MKT927: INTRO TO QUANTITATIVE MARKETING

Prof. Andrey Fradkin

Lecture 1: Introduction to Class and Quant Marketing

INTRODUCTIONS

SYLLABUS

REQUIREMENTS

- This class has students with a variety of backgrounds.
- We will focus more on topics rather than deep technical details.
- Some basic fluency in econometrics and programming is assumed. If you're not sure, talk to me.
- The key requirement is willingness to participate in class and engage with the material.
- This is the first time I'm teaching this class. Feedback is very needed. Am I going too fast? Too slow? Are the assignments too difficult? Are there things I should cover?

CLASS STRUCTURE

- We have 14 sessions this semester.
- Most sessions:
  - Lecture by me
  - Presentation by student
  - Discussion throughout
- Your responsibility: do the readings and assignments!

TOPICS WE WILL COVER

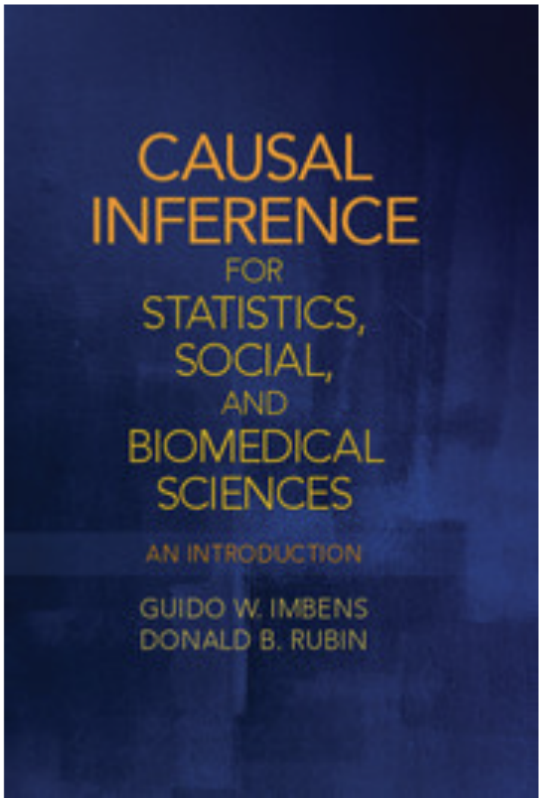
- Advertising (2 weeks)
- Demand modeling / Industry modeling (2 weeks)
- Pricing (1 week)
- Digital Platforms (4 weeks)
- AI (3 weeks)
- Methods: Experiments, Observational Causal Inference, Structural. But not deeply into each one
- The craft of research

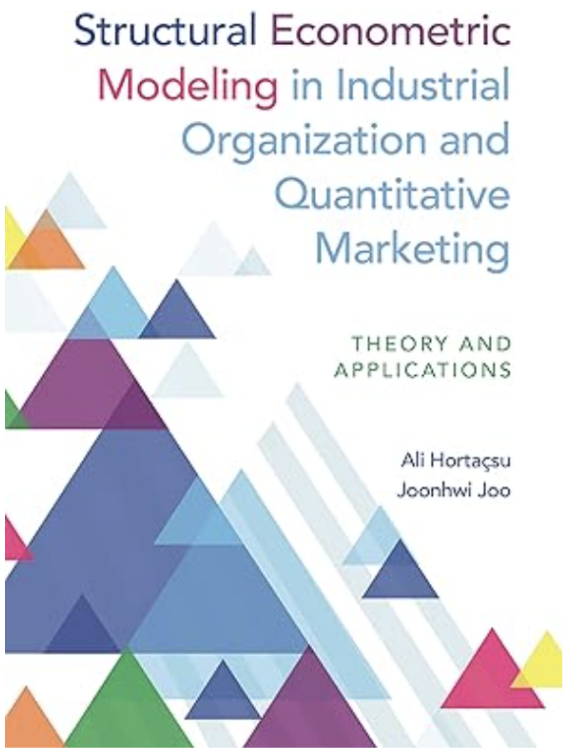
TOPICS WE WON'T COVER

- Salesforce management
- Branding
- Customer lifetime value
- Bayesian methods
- Dynamic discrete choice modeling
- Training neural networks

TEXTBOOKS

- No required textbook.
- Optional but recommended:





GRADING

- Class participation: 15%
- Two presentations of assigned papers: 22%
- Four Assignments: 28%
- Paper proposal and presentation: 35%

READING AND CLASS PARTICIPATION

- Required readings are denoted with a *
- Readings to skim or pick to present are denoted with a **. Also make sure to at least skim any paper discussed in class.
- If you'd like to present something not on the syllabus, we can discuss at least two weeks beforehand.
- Good participation occurs when you've done the reading.
- More on how to read a paper later.

PAPER PROPOSAL: TIMELINE

- Two page proposal by March 26.
- Present the proposal on the last day of class (April 30).
- Feel free to schedule office hours with me whenever (just email me).

SYLLABUS PAPER PRESENTATIONS

1. Reading and presenting papers are critical skills.
2. Each student will sign up to present twice this semester.
3. Presentations should be 30 minutes long.
4. Presentations should cover (not necessarily in this order):
   - What is the research question?
   - Why is this question interesting / important (or is it?)?
   - How was the dataset constructed?
   - Key results.
   - Methods that drive key results.
   - A critical evaluation of the above.

ASSIGNMENTS

- Goals:
  - Answer conceptual questions.
  - Implement and use methods used in quantitative marketing papers.
  - Interpret research results.
- Assignments will have a programming portion. I urge you to start this 2 weeks in advance of the due date. CB students are allowed to skip the programming question.
- Four of these.
- GenAI: You are allowed and encouraged to use LLMs. That said, there are many things that it is good to work on yourself without asking for help as these are critical skills for you as a researcher.

PAPER PROPOSAL

- Option 1: Submit a proposal with a clear research question, dataset, and methods. If you have not yet or cannot obtain the data for the course, I expect you to simulate a dataset that would look like the one you'd use for the paper.
- Option 2: Conduct a replication of a published quantitative marketing paper. Write your own code to reanalyze the main results of the paper, and augment it with anything you feel is missing.
- CB Students Only: You can conduct a literature review of a quantitative marketing topic of your choice, and suggest how a consumer behavior approach could be used to address the questions within this topic area.

QUESTIONS?

TODAY'S AGENDA

- What is quant marketing?
- What should quant marketing researchers do?
- What methods will we use?
- Where to get data.
- Start on potential outcomes if we have time.

WHAT IS MARKETING?

- It is VERY loosely defined.
- Traditionally: 4Ps (Price, Promotion, Product, Placement)
- As distinct from economics: More focus on firm and consumers versus policy and the market equilibrium.
- As distinct from psychology: More focus on applied problems (how to sell a product, etc...).
- As distinct from operations: More focus on empirical work and the consumer. Less focus on optimization of warehouse, inventory, processes, etc...
- Great news, you can study almost anything and be a quant marketing researcher.

WHY MARKETING?

- “Marketing is the new finance” - Hal Varian
- Advertising is a key revenue source for the biggest companies in the world.
- “Growth”, “Customer Engagement”, “Get to Market” -> All huge aspects of corporations. Deserve to be studied rigorously.
- For many economic phenomena, diffusion is the key constraint.

MELA, CARL F., JASON M.T. ROOS, AND YITING DENG (2013), "A KEY WORD HISTORY OF MARKETING SCIENCE," MARKETING SCIENCE, 31, 1 (JANUARY-FEBRUARY), 8-18.

Table 1 **Most Frequently Used Keywords in *Marketing Science*, 1982–2011**

Keyword	Frequency	Keyword	Frequency
▲ Pricing	122	Diffusion	37
Game theory	89	Hierarchical Bayes	37
▲ Advertising	61	Buyer behavior	27
△ Choice models	59	Forecasting	27
▲ Channel	58	Conjoint analysis	26
△ Competitive strategy	52	Customer satisfaction	25
Econometric models	50	Price discrimination	23
▲ Promotion	46	Market structure	22
▲ New products	46	Marketing mix	22
△ Brand choice	45	Marketing strategy	22
▲ Retailing	45	Structural models	22
△ Competition	43	Internet marketing	21
Bayesian analysis	38	Motion picture	21

Notes. Among the words used more than 40 times, the four P's are marked with a closed triangle (▲) and the three C's with an open triangle (△).

VIRTUAL QUANTITATIVE MARKETING SEMINAR TOPICS

WHAT SHOULD WE ATTEMPT TO DO AS RESEARCHERS?

Many Perspectives.

- Science
- Story Telling (Rubenstein)
- Engineering (Roth)
- Plumbing (Duflo)

Constant: Be rigorous.

THE SCIENCE PERSPECTIVE

- Form hypotheses, gather data, and test hypotheses. Iterate.
- Learn fundamental truths about the world.
  In our case, how humans and firms behave and how they should behave to accomplish specific objectives.
- Explain why things happen.

CONTRA THE SCIENCE PERSPECTIVE

- Much research consists of “scientism” or “cargo cult” science.
- Hypotheses are formed post-hoc, after looking at the data, but papers are written in the opposite way.
- Behavior is highly context specific, but papers claim relevance based on generalizability / external validity.
- My take: We should strive towards science, but know that finding general truths about the world by doing empirical work is hard. Need modesty and humility about research claims.

ECONOMIST AS STORY TELLER / THEORIST (RUBENSTEIN)

A good model in economic theory, like a good fable, identifies a number of themes and elucidates them. We perform thought exercises that are only loosely connected to reality and that have been stripped of most of their real-life characteristics. However, in a good model, as in a good fable, something significant remains.

ECONOMIST AS ENGINEER (THE AL ROTH PERSPECTIVE)

Economists have lately been called upon not only to analyze markets, but to design them. Market design involves a responsibility for detail, a need to deal with all of a market's complications, not just its principle features. Designers therefore cannot work only with the simple conceptual models used for theoretical insights into the general working of markets. Instead, market design calls for an engineering approach. Drawing primarily on the design of the entry level labor market for American doctors (the National Resident Matching Program), and of the auctions of radio spectrum conducted by the Federal Communications Commission, this paper makes the case that experimental and computational economics are natural complements to game theory in the work of design. The paper also

Consider the design of suspension bridges. The simple theoretical model in which the only force is gravity, and beams are perfectly rigid, is elegant and general. But bridge design also concerns metallurgy and soil mechanics, and the sideways forces of water and wind. Many questions concerning these complications can't be answered analytically, but must be explored using physical or computational models. These complications, and how they interact with the parts of the physics captured by the simple model, are the domain of the engineering literature. Engineering is often less elegant than the simple underlying physics, but it allows bridges designed on the same basic model to be built longer and stronger

AL ROTH'S CAREER IS A MODEL

- Concern for important real world problems.
- Complementarity between methodological approaches (theory, experiments, observational data, computation).
- Design is an active process where things like speed, tinkering, and politics matter.
- Details matter. In the residency match: couples, unfilled positions, etc... Can't just ignore these because model doesn't count them.
- Real world impact.

ECONOMIST AS PLUMBER (DUFLO PERSPECTIVE)

> "Plumbers try to predict as well as possible what may work in the real world, mindful that tinkering and adjusting will be necessary since our models gives us very little theoretical guidance on what (and how) details will matter. This essay argues that economists should seriously engage with plumbing, in the interest of both society and our discipline."

ECONOMIST AS PLUMBER (DUFLO PERSPECTIVE)

In the early 2000s the city of Boston decided to change how it assigned students to schools. There was an important engineering part to choosing the mechanism that would be used, initiated by Abdulkadiroglu and Sonnmez (2003) and followed by a considerable literature (see Pathak (2011) and Abdulkadiroglu and Sonnmez (2013) for reviews). But once city leaders settled on a mechanism, they still had to make many decisions. How to communicate the change to parents? How to persuade them that they could reveal their true preferences when they ranked schools? Should there be a limit on how many schools parents can (or are required to) rank? Should parents living near a school receive preferential treatment, and how should that be set up.

HOW THIS APPLIES TO MARKETING

- Research should be useful to someone. If that someone is a manager or policymaker, we have to take practical concerns seriously.
- Great research ideas often come when talking with practitioners, consuming their media (blogs, podcasts), actually working with them.
- Example: pricing.
  Suppose we calculate the "optimal" price. We still need to communicate it to the consumers and employees. We need to change internal processes. We need to think about bundles and seasonal discounts. Etc...

HOW IS RESEARCH DIFFERENT FROM CONSULTING?

- If we help a company set a price, when is that research and not consulting? Some heuristics:
  - New and non-obvious. Could be a method or finding. Non-obvious can be direction (positive or negative) or size (is the effect huge and precisely estimated? Is it a precise 0?).
  - Useful beyond the narrow setting. A simple experiment varying the price of Snickers is useful to Mars, but is unlikely to tell us much about pricing more generally.
  - Takes place in a very important market (semiconductors, big tech, schooling, etc...).
  - Studies a question of policy interest that companies may purposefully not study (regulation, discrimination, etc...).

METHODS

THEORY

- Oftentimes ignored or an afterthought in empirical work.
- But people generally have a theory in mind. Formal theory makes sure that this theory is internally consistent.
- Tracing out the logical implications of a theory can yield surprise and insight.
- "Market Design" style work typically requires theory and empirics.

EMPIRICAL WORK IN SOCIAL SCIENCE: DESCRIPTIVE

Measuring important phenomena.

- What is the average price level and how did it change over time?
- What is the total ad spending by channel?
- What are the market shares of different companies in an industry?
- How much time do people spend looking at social media per day?

Critical for science, but underappreciated in the publication process. Some of the most impactful papers are ones that introduce a new dataset and describe what it measures.

EMPIRICAL WORK IN SOCIAL SCIENCE: "REDUCED FORM"

What was the causal effect of X on Y in our dataset? Most common research style in top quantitative marketing journals.

- How does one more review affect revenue?
- Does retargeting increase sales?
- Does becoming unemployed change spending behavior?
- Did news feed algorithm A increase metrics versus algorithm B? By how much?

EMPIRICAL WORK IN SOCIAL SCIENCE: "STRUCTURAL"

An empirically grounded model of agent behavior. What would have happened if we changed factor X in a way that did not occur in the data? What is the optimal way to set X?

From my work:

- If we regulated Airbnb by banning it, how would it affect consumer surplus, host surplus, hotel prices, and hotel revenues?
- What if we changed the tax rate?
- What if we instituted caps on the number of nights per property?

A TYPICAL STRUCTURAL ABSTRACT: BARAHONA ET AL. (2023, ECMA)

Abstract

We study a regulation in Chile that mandates warning labels on products whose sugar or caloric concentration exceeds certain thresholds. We show that consumers substitute from labeled to unlabeled products—a pattern mostly driven by products that consumers mistakenly believe to be healthy. On the supply side, we find substantial reformulation of products and bunching at the thresholds. We develop and estimate an equilibrium model of demand for food and firms' pricing and nutritional choices. We find that food labels increase consumer welfare by 1.8% of total expenditure, and that these effects are enhanced by firms' responses. We then use the model to study alternative policy designs. Under optimal policy thresholds, food labels and sugar taxes generate similar gains in consumer welfare, but food labels benefit the poor relatively more.

EMPIRICAL WORK IN SOCIAL SCIENCE: PREDICTIVE

Decision makers need predictions in order to make complementary decisions. Applies statistics, machine learning, “AI” to forecast something. Examples in marketing research:

- “Can AI and AI-Hybrids Detect Persuasion Skills? Salesforce Hiring with Conversational Video Interviews”
- “Customer-Based Corporate Valuation for Publicly Traded Non-Contractual Firms”
- “When Words Sweat: Identifying Signals for Loan Default in the Text of Loan Applications.”

MODELS, MEASUREMENT, AND THE LANGUAGE OF EMPIRICAL ECONOMICS (HAILE)

Examples of a structure:

- $Y = X\beta_0 + \epsilon$, with a distribution of β_0 and $\epsilon|X$.
- Equilibrium model with parameter values and distributions.

Examples of a model:

- $Y = X\beta_0 + \epsilon$, unknown β_0 and $\epsilon \perp X$.
- Equilibrium model with demand and supply, unknown parameters and exogeneity conditions.

CAUSAL EFFECTS, COUNTERFACTUALS, AND STUCTURE (HAILE)

A causal effect is the response of an outcome Y to an intervention that altered X (but nothing else (ceteris paribus)).
A randomized control trial gives a causal effect only under certain maintained assumptions.

- For example, excludability and non-interference. We'll return to these.

Haile's point. Even what is called "Reduced Form" work requires structural assumptions. Just the structure is simpler and less driven by economic theory.

IDENTIFICATION (HAILE)

Very frequently discussed in papers. It has a precise meaning, often loosely discussed.

A data generating process implies a joint distribution of observables.

Suppose there is a subset of data generating processes that we wish to consider, and these contain the true structure.

A structural feature $\theta(S_0)$ is identified if it is uniquely determined within the set of all $\theta(S) : S \in$ Hypotheses by the joint distribution of observables.

The concept of identification requires a structure.

THE POTENTIAL OUTCOMES FRAMEWORK (NEYMAN-RUBIN CAUSAL MODEL)

Simplest setup.

- Each unit, i, can either be treated ($D_i = 1$) or not treated ($D_i = 0$) and has an observed outcome $Y_i^{D_i}$.
- The causal effect of D on Y_i is defined as $\tau_i = Y_i^1 - Y_i^0$.
- The observed outcome is: $Y_i = D_i Y_i^1 + (1 - D_i) Y_i^0$.
- The average treatment effect (ATE) is defined as $E[\tau_i]$. Note, this will vary based on the population of interest.

Key identification problem: can't see both Y_i^1 and Y_i^0,

ENDOGENEITY

- Those who are treated are on average different from those who are not treated.
- As a result, a simple difference in the average outcomes of treated and untreated does not identify the ATE.
- Similar reasoning holds for other parameters of interest (Average Treatment Effect on Treated (ATT), Price Elasticity, etc...).
- Solutions to endogeneity require arguing that something in the data is either random or as close to random as possible, and using that part of the data to identify the parameter of interest. HARD and is the challenge of most empirical papers.

WHY RANDOM ASSIGNMENT?

With simple randomization and perfect compliance, difference in means is unbiased for the ATE.

$$\begin{aligned} E[Y_i|D_i = 1] - E[Y_i|D_i = 0] \\ = E[Y_i^1] - E[Y_i^0] \\ = E[\tau_i] \end{aligned}$$

WHY RANDOMIZATION MAY NOT BE ENOUGH

- Some experiments impractical, expensive, unethical, underpowered.
- Assumptions:
  - Excludability: When treatment happens, the only thing relevant that changes is what you're interested in.
    Example of difficulty: placebo effects.
  - Non-interference, Stable Unit Treatment Value Assignment, No Spillovers
    Example of difficulty: equilibrium effects, people talking to each other, etc...

Next week, we'll talk about this in much more depth.

THE ROLE OF SIMULATION

- One way to demonstrate identification is to prove things. This is first best. For example:

> Berry, Steven T., and Philip A. Haile. “Identification in differentiated products markets using market level data.” Econometrica 82.5 (2014): 1749-1797.

- Oftentimes models are too complicated or the math is too hard. Our swiss army knife is to simulate the data generating process, and see if we can recover the parameters. LLMs make this way easier.
- Thinking through the simulation is very helpful in itself!

THE ROLE OF DATA

The future of the scientist in a world with advanced AI

by Tyler Cowen *December 27, 2024 at 12:14 am in* Science, Web/Tech

AI will know almost all of the academic literature, and will be better at modeling and solving most of the quantitative problems. It will be better at specifying the model and running through the actual statistical exercises. Humans likely will oversee these functions, but most of that will consist of nodding, or trying out some different prompts.

The humans will gather the data. They will do the lab work, or approach the companies (or the IRS?) to access new data. They will be the ones who pledge confidentiality and negotiate terms of data access. (Though an LLM might write the contract.) They will know someone, somewhere, using a telescope to track a particular quasar. They may (or may not) know that the AI's suggestion to sample the blood of a different kind of gila monster is worth pursuing. They will decide whether we should be filming dolphins or whales, so that we may talk to them using LLMs, though they will ask the LLMs for cost estimates in each case.

DATASETS IN MARKETING

Scanner Data

- IRI dataset: Data on grocery and drug chain purchases, some panel data, some advertising data. Over 400 papers have used it.
- Nielsen Dataset: Panel data on purchases, some surveys.

Digital Data

- Scraped data
- Comscore and other Web panels

CREATIVITY IN DATA COLLECTION

- If everyone is already using the dataset, there are less likely to be low-hanging fruits in research.
- Large returns about finding or collecting a previously unused dataset to answer an existing question.
- Examples:
  - Create an extension or app and recruit participants to use it (Levy 2021).
  - Run an experiment by interacting with a platform (Pallais 2014, Bai et al. 2022).
  - Combine datasets in new ways. Social media + survey experiment, etc...
  - Various company collaborations.

BIG DATA

Sutton. "The bitter lesson."

> The biggest lesson that can be read from 70 years of AI research is that general methods that leverage computation are ultimately the most effective, and by a large margin. The ultimate reason for this is Moore's law, or rather its generalization of continued exponentially falling cost per unit of computation. Most AI research has been conducted as if the computation available to the agent were constant (in which case leveraging human knowledge would be one of the only ways to improve performance) but, over a slightly longer time than a typical research project, massively more computation inevitably becomes available.

PARADOXES IN BIG DATA (MENG (2018))

> Which one should I trust more: a 1% survey with 60% response rate or a non-probabilistic dataset covering 80% of the population?" [...] Raised prior to the arrival of the era of Big Data, this question would likely be treated as an academic curiosity—how often can we get a hold of 80% of a population? Isn't the whole idea of survey sampling to learn about a population without having to record a large chunk of it?

PREPARING FOR NEXT WEEK

READINGS

- Read Chapter 4 of Causal Inference: The Mixtape
- Read “The economic analysis of advertising” by Bagwell.
- Read Blake et al. (2015) & Simonov et al. (2018).
- Volunteers for discussion of a paper.

QUESTIONS TO ASK WHEN READING A PAPER

- What is the research question?
- What is the main result?
- Does the evidence convince you of the main result? Often this relates to identification.
- If a reader remembers one thing from this paper, what should it be? Can be different from what the author wants you to remember.
- Craft: How is the data structured, how is the paper written, what exhibits are used?

BE READY TO DISCUSS EXPERIMENTAL DESIGN FOR BRANDED SEARCH ADS

1. Suppose you are a search engine, how would you design an experiment to measure the returns to brand keyword advertising for advertisers?
2. Suppose you are an advertiser, how would you design an experiment to measure the returns to advertising on your brand keyword? On your competitors brand keywords?
3. Unit of randomization, what is the treatment, how would you analyze the data?
4. Are we worried about excludability and non-interference?