ChatGPT Referrals to E-Commerce Websites: How Do LLMs Compare Against Traditional Channels?

Maximilian Kaiser* Christian Schulze†

March 2026

Abstract

We investigate organic Large Language Model traffic (oLLM) versus traditional digital channels in e-commerce. Analyzing 12 months of first-party data from 973 websites with $20 billion combined revenue, we examine over 50,000 transactions from ChatGPT referrals alongside 164 million transactions from traditional channels. Using regression models that account for data sparsity, we assess financial metrics (conversion rate, average order value, revenue per session) and engagement metrics (bounce rate, session duration, page views). Results are consistent across extensive robustness checks. One year after launch, oLLM exhibits conversion rates and revenue per session above paid social but below all other traditional channels. Product complexity moderates the effects: oLLM's financial outcomes and traffic shares are stronger in complex product categories. Engagement metrics show favorable bounce rates but lower session duration and page views. Temporal analysis shows increasing conversion rates but declining average order values, yielding only moderate revenue-per-session gains over time. Cross-website analyses support growing consumer LLM proficiency as the underlying mechanism. The descriptive study positions oLLM as a new and developing channel. With low volumes and modest revenue per session, oLLM currently serves niche informational needs of proficient consumers and does not yet function as a broad conversion channel.

Keywords:
large language models (LLMs), electronic commerce, channel analytics, performance marketing, digital economy

*University of Hamburg, Grips Intelligence; research@maximiliankaiser.org; LinkedIn
†Frankfurt School of Finance & Management; email@christian-schulze.de; LinkedIn

Acknowledgements. The authors thank Daniel Schmeh and Jan Seifried for their brilliant ideas and helpful discussions on this paper. We thank Daniel Blaseg, Alexander Bleier and Michel Clement for helpful feedback and early comments. Further, we would like to thank Grips Intelligence for supporting academic research by providing the extensive data used in this study entirely free of charge.

1 Introduction

A new digital channel emerged in August 2024 when ChatGPT, the dominant Large Language Model (LLM) platform, began providing organic outgoing links that direct users to relevant e-commerce websites upon expression of purchase intent. For instance, consumers searching for an espresso machine receive personalized product comparisons and country-specific purchase links (Figure 1). As of the time of this study, these links are fully organic and not advertiser-driven.

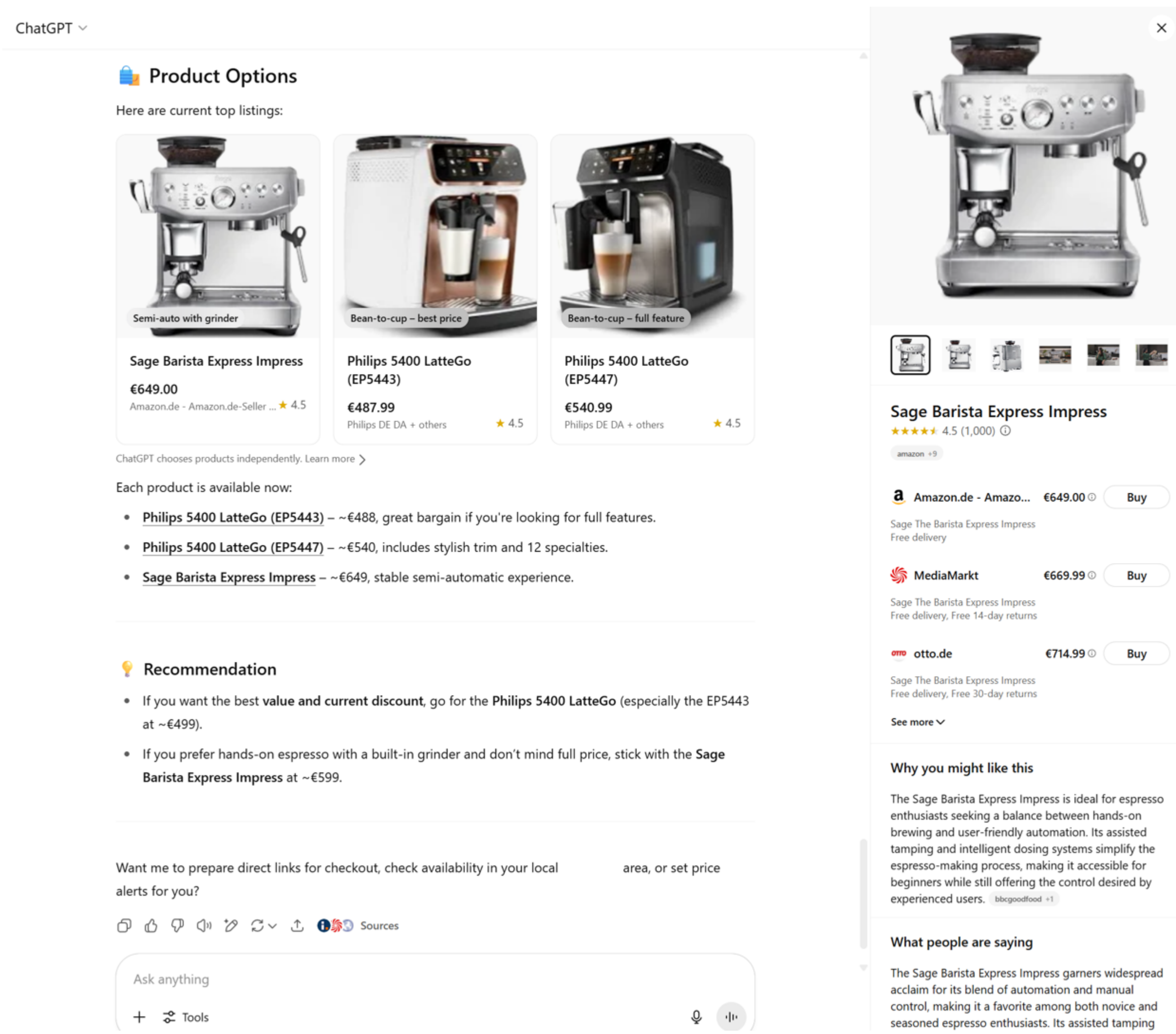


Figure 1: Product recommendations for espresso machines provided by ChatGPT, featuring outgoing links to online retailers ("Buy" links on the right).

This organic LLM traffic (oLLM) enters a landscape long shaped by organic and paid search (Ghose and Yang 2009, Li et al. 2016, Reisenbichler et al. 2022) but also referrals, email, affiliate, paid social, and direct channels (Manchanda et al. 2006, Bleier and Eisenbeiss 2015, de Haan et al. 2016, Wies et al. 2023, SimilarWeb 2024). Channels differ in where and how they match consumers to sellers along the purchase funnel. Their lower-funnel effectiveness is commonly measured by conversion rates in academia (Ghose and Yang 2009, Rutz and Bucklin 2011) and practice (Google 2024). How oLLM compares remains an open question.

Several mechanisms suggest oLLM may generate superior outcomes. LLMs access richer contextual information than traditional channels (Soviero et al. 2024), which have been constrained by privacy regulation (Johnson et al. 2020, Miller and Skiera 2024, Aridor et al. 2025). They synthesize information across attributes, reviews, and contextual factors (Yu et al. 2024) and can reason across dimensions that keyword-based search cannot (Al-Hasan et al. 2024, Luo et al. 2019, Yu et al. 2025). Conversational interfaces further enable preference clarification and iterative personalization (Jannach 2022, Al-Hasan et al. 2024), potentially reducing cognitive burden and increasing convenience and persuasion (Steyvers and Kumar 2024, Salvi et al. 2025). Early industry evidence is consistent with this logic: a study of 100 websites reports 6.7% conversion rate for oLLM versus 3.9% for organic search (ThoughtMetric 2025); another in-depth study reports 15.9% versus 1.8% (Seer Interactive 2025); and value per session from oLLM is estimated to be 4.4× higher than organic search (Semrush 2025).

Other evidence points in the opposite direction. Perceived recommendation usefulness may be hampered by inaccuracies in current LLM outputs (Search Engine Land 2025), declining response quality at scale (Huang and Rust 2025), and interface characteristics that can increase cognitive load (Nguyen et al. 2022). Consumer adoption of LLM-based shopping remains early-stage: technology anxiety and trust moderate adoption intentions (Foroughi et al. 2025), adoption and use correlate with digital sophistication and education

(Yang et al. 2025), and only 2.1% of ChatGPT conversations involve purchasable products (Chatterji et al. 2025), reflecting delayed emergence of behavioral shifts (Padilla et al. 2025). Empirical findings consistent with these indicators include Adobe (2025) reporting oLLM conversion rates 9% lower than "non-AI" channels, and SALT (2025) finding engagement levels 27% below organic search in most categories.

Understanding oLLM has broader implications. Changes in search costs and decision quality may influence consumer welfare (Anderson and Renault 2006, Lynch and Ariely 2000, Brown and Goolsbee 2002, Ursu et al. 2022, Dinerstein et al. 2018, Ellison and Ellison 2009). Recommendation bias may differ from advertiser-influenced channels (OpenAI 2025, Reuters 2025).[1] As deep-link entry grows, website design may need to adapt and retail media may become less important (Wroe 2024, eMarketer 2024). More generally, LLMs are reshaping online behavior (Padilla et al. 2025, Gholami et al. 2026) and fueling expectations of disruption to e-commerce channels (The Economist 2023).

Our descriptive study is the first large-scale empirical analysis of oLLM relative to traditional channels. Using 12 months of data (August 2024–July 2025) from 973 e-commerce websites with $20 billion in annual revenue, we observe more than 50,000 oLLM transactions and 164 million transactions from other channels. Through direct access to companies' Google Analytics accounts provided by Grips Intelligence,[2] we observe the channel that brought each consumer to the e-commerce website and their subsequent purchase behavior. Importantly, because these channel assignments rely on last-click attribution, we cannot capture upper-funnel contributions and will understate channels' roles when they primarily serve discovery functions (Li and Kannan 2014, Li et al. 2016, de Haan et al. 2016, Berman 2018). Still, last-click metrics remain the industry standard and, under current data constraints, provide the most comparable basis for cross-website channel benchmarking. We

[1] LLM platforms had not yet introduced affiliate-like systems or ads during our study period (Hermann et al. 2025).

[2] The representativeness of Grips' data relative to e-commerce in general has been previously assessed by comparing it to trusted proprietary datasets (e.g., Similarweb) and verifiable public ones (e.g., Shopify Quarterly Reports, United States E-Commerce Census). For details, see Online Appendix E in Aridor et al. (2025).

interpret all results with these limitations in mind.

We compare oLLM with traditional channels across financial metrics (conversion rate, average order value, and revenue per session) and engagement metrics (bounce rate, session duration, and page views) using both model-free and regression approaches. We test robustness across alternative data processing choices, website samples, LLM platforms, and timeframes. We also explore differential effects across websites that differ in product category complexity and consumers' LLM proficiency (Figure 2).

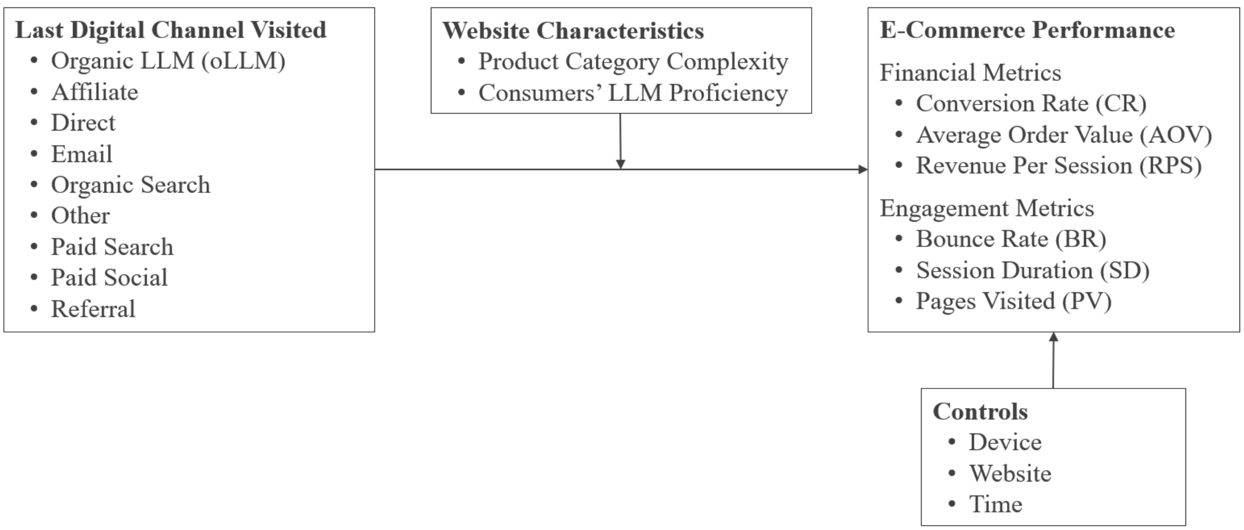


Figure 2: Conceptual model

Two patterns emerge. First, oLLM exhibits conversion rates and revenue per session above paid social but below all other traditional channels. Product complexity moderates the effects: oLLM's financial outcomes and traffic shares are stronger in complex product categories, pointing to limited perceived usefulness for simpler products. Engagement metrics reveal relatively favorable bounce rates but lower session duration and page views.

Second, temporal analysis shows increasing conversion rates. However, declining average order values partially offset these improvements, resulting in moderate revenue-per-session uplift over time. Cross-website analyses reveal that this pattern is consistent with growing consumer LLM proficiency as an underlying mechanism.

The paper proceeds as follows: We present data, metrics, and methodology, followed by descriptive evidence and regression results comparing oLLM with traditional channels. We then examine robustness, temporal dynamics, and heterogeneity in oLLM patterns across websites. We conclude with summary of key insights and limitations.

2 Data, Metrics, and Models

2.1 Traffic Volume of oLLM and Focus on ChatGPT

One year after launch, oLLM accounts for less than 0.2% of all traffic in our dataset (Figure 3), about 200 times smaller than Google's organic search. This reflects oLLM's low volume in the broader market (SE Ranking 2025, Ahrefs 2025).

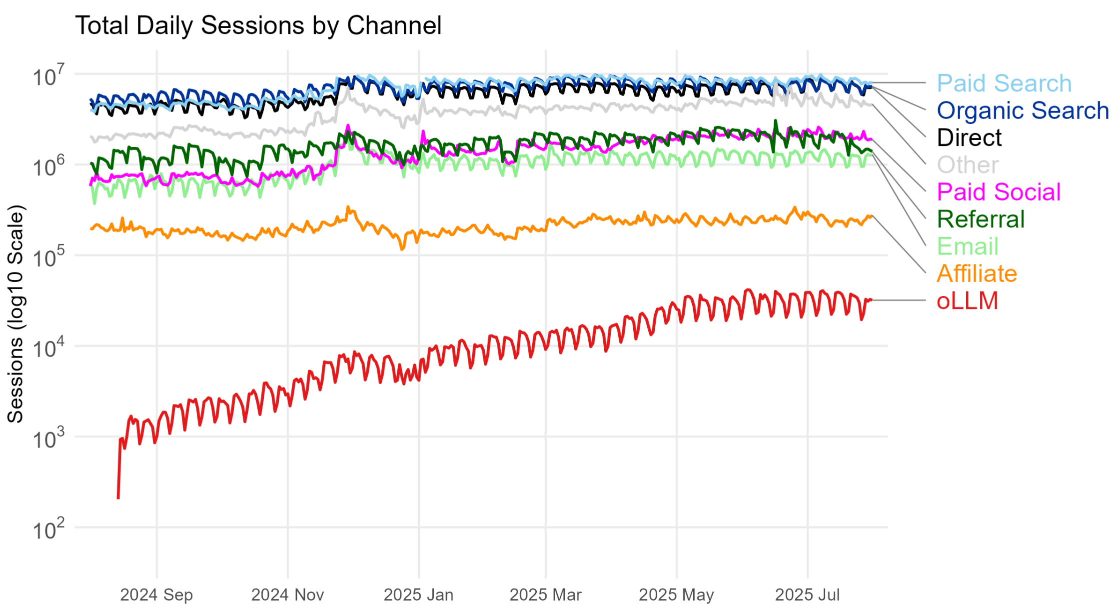


Figure 3: Organic LLM traffic (from ChatGPT only) compared to traditional channels (log scale)

ChatGPT dominates the oLLM channel, accounting for over 90% of observed sessions[3] from LLM platforms (Figure 4). Other platforms (Perplexity 4.1%, Gemini 2.6%, Copilot

[3]A session refers to a continuous period of website activity by a consumer, potentially comprising multiple page views. Sessions typically end after 30 minutes of inactivity. A single consumer may generate multiple sessions on the same website within a day.

2.1%, Deepseek 0.07%, Grok 0.02%) show negligible volume. Our main analyses therefore examine ChatGPT traffic only, but we include all platforms in a robustness check (Section 4).

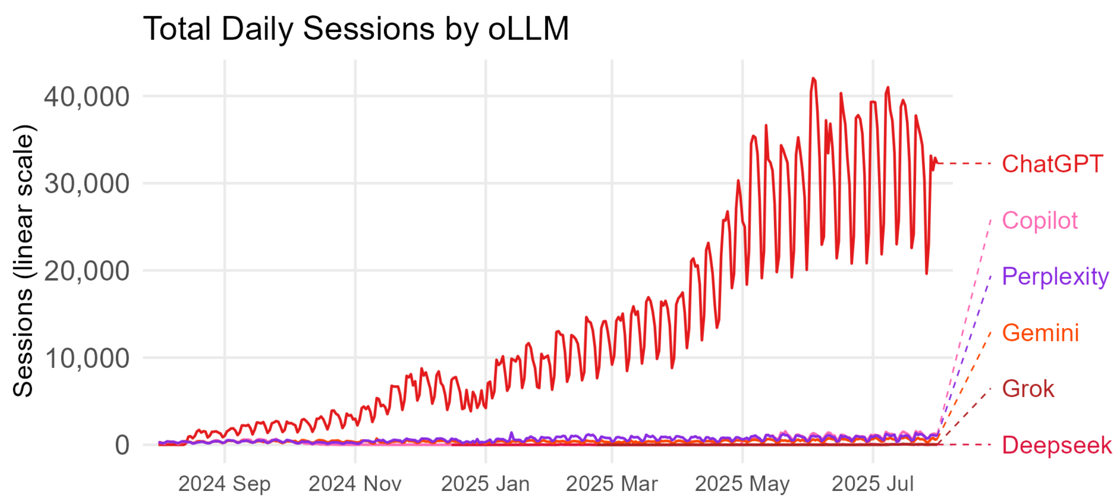


Figure 4: Observed traffic from large language models in our data (absolute scale)

2.2 Dataset

We analyze first-party e-commerce data from Google Analytics across 973 websites. Table 1 shows overlapping 3-, 6-, and 12-month data blocks. Our main analyses use the 6-month dataset, capturing most oLLM sessions while avoiding potential idiosyncratic introduction effects. We aggregate data weekly to account for sparse observations. We replicate all analyses using alternative timeframes (3 and 12 months) and aggregation periods (daily, monthly) in Section 4.

Table 1: Overview of Datasets

Metric	12 Months (robustness)	6 Months (main)	3 Months (robustness)
Date Range Start	2024-07-28	2025-02-02	2025-04-27
Date Range End	2025-08-02	2025-08-02	2025-08-02
Number of Websites	973	973	953
Total Sessions	10,513,177,108	6,010,804,542	3,271,538,643
Total Revenue	$20,602,399,030	$11,485,304,150	$6,301,491,260
Total Transactions	164,875,690	92,826,266	51,391,206
Total oLLM Transactions	50,251	44,437	33,680
Total oLLM Sessions	4,915,779	4,149,187	2,944,179
Total oLLM Revenue	$6,965,894	$6,082,834	$4,594,709
Mean Sessions per Website/week	238,627	254,189	253,509
Mean Revenue per Website/week	$467,631	$485,698	$488,298
Mean Transactions per Website/week	3,742	3,925	3,982
Mean Conversion Rate per Website/week	3.30%	3.27%	3.31%

Note: Weekly aggregated data. oLLM refers to traffic from ChatGPT referrals.

Coverage

Following Similarweb (2025), our dataset contains websites from 24 categories. The largest is "E-Commerce and Shopping" (including retailers and marketplaces) with 1.2 billion sessions, followed by "Lifestyle" (including Fashion, Beauty and Cosmetics) with 1.1 billion sessions. Detailed category descriptions and statistics appear in Online Appendix A.1.

The dataset covers all continents, with the Americas (2.7 billion sessions) and Europe (2.1 billion sessions) representing the majority. 49 countries contribute at least 10 million sessions each. Detailed statistics appear in Online Appendix A.2.

2.3 Metrics

We focus our analyses on financial metrics: conversion rate (CR), average order value (AOV), and revenue per session (RPS). CR is the most widely used financial performance metric in e-commerce, capturing the share of website sessions that result in a transaction. We also measure AOV, because the probability of purchase tends to be higher when the product is cheaper, meaning CR and AOV are confounded. RPS captures the joint effect of purchase likelihood and spending per purchase, and gives meaningful intuition on the value of sessions from an advertising perspective. Table 2 provides descriptives for oLLM and eight traditional channels.

Table 2: Channel Volume Summary Statistics (February - August 2025)

Channel	Obs	Sessions (M)	Revenue ($M)
oLLM	30,232	4.1	6.1
Paid Search	44,800	1,530.0	3,190.4
Organic Search	47,244	1,427.3	2,595.4
Direct	47,247	1,269.6	2,695.3
Other	46,941	830.5	1,581.5
Referral	46,789	353.5	605.2
Paid Social	27,019	337.5	145.5
Email	36,544	215.9	513.5
Affiliate	13,476	42.5	152.6

Notes: (M) = millions, observations refer to week-website-device-channel data points.

We investigate engagement through bounce rate (BR), session duration (SD), and page views (PV). While engagement does not directly translate to financial results, it indicates traffic quality. High BR (visitors leaving without another click) typically signals poor fit between website content and consumer interest. Complete variable definitions appear in Online Appendix A.3.

2.4 Modeling Approach

Sparse oLLM observations require careful modeling. At the week/website/device level, CRs may derive from few sessions and AOVs from single transactions. Our models address two challenges: (a) datapoints vary substantially in underlying sessions, and (b) ratio metrics exhibit overdispersion.

For CR and BR, we estimate quasibinomial models with dispersion parameters. For AOV, RPS, SD, and PV, we estimate weighted linear models with bounded, website-relative weights to reduce aggregation-induced heteroskedasticity and cap leverage. All models use:

$$\mathrm{logit}\big(p_{it}^{\mathrm{CR}}\big) = \alpha_{c(i,t)} + \gamma_i + \delta_{d(i,t)} + \mu_{m(t)}, \tag{1}$$

where α_c are channel fixed effects, γ_i website fixed effects, δ_d device fixed effects, and μ_m month fixed effects. The linear specifications follow the same fixed-effects structure, with outcomes modeled in levels rather than log-odds. Complete specifications appear in Online Appendix A.4.

3 Results

3.1 Model-Free Evidence

Model-free comparisons suggest lower CR, AOV, and RPS for oLLM relative to traditional channels (Figures 5 and 6). However, these distributions reflect sparse data: the median oLLM observation contains zero transactions. Without accounting for underlying session counts or website heterogeneity, such comparisons prove misleading—a pattern that persists across engagement metrics (Figures 7 and 8).

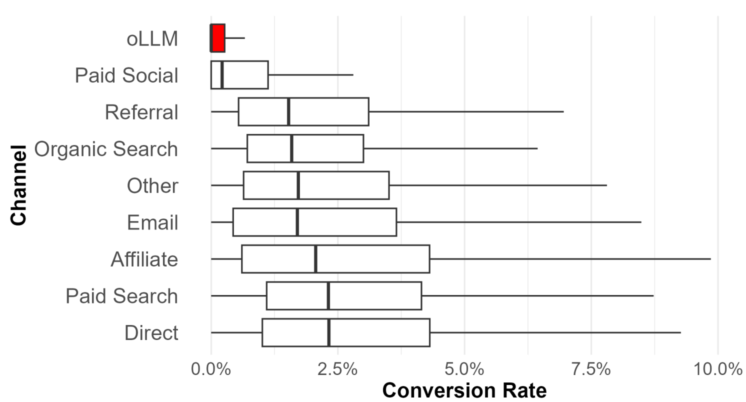
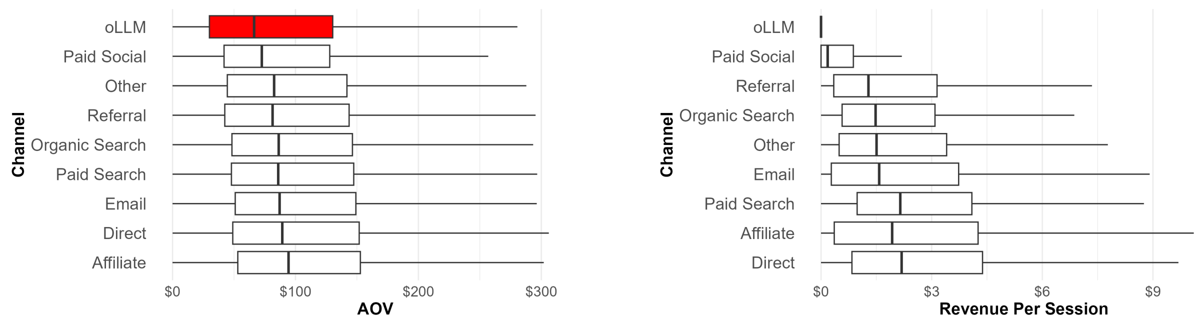


Figure 5: Model-free evidence on CR by channel



(a) AOV (b) RPS

Figure 6: Model-free evidence on AOV and RPS by channel

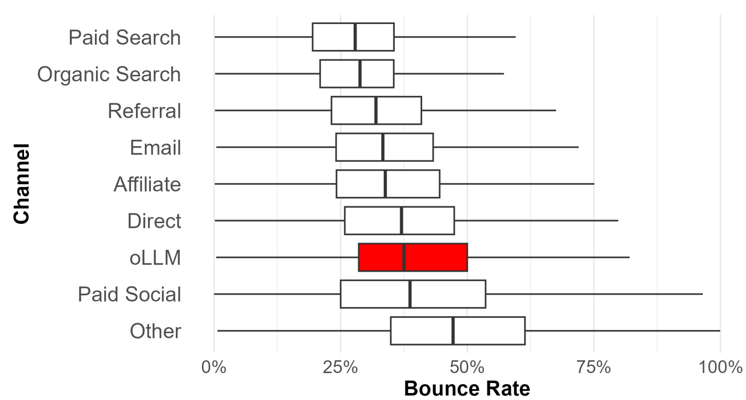
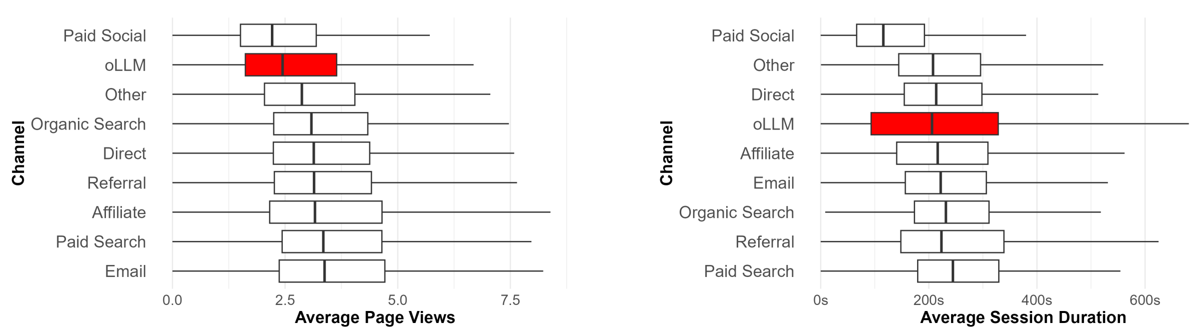


Figure 7: Model-free evidence on BR by channel



(a) PV (b) SD

Figure 8: Model-free evidence on PV and SD by channel

3.2 Regression Results

Table 3 reports regression estimates with channel fixed effects α_c measuring differences between each traditional channel and oLLM (the reference category). We highlight in **bold** coefficients with $p < 0.05$ and favoring oLLM.

Table 3: Regression Results

	Dependent variables					
	CR	AOV	RPS	BR	SD	PV
Affiliate	0.621*** (0.060)	24.7*** (3.9)	3.57*** (0.12)	**0.217***** (0.035)	**-12.2*** (6.8)	0.642*** (0.057)
Paid Search	0.370*** (0.058)	4.08 (3.10)	1.95*** (0.08)	-0.156*** (0.033)	31.2*** (4.5)	0.905*** (0.038)
Other	0.366*** (0.059)	5.51* (3.15)	2.19*** (0.08)	**0.883***** (0.033)	46.1*** (4.7)	0.622*** (0.039)
Direct	0.282*** (0.058)	13.5*** (3.1)	3.07*** (0.08)	**0.321***** (0.033)	5.86 (4.56)	0.692*** (0.038)
Email	0.276*** (0.059)	10.5*** (3.3)	2.69*** (0.09)	**0.181***** (0.033)	55.4*** (5.1)	0.876*** (0.042)
Referral	0.218*** (0.059)	12.6*** (3.2)	1.61*** (0.08)	**0.202*** (0.033)**	59.8*** (4.8)	0.776*** (0.040)
Organic Search	0.121** (0.058)	1.83 (3.08)	1.01*** (0.08)	-0.136*** (0.033)	14.6*** (4.5)	0.571*** (0.037)
Paid Social	**-0.760***** (0.059)	5.42 (3.49)	**-0.718***** (0.091)	**0.360***** (0.033)	**-131***** (5)	**-0.413***** (0.045)
Fixed Effects: Website, Device, Month	Yes	Yes	Yes	Yes	Yes	Yes
R^2 [pseudo-R^2]	[0.739]	0.472	0.284	[0.711]	0.457	0.155
Observations	340,292	289,524	340,292	340,292	340,292	340,292

*p<0.1; **p<0.05; ***p<0.01

NOTES: All coefficients relative to oLLM baseline. Standard errors in parentheses.

Financial Outcomes

Figure 9 displays channel fixed effects as log-odds coefficients. To facilitate interpretation, we convert these to percentage differences in conversion likelihood. Among all channels, paid social and organic search exhibit the smallest gaps relative to the oLLM baseline (shown as red line). Paid social sits substantially below oLLM (coefficient: -0.760), corresponding to a 53% lower likelihood of conversion. Organic search modestly exceeds oLLM (coefficient: 0.121), translating to a 13% higher conversion likelihood. Other channels show larger differences: affiliate's coefficient of 0.621 corresponds to an 86% higher likelihood of conversion relative to oLLM. These regression-adjusted estimates substantially narrow the gaps indicated by model-free evidence.

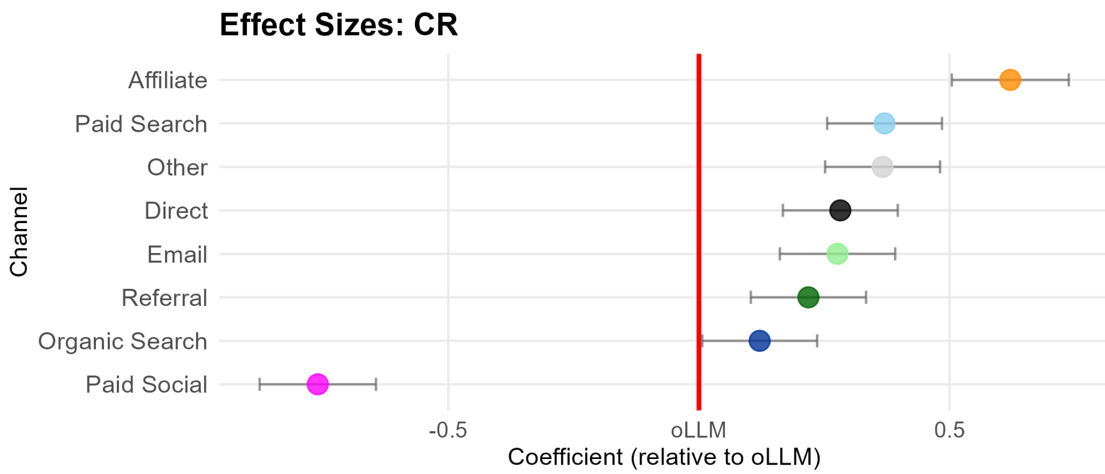


Figure 9: Regression results for CR by channel (log-odds)

AOV estimates exhibit wider confidence intervals, yielding statistically significant differences for only four of eight channels (Figure 10). Affiliate shows the largest difference at \$24.7 higher than oLLM. Differences for organic search, paid search, paid social, and other are not statistically significant.

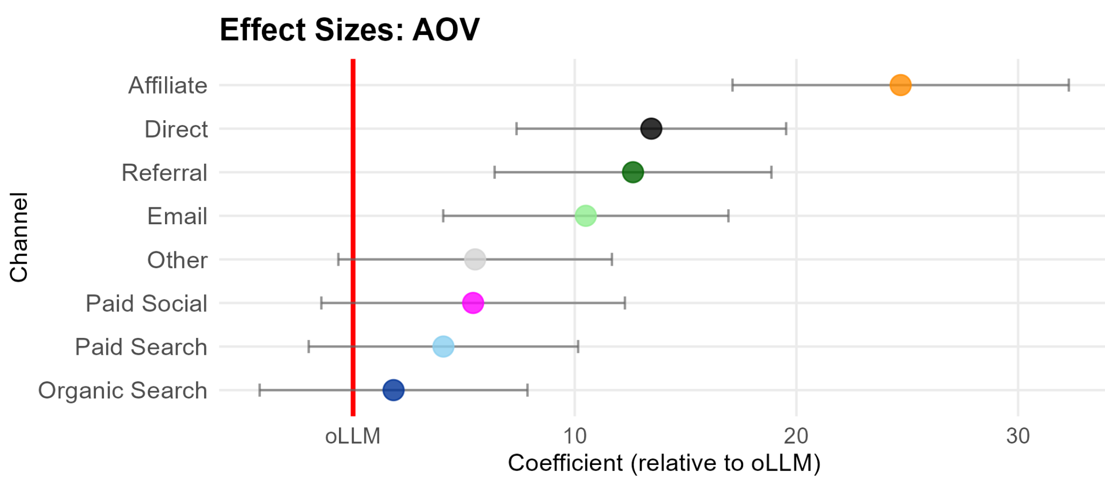


Figure 10: Regression results for AOV by channel

RPS results show tighter confidence intervals (Figure 11). oLLM exceeds paid social but falls significantly below remaining channels.

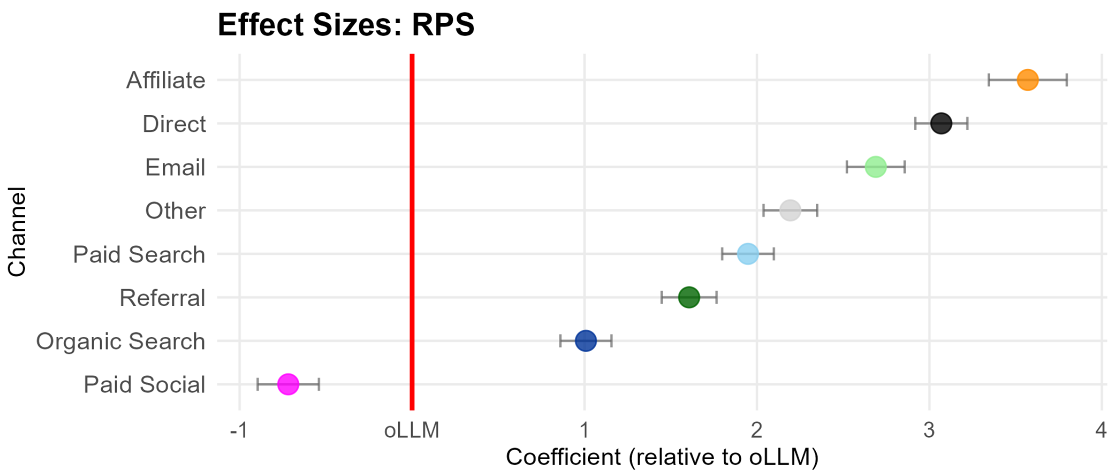


Figure 11: Regression results for RPS by channel

Engagement Outcomes

BR is the only metric where oLLM shows more favorable outcomes than most channels (Figure 12), though organic and paid search—channels optimized for low BR—still show better rates. Email shows 20% higher BR than oLLM; organic search shows 13% lower.

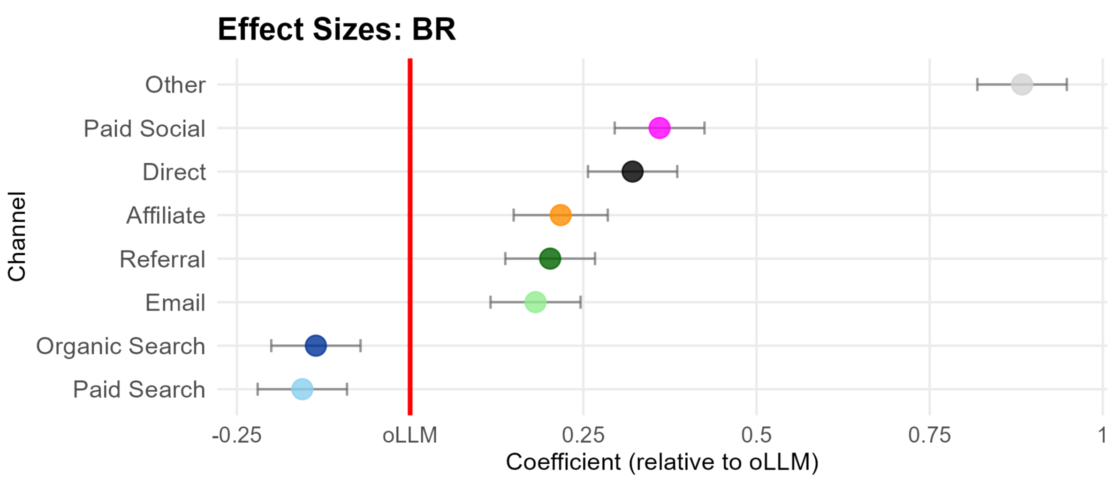
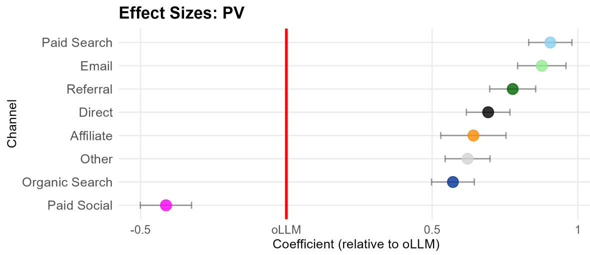
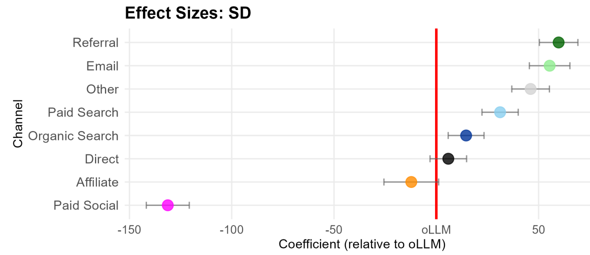


Figure 12: Regression results for BR by channel

PV analysis shows oLLM below all channels except paid social, while SD places oLLM far ahead of paid social and comparable to affiliate and direct channels (Figure 13).



(a) PV



(b) SD

Figure 13: Regression results for PV and SD by channel

4 Robustness

The main models aggregate data at the weekly level, apply no minimum observation thresholds, cover 973 websites, analyze ChatGPT as the only oLLM channel, and span 6 months. We test the robustness of our findings with respect to all of these specifications. Table 4 summarizes observations and model fit. Notably, the R^2 values of the CR model remain stable despite substantial variation in sample size.

Table 4: Robustness Checks Overview

Specification	Robustness Concern Addressed	R^2	Observations
Main Model	– (reference model)	0.739	340,292
Data Processing Choices			
Daily Aggregation	Effect of sparse observations for oLLM	0.689	3,896,527
Monthly Aggregation	Effect of sparse observations for oLLM	0.750	149,628
Min 10 Sessions	Effect of sparse observations for oLLM	0.739	320,021
Min 100 Sessions	Effect of sparse observations for oLLM	0.740	280,190
Min 1000 Sessions	Effect of sparse observations for oLLM	0.742	188,376
Winsorized	Outliers	0.727	340,292
Observed Websites			
Top 25% oLLM Sessions	Selection of websites in the data	0.700	91,385
Top 25% Total Transactions	Selection of websites in the data	0.738	91,942
Observed LLM platforms			
ChatGPT plus other LLMs	Operationalization of oLLM	0.746	408,342
ChatGPT plus mobile app traffic	Operationalization of oLLM	0.741	395,430
Observed Timeframes			
12 Months	Stability of the effects over time	0.723	614,714
3 Months	Stability of the effects over time	0.727	188,038

Notes: R^2 values shown for CR regression model denote McFadden pseudo-R^2 values.

Figure 14 reports CR channel fixed effects across all specifications. The core finding is robust: oLLM exhibits lower CR than all traditional channels except paid social. Differences relative to organic search become statistically insignificant in five specifications (monthly aggregation, minimum 100 sessions, top 25% of oLLM traffic, oLLM including mobile app, and the 3-month window) and exhibit a non-significant sign reversal in one specification (minimum 1,000 sessions). Effect sizes vary systematically: the top-25%-revenue subset shows consistently larger gaps, whereas the minimum-1,000-sessions specification yields smaller effects. Nonetheless, the overall pattern remains consistent.

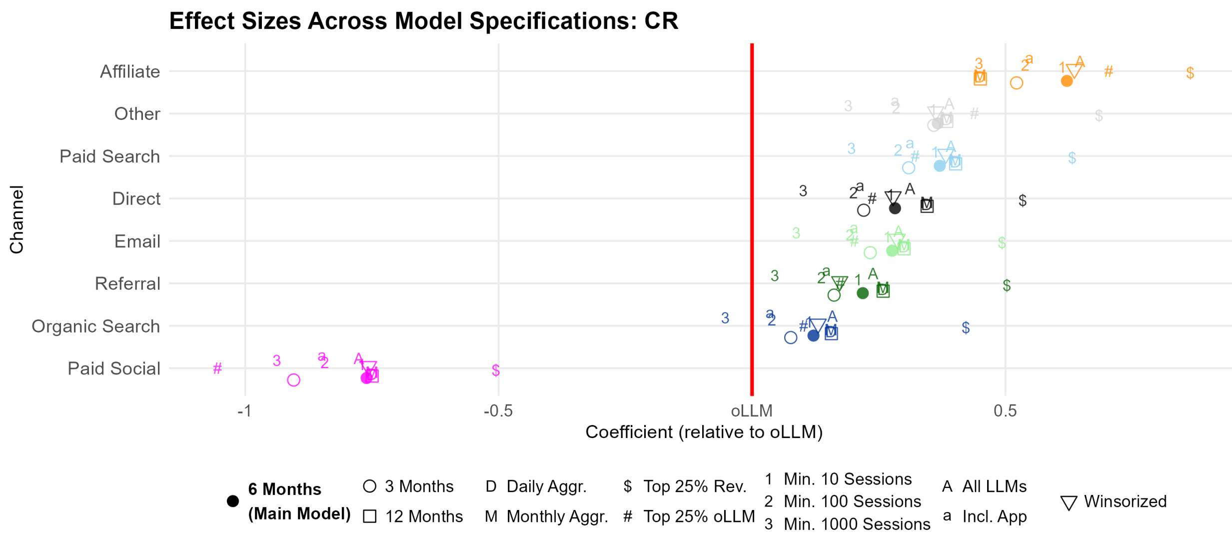


Figure 14: Illustration of robustness check results for CR by channel

AOV, RPS, and engagement metrics exhibit similar robustness, with minor variation in effect sizes. Details on all tests are available in Online Appendix B.

5 Temporal Dynamics of oLLM Outcomes

The preceding analyses document oLLM's lower CR and RPS relative to most traditional channels. Yet oLLM remains a young and rapidly developing channel, with both the underlying technology and consumer adoption still maturing. This raises a natural generalizability question: do the observed gaps persist, narrow, or widen over the study period?

To address this question, we examine temporal trends in CR, AOV, and RPS. Channel fixed effects across 12-, 6-, and 3-month windows already suggest gradual improvement. We now model these trajectories explicitly, using multiple trend specifications to assess robustness. This analysis cannot isolate specific mechanisms driving temporal changes but can characterize the direction and pace of oLLM's trajectory relative to established channels.

5.1 Model-free Evidence on Time Trend

Model-free evidence (Figure 15) shows CR increases over time for oLLM, while traditional channels remain stable except for seasonal peaks in November and December. Mean weekly AOV (Figure 16) exhibits substantially more volatility, especially for oLLM, and appears to decrease over the observation period. RPS remains stable for traditional channels but shows systematic increase for oLLM (Figure 17).

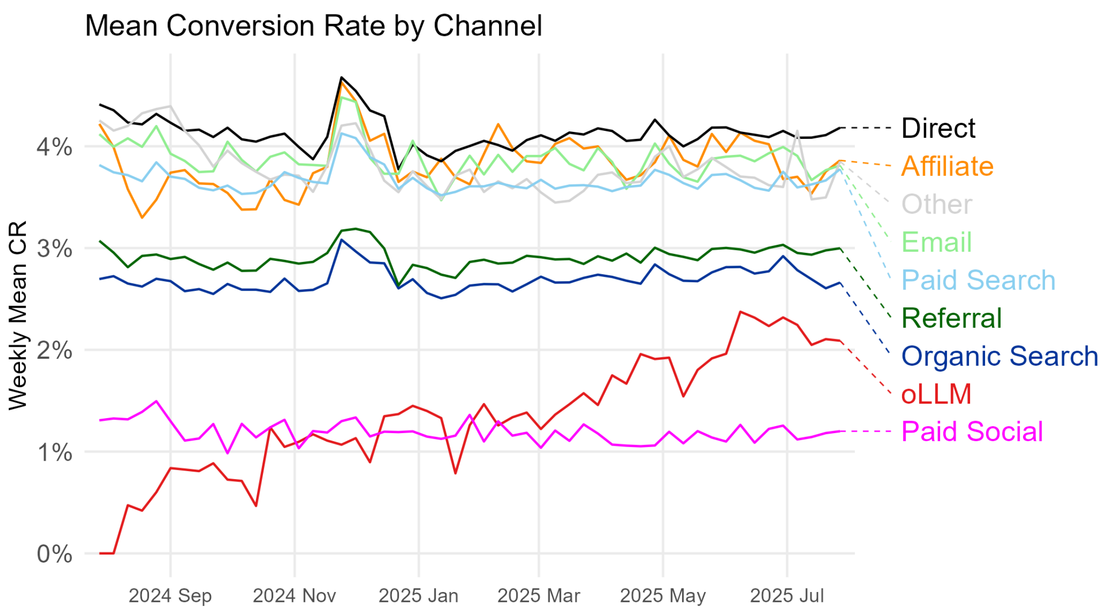


Figure 15: CR by channel over time

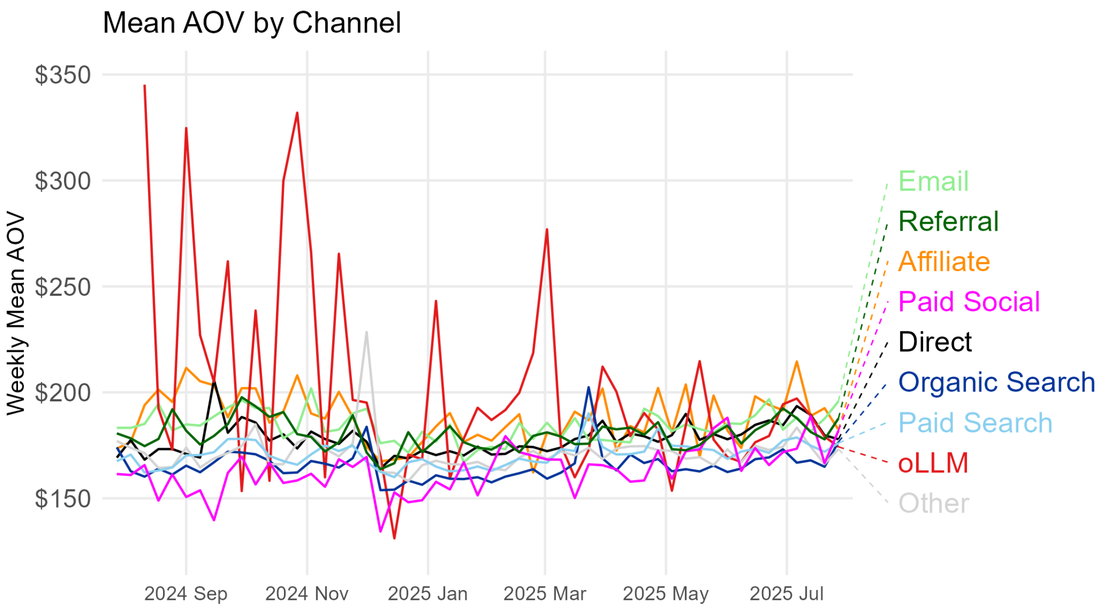


Figure 16: AOV by channel over time

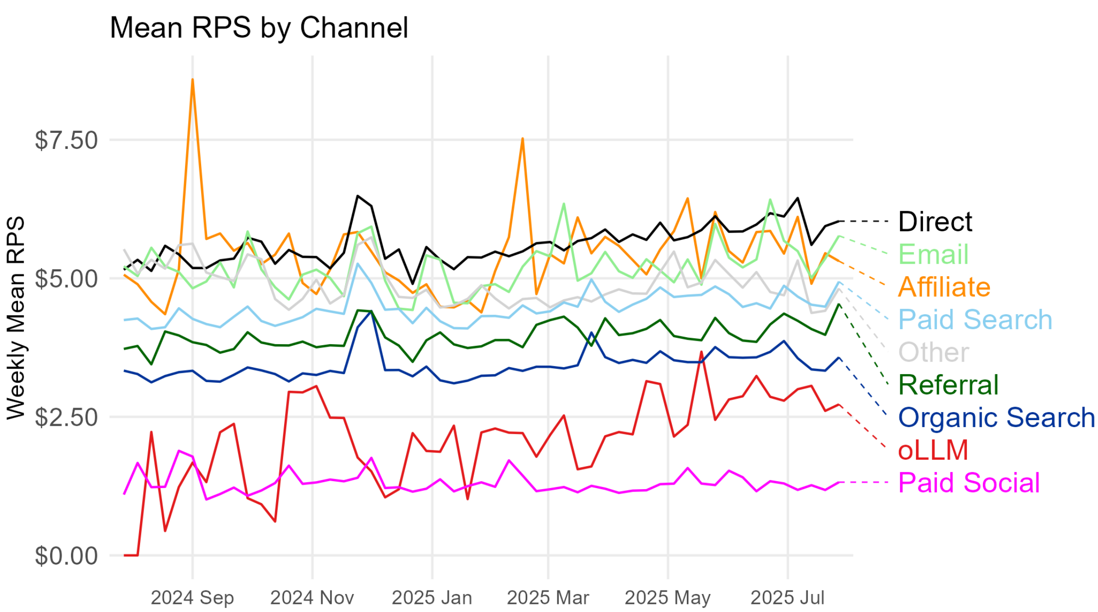


Figure 17: RPS by channel over time

5.2 Regression Analysis with Time Trend

Given data sparsity challenges with model-free evidence for oLLM, we estimate regression models mirroring main specifications but including trend variables. Trend variable specification substantially impacts predictions, so we test four alternatives: linear trend, centered (linear) model, Gompertz model (asymmetric S-curve), and Sigmoid model (logistic with hard upper limit). Complete specifications appear in Online Appendix C.1.

5.3 Projection of Financial Metrics

To facilitate interpretation, we create population estimates for all websites across time, allowing us to project oLLM outcomes in the next year. While trend continuation is highly uncertain, projections may offer insight into development speed necessary for oLLM to converge with traditional channels.

Figure 18 illustrates slow CR increase for traditional channels and comparatively steeper increase for oLLM. Three models consistently predict CR for oLLM will approach organic search but not reach parity within 12 months. The linear time model shows exponential growth due to the logit transformation of probabilities.

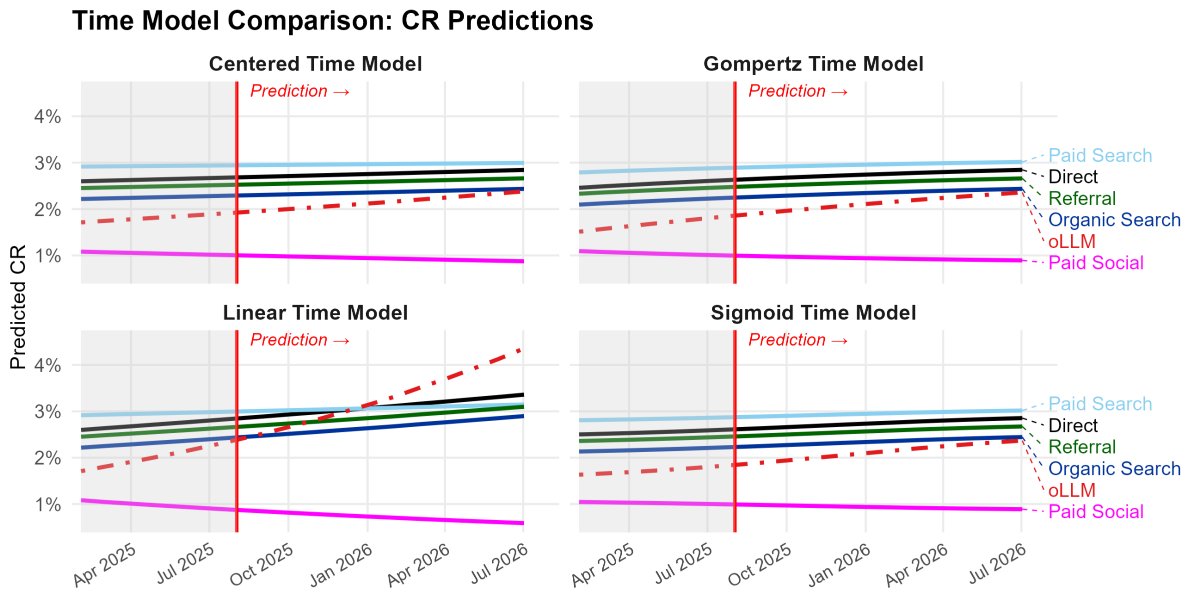


Figure 18: Prediction of CR development per channel

Figure 19 shows substantial downward trend for oLLM AOV, which counteracts the positive development in CR and raises the question, whether the CR improvement is only an artifact of lower AOVs.

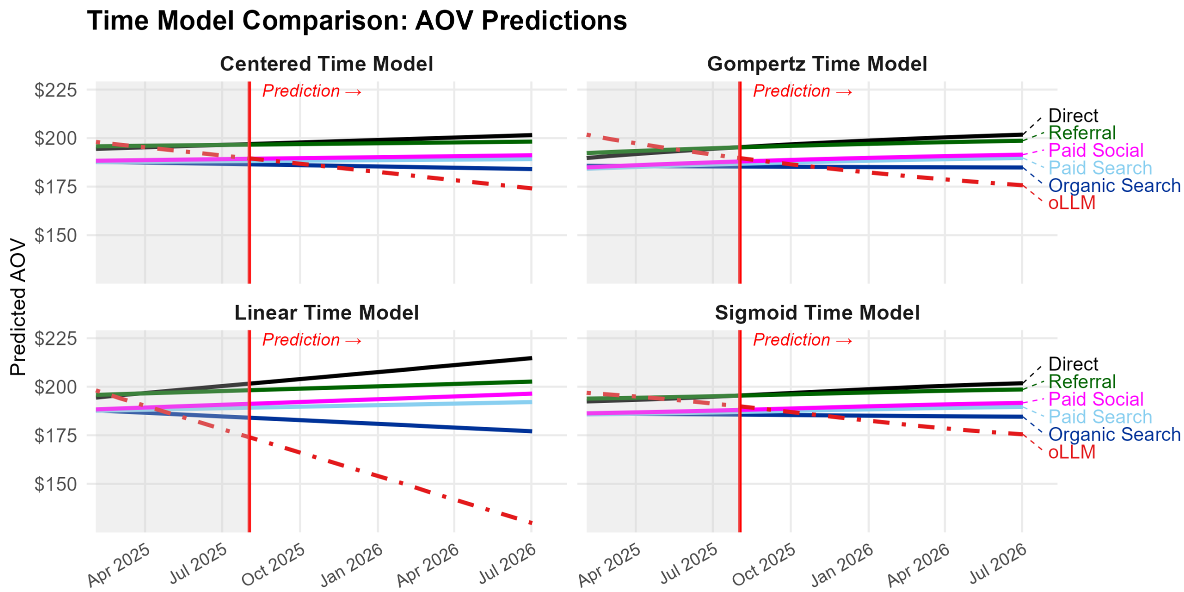


Figure 19: Prediction of AOV development per channel

For RPS, we find oLLM gradually produces more valuable sessions. Similar to CR predictions, the linear model is more optimistic about oLLM outcomes, while the other three models consistently predict oLLM will not converge with the next best channel, organic search, within 12 months (Figure 20).

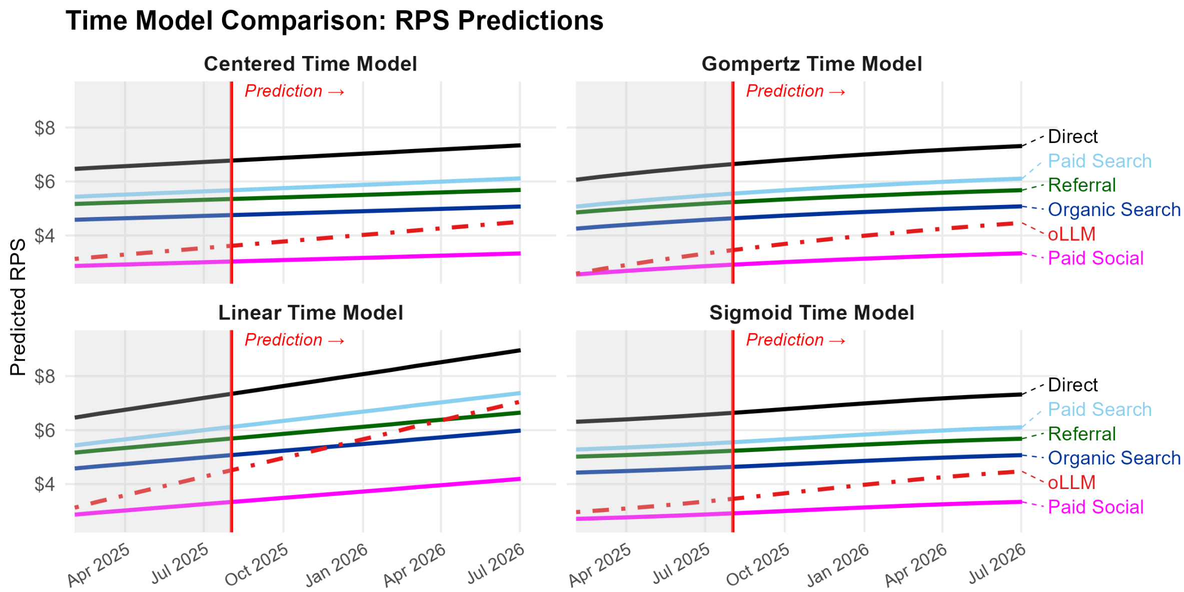


Figure 20: Prediction of RPS development per channel

6 Cross-Website Variation in oLLM Outcomes

The preceding analyses document oLLM's lower CR, AOV, and RPS relative to all traditional channels except paid social, as well as changes over time. Multiple mechanisms could explain these patterns, ranging from trust deficits and privacy concerns to differences in search intent and funnel position. We focus on two mechanisms grounded in prior research and amenable to empirical exploration with our data.

First, LLM recommendations may offer limited usefulness due to inaccuracies or interface friction (Search Engine Land 2025, Nguyen et al. 2022). Second, consumers are still developing proficiency in using LLMs (Chatterji et al. 2025, Foroughi et al. 2025, Yang et al. 2025). To investigate these conjectures, we exploit cross-sectional variation across[4] the 973 websites in our dataset. These analyses are descriptive rather than causal: they offer suggestive evidence on potential mechanisms but remain vulnerable to alternative explanations and should not be interpreted as definitive proof.

[4]Our data does not allow us to segment by product or customer characteristics *and* channels simultaneously. We therefore analyze differences across websites rather than within websites.

6.1 Product Complexity

We examine how oLLM outcomes vary with product complexity, under the assumption that LLMs are more useful when purchases require substantial information synthesis and deliberation. For simple products with straightforward decisions, conversational assistance may add little incremental value while introducing friction.

We proxy product complexity using website categories classified by three LLMs (Claude, Gemini, ChatGPT). Heavy industry and engineering, business and consumer services, vehicles, law and government, finance, and jobs and career are categorized as high-complexity categories, while reference materials, news and media, adult, sports, and arts and entertainment are categorized as low-complexity (see Table A.2 in the Online Appendix for complete rankings). Websites in high-complexity categories exhibit 4.6 times higher oLLM traffic shares on average, consistent with our usefulness assumption and classification.

Across all three financial metrics, oLLM outcomes are more favorable on high-complexity websites (see Online Appendix D for detailed results). When product complexity is high, oLLM CR exceeds paid social, referral, email, organic search, and direct channels (Figure 21); AOV differences relative to other channels narrow (Figure 22); and RPS moves closer to organic and paid search levels (Figure 23).

Taken together, these patterns are consistent with (currently) limited usefulness for simple products as one potential mechanism behind oLLM's weaker financial metrics in Section 3. However, this interpretation does not align with the temporal pattern in Section 5, which shows decreases in AOV.

A caveat is that, although all conversions in our data are financial transactions, what constitutes a conversion can still vary across product categories (see Table A.2 in the Online Appendix), so the complexity split may partly reflect these category-level differences in conversion type rather than differences in LLM assistance per se.

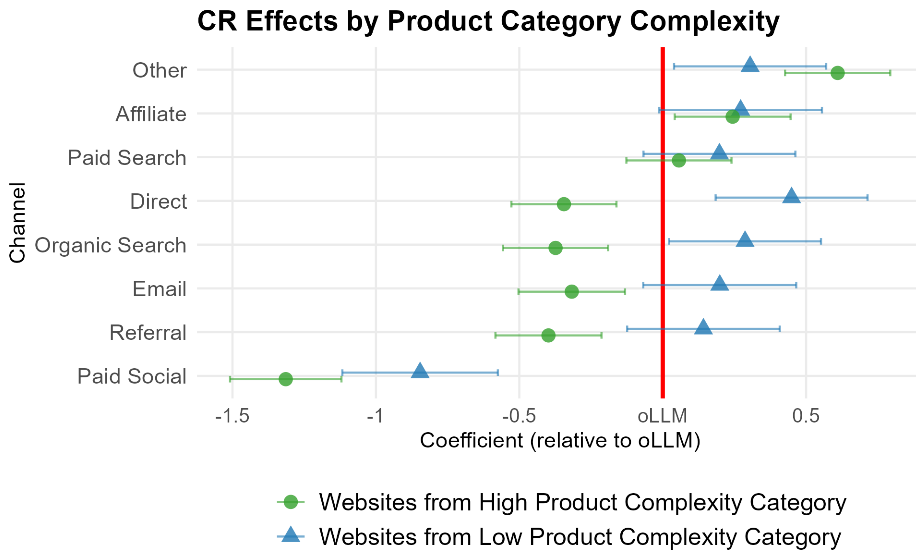


Figure 21: CR by channel (log-odds), split by websites' category complexity

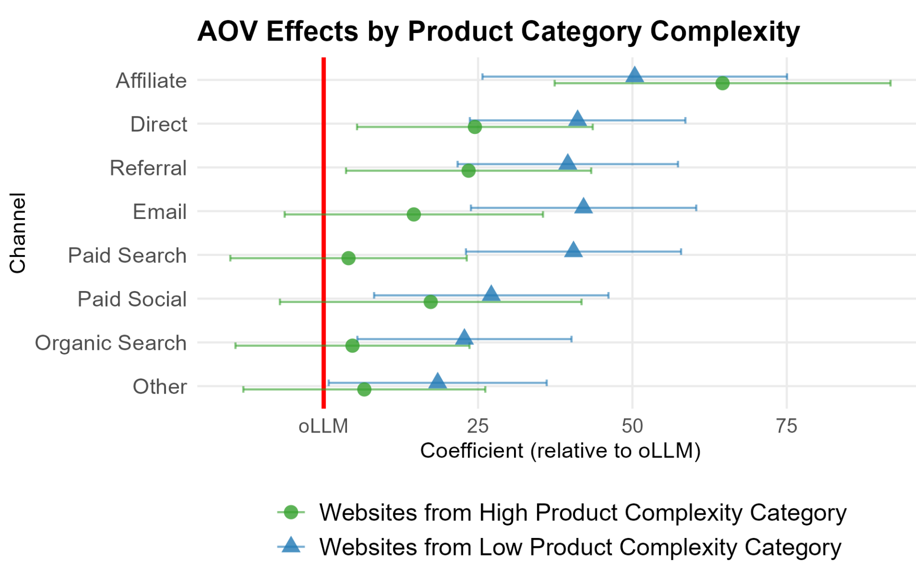


Figure 22: AOV by channel, split by websites' category complexity

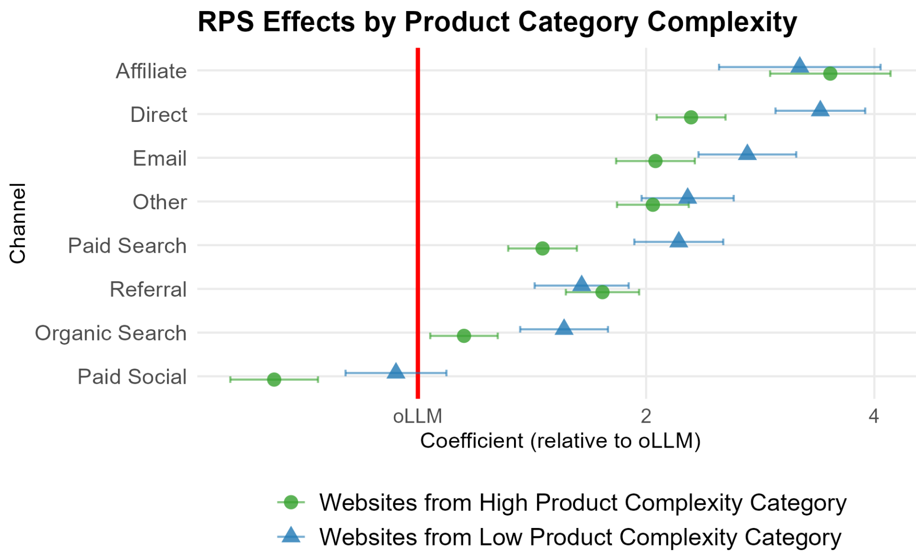
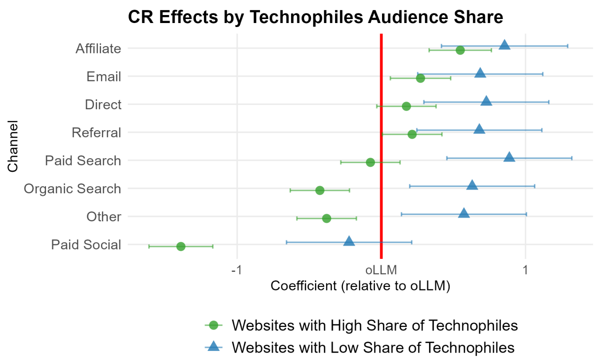


Figure 23: RPS by channel, split by websites' category complexity

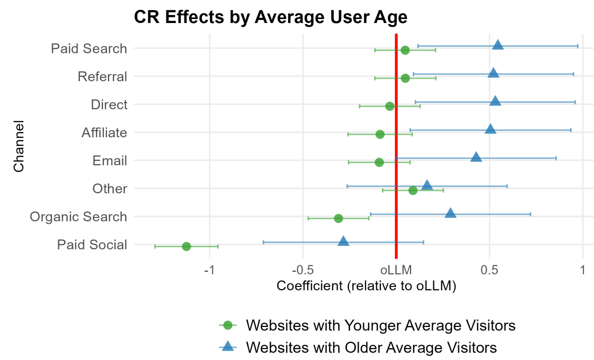
6.2 LLM Proficiency

We assess whether consumer LLM proficiency moderates channel outcomes using two website-level proxies: the share of "technophiles" among visitors (from Google's affinity segments) and average visitor age. Websites in the top technophile quartile exhibit 3.8 times greater oLLM traffic share than those in the bottom quartile. Similarly, websites with younger visitors show 5.5 times greater oLLM traffic share, supporting our proficiency proxies.

Figure 24 shows that CR gaps between oLLM and traditional channels are substantially smaller when LLM proficiency is high; oLLM CR even exceeds organic search in both analyses. In contrast, Figure 25 indicates that higher proficiency is associated with larger AOV gaps, meaning oLLM generates lower order values than traditional channels. Although the technophile split yields relatively imprecise AOV estimates, the age split produces clearly significant differences. Finally, Figure 26 shows mixed RPS patterns: oLLM gains on paid and organic search among high-technophile websites, and shows improved positioning relative to direct, affiliate, and other among younger-visitor websites (regressions in Online Appendix D).

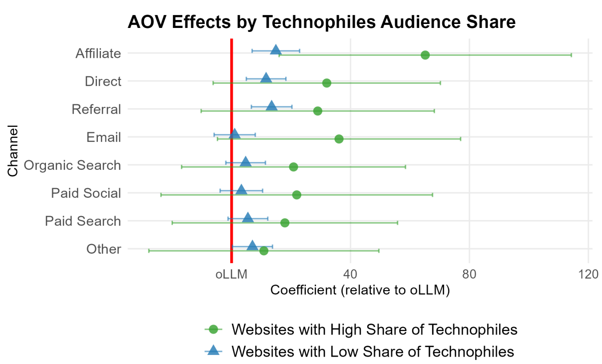


(a) Technophile split

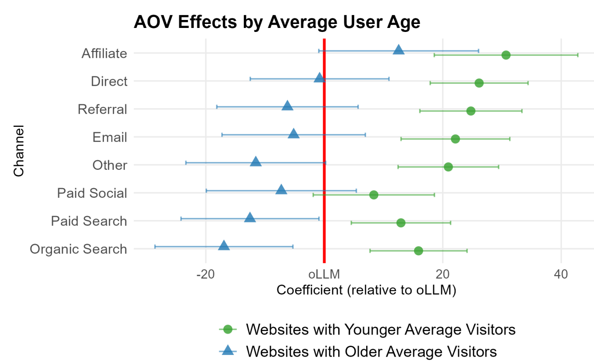


(b) Age split

Figure 24: CR by channel (log-odds), split by consumer LLM proficiency proxies

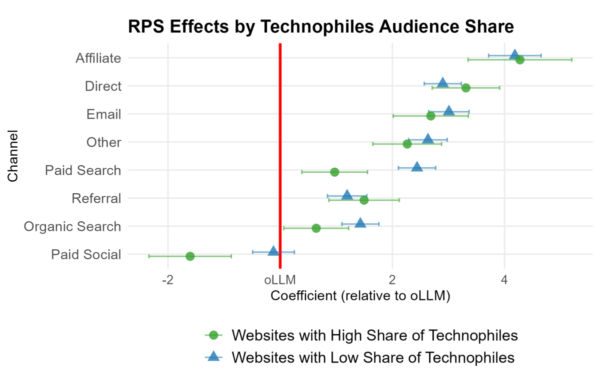


(a) Technophile split

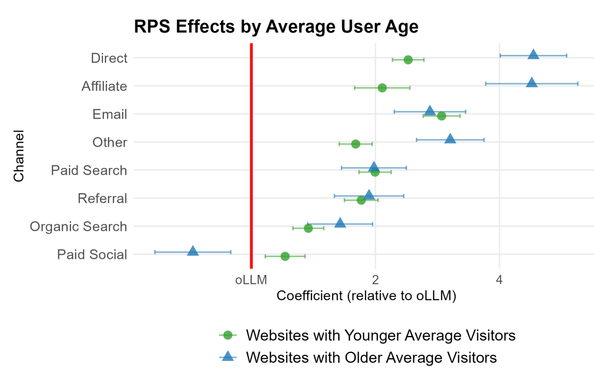


(b) Age split

Figure 25: AOV by channel, split by consumer LLM proficiency proxies



(a) Technophile split



(b) Age split

Figure 26: RPS by channel, split by consumer LLM proficiency proxies

These patterns mirror Section 5, where rising CR accompanies declining AOV,[5] yielding only moderate RPS improvements. The parallel cross-sectional and temporal findings indicate that growing consumer LLM proficiency might contribute to oLLM's development over time.

7 Key Insights and Limitations

Our research examines organic LLM traffic (oLLM) relative to traditional digital channels using 12 months of data from 973 e-commerce websites with over 50,000 oLLM transactions. Two key patterns emerge.

Current positioning. One year after launch, oLLM exhibits a higher conversion rate (CR) and revenue per session (RPS) than paid social, but lower CR and RPS than all other traditional channels. Engagement metrics show comparatively favorable bounce rates (BR) for oLLM, but fewer page views (PV) and intermediate session duration (SD). Extensive robustness checks and the geographic and category diversity of the dataset support the stability and external relevance of these findings. Online Appendix E discusses why results may differ from industry reports.

Gaps between oLLM and traditional channels are more pronounced in low-complexity product categories. In contrast, oLLM shows more favorable relative outcomes and approaches the RPS of organic search for high-complexity categories, which may benefit more from the extensive context, superior synthesis, and adaptive conversational interface LLMs can offer. While exploratory, these findings indicate that oLLM's comparatively weak financial metrics may reflect limitations in how users currently use and value LLMs as a shopping tool.

Temporal trajectory. Over the first twelve months, CR and RPS increased, while average order value (AOV) declined. Projections suggest that the CR gap relative to tra-

[5] The declining order values could reflect increased purchase targeting that reduces exploratory browsing or heightened price consciousness from LLM transparency about alternatives.

ditional channels will continue to narrow, offset by a widening AOV gap. Thus, RPS is expected to improve only moderately, reaching parity with organic search—the next-ranked channel—only under the most aggressive forecasting scenario.

This trade-off—higher conversion but lower order value—is also evident when segmenting websites by visitor LLM proficiency. Websites serving more LLM-proficient consumers display the same combination of elevated CR and reduced AOV. The robustness of this dual pattern across temporal and cross-sectional analyses points to growing consumer LLM proficiency as a likely mechanism behind the observed evolution of oLLM.

Implications

For retailers. oLLM currently offers the greatest value for retailers operating in high-complexity categories where consumers demand extensive comparison and guidance. In these settings, oLLM referrals combine comparatively strong CR and RPS with favorable BR, suggesting that they can serve as a high-intent complement to organic and paid search, even if volumes remain small. For retailers focused on low-complexity, routine purchases, current low traffic shares and weaker CR and RPS imply limited short-term gains from channel optimization alone.

For LLM platforms. Our analyses suggest that lower usefulness in low-complexity product categories aligns with the observed gap between oLLM and traditional channels. LLM platforms have launched several initiatives since the end of our observation period targeting this limitation, including instant checkout and agentic shopping. Early research indicates that agentic shopping tools may prove particularly useful for routine purchases, given a substantially higher share of observed shopping-related queries than with general-purpose LLMs (Yang et al. 2025). Such platform-based agents will compete with retailer-embedded alternatives like Amazon's Rufus, which benefit from keeping users in familiar on-site interfaces while providing optional conversational support.

Further, recent announcements on the introduction of paid advertising formats within ChatGPT introduce strategic considerations that go beyond our organic-traffic data. While oLLM is unpaid and therefore attractive for companies concerned about Return on Advertising Spend (ROAS), paid LLM traffic (pLLM) has the potential to redefine LLMs' role in the channel ecosystem (Hermann et al. 2025). Depending on implementation, sponsored placements could reduce friction for simple purchases through one-tap offers or shift the balance between organic guidance and promotional content in ways that negatively affect usefulness and trust and thus also oLLM's positioning.

Limitations and future research

All analyses in this paper are descriptive rather than causal. As such, the analysis of oLLM metrics is vulnerable to confounding factors, including systematic differences across channels' users and usage situations.

The time-trend analysis presented in Section 5 can only provide an indication of future channel outcomes. LLM platforms, retailer strategies, and consumer behaviors evolve rapidly. While our twelve-month observation period captures the initial maturation phase, the documented patterns may not persist. Disruptive developments, such as major interface redesigns, widely adopted checkout features, or paid advertising rollouts, could fundamentally alter these trajectories.

Similarly, the exploratory heterogeneity analyses in Section 6 should only be seen as initial indications of potential mechanisms. The regression framework controls for website, device, and time effects, but the analysis remains descriptive, and the product complexity and proficiency measures are constructed at the website level rather than at the level of individual products or consumers. Understanding whether the observed patterns reflect channel-specific user behavior, LLM-induced behavioral changes, or unobserved website characteristics requires further investigation.

Finally, our analyses rely on last-click attribution, which underestimates oLLM's contribution to the purchase funnel (Berman 2018, Li and Kannan 2014, Li et al. 2016). Future work using multi-touch attribution or experiments on individual websites could clarify the role of oLLM in the upper funnel and its interaction with traditional channels and supplement our broad study with more granular, in-depth insights.

Beyond these limitations, our study raises new questions on the emerging role of oLLM in the digital channel mix—including whether treating oLLM as merely another channel understates its unique ability to shape consumer decision processes through its conversational interface. In Online Appendix F, we outline promising directions for future research on how LLMs may reshape digital commerce, channel strategy, and consumer welfare.

Funding, Author Affiliations and Competing Interests

Funding. This research was supported by Grips Intelligence through the provision of proprietary data and computing resources. No direct financial support was provided to the authors. No additional third-party funding was received.

Author Affiliations and Competing Interests. The first author is employed by Grips Intelligence, a market intelligence company serving multiple clients, including various technology platforms. Grips Intelligence's business model depends on delivering accurate, unbiased market intelligence; it has no vested interest in making any particular platform, including Google, OpenAI, or others analyzed in this study, appear more or less favorable. The company's involvement was limited to: (1) providing access to data, (2) covering computing costs for data processing, and (3) offering technical guidance on data collection and segmentation methodology. Grips Intelligence did not commission the study. The company had no role in formulating the research questions, designing the study, selecting analytical methods, interpreting results, or drafting the manuscript, and it did not approve the manuscript prior to submission. The authors declare no other financial or non-financial competing interests.

References

Adobe (2025) Adobe Analytics: Traffic to U.S. retail websites from generative AI sources jumps 1,200%. Adobe Blog (March 17), https://blog.adobe.com/en/publish/2025/03/17/adobe-analytics-traffic-to-us-retail-websites-from-generative-ai-sources-jumps-1200-percent.

Ahrefs (2025) 63% of websites receive AI traffic (new study of 3,000 sites). Ahrefs Blog (February 6), https://ahrefs.com/blog/ai-traffic-study/.

Al-Hasan TM, Sayed AN, Bensaali F, Himeur Y, Varlamis I, Dimitrakopoulos G (2024) From traditional recommender systems to GPT-based chatbots: A survey of recent developments and future directions. *Big Data and Cognitive Computing* 8(4):36, http://dx.doi.org/10.3390/bdcc8040036.

Anderson SP, Renault R (2006) Advertising content. *American Economic Review* 96(1):93–113, http://dx.doi.org/10.1257/000282806776157632.

Aridor G, Che YK, Hollenbeck B, Kaiser M, McCarthy D (2025) Evaluating the impact of privacy regulation on e-commerce firms: Evidence from Apple's app tracking transparency. *Management Science* http://dx.doi.org/10.1287/mnsc.2024.06600, ePub ahead of print November 20, 2025.

Berman R (2018) Beyond the last touch: Attribution in online advertising. *Marketing Science* 37(5):771–792, http://dx.doi.org/10.1287/mksc.2018.1104.

Bleier A, Eisenbeiss M (2015) Personalized online advertising effectiveness: The interplay of what, when, and where. *Marketing Science* 34(5):669–688, http://dx.doi.org/10.1287/mksc.2015.0930.

Brown JR, Goolsbee A (2002) Does the internet make markets more competitive? evidence from the life insurance industry. *Journal of Political Economy* 110(3):481–507, http://dx.doi.org/10.1086/339714.

Chatterji A, Cunningham T, Deming DJ, Hitzig Z, Ong C, Shan CY, Wadman K (2025) How people use ChatGPT. Working Paper 34255, National Bureau of Economic Research, http://dx.doi.org/10.3386/w34255.

Chen M, Wang X, Chen K, Koudas N (2025) Generative engine optimization: How to dominate AI search. *arXiv preprint arXiv:2509.08919* .

de Haan E, Wiesel T, Pauwels K (2016) The effectiveness of different forms of online advertising for purchase conversion in a multiple-channel attribution framework. *International Journal of Research in Marketing* 33(3):491–507, http://dx.doi.org/10.1016/j.ijresmar.2015.12.001.

Dinerstein M, Einav L, Levin J, Sundaresan N (2018) Consumer price search and platform design in internet commerce. *American Economic Review* 108(7):1820–1859, http://dx.doi.org/10.1257/aer.20171218.

Ellison G, Ellison SF (2009) Search, obfuscation, and price elasticities on the internet. *Econometrica* 77(2):427–452, http://dx.doi.org/10.3982/ECTA5708.

eMarketer (2024) Retail media ad spending forecast h1 2024. Technical report, Insider Intelligence / eMarketer, https://liveramp.com/wp-content/uploads/2024/07/LiveRamp_Report_H1-2024-Retail-Media-Ad-Spending.pdf.

Foroughi B, Iranmanesh M, Yadegaridehkordi E, Wen J, Ghobakhloo M, Senali M, Annamalai N (2025) Factors affecting the use of ChatGPT for obtaining shopping information. *International Journal of Consumer Studies* 49(1):e70008, http://dx.doi.org/10.1111/ijcs.70008.

Gholami S, Firullo C, Cheyre C, Acquisti A (2026) Beyond search: LLM adoption and web traffic concentration. *SSRN* http://dx.doi.org/10.2139/ssrn.6238578.

Ghose A, Yang S (2009) An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management Science* 55(10):1605–1622, http://dx.doi.org/10.1287/mnsc.1090.1054.

Google (2024) Google Analytics 4 dimensions and metrics. https://support.google.com/analytics/table/13948007.

Hermann E, Puntoni S, Schweidel DA (2025) Conversational AI: The next frontier of digital platform monetization. *SSRN* http://dx.doi.org/10.2139/ssrn.5634270.

Huang MH, Rust RT (2025) The GenAI future of consumer research. *Journal of Consumer Research* 52(1):4–17, http://dx.doi.org/10.1093/jcr/ucaf013.

Jannach D (2022) Evaluating conversational recommender systems: A landscape of research. *Artificial Intelligence Review* http://dx.doi.org/10.1007/s10462-022-10229-x.

Johnson GA, Shriver SK, Du S (2020) Consumer Privacy Choice in Online Advertising: Who Opts Out and at What Cost to Industry? *Marketing Science* 39(1):33–51, http://dx.doi.org/10.1287/mksc.2019.1198.

Li H, Kannan PK (2014) Attributing conversions in a multichannel online marketing environment: An empirical model and a field experiment. *Journal of Marketing Research* 51(1):40–56, http://dx.doi.org/10.1509/jmr.13.0050.

Li H, Kannan PK, Viswanathan S, Pani A (2016) Attribution strategies and return on keyword investment in paid search advertising. *Marketing Science* 35(6):831–848, http://dx.doi.org/10.1287/mksc.2016.0987.

Luo X, Tong S, Fang Z, Qu Z (2019) Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science* 38(6):937–947, http://dx.doi.org/10.1287/mksc.2019.1192.

Lynch JG, Ariely D (2000) Wine online: Search costs affect competition on price, quality, and distribution. *Marketing Science* 19(1):83–103, http://dx.doi.org/10.1287/mksc.19.1.83.15183.

Manchanda P, Dubé JP, Goh KY, Chintagunta PK (2006) The effect of banner advertising on internet purchasing. *Journal of Marketing Research* 43(1):98–108, http://dx.doi.org/10.1509/jmkr.43.1.98.

Miller KM, Skiera B (2024) Economic consequences of online tracking restrictions: Evidence from cookies. *International Journal of Research in Marketing* 41(2):241–264, http://dx.doi.org/10.1016/j.ijresmar.2023.10.001.

Nguyen QN, Sidorova A, Torres R (2022) User interactions with chatbot interfaces vs. menu-based interfaces: An empirical study. *Computers in Human Behavior* 128:107093, http://dx.doi.org/10.1016/j.chb.2021.107093.

OpenAI (2025) Improved shopping results from ChatGPT search. OpenAI Help Center, https://help.openai.com/en/articles/11128490-improved-shopping-results-from-chatgpt-search.

Padilla N, Lam HT, Lambrecht A, Hollenbeck B (2025) The impact of LLM adoption on online user behavior. *SSRN* http://dx.doi.org/10.2139/ssrn.5393256.

Reisenbichler M, Reutterer T, Schweidel DA, Dan D (2022) Frontiers: Supporting content marketing with natural language generation. *Marketing Science* 41(3):441–452, http://dx.doi.org/10.1287/mksc.2022.1354.

Reuters (2025) OpenAI rolls out new shopping features with ChatGPT search update. https://www.reuters.com/business/media-telecom/openai-rolls-out-new-shopping-features-with-chatgpt-search-update-2025-04-28/.

Rutz OJ, Bucklin RE (2011) From generic to branded: A model of spillover in paid search advertising. *Journal of Marketing Research* 48(1):87–102, http://dx.doi.org/10.1509/jmkr.48.1.87.

SALT (2025) Do users really show higher intent when they click through from an LLM to a website? Accessed May 12, 2025, https://salt.agency/blog/do-users-really-show-higher-intent-when-they-click-through-from-an-llm-to-a-website/.

Salvi F, Horta Ribeiro M, Gallotti R, West R (2025) On the conversational persuasiveness of GPT-4. *Nature Human Behaviour* http://dx.doi.org/10.1038/s41562-025-02194-6.

SE Ranking (2025) Which platforms lead in AI traffic in 2025? ChatGPT, perplexity, gemini & more. SE Ranking Blog (August 27), https://seranking.com/blog/ai-traffic-research-study/.

Search Engine Land (2025) ChatGPT search vs. Google: A deep dive analysis of 62 queries. https://searchengineland.com/chatgpt-search-vs-google-analysis-449676, accessed: 2025-12-02.

Seer Interactive (2025) Case study: 6 learnings, 1 site — how traffic from ChatGPT converts. Accessed June 3, 2025, https://www.seerinteractive.com/insights/case-study-6-learnings-about-how-traffic-from-chatgpt-converts.

Semrush (2025) We studied the impact of AI search on SEO traffic: Our findings. Accessed July 21, 2025, https://www.semrush.com/blog/ai-search-seo-traffic-study/.

SimilarWeb (2024) The 2024 marketing benchmark report. Report (PDF), Accessed December 2, 2025, https://similarweb.vn/wp-content/uploads/sites/2/2024/05/The-2024-Marketing-Benchmark-Report-2.pdf.

Similarweb (2025) Industry category. Similarweb Help Center, Accessed December 2, 2025, https://support.similarweb.com/hc/en-us/articles/360000810469-Industry-Category.

Soviero B, Kuhn D, Salle A, Moreira VP (2024) ChatGPT goes shopping: LLMs can predict relevance in ecommerce search. *Advances in Information Retrieval: Proceedings of ECIR 2024*, 3–11, Lecture Notes in Computer Science (Springer), http://dx.doi.org/10.1007/978-3-031-56066-8_12.

Steyvers M, Kumar A (2024) Three challenges for AI-assisted decision-making. *Perspectives on Psychological Science* 19(5):722–734, http://dx.doi.org/10.1177/17456916231181102.

The Economist (2023) Is google's 20-year search dominance about to end? https://www.economist.com/business/2023/02/08/is-googles-20-year-search-dominance-about-to-end.

ThoughtMetric (2025) The conversion advantage: Why ChatGPT traffic outperforms Google Search for e-commerce. Accessed June 17, 2025, https://thoughtmetric.io/blog/the-conversion-advantage-why-chatgpt-traffic-outperforms-google-search-for-e-commerce.

Ursu RM, Zhang Q, Honka E (2022) Search gaps and consumer fatigue. *Marketing Science* 42(1):110–136, http://dx.doi.org/10.1287/mksc.2022.1359.

Wies S, Bleier A, Edeling A (2023) Finding goldilocks influencers: How follower count drives social media engagement. *Journal of Marketing* 87(3):383–405, http://dx.doi.org/10.1177/00222429221125131.

Wroe S (2024) How Walmart became an advertising powerhouse. *Financial Times* https://www.ft.com/content/1e2a5338-f1ae-41ee-bc08-4c86f2eb5a58.

Yang J, Yonack N, Zyskowski K, Yarats D, Ho J, Ma J (2025) The adoption and usage of AI agents: Early evidence from perplexity. *arXiv preprint* http://dx.doi.org/10.48550/arXiv.2512.07828.

Yu P, Xu Z, Wang J, Xu X (2025) The application of large language models in recommendation systems. https://arxiv.org/abs/2501.02178.

Yu Y, Ping W, Liu Z, Wang B, You J, Zhang C, Shoeybi M, Catanzaro B (2024) Rankrag: Unifying context ranking with retrieval-augmented generation in LLMs. *Advances in Neural Information Processing Systems (NeurIPS 2024)*, https://proceedings.neurips.cc/paper_files/paper/2024/file/db93ccb6cf392f352570dd5af0a223d3-Paper-Conference.pdf.

Online Appendix A Additional Data and Variable Definitions

A.1 Category Volume Details

Table A.1 provides a detailed overview of the size and relevance of oLLM across all 24 top-level categories in our dataset. The table shows sessions and revenue for all channels combined, as well as the percentage share attributable to oLLM traffic.

Table A.1: Category Volume Summary Statistics (February - August 2025)

	All Channels		oLLM Share	
Category	Sessions (M)	Revenue ($M)	% Sessions	% Revenue
E-commerce and Shopping	1,201.31	1,915.33	0.05	0.03
Lifestyle (Fashion, Cosmetics)	1,127.32	2,026.85	0.02	0.02
Business, Cons. Services	617.84	780.24	0.10	0.10
Home and Garden	531.37	1,410.57	0.03	0.03
Consumer Electronics	465.61	758.11	0.20	0.12
Vehicles	441.12	300.18	0.03	0.05
Science and Education	249.85	197.41	0.24	0.12
Arts and Entertainment	227.42	797.40	0.07	0.14
Finance	227.41	223.95	0.09	0.06
Travel and Tourism	184.67	1,279.32	0.10	0.06
Games	163.22	151.66	0.05	0.02
Health	163.10	350.79	0.04	0.04
Sports	136.05	417.42	0.02	0.02
Food and Drink	104.21	328.88	0.02	0.02
Pets and Animals	66.14	209.35	0.02	0.04
Heavy Industry	41.32	108.46	0.03	0.07
Law and Government	23.01	58.53	0.14	0.05
Gambling	8.50	85.01	>0.00	>0.00
News and Media	7.84	1.60	0.12	0.24
Hobbies and Leisure	7.44	34.66	0.06	0.07
Community and Society	6.91	31.98	0.19	0.10
Adult	6.39	15.96	0.01	0.01
Jobs and Career	2.09	1.07	2.06	2.22
Reference Materials	0.67	0.60	0.02	0.04

Note: Sessions and Revenue shown in millions for all channels combined.
oLLM Share shows percentage of total volume attributed to oLLM channel.

Table A.2: Product Category Definitions and Conversion Logic

Category	Description	Example of Conversion	Example Domains in Category
Reference Materials	Information lookup, maps, encyclopedias, dictionaries	• Premium subscription • Ad-free app upgrade	wikipedia.org, dictionary.com, google.com/maps, britannica.com
News and Media	News content, journalism, current events	• News subscription • Article unlock fee	cnn.com, bbc.com, nytimes.com, theguardian.com, reuters.com
Adult	Content intended for adult audiences only (18+)	• Monthly fan subscription • Pay-Per-View video	pornhub.com, xvideos.com, xnxx.com, xhamster.com, onlyfans.com
Sports	Sports news, team info, stats, athletic content	• News subscription • Pay-Per-View match pass	espn.com, cricbuzz.com, marca.com, espncricinfo.com
Arts and Entertainment	Streaming services, music, visual/performing arts	• Monthly streaming fee • Digital movie rental	youtube.com, netflix.com, spotify.com, twitch.tv, imdb.com
Community and Society	Social connections, dating, faith, community orgs	• Premium dating features • Virtual awards	facebook.com, reddit.com, match.com, tinder.com, meetup.com
Food and Drink	Cooking recipes, grocery delivery, restaurants	• Food delivery order • Meal kit box purchase	allrecipes.com, ubereats.com, doordash.com, foodnetwork.com
Games	Gaming platforms, info, communities, esports	• Game download purchase • Streamer subscription	twitch.tv, roblox.com, steampowered.com, ign.com
Hobbies and Leisure	Recreational activities, creative pursuits, photography	• Digital pattern download • Stock photo credits	pinterest.com, etsy.com, flickr.com, instructables.com, alltrails.com
Lifestyle	Fashion retailers, apparel, beauty, personal care	• Apparel purchase • Gift set purchase	shein.com, nike.com, macys.com, zara.com, sephora.com
Pets and Animals	Pet care, supplies, animal content, communities	• Pet food purchase • Grooming fee	chewy.com, petco.com, petsmart.com, akc.org, petmd.com
Gambling	Online casinos, sports betting, poker, lottery	• Wager placement • Lottery ticket	bet365.com, draftkings.com, fanduel.com, pokerstars.com
E-commerce and Shopping	Marketplaces, retailers, price comparison	• Retail goods purchase • Gift card purchase	amazon.com, temu.com, aliexpress.com, ebay.com, walmart.com
Travel and Tourism	Travel planning, booking platforms, accommodation	• Hotel booking • Flight ticket	booking.com, tripadvisor.com, agoda.com, trip.com, airbnb.com
Science and Education	Educational resources, academic institutions	• Course certificate fee • Digital textbook	wikipedia.org, coursera.org, khanacademy.org, duolingo.com
Home and Garden	Home improvement, furniture, gardening, DIY	• Home decor purchase • Furniture purchase	ikea.com, homedepot.com, lowes.com, wayfair.com, harborfreight.com
Consumer Electronics	Devices, gadgets, smartphones, personal tech	• Smartphone purchase • Accessory purchase	samsung.com, apple.com, bestbuy.com, huawei.com, hp.com
Health	Healthcare, wellness, medical info, pharma	• Telehealth copay • Home test kit	webmd.com, mayoclinic.org, nih.gov, cvs.com, walgreens.com
Jobs and Career	Job search, employment websites, HR platforms	• Premium profile fee • Salary report	indeed.com, linkedin.com, glassdoor.com, ziprecruiter.com
Finance	Banking, investing, insurance, payment platforms	• Insurance premium • Wire/Transfer fee	paypal.com, chase.com, bankofamerica.com, coinbase.com
Law and Government	Government services, legal resources, portals	• Passport fee • Document copy fee	usa.gov, irs.gov, gov.uk, uscis.gov, whitehouse.gov
Vehicles	Automotive marketplaces, manufacturers, car info	• Vehicle history report • Accessory purchase	cars.com, carmax.com, autotrader.com, tesla.com, ford.com
Business and Consumer Services	B2B services, shipping, real estate, professional	• Premium subscription • Transaction fee	indeed.com, linkedin.com, paypal.com, fedex.com, zillow.com
Heavy Industry and Engineering	Manufacturing, construction, agriculture, energy	• Replacement part • Safety webinar fee	caterpillar.com, ge.com, boeing.com, deere.com, siemens.com

Note: Categories and sample domains according to SimilarWeb. Domains and conversion events are examples only. Categories sorted from "most simple" to "most complex" as rated by three LLMs (Claude, Gemini, ChatGPT).

A.2 Continent and Country Volume Details

A.2.1 Continent-Level Coverage

Table A.3 summarizes traffic volume across continents. The dataset covers global e-commerce activity, with the majority of traffic stemming from the Americas and Europe.

Table A.3: Continent Volume Summary Statistics (February - August 2025)

	All Channels		oLLM Share	
Continent	Sessions (M)	Revenue ($M)	% Sessions	% Revenue
Americas	2,712.05	4,202.19	0.05	0.04
Europe	2,132.90	5,197.74	0.04	0.04
Asia	856.45	779.77	0.10	0.04
Oceania	276.74	922.58	0.04	0.02
Africa	56.25	27.76	0.20	0.04

Note: oLLM share shows percentage of total volume attributed to ChatGPT Referrals. Local currency is converted with daily average conversion rates to US dollars.

A.2.2 Country-Level Coverage

Our dataset provides broad geographic coverage at the country level. Table A.4 presents detailed statistics for all countries contributing at least 10 million sessions to our dataset. A total of 49 nations meet this threshold, demonstrating the global reach of our data.

Table A.4: Country-Level Summary 6 Months

	All Channels		oLLM Share	
Country	Sessions (M)	Revenue ($M)	% Sessions	% Revenue
United States	1,800.26	3,237.66	0.04	0.06
Brazil	696.81	757.23	0.07	0.04
United Kingdom	457.15	1,051.88	0.04	0.03
France	398.80	844.59	0.05	0.05
India	247.27	134.66	0.11	0.03
Australia	226.82	857.65	0.06	0.04
Germany	207.46	742.43	0.07	0.05
Italy	190.39	191.42	0.04	0.08
Netherlands	181.34	669.00	0.07	0.08
Thailand	121.75	114.65	0.04	0.04
Poland	120.85	307.97	0.07	0.04
South Korea	101.73	86.92	0.06	0.04
Canada	95.65	143.60	0.12	0.08
Spain	81.88	164.12	0.11	0.12
Saudi Arabia	60.34	84.04	0.05	0.02
Sweden	57.37	144.53	0.04	0.06
New Zealand	50.53	66.32	0.05	0.03
Japan	46.96	53.87	0.14	0.15
Colombia	46.24	17.39	0.07	0.04
Greece	40.32	46.37	0.05	0.07
Belgium	40.14	105.78	0.08	0.15
Vietnam	39.63	16.67	0.15	0.07
Turkey	37.09	28.59	0.11	0.08
Finland	35.58	45.65	0.03	0.04
Israel	30.16	116.38	0.08	0.04
Switzerland	29.53	98.04	0.09	0.08
Mexico	27.24	19.48	0.16	0.17
Portugal	27.11	109.27	0.11	0.05
Ireland	27.10	37.65	0.06	0.08
Russia	25.72	31.86	0.03	0.01
United Arab Emirates	24.98	36.47	0.18	0.06
Denmark	24.92	98.21	0.06	0.06
Romania	23.72	47.49	0.11	0.13
Austria	22.62	108.63	0.07	0.07
China	21.37	21.54	0.05	0.01
Malaysia	20.98	4.53	0.17	0.07
Ukraine	20.44	17.79	0.13	0.21
Hungary	20.32	69.02	0.06	0.06
Philippines	17.40	11.20	0.25	0.35
Norway	16.75	33.29	0.06	0.10
Indonesia	16.58	27.13	0.28	0.02
Czechia	15.88	40.24	0.10	0.13
Slovakia	15.73	35.72	0.05	0.08
Egypt	14.02	5.86	0.21	0.05
South Africa	13.04	6.61	0.23	0.13
Pakistan	12.94	4.00	0.32	0.27
Bulgaria	12.32	16.48	0.08	0.07
Croatia	11.78	96.77	0.06	0.03
Chile	10.17	4.10	0.20	0.13
Taiwan	8.49	8.07	0.18	0.15

Note: Omitted 141 countries with lower coverage.

A.3 Complete Variable Definitions

Table A.5 provides comprehensive definitions for all dependent and independent variables used in our regression models. This includes focal financial metrics (CR, AOV, RPS), secondary engagement metrics (BR, SD, PV), and control variables.

Table A.5: Variable Definitions

Variable Type	**Variable Name**	**Definition**
Dependent Variables: Focal	Conversion Rate (CR)	Most widely used financial metric in e-commerce, defined as (Number of Transactions/Number of Sessions) per time unit, if (Number of Sessions > 0)
	Average Order Value (AOV)	Captures size of the typical shopping basket in USD, defined as (Total Revenue/Number of Transactions) per time unit, if (Number of Transactions > 0)
	Revenue Per Session (RPS)	Combines CR and AOV into one metric, defined as (Total Revenue/Number of Sessions) per time unit, if (Number of Sessions > 0)
Dependent Variables: Secondary	Bounce Rate (BR)	Focal engagement metric, measures share of visitors that leave target website without further action, capturing content relevance.
	Session Duration (SD)	Total time spent on websites from entry to exit, measured in seconds. Reflects engagement depth and deliberation time in purchase decisions.
	Page Views (PV)	Measures number of distinct pages viewed during website sessions, indicates relevance of the destination website.
Independent Variable: Focal	Channel	Categorical variable identifying visitor source, with oLLM (i.e., organic referral traffic from ChatGPT) serving as baseline comparison category in regression analysis. Other channels include Affiliate, Direct, Email, Organic Search, Paid Search, Paid Social, Referral, and Other.
Independent Variables: Controls	Website	E-commerce website identifier (website fixed effect) capturing shop-specific characteristics, as well as industry effects.
	Device	Device used by website visitors, accounting for device-specific conversion and engagement patterns.
	Month	Month fixed effects capturing seasonal patterns and temporal trends.

A.4 Complete Technical Model Specifications

This section provides complete technical specifications for all regression models used in our analysis, including notation, data structure, and detailed model equations.

A.4.1 Notation and Data Structure

Reflecting the data structure, we index websites, weeks, channels, device types, and months as shown in Table A.6.

Table A.6: Indices

Symbol	Description
$i \in \{1, \ldots, I\}$	Website
$t \in \{1, \ldots, T\}$	Week
$c(i,t)$	Channel (e.g., oLLM, organic search)
$d(i,t)$	Device (desktop or mobile)
$m(t)$	Calendar month of observation

Table A.7 summarizes notation for derived metrics and outcome variables.

Table A.7: Outcomes

Symbol	Description
n_{it}^{sess}	Number of sessions for i,t
y_{it}^{txn}	Number of transactions for i,t
y_{it}^{rev}	Total revenue for i,t
y_{it}^{bounce}	Number of bounces for i,t
y_{it}^{SD}	Total session duration for i,t (seconds)
y_{it}^{PV}	Total page views for i,t
$\text{CR}_{it} = y_{it}^{\text{txn}}/n_{it}^{\text{sess}}$	Conversion rate for i,t
$\text{AOV}_{it} = y_{it}^{\text{rev}}/y_{it}^{\text{txn}}$	Average order value for i,t (defined for $y_{it}^{\text{txn}} > 0$)
$\text{RPS}_{it} = y_{it}^{\text{rev}}/n_{it}^{\text{sess}}$	Revenue per session for i,t (includes zeros)
$\text{BR}_{it} = y_{it}^{\text{bounce}}/n_{it}^{\text{sess}}$	Bounce rate for i,t
$\text{SD}_{it} = y_{it}^{\text{SD}}/n_{it}^{\text{sess}}$	Average session duration for i,t
$\text{PV}_{it} = y_{it}^{\text{PV}}/n_{it}^{\text{sess}}$	Average pageviews per session for i,t

Note: Outcome variables are only defined for $n_{it}^{\text{sess}} > 0$.

We include several fixed effects, with notation summarized in Table A.8. The list contains a time-trend variable (equal to the count of months since the start of the observation period) that is used in Section 5.

Table A.8: Fixed Effects and Parameters

Symbol	Description
α_c	Channel fixed effects
γ_i	Website fixed effects
δ_d	Device fixed effects
μ_m	Month fixed effects
β_0^X	Intercept for outcome $X \in \{\mathrm{AOV}, \mathrm{RPS}, \mathrm{SD}, \mathrm{PV}\}$
$trend_t$	Linear monthly time trend

A.4.2 Model Specification for CR and BR

By definition, $\mathrm{CR}_{it}, \mathrm{BR}_{it} \in [0, 1]$ for all i, t, as they measure the probability of success (a transaction or a bounce) for a given number of trials (sessions). Both metrics exhibit overdispersion in our data. For CR, the variance of the transaction counts, $\mathrm{Var}(y_{it}^{\mathrm{txn}})$, exceeds that implied by a standard binomial model. Similarly, BR shows variance in bounce counts, $\mathrm{Var}(y_{it}^{\mathrm{bounce}})$, that cannot be adequately captured by standard binomial assumptions. We therefore employ quasibinomial specifications with dispersion parameters θ^{CR} and θ^{BR} for both metrics:

$$y_{it}^{\mathrm{txn}} \mid n_{it}^{\mathrm{sess}}, p_{it}^{\mathrm{CR}}, \theta^{\mathrm{CR}} \sim \mathrm{QuasiBinomial}\big(n_{it}^{\mathrm{sess}}, p_{it}^{\mathrm{CR}}, \theta^{\mathrm{CR}}\big), \tag{2}$$

$$\mathrm{logit}\big(p_{it}^{\mathrm{CR}}\big) = \alpha_{c(i,t)}^{CR} + \gamma_i^{CR} + \delta_{d(i,t)}^{CR} + \mu_{m(t)}^{CR}, \tag{3}$$

$$\mathrm{Var}\big(y_{it}^{\mathrm{txn}}\big) = \theta^{\mathrm{CR}} \cdot n_{it}^{\mathrm{sess}} \cdot p_{it}^{\mathrm{CR}} \cdot \big(1 - p_{it}^{\mathrm{CR}}\big), \tag{4}$$

$$y_{it}^{\mathrm{bounce}} \mid n_{it}^{\mathrm{sess}}, p_{it}^{\mathrm{BR}}, \theta^{\mathrm{BR}} \sim \mathrm{QuasiBinomial}\big(n_{it}^{\mathrm{sess}}, p_{it}^{\mathrm{BR}}, \theta^{\mathrm{BR}}\big), \tag{5}$$

$$\mathrm{logit}\big(p_{it}^{\mathrm{BR}}\big) = \alpha_{c(i,t)}^{BR} + \gamma_i^{BR} + \delta_{d(i,t)}^{BR} + \mu_{m(t)}^{BR}, \tag{6}$$

$$\mathrm{Var}\big(y_{it}^{\mathrm{bounce}}\big) = \theta^{\mathrm{BR}} \cdot n_{it}^{\mathrm{sess}} \cdot p_{it}^{\mathrm{BR}} \cdot \big(1 - p_{it}^{\mathrm{BR}}\big). \tag{7}$$

A.4.3 Model Specification for AOV, RPS, SD, and PV

The remaining outcome metrics, $\mathrm{AOV}_{it}, \mathrm{RPS}_{it}, \mathrm{SD}_{it}, \mathrm{PV}_{it}$, are modeled with standard linear specifications. We begin with the monetary variables:

$$\mathrm{AOV}_{it} = \beta_0^{\mathrm{AOV}} + \alpha_{c(i,t)}^{\mathrm{AOV}} + \gamma_i^{\mathrm{AOV}} + \delta_{d(i,t)}^{\mathrm{AOV}} + \mu_{m(t)}^{\mathrm{AOV}} + \varepsilon_{it}^{\mathrm{AOV}}, \tag{8}$$

$$\varepsilon_{it}^{\mathrm{AOV}} \sim \mathcal{N}(0, \sigma_{\mathrm{AOV}}^2), \tag{9}$$

$$\mathrm{RPS}_{it} = \beta_0^{\mathrm{RPS}} + \alpha_{c(i,t)}^{\mathrm{RPS}} + \gamma_i^{\mathrm{RPS}} + \delta_{d(i,t)}^{\mathrm{RPS}} + \mu_{m(t)}^{\mathrm{RPS}} + \varepsilon_{it}^{\mathrm{RPS}}, \tag{10}$$

$$\varepsilon_{it}^{\mathrm{RPS}} \sim \mathcal{N}(0, \sigma_{\mathrm{RPS}}^2). \tag{11}$$

By definition of AOV_{it}, we exclude observations with zero transactions. In contrast, RPS_{it} includes realized zeros, as observations with zero revenue are informative about channel quality. For the engagement metrics, we specify SD and PV per session analogously:

$$\mathrm{SD}_{it} = \beta_0^{\mathrm{SD}} + \alpha_{c(i,t)}^{\mathrm{SD}} + \gamma_i^{\mathrm{SD}} + \delta_{d(i,t)}^{\mathrm{SD}} + \mu_{m(t)}^{\mathrm{SD}} + \varepsilon_{it}^{\mathrm{SD}}, \tag{12}$$

$$\varepsilon_{it}^{\mathrm{SD}} \sim \mathcal{N}(0, \sigma_{\mathrm{SD}}^2), \tag{13}$$

$$\mathrm{PV}_{it} = \beta_0^{\mathrm{PV}} + \alpha_{c(i,t)}^{\mathrm{PV}} + \gamma_i^{\mathrm{PV}} + \delta_{d(i,t)}^{\mathrm{PV}} + \mu_{m(t)}^{\mathrm{PV}} + \varepsilon_{it}^{\mathrm{PV}}, \tag{14}$$

$$\varepsilon_{it}^{\mathrm{PV}} \sim \mathcal{N}(0, \sigma_{\mathrm{PV}}^2). \tag{15}$$

A.4.4 Weighting Scheme

All linear models are estimated with weights to handle heteroskedasticity from aggregation, as averages based on more sessions (or transactions in the case of AOV) are systematically less noisy. Unweighted OLS would give too much weight to small, volatile cells, which would particularly impact channels with sparse observations, such as oLLM. For website i and week t, let s_{it} be sessions, $\bar{s}_i = \frac{1}{T_i} \sum_t s_{it}$, $\sigma_{s,i} = \sqrt{\frac{1}{T_i - 1} \sum_t (s_{it} - \bar{s}_i)^2}$, $z_{it} = (s_{it} - \bar{s}_i)/\sigma_{s,i}$, and define a bounded, site-relative weight

$$w_{it} = \frac{1}{1 + \exp(-z_{it})} \in (0, 1).$$

This monotone proxy for inverse variance (i) favors more precise cells, (ii) caps leverage so that no website dominates, and (iii) avoids raw-size bias from very large retailers.

A simpler alternative would be to weight by

$$w_{it} = \sqrt{sessions_{it}}.$$

This squareroot scheme concentrates weight on the largest websites and their most-visited channels (Figure A.1): the largest 12.5% of websites carry approximately one-third of total weight for each channel, while the largest 50% of websites carry over three-quarter. In contrast, the sigmoid approach balances weight more evenly across websites while preserving within-website variation that reflects session counts.

Figure A.1: Cumulative Weight Comparison

We *do not* weight the quasibinomial CR and BR models because n_{it}^{sess} already enters the likelihood and additional weights would therefore double-count precision.

Online Appendix B Detailed Robustness Check Results

This section presents detailed regression results on all robustness checks, covering data processing choices, observed websites, observed LLM platforms, and observed timeframes. It further contains result graphs on financial metrics AOV and RPS, as well as on engagement metrics BR, PV, and SD.

B.1 Robustness Regressions: Data Processing Choices

We vary the aggregation scheme by testing daily and monthly data as alternatives to weekly aggregation. Daily aggregation increases granularity but exacerbates sparsity for oLLM. Monthly aggregation reduces sparsity and improves estimates of ratio metrics, at the cost of fewer observations and potentially wider confidence intervals. We then apply filters requiring minimum observations of 10, 100, or 1000 per week/website/device/channel combination. These thresholds increase data robustness but disproportionately reduce oLLM observations, potentially biasing the sample toward larger websites. Finally, we winsorize dependent variables at the 2.5% and 97.5% levels to reduce outlier influence.

Table B.1: Channel Metrics: Daily Aggregated Data

	Dependent variables					
	CR	AOV	RPS	BR	SD	PV
Affiliate	0.449*** (0.025)	22.9*** (4.3)	3.62*** (0.06)	0.267*** (0.014)	-15.9*** (3.6)	0.674*** (0.038)
Paid Search	0.401*** (0.025)	3.59 (3.98)	2.13*** (0.05)	-0.168*** (0.013)	18.1*** (2.7)	0.937*** (0.029)
Other	0.384*** (0.025)	8.98** (4.01)	2.55*** (0.05)	0.892*** (0.013)	39.2*** (2.8)	0.794*** (0.029)
Direct	0.345*** (0.025)	12.8*** (4.0)	3.25*** (0.05)	0.352*** (0.013)	-3.77 (2.74)	0.733*** (0.029)
Email	0.300*** (0.025)	10.6*** (4.1)	3.05*** (0.05)	0.144*** (0.013)	63.8*** (2.9)	0.971*** (0.031)
Referral	0.258*** (0.025)	13.2*** (4.0)	1.77*** (0.05)	0.208*** (0.013)	34.5*** (2.8)	0.759*** (0.030)
Organic Search	0.156*** (0.025)	2.02 (3.98)	1.21*** (0.04)	-0.126*** (0.013)	-3.46 (2.72)	0.561*** (0.029)
Paid Social	-0.750*** (0.025)	4.71 (4.18)	-0.371*** (0.051)	0.278*** (0.013)	-130*** (3)	-0.346*** (0.032)
Fixed Effects: Website, Device, Month	Yes	Yes	Yes	Yes	Yes	Yes
R^2 [pseudo-R^2]	[0.689]	0.170	0.149	[0.639]	0.262	0.052
Observations	3,896,527	2,953,982	3,896,527	3,896,527	3,896,527	3,896,527

*p<0.1; **p<0.05; ***p<0.01

Notes: All coefficients relative to oLLM baseline. Standard errors in parentheses.

Table B.2: Channel Metrics: Monthly Aggregated Data

	Dependent variables					
	CR	AOV	RPS	BR	SD	PV
Affiliate	0.449*** (0.109)	18.4*** (3.7)	3.20*** (0.13)	0.267*** (0.064)	-1.49 (9.87)	0.627*** (0.045)
Paid Search	0.401*** (0.108)	1.00 (2.88)	2.01*** (0.09)	-0.168*** (0.062)	45.2*** (6.5)	0.951*** (0.030)
Other	0.384*** (0.108)	3.99 (2.93)	2.36*** (0.09)	0.892*** (0.062)	66.3*** (6.8)	0.721*** (0.031)
Direct	0.345*** (0.108)	10.1*** (2.9)	2.93*** (0.09)	0.352*** (0.062)	14.6** (6.6)	0.693*** (0.030)
Email	0.300*** (0.108)	8.45*** (3.06)	2.46*** (0.10)	0.144** (0.062)	72.5*** (7.3)	0.931*** (0.033)
Referral	0.258** (0.108)	10.8*** (3.0)	1.63*** (0.09)	0.208*** (0.062)	66.3*** (7.0)	0.787*** (0.032)
Organic Search	0.156 (0.108)	-1.71 (2.86)	1.07*** (0.09)	-0.126** (0.062)	27.0*** (6.5)	0.599*** (0.030)
Paid Social	-0.750*** (0.109)	2.48 (3.26)	-0.666*** (0.104)	0.278*** (0.062)	-117*** (8)	-0.381*** (0.035)
Fixed Effects: Website, Device, Month	Yes	Yes	Yes	Yes	Yes	Yes
R^2 [pseudo-R^2]	[0.750]	0.645	0.366	[0.681]	0.432	0.384
Observations	149,628	132,458	149,628	149,628	149,628	149,628

*p<0.1; **p<0.05; ***p<0.01

Notes: All coefficients relative to oLLM baseline. Standard errors in parentheses.

Table B.3: Channel Metrics: Min. 10 Sessions

	Dependent variables					
	CR	AOV	RPS	BR	SD	PV
Affiliate	0.612*** (0.062)	24.1*** (4.0)	3.61*** (0.11)	0.225*** (0.036)	-36.1*** (7.5)	0.528*** (0.060)
Paid Search	0.361*** (0.061)	3.41 (3.28)	1.94*** (0.08)	-0.148*** (0.034)	0.465 (5.432)	0.763*** (0.043)
Other	0.357*** (0.061)	4.88 (3.33)	2.20*** (0.08)	0.891*** (0.034)	14.2** (5.6)	0.454*** (0.044)
Direct	0.273*** (0.061)	12.8*** (3.3)	3.05*** (0.08)	0.329*** (0.034)	-25.4*** (5.4)	0.546*** (0.043)
Email	0.267*** (0.061)	9.63*** (3.46)	2.59*** (0.09)	0.189*** (0.034)	28.3*** (5.9)	0.739*** (0.047)
Referral	0.209*** (0.061)	12.0*** (3.4)	1.62*** (0.08)	0.210*** (0.034)	30.0*** (5.7)	0.639*** (0.045)
Organic Search	0.112* (0.061)	1.14 (3.27)	0.999*** (0.079)	-0.128*** (0.034)	-16.5*** (5.4)	0.429*** (0.043)
Paid Social	-0.770*** (0.062)	4.97 (3.66)	-0.631*** (0.092)	0.368*** (0.034)	-158*** (6)	-0.548*** (0.049)
Fixed Effects: Website, Device, Month	Yes	Yes	Yes	Yes	Yes	Yes
R^2 [pseudo-R^2]	[0.739]	0.472	0.362	[0.711]	0.481	0.182
Observations	320,021	287,409	320,021	320,021	320,021	320,021

*p<0.1; **p<0.05; ***p<0.01

Notes: All coefficients relative to oLLM baseline. Standard errors in parentheses.

Table B.4: Channel Metrics: Min. 100 Sessions

	Dependent variables					
	CR	AOV	RPS	BR	SD	PV
Affiliate	0.539*** (0.074)	25.6*** (5.4)	3.20*** (0.15)	0.252*** (0.041)	-47.3*** (10.9)	0.401*** (0.045)
Paid Search	0.288*** (0.073)	4.20 (4.81)	1.50*** (0.13)	-0.121*** (0.039)	-12.4 (9.4)	0.659*** (0.039)
Other	0.284*** (0.073)	6.29 (4.84)	1.68*** (0.13)	0.918*** (0.039)	-1.20 (9.47)	0.172*** (0.039)
Direct	0.200*** (0.073)	13.9*** (4.8)	2.63*** (0.13)	0.356*** (0.039)	-40.2*** (9.4)	0.412*** (0.039)
Email	0.193*** (0.073)	9.81** (4.95)	1.93*** (0.14)	0.216*** (0.039)	12.0 (9.8)	0.579*** (0.040)
Referral	0.137* (0.073)	13.0*** (4.9)	1.27*** (0.13)	0.237*** (0.039)	6.10 (9.54)	0.492*** (0.039)
Organic Search	0.039 (0.073)	2.15 (4.79)	0.569*** (0.132)	-0.101*** (0.039)	-30.9*** (9.3)	0.317*** (0.038)
Paid Social	-0.843*** (0.074)	6.38 (5.10)	-0.944*** (0.141)	0.395*** (0.039)	-165*** (10)	-0.683*** (0.041)
Fixed Effects: Website, Device, Month	Yes	Yes	Yes	Yes	Yes	Yes
R^2 [pseudo-R^2]	[0.740]	0.486	0.410	[0.712]	0.513	0.489
Observations	280,190	267,602	280,190	280,190	280,190	280,190

*p<0.1; **p<0.05; ***p<0.01

Notes: All coefficients relative to oLLM baseline. Standard errors in parentheses.

Table B.5: Channel Metrics: Min. 1000 Sessions

	Dependent variables					
	CR	AOV	RPS	BR	SD	PV
Affiliate	0.447*** (0.119)	39.2*** (11.9)	2.65*** (0.31)	0.293*** (0.059)	-21.1 (16.2)	0.481*** (0.076)
Paid Search	0.196* (0.118)	22.0* (11.3)	1.43*** (0.29)	-0.086 (0.057)	10.3 (15.2)	0.915*** (0.071)
Other	0.190 (0.118)	24.8** (11.3)	1.35*** (0.29)	0.955*** (0.057)	-3.05 (15.30)	0.193*** (0.072)
Direct	0.101 (0.118)	29.5*** (11.3)	2.25*** (0.29)	0.392*** (0.057)	-21.8 (15.2)	0.547*** (0.071)
Email	0.088 (0.119)	24.1** (11.4)	1.40*** (0.30)	0.255*** (0.058)	28.2* (15.5)	0.639*** (0.072)
Referral	0.045 (0.119)	28.5** (11.3)	1.13*** (0.30)	0.278*** (0.058)	13.2 (15.4)	0.580*** (0.072)
Organic Search	-0.053 (0.118)	19.2* (11.2)	0.553* (0.292)	-0.066 (0.057)	-4.59 (15.22)	0.586*** (0.071)
Paid Social	-0.937*** (0.119)	27.5** (11.5)	-0.799*** (0.299)	0.431*** (0.058)	-141*** (16)	-0.550*** (0.073)
Fixed Effects: Website, Device, Month	Yes	Yes	Yes	Yes	Yes	Yes
R^2 [pseudo-R^2]	[0.742]	0.491	0.479	[0.713]	0.705	0.676
Observations	188,376	185,057	188,376	188,376	188,376	188,376

*p<0.1; **p<0.05; ***p<0.01

Notes: All coefficients relative to oLLM baseline. Standard errors in parentheses.

Table B.6: Channel Metrics: Winsorized Data

	Dependent variables					
	CR	AOV	RPS	BR	SD	PV
Affiliate	0.636*** (0.058)	29.9*** (1.4)	3.27*** (0.05)	0.208*** (0.032)	9.91*** (2.32)	0.673*** (0.016)
Paid Search	0.382*** (0.057)	18.5*** (1.1)	2.35*** (0.03)	-0.161*** (0.031)	55.2*** (1.5)	0.913*** (0.010)
Other	0.362*** (0.057)	19.7*** (1.1)	2.23*** (0.03)	0.829*** (0.031)	27.9*** (1.6)	0.388*** (0.011)
Direct	0.278*** (0.057)	28.7*** (1.1)	3.14*** (0.03)	0.306*** (0.031)	26.4*** (1.6)	0.674*** (0.010)
Email	0.286*** (0.057)	19.3*** (1.2)	2.51*** (0.03)	0.174*** (0.031)	43.7*** (1.7)	0.885*** (0.012)
Referral	0.173*** (0.057)	24.4*** (1.2)	1.84*** (0.03)	0.191*** (0.031)	56.7*** (1.6)	0.764*** (0.011)
Organic Search	0.130** (0.057)	16.6*** (1.1)	1.56*** (0.03)	-0.138*** (0.031)	39.5*** (1.5)	0.617*** (0.010)
Paid Social	-0.757*** (0.057)	12.3*** (1.3)	-0.218*** (0.036)	0.352*** (0.031)	-107*** (2)	-0.387*** (0.012)
Fixed Effects: Website, Device, Month	Yes	Yes	Yes	Yes	Yes	Yes
R^2 [pseudo-R^2]	[0.727]	0.788	0.493	[0.729]	0.739	0.625
Observations	340,292	289,524	340,292	340,292	340,292	340,292

*p<0.1; **p<0.05; ***p<0.01

Notes: All coefficients relative to oLLM baseline. Standard errors in parentheses.

B.2 Robustness Regressions: Observed Websites

We restrict the sample to two subsets: the top 25% of websites by oLLM sessions, testing whether results are driven by sites with low oLLM traffic, and the top 25% by transactions, assessing whether larger websites exhibit systematically different oLLM outcomes. The 40% overlap between subsets indicates that current oLLM traffic may disproportionately favor smaller websites, despite LLMs reportedly favoring short tail sources in other settings (Padilla et al. 2025, Gholami et al. 2026).

Table B.7: Channel Metrics: Top 25% oLLM

	Dependent variables					
	CR	AOV	RPS	BR	SD	PV
Affiliate	0.704*** (0.101)	39.8*** (4.1)	3.10*** (0.12)	0.145** (0.061)	-18.5 (11.6)	0.447*** (0.040)
Paid Search	0.322*** (0.098)	7.72** (3.07)	1.38*** (0.08)	-0.143** (0.057)	32.6*** (7.8)	0.734*** (0.027)
Other	0.439*** (0.098)	8.10*** (3.13)	1.47*** (0.09)	0.935*** (0.057)	35.2*** (8.0)	0.209*** (0.027)
Direct	0.237** (0.098)	15.9*** (3.0)	1.67*** (0.08)	0.326*** (0.056)	-2.82 (7.69)	0.361*** (0.026)
Email	0.202** (0.099)	12.2*** (3.4)	1.66*** (0.09)	0.238*** (0.057)	32.0*** (8.7)	0.434*** (0.030)
Referral	0.174* (0.099)	8.28*** (3.19)	1.09*** (0.09)	0.210*** (0.057)	53.7*** (8.3)	0.441*** (0.028)
Organic Search	0.102 (0.098)	2.41 (3.01)	0.607*** (0.080)	-0.121** (0.056)	19.3** (7.5)	0.481*** (0.026)
Paid Social	-1.05*** (0.10)	13.9*** (3.7)	-0.391*** (0.099)	0.261*** (0.057)	-148*** (9)	-0.816*** (0.032)
Fixed Effects: Website, Device, Month	Yes	Yes	Yes	Yes	Yes	Yes
R^2 [pseudo-R^2]	[0.700]	0.719	0.342	[0.724]	0.510	0.604
Observations	91,385	76,686	91,385	91,385	91,385	91,385

*p<0.1; **p<0.05; ***p<0.01

NOTES: All coefficients relative to oLLM baseline. Standard errors in parentheses.

Table B.8: Channel Metrics: Top 25% Revenue

	Dependent variables					
	CR	AOV	RPS	BR	SD	PV
Affiliate	0.865*** (0.150)	29.8*** (9.6)	6.35*** (0.28)	0.203** (0.086)	8.01 (16.27)	0.724*** (0.035)
Paid Search	0.632*** (0.149)	1.16 (7.96)	3.37*** (0.20)	-0.159* (0.083)	29.9** (12.0)	1.03*** (0.03)
Other	0.685*** (0.149)	7.21 (8.10)	3.71*** (0.21)	0.893*** (0.083)	95.6*** (12.4)	0.260*** (0.027)
Direct	0.535*** (0.149)	9.68 (7.99)	4.81*** (0.21)	0.356*** (0.083)	-43.8*** (12.1)	0.562*** (0.026)
Email	0.494*** (0.149)	5.14 (8.37)	4.13*** (0.22)	0.224*** (0.084)	25.3* (13.1)	0.652*** (0.028)
Referral	0.503*** (0.149)	9.05 (8.24)	2.67*** (0.22)	0.221*** (0.084)	22.3* (12.9)	0.619*** (0.028)
Organic Search	0.422*** (0.149)	7.28 (7.96)	2.33*** (0.20)	-0.143* (0.083)	25.6** (12.0)	0.897*** (0.026)
Paid Social	-0.505*** (0.150)	5.56 (8.80)	-0.989*** (0.231)	0.338*** (0.084)	-151*** (14)	-0.647*** (0.029)
Fixed Effects: Website, Device, Month	Yes	Yes	Yes	Yes	Yes	Yes
R^2 [pseudo-R^2]	[0.738]	0.423	0.338	[0.727]	0.386	0.619
Observations	91,942	81,983	91,942	91,942	91,942	91,942

*p<0.1; **p<0.05; ***p<0.01

NOTES: All coefficients relative to oLLM baseline. Standard errors in parentheses.

B.3 Robustness Regressions: Observed LLM Platforms

We explore two alternative definitions of oLLM. First, we expand oLLM to include not only ChatGPT, but also Perplexity, Gemini, Copilot, Deepseek, and Grok. Second, we expand oLLM to include not only mobile and desktop website traffic, but also mobile *app* traffic from ChatGPT.

Table B.9: Channel Metrics: All LLMs

	Dependent variables					
	CR	AOV	RPS	BR	SD	PV
Affiliate	0.647*** (0.055)	26.0*** (4.1)	4.10*** (0.10)	0.202*** (0.001)	2.91 (6.22)	0.855*** (0.054)
Paid Search	0.392*** (0.054)	4.50 (3.30)	2.31*** (0.06)	-0.178*** (0.001)	52.6*** (3.8)	1.12*** (0.03)
Other	0.389*** (0.054)	6.31* (3.35)	2.51*** (0.06)	0.857*** (0.001)	63.9*** (4.0)	0.814*** (0.035)
Direct	0.312*** (0.054)	16.1*** (3.3)	3.45*** (0.06)	0.305*** (0.001)	30.5*** (3.9)	0.890*** (0.034)
Email	0.290*** (0.054)	11.3*** (3.5)	2.99*** (0.07)	0.162*** (0.001)	72.9*** (4.4)	1.12*** (0.04)
Referral	0.239*** (0.054)	13.2*** (3.4)	1.91*** (0.07)	0.174*** (0.001)	79.1*** (4.1)	1.00*** (0.04)
Organic Search	0.159*** (0.054)	2.88 (3.28)	1.39*** (0.06)	-0.154*** (0.001)	33.8*** (3.8)	0.767*** (0.033)
Paid Social	-0.776*** (0.055)	6.42* (3.72)	-0.339*** (0.076)	0.356*** (0.001)	-116*** (5)	-0.219*** (0.041)
Fixed Effects: Website, Device, Month	Yes	Yes	Yes	Yes	Yes	Yes
R^2 [pseudo-R^2]	[0.746]	0.398	0.263	[0.710]	0.436	0.146
Observations	408,342	333,657	408,342	408,342	408,342	408,342

*p<0.1; **p<0.05; ***p<0.01

NOTES: All coefficients relative to oLLM baseline. Standard errors in parentheses.

Table B.10: Channel Metrics: Including ChatGPT Mobile App

	Dependent variables					
	CR	AOV	RPS	BR	SD	PV
Affiliate	0.547*** (0.041)	26.0*** (3.3)	3.05*** (0.12)	0.170*** (0.025)	-7.67 (5.98)	0.609*** (0.051)
Paid Search	0.312*** (0.040)	6.16*** (2.37)	1.52*** (0.07)	-0.193*** (0.023)	34.7*** (3.6)	0.878*** (0.031)
Other	0.282*** (0.040)	10.3*** (2.4)	1.73*** (0.07)	0.850*** (0.023)	48.1*** (3.8)	0.624*** (0.033)
Direct	0.213*** (0.040)	15.8*** (2.4)	2.68*** (0.07)	0.277*** (0.023)	12.5*** (3.7)	0.658*** (0.032)
Email	0.201*** (0.040)	12.4*** (2.6)	2.42*** (0.08)	0.137*** (0.023)	61.6*** (4.2)	0.872*** (0.036)
Referral	0.147*** (0.040)	17.0*** (2.5)	1.19*** (0.08)	0.159*** (0.023)	62.2*** (3.9)	0.754*** (0.034)
Organic Search	0.036 (0.040)	2.18 (2.35)	0.545*** (0.070)	-0.179*** (0.023)	19.1*** (3.6)	0.529*** (0.031)
Paid Social	-0.849*** (0.041)	7.09** (2.85)	-1.13*** (0.09)	0.318*** (0.023)	-124*** (4)	-0.417*** (0.038)
Fixed Effects: Website, Device, Month	Yes	Yes	Yes	Yes	Yes	Yes
R^2 [pseudo-R^2]	[0.741]	0.442	0.233	[0.717]	0.461	0.175
Observations	395,430	331,870	395,430	395,430	395,430	395,430

*p<0.1; **p<0.05; ***p<0.01

NOTES: All coefficients relative to oLLM baseline. Standard errors in parentheses.

B.4 Robustness Regressions: Observed Timeframes

We re-estimate models on 3- and 12-month datasets. The 12-month dataset provides longer horizon but may capture idiosyncratic introduction effects. The 3-month dataset contains fewer observations but reflects the most recent period of relative stability.

Table B.11: Channel Metrics: 3 Months Data

	Dependent variables					
	CR	AOV	RPS	BR	SD	PV
Affiliate	0.522*** (0.072)	30.7*** (3.4)	3.29*** (0.15)	0.221*** (0.001)	-35.3*** (9.2)	0.545*** (0.056)
Paid Search	0.309*** (0.070)	9.78*** (2.59)	1.73*** (0.10)	-0.135*** (0.001)	16.2*** (6.0)	0.772*** (0.036)
Other	0.359*** (0.070)	10.5*** (2.6)	2.04*** (0.10)	0.871*** (0.001)	26.9*** (6.3)	0.393*** (0.038)
Direct	0.220*** (0.070)	20.6*** (2.6)	2.92*** (0.10)	0.287*** (0.001)	-6.85 (6.08)	0.572*** (0.037)
Email	0.233*** (0.071)	17.7*** (2.8)	2.50*** (0.11)	0.192*** (0.001)	42.5*** (6.8)	0.781*** (0.041)
Referral	0.162** (0.071)	18.2*** (2.7)	1.36*** (0.10)	0.202*** (0.001)	46.1*** (6.4)	0.640*** (0.039)
Organic Search	0.076 (0.070)	5.53** (2.58)	0.768*** (0.095)	-0.145*** (0.001)	3.92 (5.96)	0.480*** (0.036)
Paid Social	-0.904*** (0.071)	10.2*** (3.0)	-0.973*** (0.114)	0.372*** (0.001)	-147*** (7)	-0.531*** (0.043)
Fixed Effects: Website, Device, Month	Yes	Yes	Yes	Yes	Yes	Yes
R^2 [pseudo-R^2]	[0.727]	0.663	0.312	[0.727]	0.457	0.247
Observations	188,038	159,659	188,038	188,038	188,038	188,038

*p<0.1; **p<0.05; ***p<0.01

Notes: All coefficients relative to oLLM baseline. Standard errors in parentheses.

Table B.12: Channel Metrics: 12 Months Data

	Dependent variables					
	CR	AOV	RPS	BR	SD	PV
Affiliate	0.450*** (0.058)	21.3*** (3.6)	3.56*** (0.08)	0.267*** (0.033)	-1.07 (5.51)	0.704*** (0.038)
Paid Search	0.402*** (0.057)	1.60 (3.08)	2.16*** (0.06)	-0.169*** (0.032)	38.2*** (3.8)	0.992*** (0.027)
Other	0.384*** (0.057)	5.66* (3.11)	2.57*** (0.06)	0.891*** (0.032)	58.5*** (4.0)	0.779*** (0.027)
Direct	0.345*** (0.057)	11.0*** (3.1)	3.21*** (0.06)	0.351*** (0.032)	12.1*** (3.9)	0.763*** (0.027)
Email	0.300*** (0.057)	8.98*** (3.21)	2.84*** (0.06)	0.143*** (0.032)	68.7*** (4.2)	0.984*** (0.029)
Referral	0.259*** (0.057)	11.9*** (3.1)	1.81*** (0.06)	0.208*** (0.032)	60.1*** (4.0)	0.830*** (0.028)
Organic Search	0.156*** (0.057)	-0.494 (3.067)	1.23*** (0.06)	-0.127*** (0.032)	18.2*** (3.8)	0.624*** (0.026)
Paid Social	-0.750*** (0.057)	2.45 (3.37)	-0.468*** (0.069)	0.277*** (0.032)	-118*** (4)	-0.324*** (0.031)
Fixed Effects: Website, Device, Month	Yes	Yes	Yes	Yes	Yes	Yes
R^2 [pseudo-R^2]	[0.723]	0.419	0.289	[0.658]	0.405	0.196
Observations	614,714	526,432	614,714	614,714	614,714	614,714

*p<0.1; **p<0.05; ***p<0.01

Notes: All coefficients relative to oLLM baseline. Standard errors in parentheses.

B.5 Robustness Graphs: Financial Metrics beyond Conversion Rate

AOV robustness checks exhibit greater variation in effect sizes (Figure B.1), with two non-significant sign changes for organic search (monthly aggregation and 12-month timeframe, see Tables B.2 and B.12). Results remain largely consistent with the main model, which found non-significant differences between oLLM and four traditional channels.

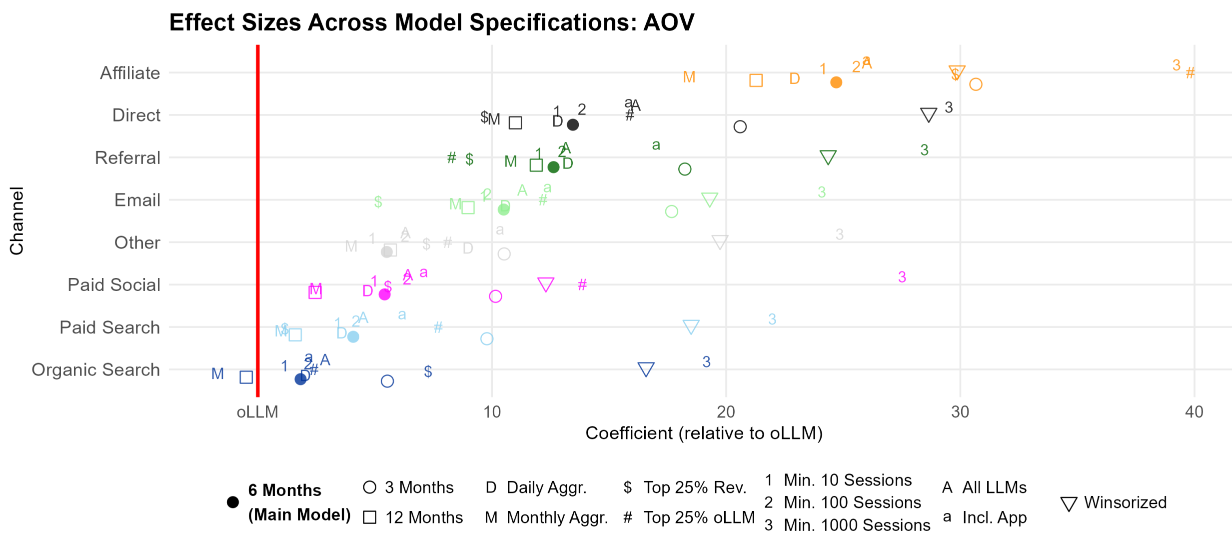


Figure B.1: AOV by channel - All specifications

RPS robustness checks shown in Figure B.2 confirm the main analysis. In contrast to the CR results, RPS for oLLM remains significantly below organic search across all 12 specifications. The top-25%-revenue specification again exhibits systematically larger gaps, mirroring the patterns observed for CR in Figure 14.

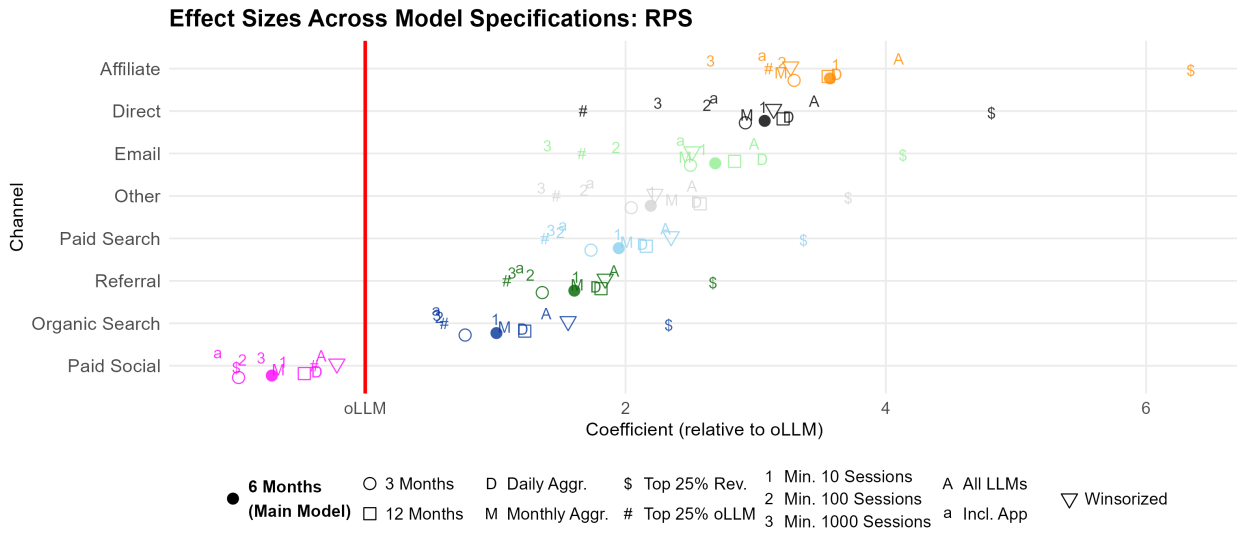


Figure B.2: RPS by channel - All specifications

B.6 Robustness Graphs: Engagement Metrics

Figures B.3 through B.5 present comprehensive robustness checks for engagement metrics (BR, PV, SD) across all specifications tested.

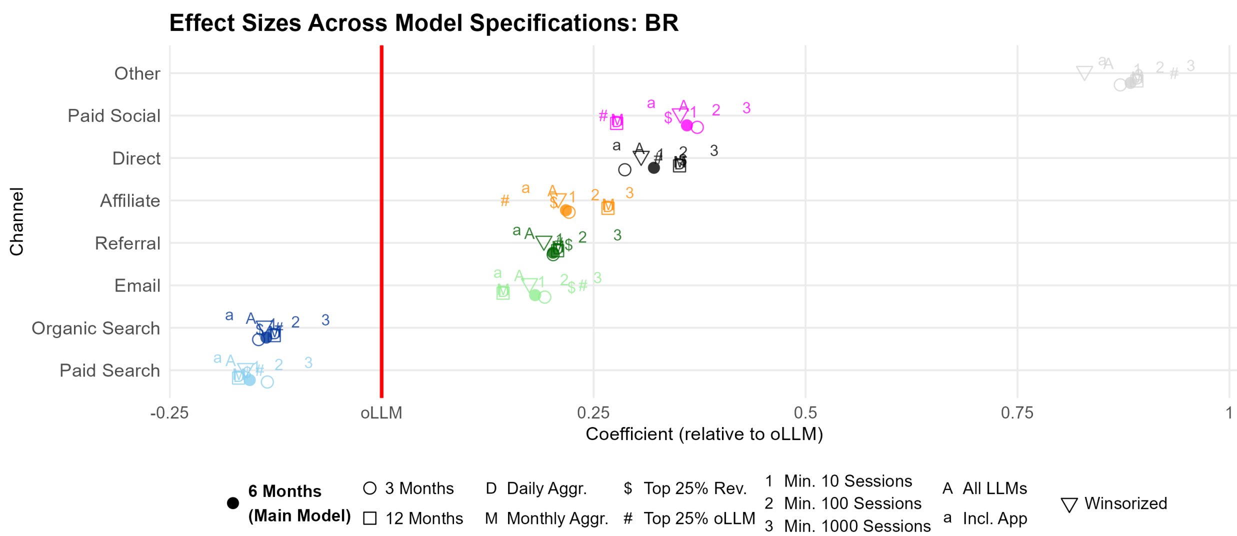


Figure B.3: BR by channel - All specifications

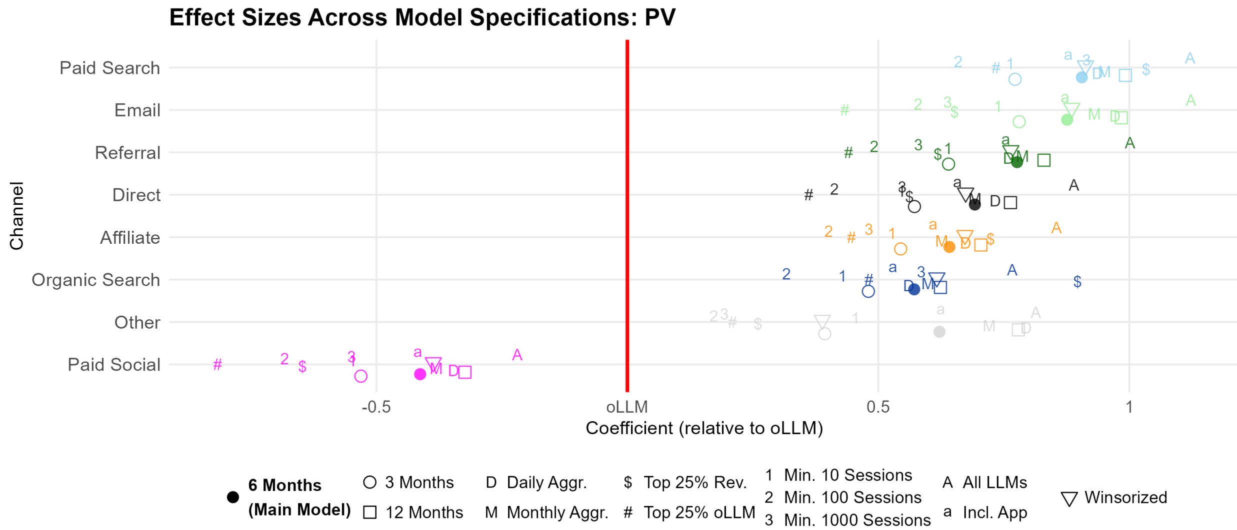


Figure B.4: PV by channel - All specifications

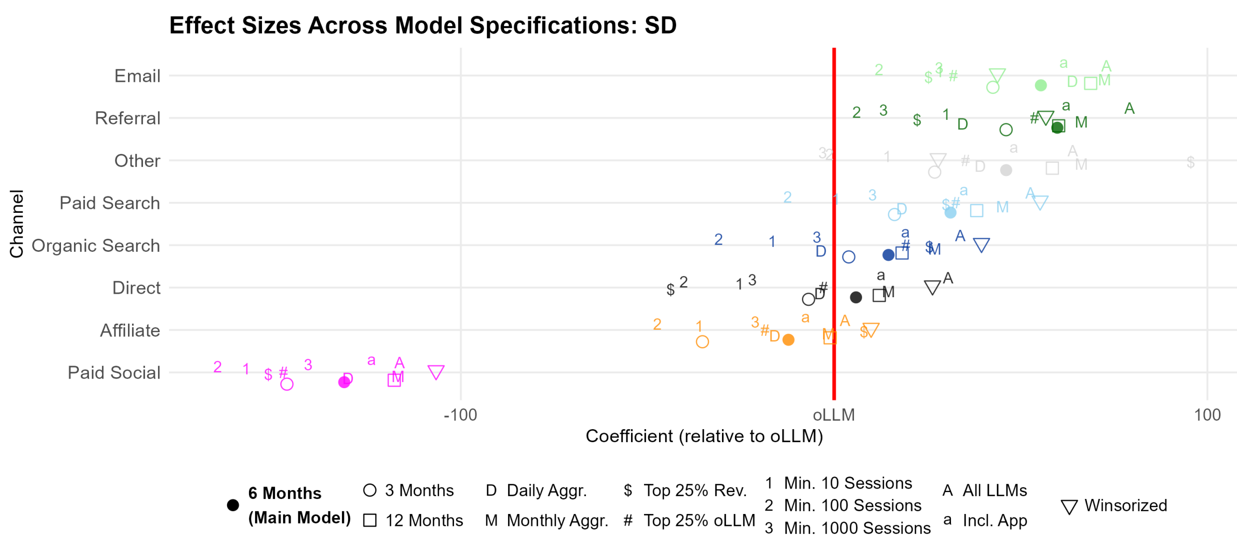


Figure B.5: SD by channel - All specifications

Online Appendix C Time Trend Technical Specifications

C.1 Complete Time Trend Regression Equations

This section provides complete regression specifications for all time trend models used in Section 5. We present equations for all dependent variables with time trend interactions, including the four alternative trend specifications tested: linear, centered, Gompertz, and Sigmoid.

C.1.1 Trend Specifications for CR and BR Models

The quasibinomial CR models are specified with four alternative trend transformations:

Linear trend:

$$\operatorname{logit}\left(p_{it}^{\mathrm{CR}}\right) = \alpha_{c(i,t)}^{CR} + \gamma_i^{CR} + \delta_{d(i,t)}^{CR} + \mu_{m(t)}^{CR} \tag{16}$$

$$+ \operatorname{trend}_t + \alpha_{c(i,t)}^{CR} \cdot \operatorname{trend}_t. \tag{17}$$

We note that this translates into a seemingly exponential effect for the quasibinomial regressions due to the logit transformation for predictions, yet remains a standard linear effect in the regular regressions.

Centered transformation , where $\operatorname{trend}_t^{\mathbb{C}} = \frac{\operatorname{trend}_t - \operatorname{median}(\operatorname{trend})}{\operatorname{sd}(\operatorname{trend})}$:

$$\operatorname{logit}\left(p_{it}^{\mathrm{CR}}\right) = \alpha_{c(i,t)}^{CR} + \gamma_i^{CR} + \delta_{d(i,t)}^{CR} + \mu_{m(t)}^{CR} + \tag{18}$$

$$+ \ \operatorname{trend}_t^{\mathbb{C}} + \ \alpha_{c(i,t)}^{CR} \cdot \operatorname{trend}_t^{\mathbb{C}}, \tag{19}$$

This results in a simple intercept shift for the linear regression models and reduces the perceived exponential effect for the quasibinomial models.

Gompertz transformation , where $\text{trend}_t^{\mathbb{G}} = -\exp\left(-\frac{\text{trend}_t}{\max(\text{trend})}\right)$:

$$\text{logit}\left(p_{it}^{\text{CR}}\right) = \alpha_{c(i,t)}^{CR} + \gamma_i^{CR} + \delta_{d(i,t)}^{CR} + \mu_{m(t)}^{CR} + \tag{20}$$

$$+ \ \text{trend}_t^{\mathbb{G}} + \ \alpha_{c(i,t)}^{CR} \cdot \text{trend}_t^{\mathbb{G}}, \tag{21}$$

Sigmoid transformation , where $\text{trend}_t^{\mathbb{S}} = \frac{1}{1+\exp(-\text{trend}_t^{\text{scaled}})}$ and $\text{trend}_t^{\text{scaled}}$ is the standardized trend_t:

$$\text{logit}\left(p_{it}^{\text{CR}}\right) = \alpha_{c(i,t)}^{CR} + \gamma_i^{CR} + \delta_{d(i,t)}^{CR} + \mu_{m(t)}^{CR} + \tag{22}$$

$$+ \ \text{trend}_t^{\mathbb{S}} + \ \alpha_{c(i,t)}^{CR} \cdot \text{trend}_t^{\mathbb{S}}, \tag{23}$$

C.1.2 Trend Specifications for Linear Models (AOV and RPS)

For the linear models, the trend specifications are analogous. We provide the complete equations below:

Average Order Value (AOV):

$$\text{AOV}_{it} = \beta_0^{\text{AOV}} + \alpha_{c(i,t)}^{\text{AOV}} + \gamma_i^{\text{AOV}} + \delta_{d(i,t)}^{\text{AOV}} + \mu_{m(t)}^{\text{AOV}} + \text{trend}_t + \ \alpha_{c(i,t)}^{\text{AOV}} \cdot \text{trend}_t + \varepsilon_{it}^{\text{AOV}}, \tag{24}$$

$$\varepsilon_{it}^{\text{AOV}} \sim \mathcal{N}(0, \sigma_{\text{AOV}}^2). \tag{25}$$

Revenue per Session (RPS):

$$\text{RPS}_{it} = \beta_0^{\text{RPS}} + \alpha_{c(i,t)}^{\text{RPS}} + \gamma_i^{\text{RPS}} + \delta_{d(i,t)}^{\text{RPS}} + \mu_{m(t)}^{\text{RPS}} + \text{trend}_t + \ \alpha_{c(i,t)}^{\text{RPS}} \cdot \text{trend}_t + \varepsilon_{it}^{\text{RPS}}, \tag{26}$$

$$\varepsilon_{it}^{\text{RPS}} \sim \mathcal{N}(0, \sigma_{\text{RPS}}^2). \tag{27}$$

All other definitions remain unchanged from the base model specifications in Online Appendix A.4. The four alternative trend transformations (linear, centered, Gompertz, Sigmoid) can be applied to any of the linear models above by replacing trend_t with the appropriate transformation.

Online Appendix D Heterogeneity Regressions: Quartile Splits

This section presents regression results underlying Figures 21 – 26. Each column presents a separate regression on a subset of websites, based on websites' "product category complexity", "share of technophiles" (based on Google's affinity segmentation), and "average user age".

Table D.1: Channel Effects by Product Category Complexity (Quartile Split)

	CR		AOV		RPS	
	Bottom Quartile (<Q1)	Top Quartile (>Q3)	Bottom Quartile (<Q1)	Top Quartile (>Q3)	Bottom Quartile (<Q1)	Top Quartile (>Q3)
Affiliate	0.272* (0.145)	0.244** (0.103)	50.4*** (12.6)	64.6*** (13.9)	3.35*** (0.36)	3.61*** (0.27)
Paid Search	0.198 (0.135)	0.056 (0.093)	40.5*** (8.9)	4.02 (9.77)	2.29*** (0.20)	1.09*** (0.15)
Other	0.305** (0.135)	0.610*** (0.094)	18.5** (9.0)	6.58 (10.00)	2.36*** (0.21)	2.06*** (0.16)
Direct	0.449*** (0.135)	-0.345*** (0.093)	41.1*** (8.9)	24.5** (9.7)	3.53*** (0.20)	2.39*** (0.15)
Email	0.199 (0.136)	-0.317*** (0.095)	42.1*** (9.3)	14.6 (10.7)	2.89*** (0.22)	2.08*** (0.18)
Referral	0.142 (0.136)	-0.398*** (0.094)	39.5*** (9.1)	23.5** (10.1)	1.44*** (0.21)	1.62*** (0.16)
Organic Search	0.287** (0.135)	-0.374*** (0.093)	22.8*** (8.8)	4.66 (9.67)	1.28*** (0.20)	0.404*** (0.151)
Paid Social	-0.846*** (0.138)	-1.31*** (0.10)	27.1*** (9.7)	17.3 (12.5)	-0.192 (0.225)	-1.26*** (0.20)
R^2 [pseudo-R^2]	[0.680]	[0.782]	0.523	0.281	0.222	0.349
Observations	24,816	85,281	20,946	69,402	24,816	85,281
Fixed Effects: Device, Website, Month	Yes	Yes	Yes	Yes	Yes	Yes

*p<0.1; **p<0.05; ***p<0.01

Notes: All coefficients relative to oLLM baseline. Standard errors in parentheses.

Table D.2: Channel Effects by Technophiles Audience Share (Quartile Split)

	CR		AOV		RPS	
	Bottom Quartile (<Q1)	Top Quartile (>Q3)	Bottom Quartile (<Q1)	Top Quartile (>Q3)	Bottom Quartile (<Q1)	Top Quartile (>Q3)
Affiliate	0.855*** (0.223)	0.547*** (0.110)	14.9*** (4.1)	65.1*** (25.1)	4.18*** (0.24)	4.27*** (0.47)
Paid Search	0.888*** (0.221)	-0.076 (0.105)	5.49 (3.41)	17.9 (19.3)	2.44*** (0.17)	0.971*** (0.298)
Other	0.573*** (0.221)	-0.379*** (0.105)	7.01** (3.44)	10.9 (19.7)	2.63*** (0.17)	2.26*** (0.31)
Direct	0.728*** (0.221)	0.174* (0.105)	11.6*** (3.4)	32.0 (19.5)	2.90*** (0.17)	3.31*** (0.31)
Email	0.686*** (0.221)	0.272** (0.107)	1.05 (3.53)	36.1* (20.9)	3.01*** (0.18)	2.68*** (0.34)
Referral	0.680*** (0.221)	0.214** (0.105)	13.5*** (3.5)	29.0 (20.0)	1.19*** (0.18)	1.50*** (0.32)
Organic Search	0.630*** (0.221)	-0.426*** (0.105)	4.73 (3.40)	20.8 (19.2)	1.43*** (0.17)	0.642** (0.295)
Paid Social	-0.224 (0.222)	-1.39*** (0.11)	3.29 (3.64)	21.9 (23.3)	-0.119 (0.190)	-1.61*** (0.37)
R^2 [pseudo-R^2]	[0.666]	[0.840]	0.688	0.185	0.235	0.200
Observations	41,719	40,321	37,026	33,380	41,719	40,321
Fixed Effects: Device, Website, Month	Yes	Yes	Yes	Yes	Yes	Yes

*p<0.1; **p<0.05; ***p<0.01

Notes: All coefficients relative to oLLM baseline. Standard errors in parentheses. Q1 = bottom quartile; Q3 = top quartile.

Table D.3: Channel Effects by Average User Age (Quartile Split)

	CR		AOV		RPS	
	Bottom Quartile (<Q1)	Top Quartile (>Q3)	Bottom Quartile (<Q1)	Top Quartile (>Q3)	Bottom Quartile (<Q1)	Top Quartile (>Q3)
Affiliate	-0.087 (0.088)	0.505** (0.220)	30.7*** (6.2)	12.6* (6.9)	2.11*** (0.23)	4.52*** (0.38)
Paid Search	0.048 (0.083)	0.545** (0.219)	12.9*** (4.3)	-12.5** (5.9)	1.99*** (0.13)	1.98*** (0.27)
Other	0.090 (0.083)	0.165 (0.219)	20.9*** (4.3)	-11.6* (6.0)	1.68*** (0.14)	3.21*** (0.28)
Direct	-0.035 (0.083)	0.531** (0.219)	26.1*** (4.2)	-0.791 (5.975)	2.53*** (0.13)	4.55*** (0.27)
Email	-0.091 (0.084)	0.428* (0.219)	22.1*** (4.7)	-5.17 (6.17)	3.07*** (0.15)	2.88*** (0.29)
Referral	0.049 (0.084)	0.521** (0.219)	24.8*** (4.4)	-6.22 (6.09)	1.77*** (0.14)	1.90*** (0.28)
Organic Search	-0.310*** (0.083)	0.291 (0.219)	15.9*** (4.2)	-16.9*** (5.9)	0.918*** (0.127)	1.43*** (0.27)
Paid Social	-1.13*** (0.09)	-0.283 (0.219)	8.36 (5.22)	-7.26 (6.46)	0.544*** (0.164)	-0.943*** (0.312)
R^2 [pseudo-R^2]	[0.792]	[0.779]	0.684	0.650	0.435	0.241
Observations	41,229	42,957	33,981	37,060	41,229	42,957
Fixed Effects: Device, Website, Month	Yes	Yes	Yes	Yes	Yes	Yes

*p<0.1; **p<0.05; ***p<0.01

Notes: All coefficients relative to oLLM baseline. Standard errors in parentheses. Q1 = bottom quartile; Q3 = top quartile.

Online Appendix E Reconciling Differences with Prior Studies

This appendix discusses potential explanations for why our findings differ from some industry reports and from expectations based on prior research. We examine both methodological factors and substantive mechanisms that may account for these discrepancies.

E.1 Potential Explanations for Differences vis-à-vis Industry Study Results

Two recent industry studies report substantially higher CR for oLLM than for organic search, in direct contrast to our findings. While we cannot definitively reconcile these results, two factors likely contribute.

E.1.1 Website-level heterogeneity

Seer Interactive (2025) analyze a single website and report an oLLM CR of 15.9%. Our broader sample indicates that such outcomes are atypical. We observe one website with an even higher oLLM CR of 26.0%, but this is a clear outlier; the second-highest oLLM CR drops to 12.9%, suggesting that exceptionally high CRs are rare rather than representative. The multi-site study by ThoughtMetric (2025) reports a 6.7% average oLLM CR across 100 websites, yet in our data only seven websites achieve oLLM CRs of 6.7% or higher. Given this breadth, sample selection alone seems unlikely to explain the divergence.

E.1.2 Methodological upward bias

Differences in analytic practice between industry reports and academic work may also matter. Industry analyses often operate on daily aggregates, which present serious challenges when evaluating a channel like oLLM that represents less than 0.2% of total sessions (see

Section 2.4). The sizable gaps between our model-free evidence and our regression estimates (Section 3) indicate that these challenges can produce genuinely misleading interpretations, and do not simply reflect academic overcautiousness. While ThoughtMetric (2025) provides limited methodological transparency, improper aggregation or data filtering practices—such as applying minimum transaction thresholds—might introduce upward bias in reported oLLM CR.

E.2 Potential Explanations for Differences vis-à-vis Research Study Expectations

Some prior research has posited that oLLM should provide stronger financial and engagement outcomes than traditional channels due to greater context, better synthesis, and a superior interface. Our evidence indicates this is not (yet) realized. We highlight three contributing factors.

E.2.1 Platform architecture and commercial optimization gaps

LLM technology remains early-stage. Platforms have prioritized core model quality and general UX over commerce-specific capabilities. Because only 2.1% of conversations involve purchasable products (Chatterji et al. 2025), features critical for retail, such as comprehensive product catalogs, current price and availability, and reliable deep-linking, may not be fully built out (Semrush 2025). Consistent with this view, oLLM-driven traffic volumes have been stagnant. Although BR is comparatively favorable (suggesting relevance when clicks occur), oLLM lags mature channels on session depth (PV) and duration (SD).

E.2.2 Website infrastructure and investment coordination

Given low volumes and relatively weak financial outcomes to date, many retailers likely have not fully optimized landing pages, site structure, or merchandising for oLLM. This creates a coordination problem: weak outcomes depress incentives to invest, and limited investment sustains weak outcomes. Our evidence is consistent with such dynamics. Improvements in CR coincide with declining AOV, suggesting early optimization might have focused on completion probability rather than basket expansion. The fact that large sites are not (yet) leading in oLLM effectiveness further suggests that dedicated optimization is not currently an industry priority.

E.2.3 Consumer adoption and channel integration

While LLMs have reshaped online behavior with respect to many major websites (Gholami et al. 2026), shopping via LLMs remains nascent: only a small fraction of conversations concern purchasable products (Chatterji et al. 2025). Users may not view oLLM as a default entry point for shopping, in particular for low-complexity products (see Section 6.1). Early-stage familiarity and trust plausibly induce verification behavior. Users confirm recommendations via search engines or retailer sites—which, under last-click attribution, assigns conversions to other channels. Engagement patterns align with this interpretation: low BR on LLM-provided links indicates relevance, yet shorter SD and fewer PVs imply high task orientation; combined with declining AOV, these signals suggest users are still building proficiency and developing habits (see Section 6.2, but also Padilla et al. 2025).

Online Appendix F Directions for Future Research

Ours is the first large-scale empirical study describing financial and engagement outcomes of organic LLM traffic as an e-commerce channel. Following an empirics first approach, it provides solid empirical evidence, but also raises many important questions for future research.

F.1 Outcome Gap Mechanisms and Improvement Trajectories

Our finding that oLLM shows improving CRs alongside declining AOVs warrants further investigation. Future research should expand on our cross-website analysis of consumers' LLM proficiency and product category complexity, and examine more deeply whether this pattern reflects platform learning effects, consumer learning effects, or systematic differences in the types of products oLLM facilitates. Investigation of these mechanisms could inform predictions about convergence trajectories and assess whether current outcomes represent temporary growing pains or structural limitations of LLMs.

F.2 Platform Friction and Cross-Platform Comparison

Our cross-website analysis in Section 6.1 indicates that oLLM outcomes might reflect perceived tradeoffs between LLM usefulness and convenience. The current oLLM experience requires consumers to leave the LLM interface and navigate to a retailer website, introducing friction that may partly explain weaker financial metrics for simpler products. Different approaches to reducing this friction—such as in-platform shopping agents, specialized shopping LLMs, and on-site conversational assistants—offer different tradeoffs in convenience and usefulness. Understanding which friction points matter most for conversion and how they can be overcome represents a promising research direction. Comparative analyses examining different providers (ChatGPT, Claude, Gemini, Perplexity, Grok, Deepseek) could further reveal how interface and algorithm differences affect consumer-product matching.

F.3 Generative Engine Optimization and Retailer Visibility

oLLM's growing volume and improving financial metrics justify companies' increasing interest in Generative Engine Optimization (GEO), yet many aspects of this nascent field remain unexplored. Our understanding along the entire funnel is limited: it is unclear how companies can increase their LLM exposure, click-through, and subsequent conversions, and how these stages are interrelated. Additionally, the coordination problem between weak oLLM outcomes and limited retailer optimization suggests research opportunities examining how retailers should modify website architecture and merchandising strategies for direct product-page entry (Chen et al. 2025).

F.4 Moderating Factors in Channel Outcomes

Future research should investigate contextual factors explaining current gaps. Competition intensity represents a promising avenue, as high-competition categories may benefit more from preference elicitation. Brand strength effects present competing predictions worth investigating, as oLLM recommendations may exhibit less advertising bias, but may also favor brands that are more prominent in retrieved and training data (Chen et al. 2025, Padilla et al. 2025).

F.5 Consumer Adoption and Market Implications

Research examining how consumers develop LLM proficiency and shopping habits in LLM environments would inform understanding of adoption dynamics. The finding that only a small fraction of LLM conversations concern purchasable products (Chatterji et al. 2025) suggests substantial room for growth. Investigation of oLLM's impact on broader market outcomes deserves attention despite current low volumes. Key areas include the introduction of advertising in LLM platforms (pLLM) and potential disruption to the retail media ecosystem as consumers bypass traditional homepage advertising.