When AI Disclosure Backfires:
The Economic Consequences of Labeling AI-Generated Review Summaries

Presenter: Heeseung Andrew Lee

Angela Aerry Choi, Korea University
Anindya Ghose, New York University
Heeseung Andrew Lee, University of Texas at Dallas
Wonseok Oh, Korea Advanced Institute of Science and Technology

Third Federal Trade Commission Conference on Marketing and Public Policy

Policy Background

Policymakers across jurisdictions are **tightening transparency and accountability requirements** for **AI-generated content** in consumer-facing contexts:

US

FTC Consumer Reviews Rule

16 CFR Part 465

Targets fake or false consumer reviews and testimonials, including deceptive AI-enabled practices.

Not yet a general AI labeling mandate

Effective Oct 2024

CN

China AI Labeling

CAC Labeling Measures

Mandates explicit and implicit **labeling** of AI-generated synthetic content for covered network information service providers.

Effective Sep 2025

EU

EU AI Act

Regulation 2024/1689

Article 50 establishes transparency obligations for AI-generated or AI-manipulated content, including **marking** and detectability requirements.

Effective Aug 2026

While intended to preserve trust, these mandates raise an underexplored concern: **the potential unintended economic consequences of AI disclosure for digital platforms.**

Research Motivation

Review Summaries in E-commerce Platform

AI-generated from the text of customer reviews

Customers like the insulation, color, quality and size of the drinking cup. For example, they mention it keeps ice cold for a long time, it stands out well and that it's heavy duty. Some appreciate the size, saying it fits perfectly in their cup holders. That said, opinions are mixed on value.

Insulation | Quality | Color | Size | Value | Condition

AI Labeling & Both Positive and Negative

CARMAX

4.3 out of 5 51 customer reviews

The 2012 Jeep Liberty is a car that has been well-received by customers. Customers like the car's design, the car's handling, and the car's performance. Customers also like the storage capacity and fuel efficiency.

No Labeling & Positive-Only

Research Questions

1. **How does the introduction of review summaries affect consumer shopping decisions across the purchase funnel?**
   - Consideration (e.g., dealer appointment)
   - Conversion (e.g., purchase)

2. **To what extent does source labeling in review summaries influence consumer behavior?**
   - None
   - AI attribution (e.g., “AI-Generated”)
   - Human attribution (e.g., “Human Editor-Generated”)

3. **How do content framing (balanced vs. positive-only) and source labeling in review summaries jointly affect consumer decisions?**

Experiment Context

We collaborated with a leading **automotive e-commerce platform** in Asia to conduct a **large-scale randomized field experiment**.

High Financial Stakes

Average vehicle price ~$62,000.
Consumers scrutinize every information cue, including **source labels** of review summary.

Multi-Stage Decision Funnel

Consumers go through [product detail page] → [dealer consultation] → [purchase].

Clear separation of **consideration** and **conversion stages**.

High Information Asymmetry

This makes consumers highly **dependent on review summaries** to evaluate products.

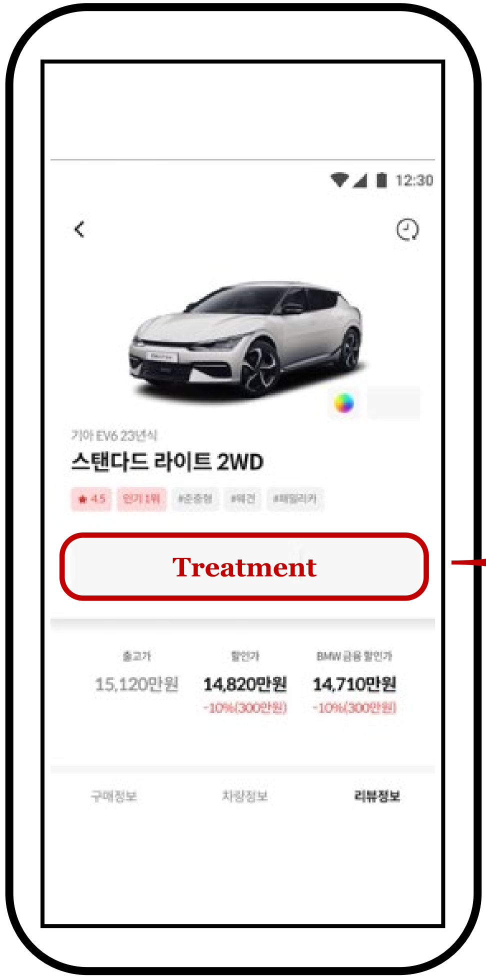
152,634
unique users

9 months
experiment period (Nov 2024 – Aug 2025)

7 conditions
2×3 factorial design + control

Field Experiment Design

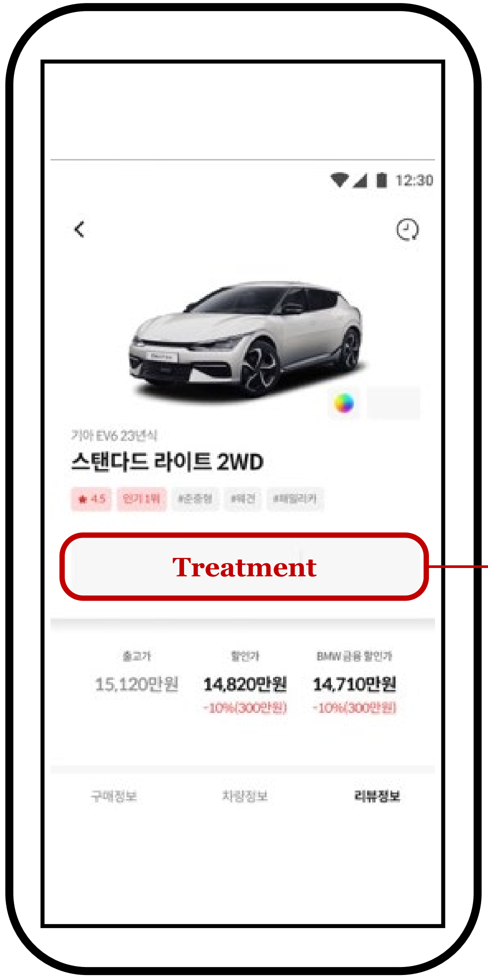
Product Page



Group	Condition	Source Labeling	Content Framing	Example
C	Control	**No Review Summary**		Carvana
G1	Treatment	None	Both Positive and Negative	Apple (App Store)
G2	Treatment	Human ("Human Editor-Generated")	Both Positive and Negative	
G3	Treatment	AI ("AI-Generated")	Both Positive and Negative	Amazon
G4	Treatment	None	Positive Only	CarMax
G5	Treatment	Human ("Human Editor-Generated")	Positive Only	
G6	Treatment	AI ("AI-Generated")	Positive Only	Newegg

Field Experiment Design

Product Page



Group	Source Labeling	Content Framing
G3	AI	Both Positive and Negative

AI-Generated Review Summary

- Strengths:

The vehicle offers **spacious interiors** and **comfortable rides**, making it great for families. Its design provides **more room** and **steady comfort** even on long trips.

- Weaknesses:

The vehicle has **high purchase** and **maintenance costs**, reducing affordability. Its large size makes **driving** and **parking difficult** in city areas.

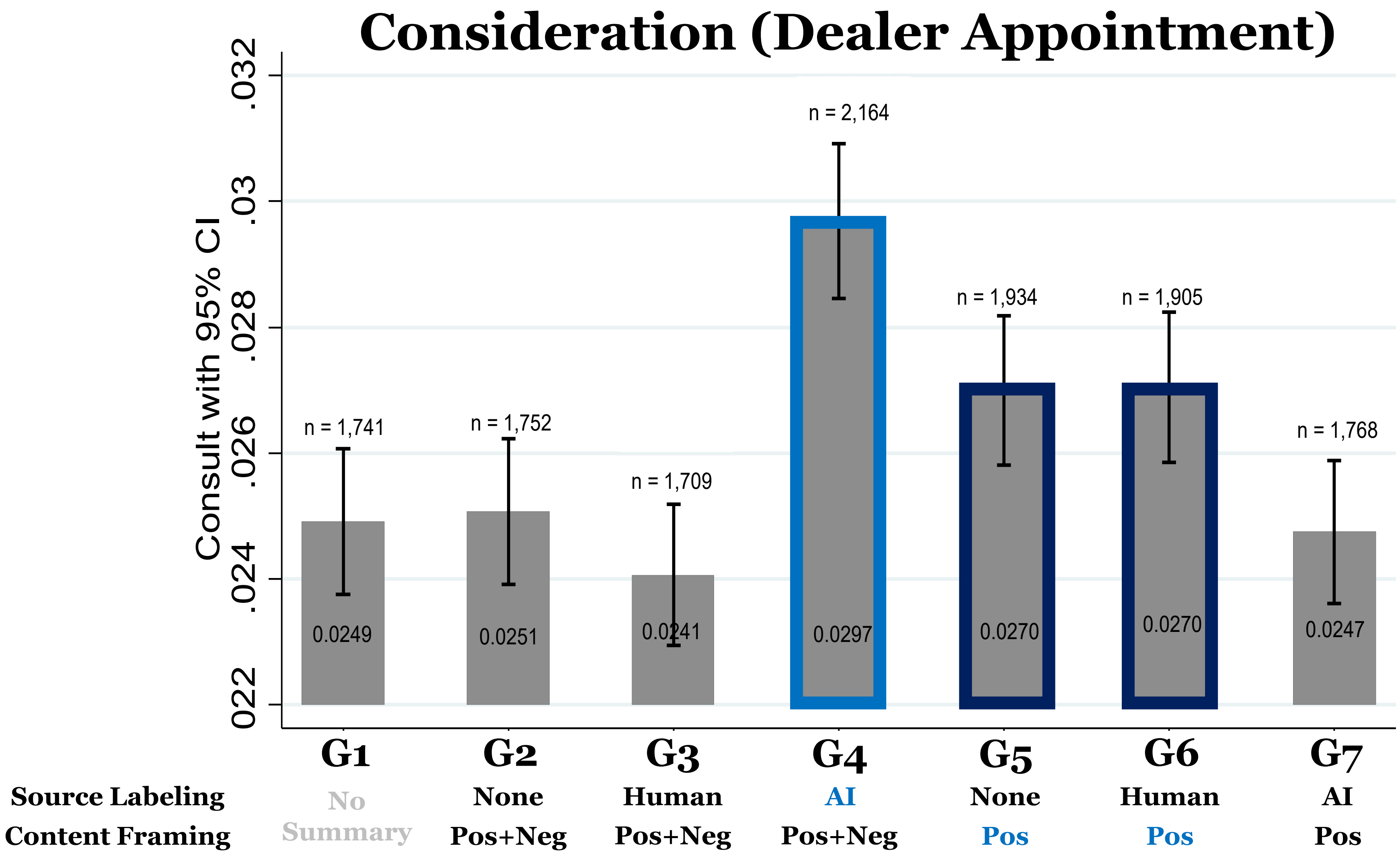
Data

- ***Nov 20, 2024 ~ Aug 30, 2025***
- ***152,634 individuals***

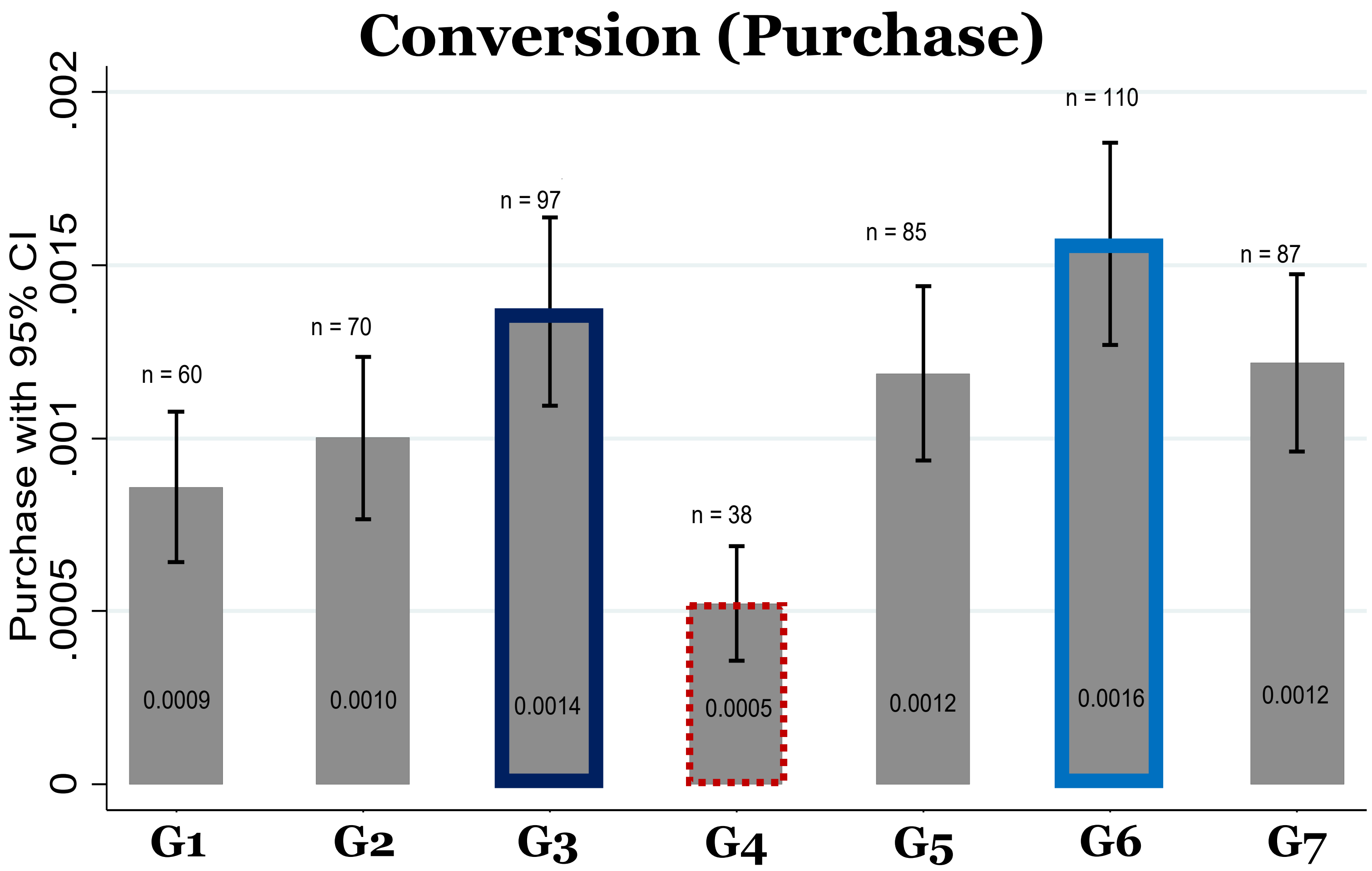
(1) randomized field experiment + (2) individual shopping click stream + (3) individual transaction

Category	Variable	Description	Mean	Min	Max
Dependent Variables	***Dealer Appointment***	A binary variable indicating whether a consumer (i) requested consultation for a car (j)	0.026	0	1
	Purchase	A binary variable indicating whether a consumer (i) purchased a car (j)	0.001	0	1
Control Variables	*Review Valence*	Average valence score of consumer reviews for a car (j) when a consumer (i) visited product page.	4.380	1	5
	Review Volume	The number of consumer reviews of a car (j) when a consumer (i) visited product page.	63.520	0	509
	Price	Fixed price of a car (j) (USD)	62,135.2	7,969.231	451,292.3
	Brand	A categorical variable of car (j) brand	12.435	0	34
	Historical Browsing	The number of cars searched by a consumer (i) prior to 6 months of experiment date (Nov 20, 2024)	3.643	0	73
	Historical Consult Frequency	The number of consult requests made by a consumer (i) prior to 6 months of experiment date (Nov 20, 2024)	0.014	0	4
	Historical Average Price	Average price of cars searched by a consumer (i) prior to 6 months of experiment date (Nov 20, 2024)	64,222.57	9,384.615	267,692.3

Model-Free Analysis: Effects on Dealer Appointment



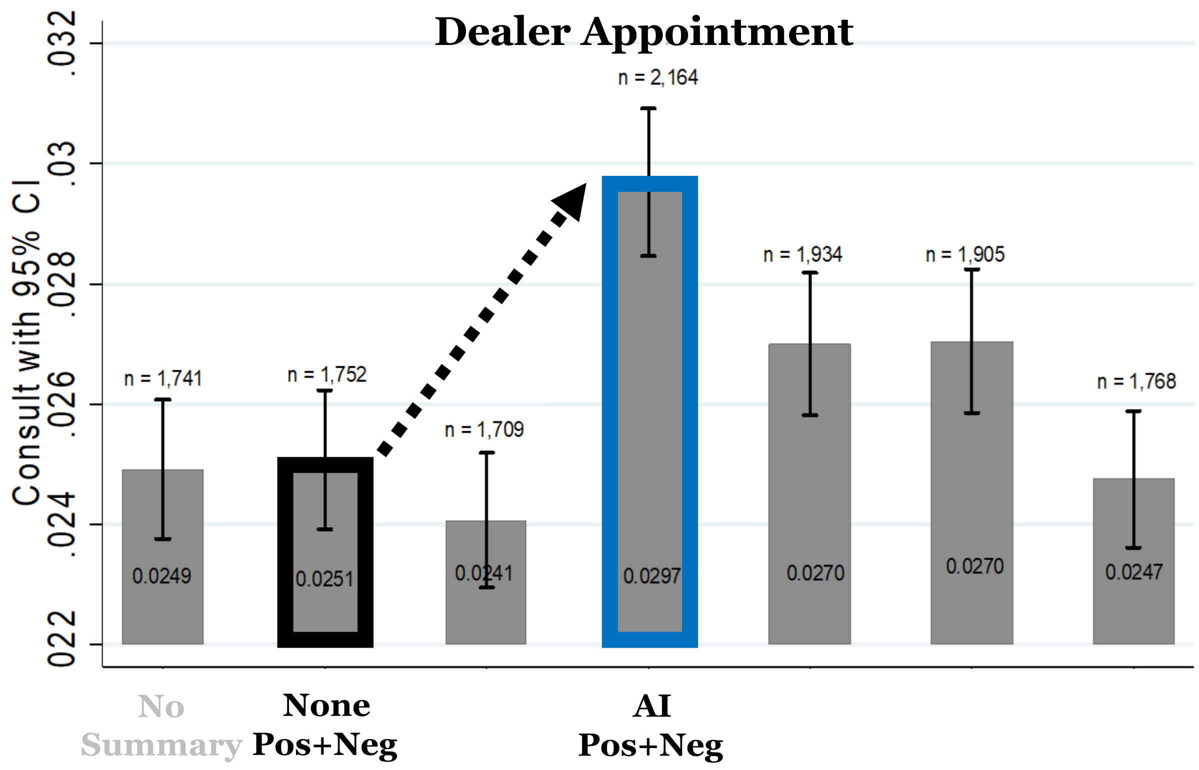
Model-Free Analysis: Effects on Purchase

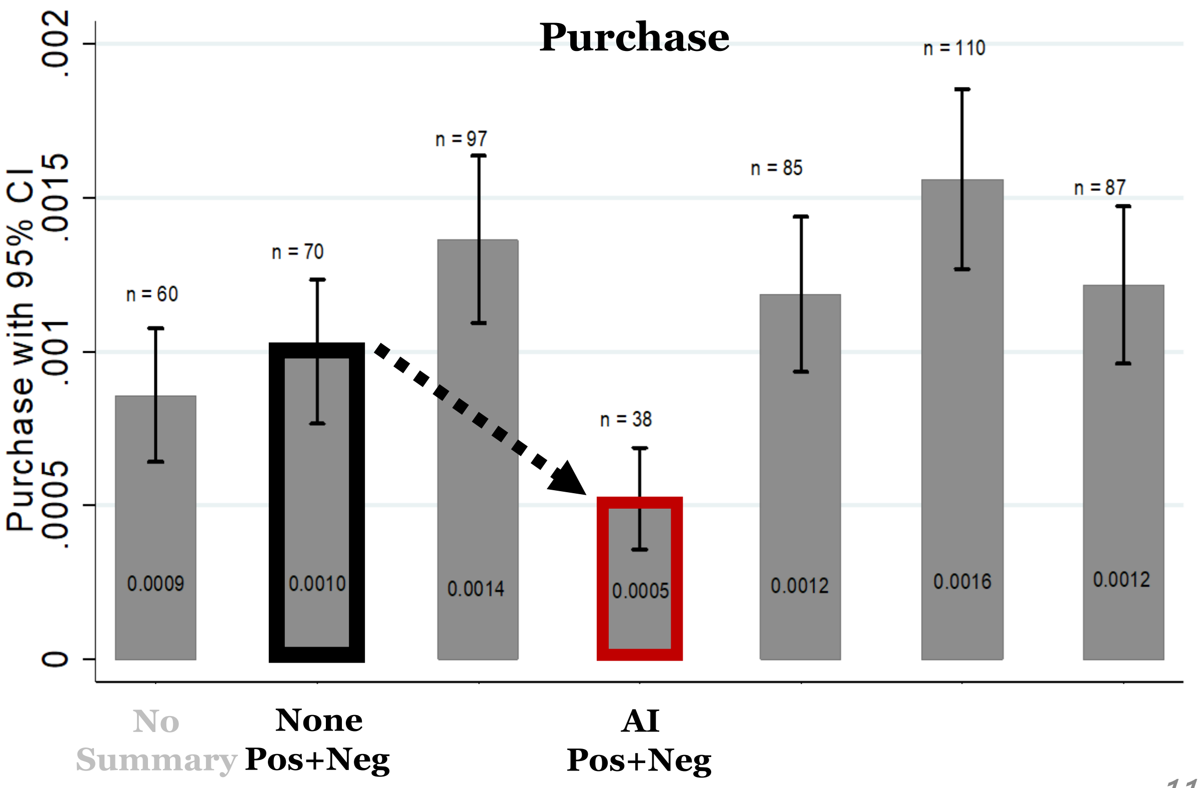


	G1	G2	G3	G4	G5	G6	G7
Source Labeling	No Summary	None	Human	AI	None	Human	AI
Content Framing		Pos+Neg	Pos+Neg	Pos+Neg	Pos	Pos	Pos

Key Finding (1): AI Disclosure Paradox

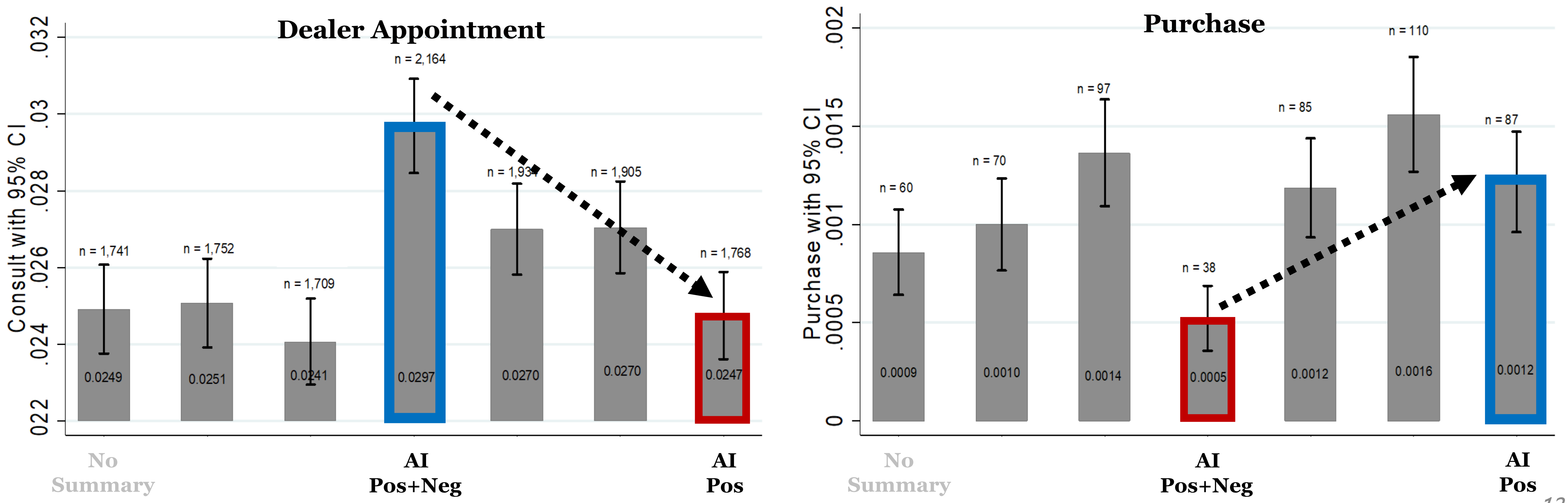
AI Disclosure Paradox: Disclosing AI authorship drives engagement in early decision stages (e.g., consideration) but reduces purchase at later stages (e.g., conversion).





Key Finding (2): Interaction Between AI Disclosure and Content Framing

AI disclosure effects differ depending on content framing: When the summary is positive-only, AI disclosure reduces early-stage engagement but boosts purchase at later stages.



Empirical Analyses: Effects on Consultation and Purchase

Table 1. Estimated Results for the Effectiveness of Review Summary (Baseline: Group 2)

Content Framing	Group	*Consideration (Dealer Appointment)*	*Conversion (Purchase)*
Both Positive and Negative	*Group 3 (Human-labeled Review Summary)*	-0.019	**0.085***
		(0.019)	(0.050)
	Group 4 (AI-labeled Review Summary)	**0.089*****	**-0.204*****
		(0.018)	(0.060)
Positive-only	*Group 5 (Positive-only Review Summary)*	**0.041****	0.040
		(0.019)	(0.051)
	Group 6 (Human-labeled & Positive-only Review Summary)	**0.031***	**0.129*****
		(0.018)	(0.049)
	Group 7 (AI-labeled & Positive-only Review Summary)	-0.007	0.057
		(0.019)	(0.051)
Control Variables		YES	YES
Constant		-2.595***	-1.963*
Log-likelihood		-48,155	-48,155
Observations		427,368	427,368

*Notes. Bivariate probit with correlated errors; jointly estimated. Standard errors are clustered at the individual level; Significant values are in boldface and * indication.*

Why Does the AI Disclosure Paradox Occur?

AI-
Generated

Common Starting Point

Consumers perceive AI as having systematically processed a large volume of consumer reviews.
This increases the **perceived credibility** of the entire summary, for both **strengths** and **weaknesses**.

Consideration

Dealer Appointment

AI credibility makes the product's **strengths** compelling enough to **consider seriously**, while the credibly reported **weaknesses** motivate consumers to **seek more information** before deciding.

Conversion

Purchase

At the point of committing ~$62,000, **negative evaluations** weighs disproportionately more (**loss aversion**). AI-flagged **weaknesses** are seen as **widespread, systematic issues**, amplifying downside risk and anticipated regret.

Key Insight: The same AI credibility that makes the product **worth exploring** also makes its weaknesses feel **too risky** to accept at the point of purchase.

Supporting Evidence: Moderating Effects by Review Rating and Volume

Higher **review ratings** and higher **review volume** **intensify** the **AI disclosure paradox**, strengthening both the positive effect on consideration and the negative effect on purchase.

Review Rating

For AI-labeled balanced summaries (G3)

The estimated coefficient of [G3 × ln(ReviewRating)]:

Consultation
+0.220**

Purchase
−1.269***

Higher ratings *create* ***a sharper contrast*** *between the product's strong reputation and AI-flagged weaknesses.*

Credibly presented strengths become even more compelling to explore (consideration ↑), while ***credibly presented weaknesses feel more consequential*** *against an otherwise positive backdrop (purchase ↓).*

Review Volume

For AI-labeled balanced summaries (G3)

The estimated coefficient of [G3 × ln(ReviewVolume)]:

Consultation
+0.023**

Purchase
−0.204***

More reviews *reinforce* ***the belief*** *that* ***AI has processed a larger, more comprehensive dataset****.*

This makes both the strengths and weaknesses in the summary feel more ***representative and data-supported****, strengthening the credibility mechanism in both directions.*

Policy Implications for Regulators

Blanket Mandates May Backfire

AI labeling increases product consideration but suppresses purchase.

A one-size-fits-all disclosure requirement may hurt platform economics.

Principle-Based Flexibility

Regulators should consider flexibility in *when* and *how* disclosure is phrased.

Hybrid formats (e.g., "*AI-generated* and *human editor-curated*") merit further investigation.

Stage-Dependent Disclosure Design

Strategies effective at the consideration stage may not work at the purchase stage.

Policy design should account for the *consumer's position* in the *decision funnel*.

We suggest that well-meaning disclosure mandates can produce ***unintended economic consequences*** *that deserve careful consideration in policy design.*

End of Presentation

THANK YOU!

andrewlee@utdallas.edu