

# Using GPT for Market Research

Ayelet Israeli

Harvard Business School

September 2023

with James Brand and Donald Ngwe

Office of the Chief Economist, Microsoft

# Generative Artificial Intelligence (AI) is becoming ubiquitous

- Large Language Models (LLMs)
  - Designed to understand and generate human-like language
  - Trained on vast amounts of text data
  - Generate text based on the most likely next words
    - E.g., Gmail sentence completion
  - Wide range of applications:
    - language translation, content generation
    - coding (Github Copilot), search (Bing, Bard)
  - OpenAI's Generative Pre-trained Transformer (GPT)

# GPT can improve productivity

AI

write the first slide of a presentation about "using gpt for market research"

## GPT can improve productivity

- Github Copilot improves the productivity (speed) of programmers by 55% (Peng et al., 2023)
- ChatGPT writes faster (by 0.8 SD), higher quality (by 0.4 SD) responses in professional writing tasks (Noy and Zhang 2023)

## GPT can simulate humans

- Replicates results of classic experiments in social psychology and economics (Aher et al., 2022)

Research shows how inputs impact GPT outputs:

- Partisan text generation and voting behavior, based on ideology inputs (Argyle et al., 2022)
- Results of classic behavioral economic experiments, with heterogeneity based on “beliefs” (Horton 2023)



GPT may be useful for social science research

# Research Questions

Can GPT be used as a tool for market research?

A. Are GPT responses aligned with fundamental properties of consumer demand?

B. Are GPT's responses (about demand) realistic and meaningful?

# Preview of the results

A. Are GPT responses aligned with fundamental properties of consumer demand?

- Yes: we document downward-sloping demand curves, impact of income on demand, and state dependence

B. Are GPT responses (about demand) realistic and meaningful?

- Realistic WTP distributions
- WTP for attributes consistent with recent conjoint studies with human consumers

# CONTENTS

- Introduction
- Empirical Approach
- Main Studies
  - Testing predictions from economic theory
  - Contextualizing GPT responses
- Conclusion
- Small Group Activity



# How can we learn about customer preferences?

- To conduct market research, we typically utilize data from many (randomly sampled) consumers
  - Focus groups, surveys, observational studies
  - Costly, time-consuming, labor-intensive, ideal data may not exist
- First question: can we use GPT to emulate this type of data collection?
- Answer: we can query GPT hundreds of times for each bundle of goods we consider, generating responses from a “random sample of consumers”
  - We maximize variation using GPT’s randomness input (“temperature”)
  - Training data likely include product websites and UGC such as blogs, reviews

# Our prompts aim to emulate a customer survey

- Ask GPT to fill in the response to a survey question as if it were a customer that was shopping in the category of interest and was randomly selected to participate in a survey
- Describe relevant features of the customer (e.g., annual income)
- Describe one or two products, as well as the option to opt-out
- Ask GPT to either make a choice or to provide WTP

We repeat each prompt hundreds of times to generate our sample

# We design six studies to answer our research questions

A. Are GPT responses aligned with fundamental properties of consumer demand?

- 1) Downward-sloping demand curve
- 2) Impact of income on demand
- 3) State dependence
- 4) Diminishing marginal utility of consumption

B. Are GPT responses (about demand) realistic and meaningful?

- 1) WTP for products
- 2) WTP for product attributes

# CONTENTS

- Introduction
- Empirical Approach
- Main Studies
  - Testing predictions from economic theory
  - Contextualizing GPT responses
- Conclusion
- Small Group Activity

# Study A1: Does GPT generate a downward-sloping demand curve?

- Binary choice: laptop or no purchase
- Choice among multiple options: two goods versus no purchase

# Study A1: First, we examine GPT's response to a binary choice

“A customer is randomly selected while shopping for laptops. Their annual income is \$70,000.

While shopping, the customer sees a Surface Laptop 3, Price: **PRICE**, Processor: Intel Core i5, RAM: 8GB, Screen Size: 13.5in, SD: 128GB

The customer is asked, after they finish shopping: Did you purchase any laptop? If so, which one?

Customer:”

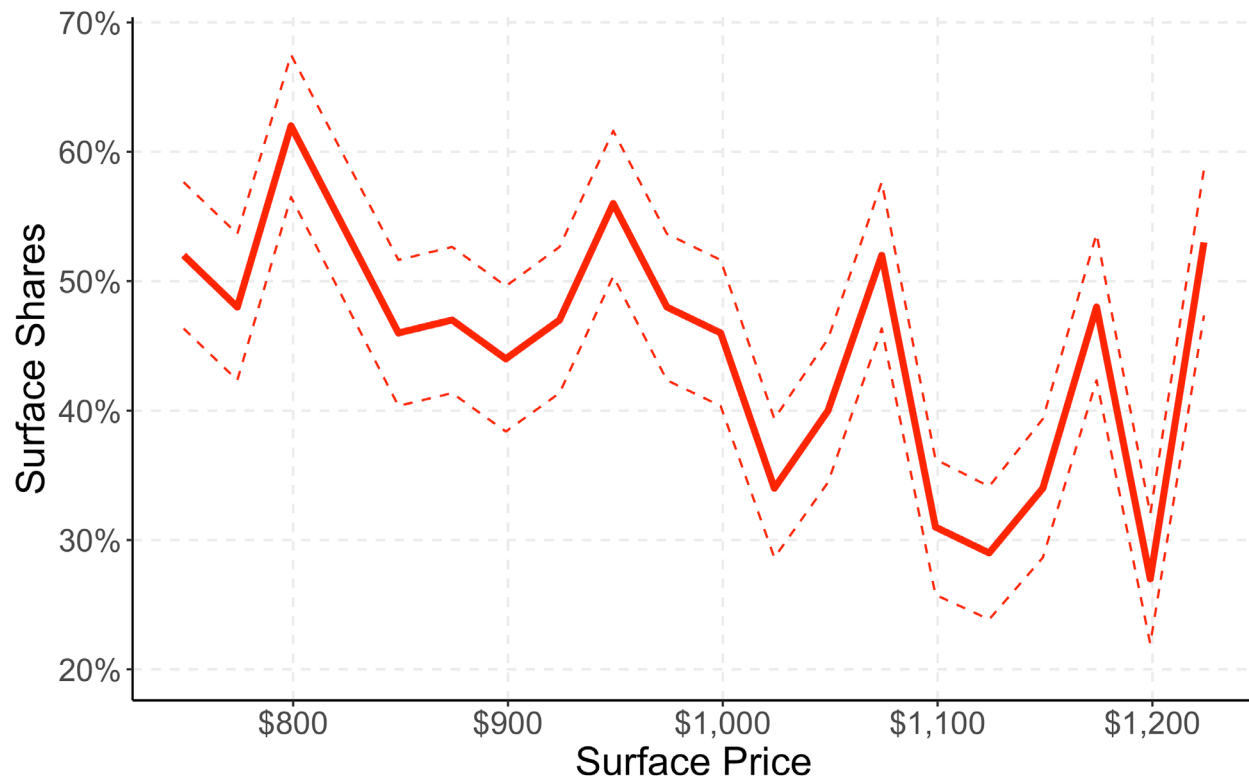
Where: **PRICE**  $\in$  [\$750...\$1,250]

We use OpenAI's GPT API with the “text-davinci-003” text completion model

# GPT generates hundreds of responses

- Yes, I purchased a Surface Laptop 3.
- I did purchase the Surface Laptop 3. It had a great processor, good RAM, and the screen size was perfect for my needs. The 128GB SD was a bonus too.
- No, I did not purchase any laptop. I was looking at the Surface Laptop 3 but ultimately decided it was not the right one for me.
- No, I did not purchase a laptop. I found the Surface Laptop 3 to be a bit too expensive for my budget.
- No, I didn't purchase any laptop. I was looking at the Surface Laptop 3, but I wasn't convinced it was a good value considering its specs.

# We aggregate the responses and plot choice shares



## INSIGHT:

GPT's demand is downward trending



## Study A1: We then include an alternative option in the prompt

“A customer is randomly selected while shopping for laptops. Their annual income is \$70,000.

While shopping, the customer has three options:

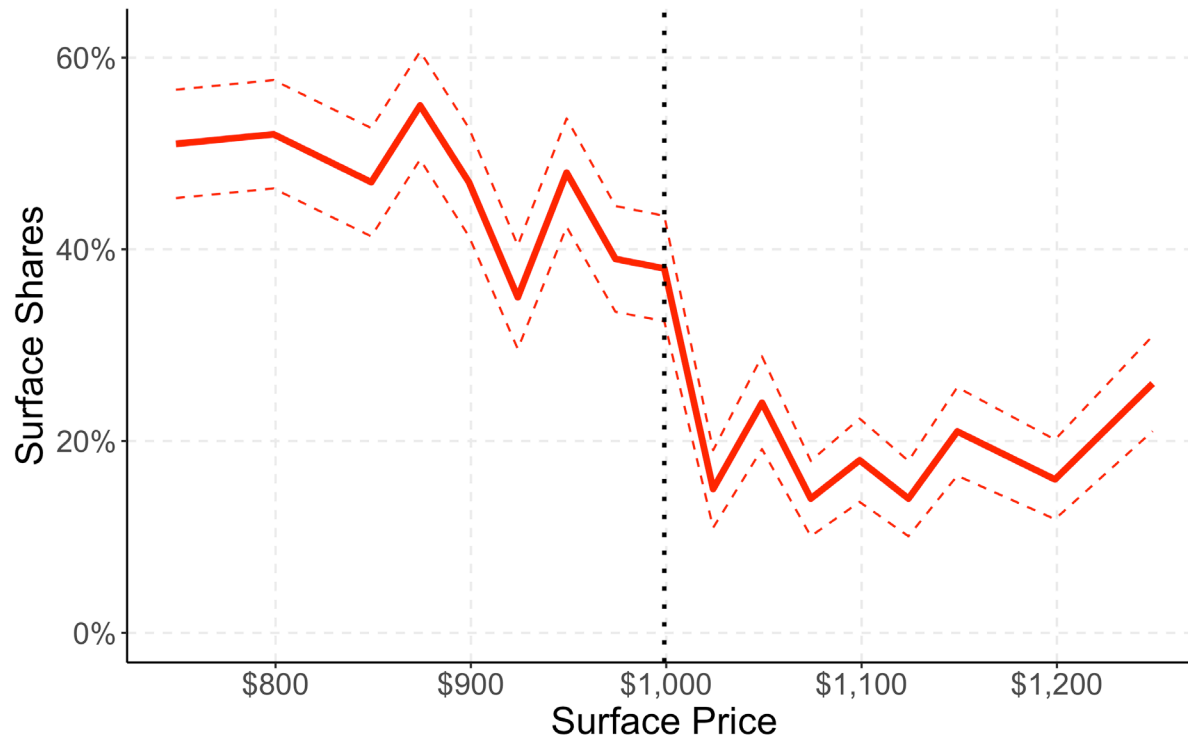
- a Surface Laptop 3, Price: **PRICE**, Processor: Intel Core i5, RAM: 8GB, Screen Size: 13.5in, SD: 128GB
- A Macbook Air (2019) Laptop, Price: \$999, Processor: Intel Core i5, RAM: 8GB, Screen Size: 13.3in, SD: 128GB

They also have the option not to purchase a laptop. The customer is asked, after they finish shopping: Which laptop, if any, did you purchase?

Customer:”

Where: **PRICE** ∈ [\$750...\$1,250]

# GPT is consistent with human behavior



## INSIGHTS:

- GPT is price sensitive
- Context matters to GPT
- Reference price matters to GPT
- It has brand preferences
- GPT gravitates toward the first option presented

## Study A1: We also examine supermarket goods

“A customer is randomly selected while shopping at the supermarket. Their annual income is \$70,000.

While shopping, the customer passes by the toothpaste aisle and sees two options:

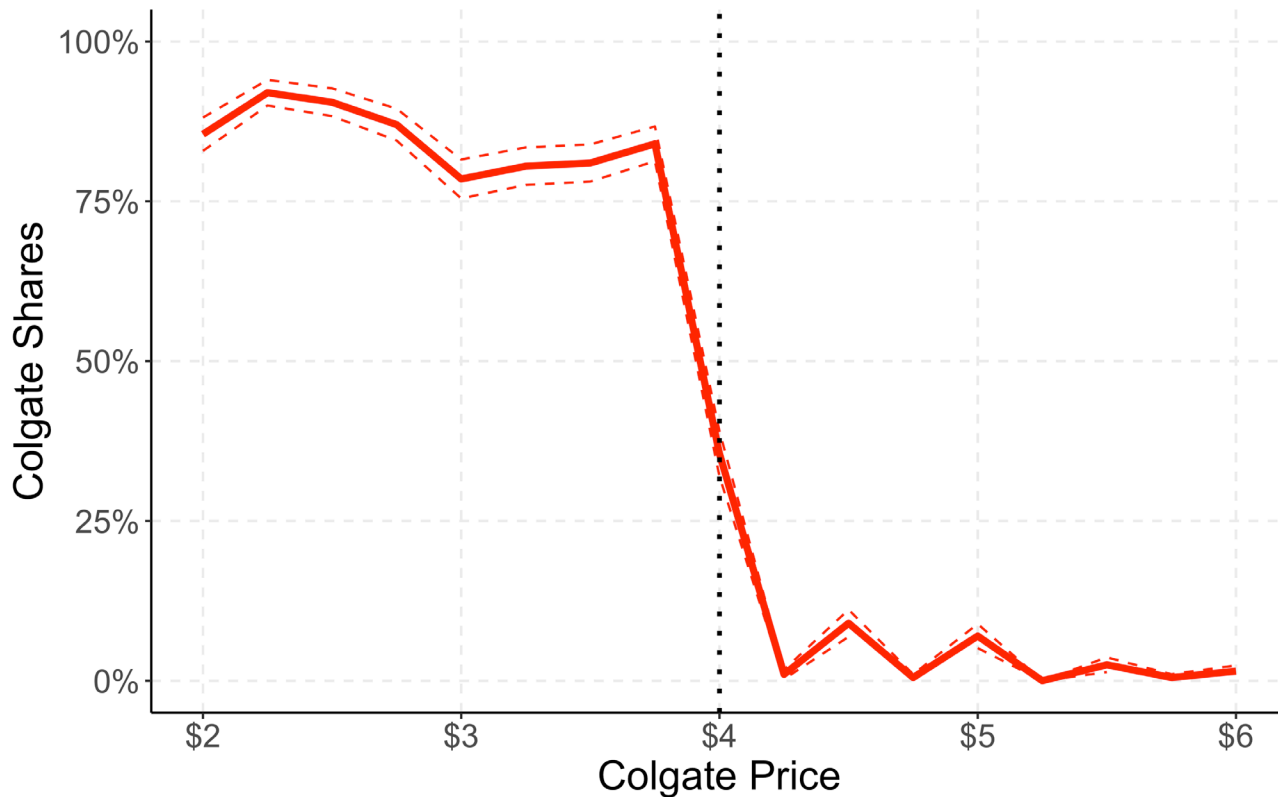
- Colgate whitening toothpaste with fluoride, price **PRICE**.
- Crest whitening toothpaste with fluoride, price \$4.

They also have the option not to purchase toothpaste. The customer is asked, after they finish shopping: Which toothpaste, if any, did you purchase?

Customer:”

Where: **PRICE**  $\in$  [\$2...\$6]

# Multiple options results – Colgate (vs. \$4 Crest)



## INSIGHTS:

- Toothpaste is not as differentiated as laptops for GPT
- GPT has diminishing sensitivity

## Study A2: Does income impact GPT's choice?

“A customer is randomly selected while shopping for laptops. Their annual income is **INCOME**.

While shopping, the customer has three options:

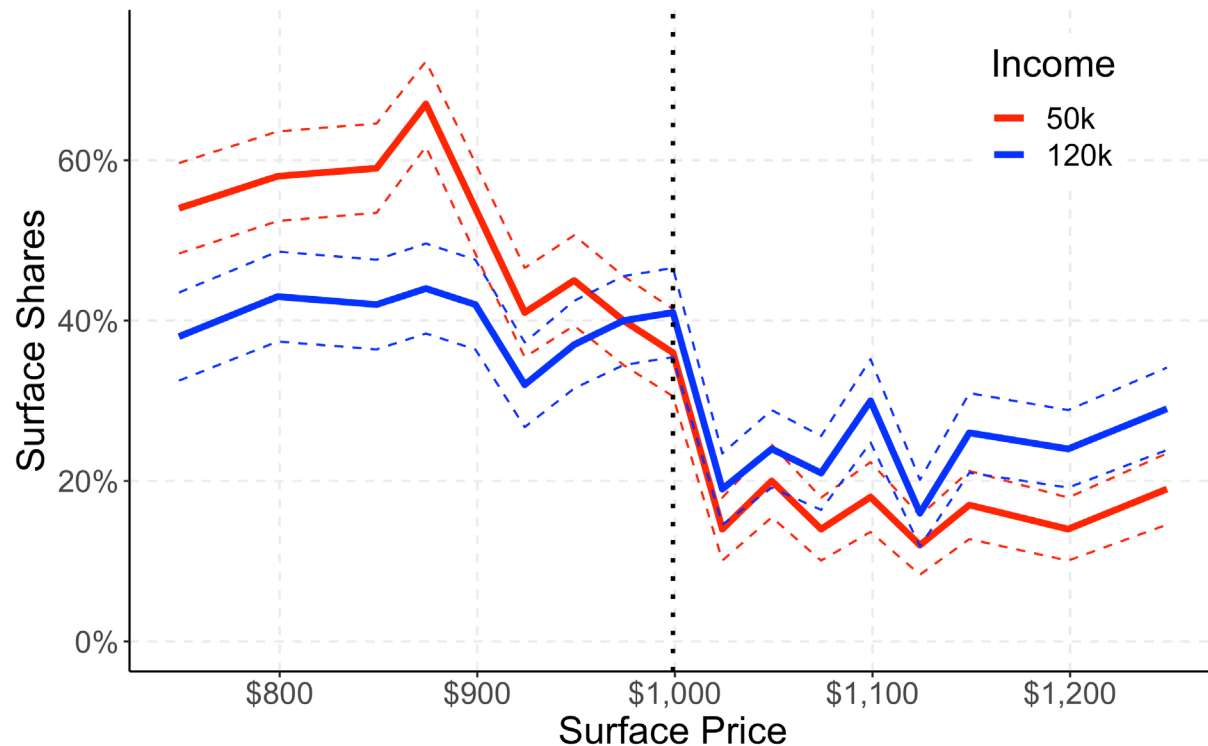
- a Surface Laptop 3, Price: **PRICE**, Processor: Intel Core i5, RAM: 8GB, Screen Size: 13.5in, SD: 128GB
- A Macbook Air (2019) Laptop, Price: \$999, Processor: Intel Core i5, RAM: 8GB, Screen Size: 13.3in, SD: 128GB

They also have the option not to purchase a laptop. The customer is asked, after they finish shopping: Which laptop, if any, did you purchase?

Customer:”

Where: **INCOME**  $\in \{\$50,000, \$120,000\}$  **PRICE**  $\in [\$750... \$1,250]$

# Study A2: Does income impact GPT's choice?



## INSIGHT:

GPT's price sensitivity depends on "its" income

## Study A3: Does past choice impact future choice?

“A customer is randomly selected while shopping at the supermarket. Their annual income is \$70,000.

While shopping, the customer passes by the toothpaste aisle and sees two options:

- Colgate whitening toothpaste with fluoride, price **PRICE**.
- Crest whitening toothpaste with fluoride, price \$4.

They also have the option not to purchase toothpaste. **This customer bought the Colgate whitening toothpaste last time they shopped for toothpaste.**

The customer is asked, after they finish shopping: Which toothpaste, if any, did you purchase?

Customer:”

## Study A3: Does past choice impact future choice?

“A customer is randomly selected while shopping at the supermarket. Their annual income is \$70,000.

While shopping, the customer passes by the toothpaste aisle and sees two options:

- Colgate whitening toothpaste with fluoride, price **PRICE**.
- Crest whitening toothpaste with fluoride, price \$4.

They also have the option not to purchase toothpaste. **The customer says that last time they shopped for toothpaste they purchased the Colgate whitening toothpaste.**

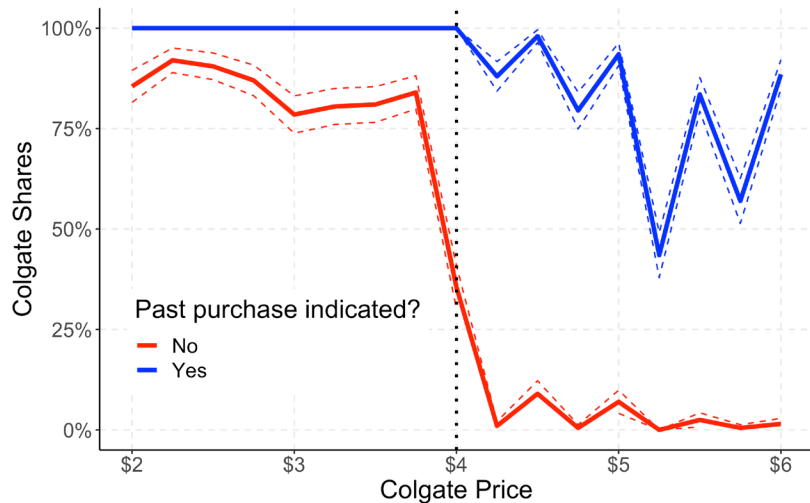
The customer is asked, after they finish shopping: Which toothpaste, if any, did you purchase?

Customer:”

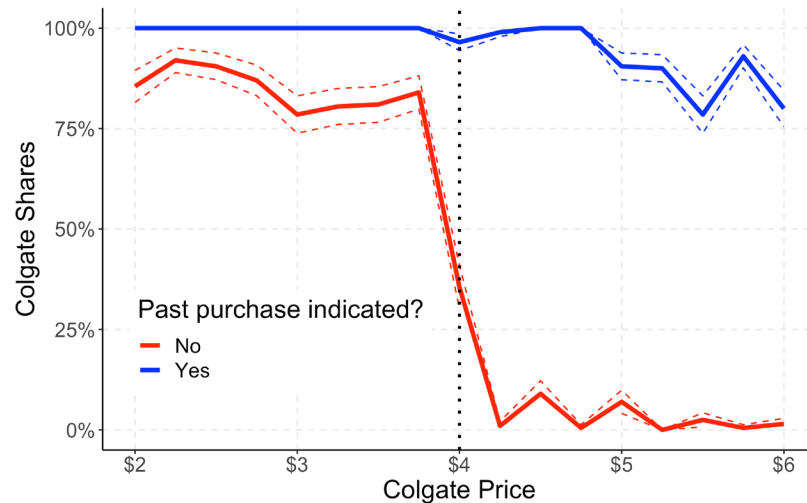


# GPT adapts to past choice

Stated as fact



Stated as “customer says”



# Overall, GPT responses correspond with fundamental properties of consumer demand

- 1) Downward-sloping demand curve
- 2) Impact of income on demand
- 3) State dependence
- 4) Minimal evidence of diminishing marginal utility of consumption

# CONTENTS

- Introduction
- Empirical Approach
- Main Studies
  - Testing predictions from economic theory
  - Contextualizing GPT responses
- Conclusion
- Small Group Activity

# Are GPT's responses realistic and meaningful?

- Study B1: Recovering realistic Willingness to Pay (WTP) for **products**
- Study B2: Recovering realistic Willingness to Pay (WTP) for **attributes**

## Study B1: Recovering realistic WTP for products

“A customer is randomly selected while shopping for laptops. Their annual income is \$70,000.

While shopping, the customer sees a Surface Laptop 3, Processor: Intel Core i5, RAM: 8GB, Screen Size: 13.5in, SD: 128GB

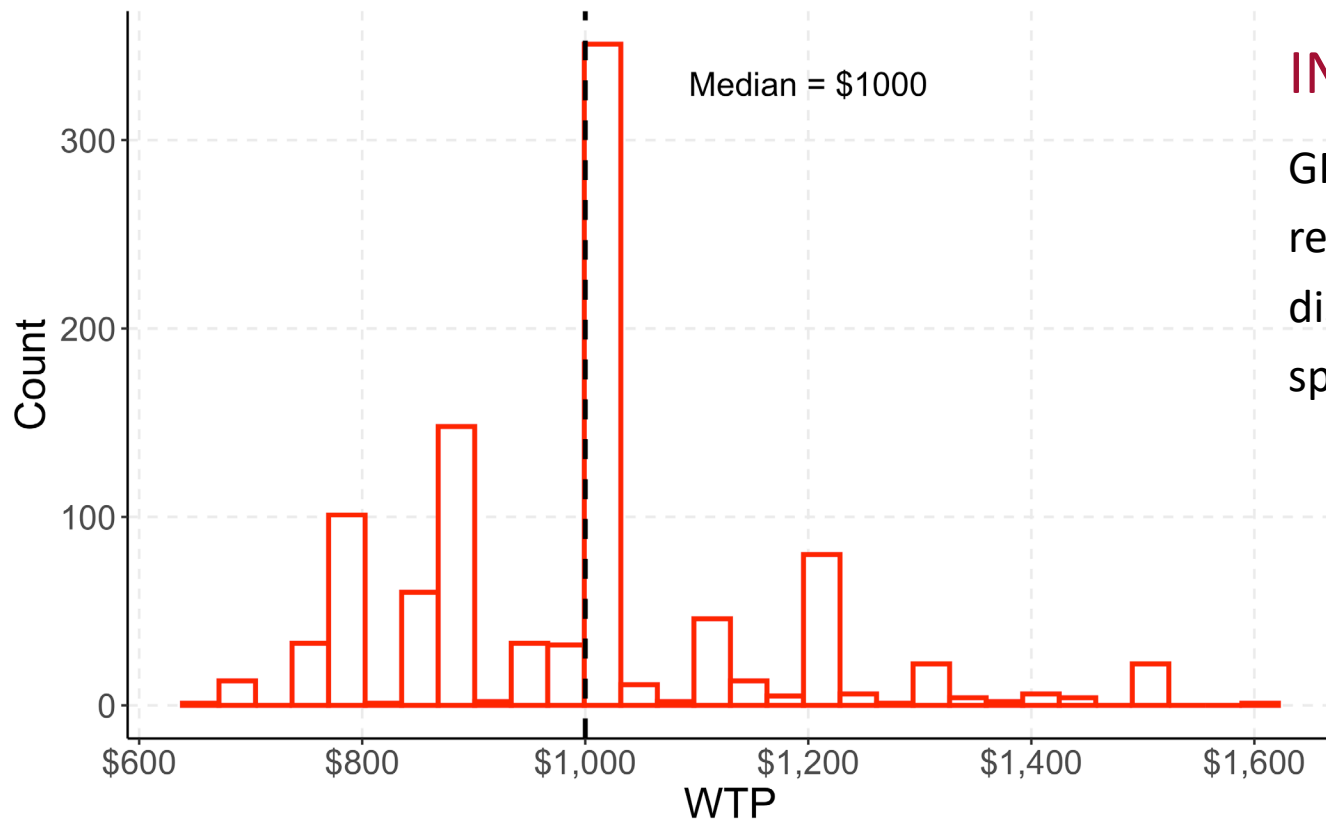
The customer is asked: what is the maximum price you would be willing to pay for this Surface laptop? please give a single price as your answer.

Customer:\$”

# GPT generates hundreds of responses

- 1,249.99
- 1,500. I think a laptop at this price range would have enough features to meet my needs without breaking the bank.
- 1,000-\$1,200
- 1,600 or less. My annual income is \$70k, so this laptop needs to be within my budget. If it is more expensive than that, I will look for something more affordable.
- 1,500 is the maximum price I would be willing to pay for this Surface laptop. I understand that this is a high-end laptop with excellent specifications and I am confident that I can acquire it for that price or lower. I will check online to compare prices and see if I can find a better deal.

# Recovering realistic WTP for products results



## INSIGHT:

GPT recovers a reasonable price distribution for a specific product

# Study B1: Recovering realistic WTP for general products

“A customer is randomly selected while **LOCATION**. Their annual income is \$70,000.

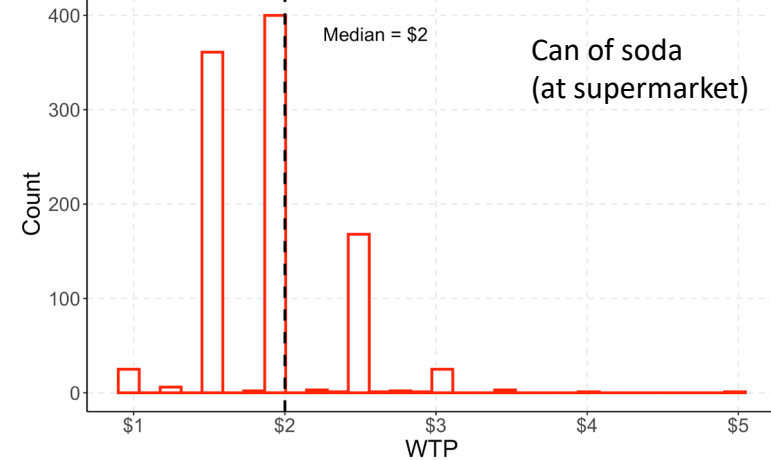
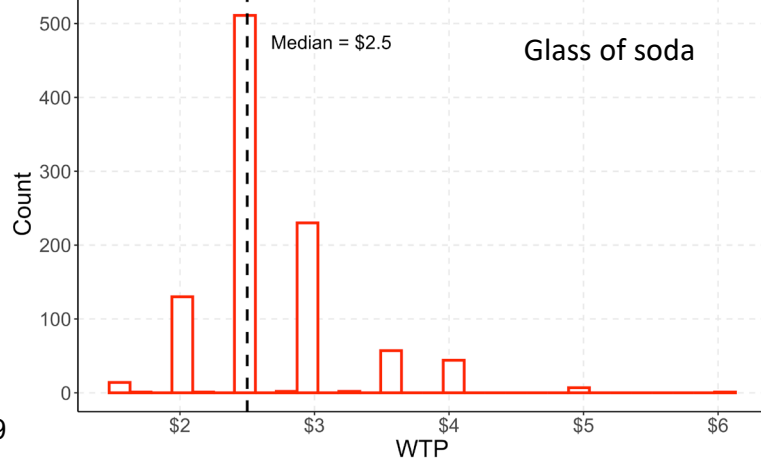
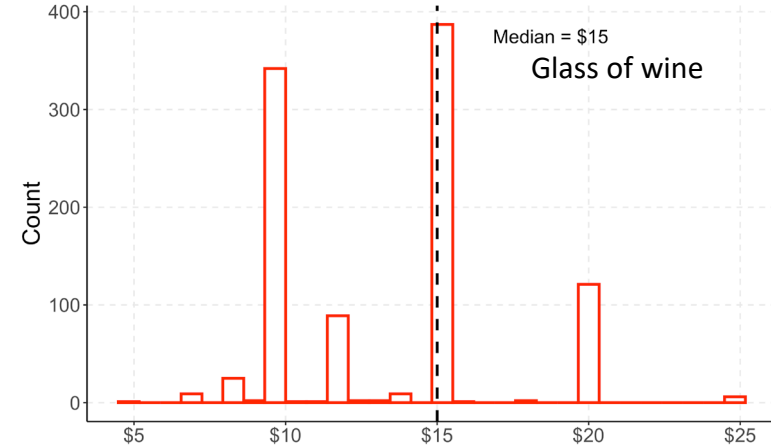
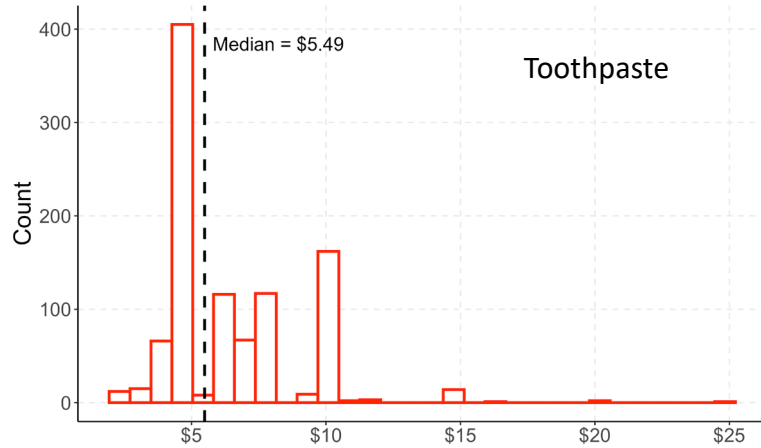
The customer is asked: What is the maximum price you would be willing to pay for one **unit/glass/can** of **PRODUCT**? please give a single price as your answer.

Customer:\$”

Where: **LOCATION**  $\in$  {Shopping at the supermarket, sitting at a restaurant};  
**PRODUCT**  $\in$  {toothpaste, wine, soda}



# Recovering realistic WTP for products results



# Study B2: Recovering realistic WTP for attributes

- Can we recover WTP for attributes of products?
  - Not likely to appear directly in GPT's corpus
  - Three approaches:
    - Direct solicitation (WTP prompts)
    - Indirect solicitation (via demand curves)
    - Conjoint analysis
- Comparing to Fong et al (JMR, forthcoming)'s results
  - In a different context, study WTP for fluoride in toothpaste, aluminum in deodorant
  - Conduct conjoint studies and thoroughly validate outcomes (prices, shares) with real-world data

## Study B2: WTP for attributes – direct solicitation

“A customer is part of a survey meant to elicit their willingness to pay for different attributes of goods. Their annual income is \$70,000.

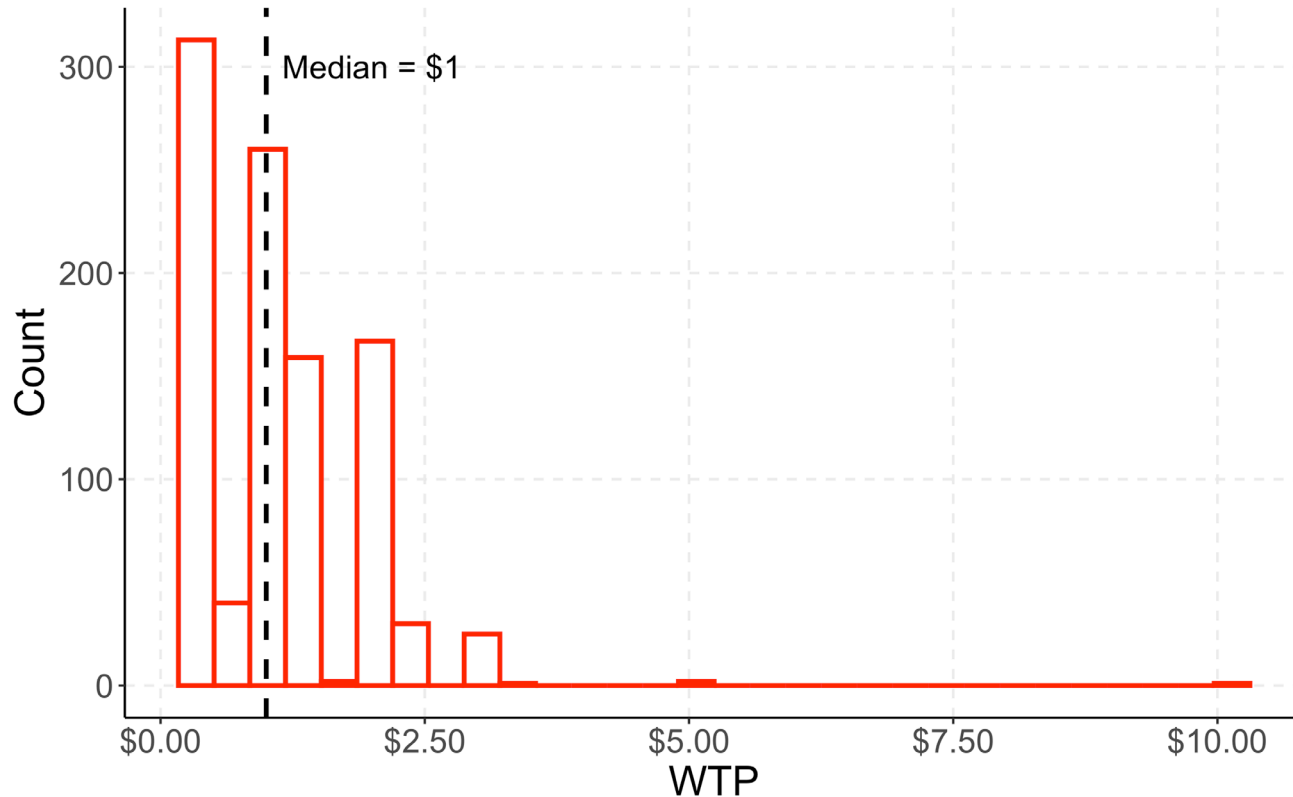
The customer is asked to consider two options:

- Option 1: Colgate toothpaste, **without fluoride**, whitening
- Option 2: Colgate toothpaste, **with fluoride**, whitening

The customer is then asked: ‘how much more would you be willing to pay for Option 2 than for Option 1?’ Please answer by giving an amount in dollars and cents.

Customer:\$”

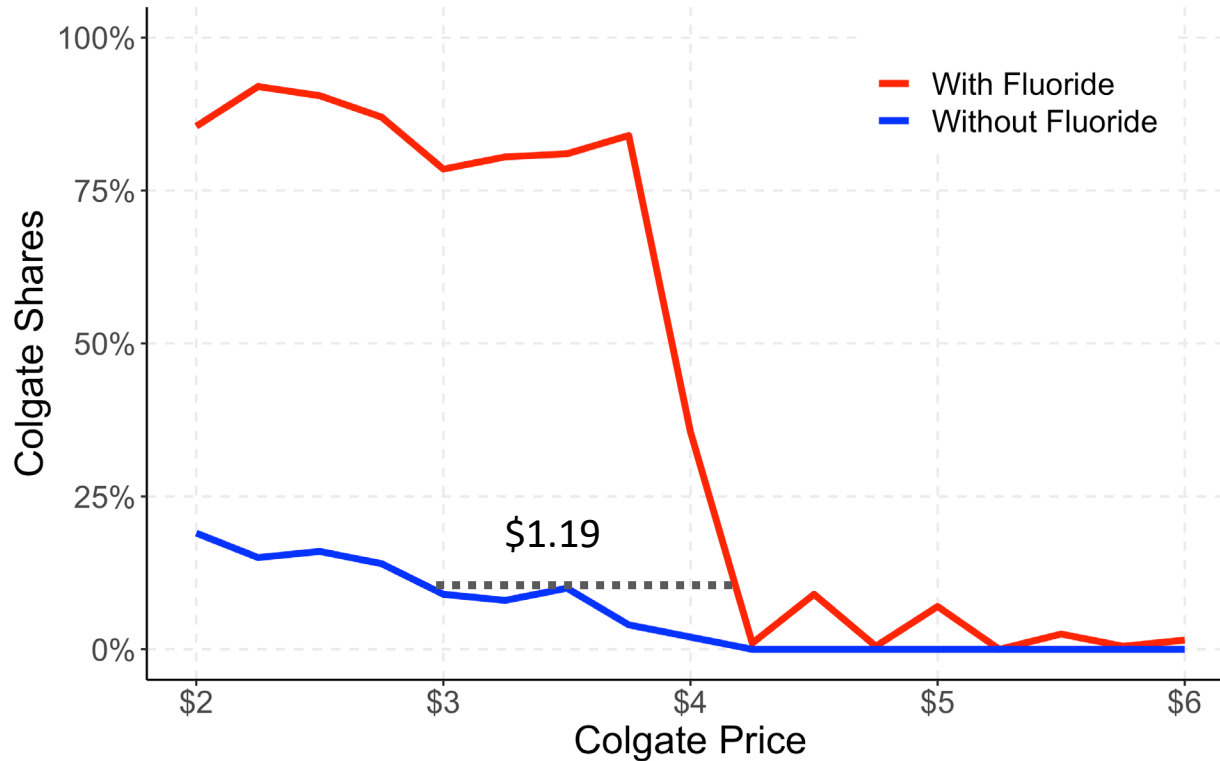
# WTP for attributes: direct solicitation results



# WTP for attributes: indirect solicitation via choice

- Two-step approach:
  - Estimate demand for Colgate with and without fluoride using the same paradigm as earlier, where the alternative option, Crest, has fluoride and the reference price is \$4
  - Compare the demand curve for Colgate without fluoride to the original demand curve (with fluoride) to derive the WTP for fluoride

The demand for Colgate without fluoride is much lower than the demand for Colgate with fluoride



# WTP for attributes: recovering preferences via conjoint

- We test whether the effect of price and non-price attributes on choice probabilities is consistent with economic predictions
- We use the responses to estimate a multinomial logit to evaluate the realism of the model-based WTP
- Instead of having to derive a few choice sets which are orthogonal and balanced across configurations, we use the full set of options

## Study B2: Recovering preferences for attributes via conjoint

“A customer is randomly selected while shopping at the supermarket. Their annual income is \$70,000.

While shopping, the customer passes by the toothpaste aisle and sees two options:

- Colgate whitening toothpaste **COLGATE\_WITH** fluoride, price **COLGATE\_PRICE**.
- Crest whitening toothpaste **CREST\_WITH** fluoride, price **CREST\_PRICE**.

They also have the option not to purchase toothpaste. The customer is asked, after they finish shopping: Did you purchase any toothpaste, if so, which one?

Customer:”

Where: **\_PRICE**  $\in \{0.99, 1.99, 2.99\}$ ; **\_WITH**  $\in \{\text{with}, \text{without}\}$ ;

We use a total of  $6 \times 6 = 36$  combinations



# Our results are consistent with human studies

- First, we corroborate that choices are consistent with economic theory and substitution patterns:
  - When Colgate price is low / Crest price is high: more likely to choose Colgate
  - When Colgate contains fluoride / Crest doesn't: more likely to choose Colgate
- We use the responses to estimate a multinomial logit to evaluate the realism of the model-based WTP
  - Our implied measure is \$3.4, consistent with Fong et al (2023) \$3.27.
  - We also confirm our approach for aluminum in deodorant.

# Comparing our approach to real-world conjoint studies

GPT-based conjoint	Real-world conjoint
Run all possible configurations	Run sample configurations, simulate additional configurations
1 GPT	100s+ humans
\$3 to collect <b>all</b> 10,800 responses 35 minutes <b>total</b>	>\$3 <b>per participant</b> >15 minutes <b>per participant</b>
The script that parses responses and generates csv files takes a couple of more minutes to run and verify	Data collected immediately for each participant

# CONTENTS

- Introduction
- Empirical Approach
- Main Studies
  - Testing predictions from economic theory
  - Contextualizing GPT responses
- Conclusion
- Small Group Activity

# Summary

- We used GPT as a random customer sampling mechanism and examined the distribution of its responses
- Our first set of results highlights that when prompted as a random customer, GPT exhibits behaviors consistent with economic theory
- Our second set of results demonstrates that GPT-based estimates are realistic and consistent with values obtained from existing research



GPT can serve as a powerful tool for understanding customer preferences

# What does this mean?

- At a minimum, GPT can serve as a realistic simulator of consumer choice
- GPT can be provided with additional “knowledge” (product attributes, customer personas, representative demographics, proprietary data), or fine-tuning to generate results
- As LLMs advance, we expect them to become more useful

## Limitations and words of caution

- GPT “hallucinates” and responses need to be examined critically
- Limitations of the model and training data: Pre-trained; bias;
- Use of GPT should be disclosed for ethical and external validity reasons



Slide:

### Conclusion

- In conclusion, GPT models have tremendous potential for revolutionizing market research by providing faster, more accurate insights into customer behavior and preferences.
- By leveraging GPT-based tools, businesses can gain a deeper understanding of their customers, identify trends and patterns, and develop more effective marketing strategies.
- However, it's essential to recognize the limitations and potential biases of the technology and use it in conjunction with other research methods to achieve the most accurate and comprehensive results.
- As GPT models continue to evolve and improve, we encourage businesses to explore their use in market research and stay abreast of the latest developments in this exciting and rapidly advancing field.

# CONTENTS

- Introduction
- Empirical Approach
- Main Studies
  - Testing predictions from economic theory
  - Contextualizing GPT responses
- Conclusion
- Small Group Activity

## Behind the research...

- How did we come up with our prompts?

## and beyond! [small group activity]

- What would you like to learn about your customers?
- How would you learn that using GPT?



# Discussion Questions

- When should we use GPT for market research questions? (and when not)
- What are the limitations and challenges that exist?
- What are the opportunities it presents?
- What should we consider when designing prompts (and evaluating GPT's responses)?

# Reflection Questions

- What would you like to learn about your customers?
- How would you learn that using GPT?
- Who in your organization can help?
  - Implement
  - Iterate
  - Evaluate the responses

# Thank You!

[aisraeli@hbs.edu](mailto:aisraeli@hbs.edu)

## Questions / Comments?