EC6322 Advanced Industrial Organization
Introduction

Hsin-Tien Tiffany Tsai

National University of Singapore

Me - TSAI Hsin-Tien (Tiffany)

- Ph.D., Economics, University of California, Berkeley in 2019.
- I joined NUS in 2019 and have taught EC4305 since 2019, and EC6322 since this year.
- My research focuses on Industrial Organization, Digital Economy, Fintech, especially related to market structure, competition, and pricing.
  - How Amazon steers consumers towards its own products using product recommendations.
  - How mortgage subsidy might lead to selection and inefficiency.
  - Whether Amazon has an information advantage regarding demand uncertainty and what the implications are for competition fairness.

Timeline

Week	Topics	Notes
Week 1: Aug. 13	Introduction to Course and IO Overview	
Week 2: Aug. 20	Demand estimation: overview	
Week 3: Aug. 27	Demand estimation	Problem set 1 will be handed-out
Week 4: Sep. 3	Demand estimation: practical guideline	
Week 5: Sep. 10	IO topics in demand estimation	Problem set 2 will be handed-out
Week 6-9	Instructed by Sorawoot (Tang) Srisuma	Problem set 1 due at 11:59pm Sep. 27
Week 10: Oct. 22	Vertical markets	
Week 11: Oct. 29	Two-sided markets	Problem set 2 due at 11:59pm Oct. 29
Week 12: Nov. 5	More IO topics	
Week 13: Nov. 12	Class Presentation	

Evaluations

- Problem Set
- Presentation
- Class Participation (attendance)

There will be no exams.

Problem set: I will provide you with a dataset. Kindly follow the instructions to complete the problem set and submit both the output and the code.

Class Participation: I will take attendance, and if you are unable to attend any lectures during the week, please feel free to email me. I strongly encourage questions during class.

Lectures

Lecture (Seminar Style)	Week day	Start	End	Locations
EC6322	Tuesday	9:00	13:00	AS1-0201

- Starts at 9:00 sharp
- Break at 10:30 for 30 minutes
- Ends at 12:30 (or earlier)
- Attendance and participation (Missing few classes is fine).
- Email to schedule office hours.
- Slides/notes will be posted by Monday.

New empirical IO

- Bresnahan (1989 handbook chapter) coined the phrase "New Empirical IO"
  - Individual industries are distinct, and industry details are important (variations across industries problematic)
  - Goal is to understand institutional details and use this knowledge to test hypotheses of consumer/firm behavior or conduct counterfactual
  - Economic theory central in the analysis

- Influential IO papers focus on:
  - theoretical guidance
  - institutional specifics
  - measurement of key variables
  - econometric identification

Limitations of industry based studiest

- Difficult to generalize findings to other industries
- Broad takeaways tend to be qualitative or empirical methods

Road map: demand

- Demand estimation: important in its own right
  - consumer behavior
  - welfare effects of new products, etc
- Demand estimation: critical in firm studies
  - revenue
  - market power
- Consumer choice of differentiated products
  - examples: electronics, autos, pharmaceutical products, means of transport
  - n^2 price coefficients
  - exogenous variation in prices to identify n^2 parameters difficult to find
- Discrete-choice models

Road map: demand II

- Extensively used in antitrust analysis: merger (defining market scope), litigation (quantifying damage)
- Diffused into other fields: trade, health, development, urban (housing), environmental (land use), education
- Public policy: consumer price index, government regulation

Why do we learn structural method?

- Always be flexible in terms of methods, driven by research questions.
- Always think: Why do we need a structural model in this paper? What does it bring to you?
- Recognize the limitations of both reduced-form evidence and structural methods (revisit in Week 5!)
- Apply economic models to the data.
- Understand the key parameters, e.g., price elasticity.
- Understand counterfactual policies.

Static Analysis

Static analysis conditions on

- The goods marketed (or their characteristics) and their cost functions
- Consumer's preferences over goods (or over characteristics tuples)
- "Institutional" features like; the type of equilibrium, structure of ownership, regulatory rules

- We then analyzes how prices, quantities, and the distribution of profits and consumer surplus, are determined.
- Comparative statics changes one of the elements in the conditioning set, and compares the prices, quantities, and the distribution of profits and consumer surplus, from the new and old situations.

Static Analysis

To do the static analysis we need the following primitives.

- The demand system
- The cost system
- An equilibrium assumption

We will learn how to estimate (1) and (2), and we will, at least initially take the equilibrium assumptions as known from knowledge of institutions.

Use of Theory

- **Role of Theory**
  - Provides notions of equilibria for different environments and problems.
  - Basis for most analyses, including basic concepts of equilibria.
  - Use of learning theory to guide transitions between equilibria.
- **Intuition for "What Can Happen"**
  - Theory derives results for stylized environments.
  - Helps understand complex environments.
  - Analytic results may not hold when mimicking real-world detail.

Use of Theory: Stylized Environments and Their Limitations

- **Stylized Environments**
  - Sometimes too simplified to get accurate results.
  - Rarely indicate the quantitative importance of effects.
- **Example: Strategic Substitutes vs. Complements**
  - Simple demand functions often suggest prices are strategic substitutes.
  - Empirical work shows this is not always true.
  - Correct understanding of demand curves is crucial for policy issues (e.g., horizontal mergers, tariffs).
- **Importance of Accurate Estimation**
  - The nature of demand curves varies across industries.
  - Accurate estimation of industry-specific demand curves is essential.

Uses of Econometrics

- **Descriptive Statistics**
  - Provides an adequate description of phenomena or industries.
  - Essential preliminary step before detailed analysis.
  - Example: Demonstrating that prices today depend on prices yesterday.
- **Intuitive Evidence**
  - Evidence on entry deterrence before theoretical explanations.
  - Discussed in context of specific examples.

Parameter Estimation and Empirical Work

- **Parameter Estimation**
  - Using theory to generate estimating equations.
  - Using data to infer parameter values.
- **Method of Moments Estimators**
  - Either moment equalities or moment inequalities.
  - Intuition and guidance provided in class.
  - Method of moments with equality constraints will be taught and used in problem sets.
- **Computational Skills**
  - Computing estimators and equilibria for applied problems.
  - Builds self-confidence for handling empirical problems.
  - Problem sets form the basis for most of the course grade.

The Role and Usefulness of Models

- **Inherent Imperfection of Models**
  - No model will be entirely correct; the world is too complex.
- **Usefulness of Models**
  - Models don't need to be perfect to be useful.
  - Can explain past events or predict reactions to changes.
- **Evaluating Models**
  - Question of immediate usefulness depends on whether models are better than available alternatives.
  - Important for informing decisions by firms, regulatory agencies, or voters.
- **Progression to Better Models**
  - Even imperfect models can be stepping stones to better ones.

Introduction and Review for Demand Estimation

Demand Equations

- Demand: $Q = \alpha_d P + \beta_d' X + \epsilon_d$

- How much is consumed at a given price?

- Two main challenges in using aggregate data
  - Heterogeneity: To differentiate products that appeal differentially to different segments to reduce competition and increase margins; there is a need to account for this.
  - Endogeneity: Researchers typically do not know (or have data on) all factors that firms offer and consumers value in a product at a given time or market. Firms account for this in setting the marketing mix, creating a potential endogeneity problem.

Review: IV and Causality

Review: IV and Causality

- A causal constant-effect model of schooling on earning:

$$Y_i = \alpha + \rho s_i + \eta_i \quad \text{(short regression)}$$

- Suppose there is a control variables, A_i, called "ability" that determines η_i:

$$\eta_i = \gamma A_i + \nu_i$$

  where A_i and ν_i are uncorrelated by construction

- Assume A_i is the only reason why η_i and s_i are correlated, then

$$s_i \perp\!\!\!\perp \nu_i | A_i \quad \text{(CIA)}$$

Review: IV and Causality

- If A_i is observed, we estimate the causal effect of schooling on wage by estimating the "long" regression

$$Y_i = \alpha + \rho s + \gamma A_i + \nu_i \quad \text{(long regression)} \tag{1}$$

- However, most often, A_i is unobserved.
- If we just estimate the short regression, OLS estimator is biased. The omitted variables bias formula is:

$$\begin{aligned}
\rho^{OLS} &= \frac{Cov(Y_i, s_i)}{Var(s_i)} \\
&= \frac{Cov(\alpha + \rho s + \gamma A_i + \nu_i, s_i)}{Var(s_i)} \\
&= \rho + \gamma \frac{Cov(A_i, s_i)}{Var(s_i)}
\end{aligned}$$

Review: IV and Causality

- We can use the OVB formula to get a sense of the likely consequence of omitting ability for schooling coefficients. Is bias term positive or negative?

Table: Estimates of the returns to education for men in the NLSY

	(1)	(2)	(3)	(4)
Controls:	None	Age dummies	Col. (2) and additional controls*	Col. (3) and AFQT score
	0.132	0.131	0.114	0.087
	(0.007)	(0.007)	(0.007)	(0.009)

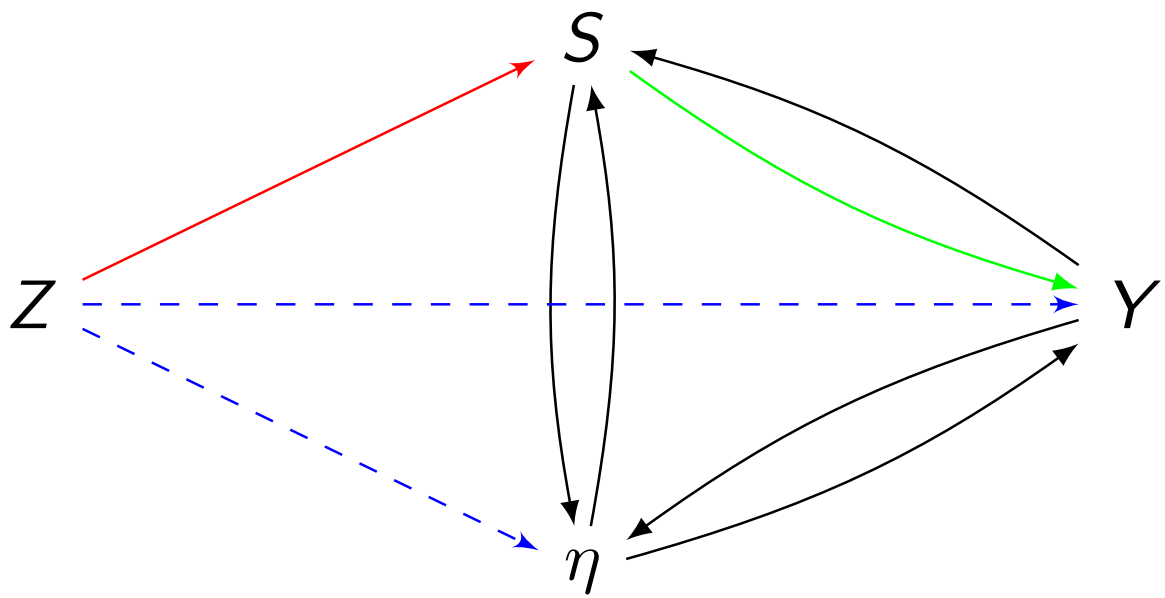
Notes: Additional controls are mother's and father's years of schooling and dummy variables for race and Census region.

Review: IV and Causality

- Instrumental variables methods allow us to estimate the long-regression coefficient, ρ, when A_i is unobserved.
- IV methods solve the problem of missing or unknown control variables, much as a randomized trial obviates extensive controls in a regression.
- A valid instrumental variable (IV) satisfies two conditions:
  (1) Relevance condition: $Cov(s_i, Z_i) \neq 0$
  (2) Exclusion restriction: $Cov(\eta_i, Z_i) = 0$

Review: IV and Causality

- Exclusion restriction:
  - The instrument is as good as randomly assigned, i.e., independent of potential outcomes (conditional on covariates.)
  - Z has no effect on outcomes other than through S.



Review: IV and Causality

- Given relevance condition and exclusion restriction, it follows:

$$\rho = \frac{Cov(Y_i, Z_i)}{Cov(s_i, Z_i)} = \frac{\overbrace{Cov(Y_i, Z_i)/V(Z_i)}^{\text{Reduced Form}}}{\underbrace{Cov(s_i, Z_i)/V(Z_i)}_{\text{First Stage}}}$$

- Proof:

$$\begin{aligned}\frac{Cov(Y_i, Z_i)}{Cov(s_i, Z_i)} &= \frac{Cov(\alpha + \rho s + \gamma A_i + \nu_i, Z_i)}{Cov(s_i, Z_i)} \\ &= \rho \frac{Cov(s_i, Z_i)}{Cov(s_i, Z_i)} = \rho\end{aligned}$$

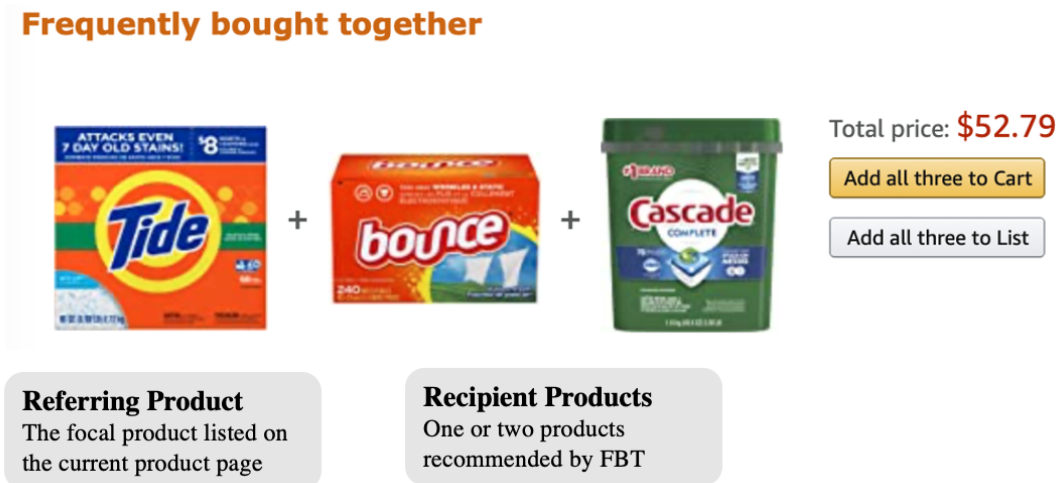
Endogeneity

- Now, let's return to the demand equation.
- Demand: $Q = \alpha_d P + \beta_d' X + \epsilon_d$
- Endogenous Variable: P
- Unobserved Variable: Quality
- If $\text{Cov}(P, \text{Quality}) > 0$, is α_d biased upward or downward?
- Solution: Find an instrument for P.
  - What are some variables that are typically excluded from Demand Equations? Cost shocks

Example: Demand and Supply Shocks

Chen and Tsai (Forthcoming, RAND)

- Dual role: Amazon owns the marketplace and guides consumers using product recommendations; at the same time, it also sells products directly (private brand and non-private brand).
- Amazon uses recommendation system to guide consumers: 30% page views on Amazon were from recommendations (Sharma et al 2015).
- We focus on the classic "Frequently Bought Together" recommendations (FBT).
  - Each product has only two slots to recommend other products. A product can receive many FBTs.



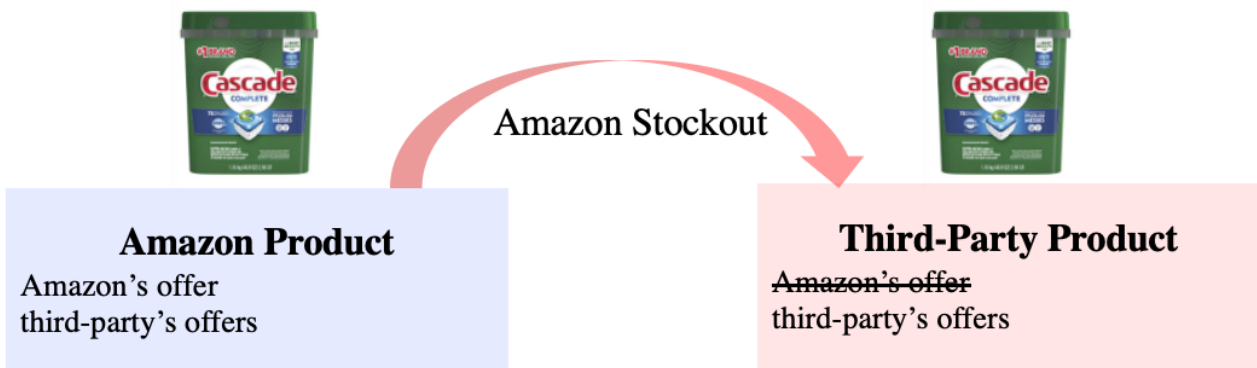
Amazon: Incentive and Ability to Steer

- Third party (referral fee of $\approx 15\%$) versus Amazon (full retail margin)
- In some scenarios, it might be more profitable to recommend the one that deviates from consumers' most preferred option.
  - The recommendation depends on seller identify.
- NOT always recommend itself: this depends on the alternative recipient, amazon's cost, consumer preference (conversion rate), price, and so on.

Our Approach

- Collect large data in high-frequency.
  - Cover the top 6M popular products for 5 rounds over three months.
  - Observe prices, sales ranks, Frequently Bought Together recommendations, whether Amazon sells, and other product-level information.

- Research design:
  - Temporary variations in Amazon's presence due to **stock-out events**.
  - Robustness checks to rule out alternative explanations.



Recommending Amazon-Selling Products

We leverage the real-time changes in recommendations patterns and Amazon's presence to identify whether Amazon affects recommendation patterns.

A pair of referring and recipient products n.

$$\text{FBT}_{nt} = \theta \times \text{PLAT\_Recipient}_{nt} + \text{Pair\_FE}_{n} + \text{Cat\_Day}_{nt} + \epsilon_{nt}.$$

$\text{PLAT\_Recipient}_{nt}$: an indicator of whether Amazon sells in the recipient product's market,

Pair_FE_{n}: pair fixed effects (i.e., recipient product-referring product fixed effects).

Cat_Day_{nt}: the category–day fixed effects.

$$\theta = [\text{FBT}|\text{PLAT\_Recipient} = 1, n, t] - [\text{FBT}|\text{PLAT\_Recipient} = 0, n, t]$$

Recommending Amazon-Selling Products

	Dependent Var=FBT$_t$		
	(1)	(2)	(3)
θ: PLAT_Recipient$_t$	0.087***	0.083***	0.080***
	(0.016)	(0.015)	(0.014)
log(Q_Recipient$_t$)		0.016***	0.016***
		(0.003)	(0.003)
log(Market Price$_t$)			-0.014***
			(0.004)
Product Pair Fixed Effects	Y	Y	Y
Category–Day Fixed Effects	Y	Y	Y
No. of Observations	32,375,268	32,375,268	32,375,268
Adjusted R-squared	0.397	0.397	0.397

Note: Table reports coefficient estimates of interest. The dependent variable is an indicator of recommendation received. Other coefficients and fixed effects are omitted for brevity. Robust standard errors in parentheses are clustered at the category level. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01).

The coefficient of PLAT_Recipient measures **the extent of steering**.

Alternative Explanation: Potential Omitted Variables

1. **Price or shipping increases** due to Amazon stockouts.
2. **Sales decrease** due to Amazon stockouts.

- Prices/sales are omitted variables that correlate with both stockouts and FBT.

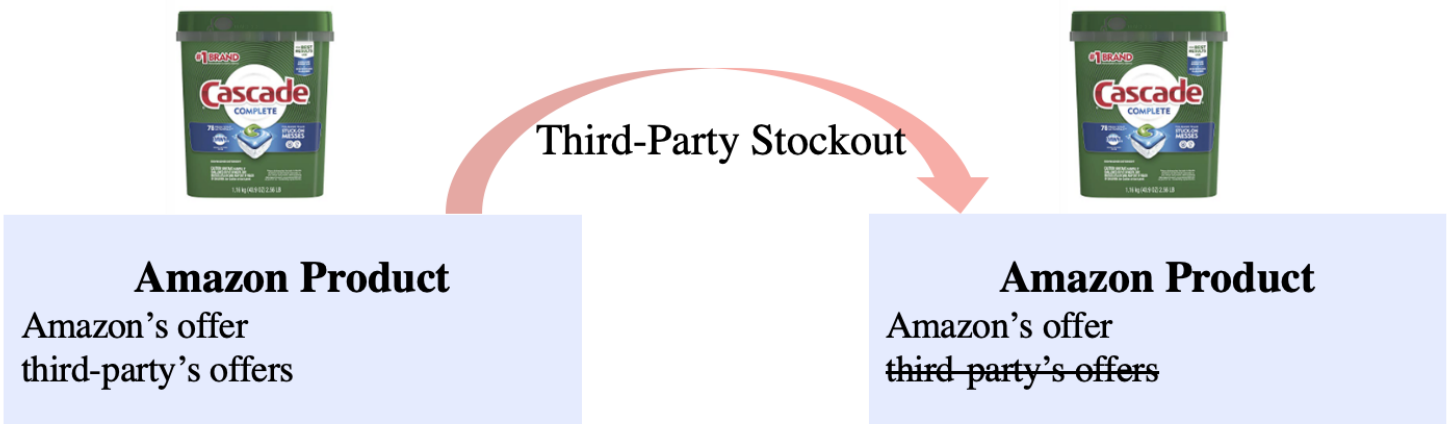
$$\text{FBT}_{nt} = \theta \times \text{PLAT\_Recipient}_{nt} + \text{Pair\_FE}_{n} + \text{Cat\_Day}_{nt} + \underbrace{\epsilon_{nt}}_{=\text{Price, Sales}+\nu_{nt}}.$$

- We control for real-time price and sales (sales rank) and the results do not change much.
- **THINK!**
- demand shocks that correlated with Amazon stockouts.
- supply shocks that correlated with Amazon stockouts.

Alternative Explanation: Stockouts are Endogenous?

$$\text{FBT}_{nt} = \theta \times \text{PLAT\_Recipient}_{nt} + \text{Pair\_FE}_n + \text{Cat\_Day}_{nt} + \underbrace{\epsilon_{nt}}_{=\text{demand, supply} + \nu_{nt}}.$$

- Demand shocks that cause the stockout, FBT ↑
- **Supply shocks that cause the stockout, FBT ↓**



There is no effect for third-party's stockouts; supply shocks are unlikely to be the omitted variables that correlate with a seller's presence.

Next Week

- Discrete Choice Model
- To do: Install Matlab (see `https://nusit.nus.edu.sg/services/software_and_os/software/software-for-student/`)