collection-batch fixed effects, which are defined separately within each wave. The squared source-quality terms are reported in units of 10^{-4}.

6,130 commercial-cell SERPs with an observed CPC.

Table D2: Advertising-value slope by region and by YMYL status, commercial cell.

	(1) CPC × EU sample	(2) CPC × YMYL
ln(CPC)	0.381*** (0.109)	0.744*** (0.185)
ln(CPC) × EU sample	0.617** (0.212)	
ln(CPC) × YMYL		-0.377 (0.217)
Wave 2026	0.163 (0.096)	0.161 (0.097)
Organic fulfillment	-0.038 (0.057)	-0.025 (0.058)
Source quality	0.021*** (0.005)	0.022*** (0.005)
ln(search volume)	-0.054 (0.032)	-0.057 (0.033)
Mobile device	0.489*** (0.072)	0.486*** (0.071)
YMYL	0.251 (0.200)	0.717* (0.324)
Question query	0.365 (0.225)	0.379 (0.227)
EU sample	-1.536*** (0.281)	-0.888*** (0.181)
Google ecosystem score	0.023 (0.482)	0.020 (0.471)
Local query	-1.273*** (0.185)	-1.235*** (0.190)
Constant	-1.997*** (0.482)	-2.407*** (0.514)
SERPs	6,130	6,130
Query clusters	425	425
Pseudo R^2	0.176	0.172

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Logistic regression for the probability that an AI Overview is deployed, on commercial-cell SERPs with an observed CPC, as in column (3) of Table 4. Standard errors clustered on query and location. Marginal effects are reported in the text of Section 5.7.

D.3 Local placebo

Table D3 adds interactions between local status and the wave, intent ambiguity, and organic fulfillment to the specification of column (2) of Table 4. The wave increase is 11.2 points on non-local queries (CI $[9.1, 13.2]$) and 6.2 points on local ones (CI $[3.2, 9.2]$), a difference of -4.9 points (CI $[-8.7, -1.2]$). Within non-local queries the incidence across the monetization dimension is unchanged, at 16.5 points for non-monetizable and 5.7 points for monetizable queries, a difference of -10.8 points (CI $[-14.5, -7.1]$).

Table D3: Local placebo: wave growth interacted with local status.

	Pr(AI Overview)
Exploratory	3.328** (1.236)
Monetizable	0.542* (0.264)
Exploratory × Monetizable	-1.189*** (0.292)
Wave 2026	1.885*** (0.171)
Exploratory × Wave 2026	-1.268*** (0.193)
Monetizable × Wave 2026	-1.051*** (0.235)
Exploratory × Monetizable × Wave 2026	0.548* (0.264)
Local query × Wave 2026	-0.075 (0.135)
Local query × intent ambiguity	0.581 (0.303)
Local query × organic fulfillment	-0.247* (0.103)
Organic fulfillment	-0.349*** (0.089)
Exploratory × organic fulfillment	0.400*** (0.100)
Source quality	0.035 (0.031)
Source quality squared	-1.50 (2.31)
Exploratory × source quality	0.013 (0.038)
Exploratory × source quality squared	-1.27 (2.85)
Intent ambiguity	0.534*** (0.154)
Exploratory × intent ambiguity	-1.093*** (0.201)
ln(search volume)	-0.106*** (0.017)
Mobile device	0.364*** (0.036)
YMYL	0.950*** (0.085)
Question query	0.643*** (0.104)
EU sample	-0.966*** (0.094)
Google ecosystem score	-0.222 (0.178)
Local query	-1.935*** (0.207)
Constant	-3.949*** (1.027)
SERPs	24,217
Query clusters	1,575
Pseudo R^2	0.290

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Logistic regression for the probability that an AI Overview is deployed, with local status interacted with the wave, intent ambiguity, and organic fulfillment. Standard errors clustered on query and location. The squared source-quality terms are reported in units of 10^{-4}. Marginal effects are reported in the text of Section 5.7.