How and Why to Use Models

Lecture Notes

Introduction

- Before diving into methods, review important use cases for quantitative economic models
- Why use a model? A legitimate question given that many applied papers don't need one
- Central question: *What is my model for? How is it central to answering my research question?*

When You Possibly Don't Need a Model

Causal Inference Without Deep Structure

- Many first papers pose simple questions: "What is the effect of policy X on outcome Y?"
- Often answerable with simple statistical models of causality:
  - Potential Outcomes Model (Rubin Causal Model)
  - Generalized Roy Model
- These formalize causality in terms of potential outcomes
- No need to specify preferences, technology, or equilibrium
- **Key requirement:** observed variation directly identifies the counterfactual of interest

Three Examples from Prototype Models

Example 1: Social Security

Setup

- Life-cycle savings model with a social security system
- Budget constraint:

$$c_t + a_{t+1}/(1+r) \leq (1-\tau)y_t + \mathbf{1}\{t \geq 65\}b$$

- Individuals receive benefit b after age 65, pay proportional taxes τ

The Quasi-Experiment

- Suppose eligibility age is unexpectedly lowered from 65 to 60
- Have repeated cross-sectional data on consumption by age, from before and after the reform
- Let t^* index cohorts by their age at the time of the policy announcement

- Note: t^* indicates both a cohort *and* a treatment

Research Question

- *"What was the effect of this eligibility expansion on consumption at age t for a cohort aged t^* at the time of expansion?"*

Potential Outcomes Framework

- $C_{t^*,t}(1)$: potential consumption under the policy announcement
- $C_{t^*,t}(0)$: potential consumption under the counterfactual (no announcement)
- Target parameter:

$$\alpha_{t^*,t} = \mathbb{E}[C_{t^*,t}(1) - C_{t^*,t}(0)]$$

Identification via Difference-in-Differences

- Use cohort aged 65 at announcement (unaffected) as control group
- **Parallel trends assumption:** for $t^* < 65$, ages $t \geq t^*$ and $s < t^*$:

$$\mathbb{E}[C_{t^*,t}(0)] - \mathbb{E}[C_{65,t}(0)] = \mathbb{E}[C_{t^*,s}(0)] - \mathbb{E}[C_{65,s}]$$

- The DD estimator constructs counterfactual from observables:

$$\alpha_{t^*,t} = \mathbb{E}[C_{t^*,t}(1)] - \mathbb{E}[C_{65,t}(0)] - (\mathbb{E}[C_{t^*,s}(0)] - \mathbb{E}[C_{65,s}(0)])$$

Regression Specification

- Common regression with many cohorts and ages:

$$\mathbb{E}[C_{t^*,t}] = \gamma_{t^*} + \mu_t + \alpha_{t^*,t}\mathbf{1}\{t \geq t^*\}$$

- Layers in stricter, interconnected parallel trends assumptions
- Only need to defend parallel trends, not utility functions or savings technology
- Farther-apart cohorts make parallel trends stronger (potentially less credible)
- Regression cannot select "more ideal" controls for each cohort (Goodman-Bacon 2021)

Example 2: Firm Entry and Minimum Wage

Setup

- Panel data $(A_{m,t,1}, A_{m,t,2}, Z_{m,t})$ across markets m and periods t
- $A_{m,t,j}$: whether firm j is active in market m at time t
- Entry occurs when $A_{m,t+1,j} - A_{m,t,j} = 1$
- X_m: unobservable market-level profitability factor
- $Z \in \{0, 1\}$: local minimum wage policy indicator

Payoff Specification

- Payoff to being active:

$$u_1(x, a, d') = \phi_0 + \phi_1 x - \phi_2 d' - \phi_3(1 - a) + \phi_4 z$$

- ϕ_4 captures the minimum wage effect on firm payoffs

The Quasi-Experiment

- $Z_{m,1} = 0$ for all markets initially
- At t^*, a subset adopts the policy permanently, unanticipated

Research Question

- *"What is the effect of the minimum wage on firm entry?"*

Target Parameter

- $N_{m,t} = D_{m,t,1} + D_{m,t,2}$: number of active firms
- $N_{\tau,m}(z)$: potential outcome of N in market m, τ periods after adoption
- Dynamic treatment effect on the treated:

$$\alpha_\tau = \mathbb{E}[N_\tau(1) - N_\tau(0)|Z = 1] = \mathbb{E}[\Delta_\tau|Z = 1]$$

Heterogeneity Matters

- Model highlights that effects differ across markets
- Selection into adoption implies: $\mathbb{E}[\Delta_\tau|Z = 1] \neq \mathbb{E}[\Delta_\tau]$

Parallel Trends for Event Study

- Justifying assumption:

$$\mathbb{E}[N_t(0)|Z = 1] - \mathbb{E}[N_t(0)|Z = 0] = \text{constant}$$

- If markets are in the ergodic distribution of the MPE, N is stationary absent intervention
- Event-study specification:

$$D_{m,t} = \gamma_m + \mu_t + \mathbf{1}\{t \geq t^*\}\alpha_{t-t^*} + \epsilon_{m,t}$$

- Robustly identifies average effect among adopters (with uniform timing)
- Staggered adoption + heterogeneous effects: regression gives hard-to-interpret weighted average (Goodman-Bacon 2021)

Example 3: Bundles of Tax Reforms

Setup

- Three states (A, B, C) run different experiments with taxes and transfers
- Net income functions:

$$\mathcal{Y}_A(W, H) = b_A + (1 - \tau_A)WH \tag{1}$$

$$\mathcal{Y}_B(W, H) = WH + \sum_{k=0}^{5} \tau_k(WH - \overline{E}_k)\mathbf{1}\{WH > \overline{E}_k\} \tag{2}$$

$$\mathcal{Y}_C(W, H) = WH(1 - \tau_C) + \mathbf{1}\{H > 20\}b_C \tag{3}$$

- Participants don't anticipate assignment; experiment lasts 3 periods
- $Z_j \in \{0, 1\}$: assignment to treatment or control in state j

Research Question

- *"What is the effect of each unannounced, temporary, tax reform on labor supply?"*

Identification

- Compare treatment and control means in each state:

$$\mathbb{E}[H|Z_j = 1, j] - \mathbb{E}[H|Z_j = 0, j]$$

- Random assignment identifies the causal effect — no further structure needed

---

Reasons to Use a Model

When the Question Can't Be Articulated Without One

- Most obvious reason: research question cannot be stated without a model
- Most useful modeling insights concern:
  - Economic efficiency
  - Potential for policies to resolve market inefficiencies
  - Policy design

Examples

1. In the labor supply model, given a welfare objective, what is the optimal system of taxes and transfers?
2. What is the cheapest way to incentivize competition between firms in dynamic duopoly?
3. What are the welfare costs of incomplete markets in the life-cycle savings model?

To Make Welfare Calculations

- **Revealed preference:** treating individuals as knowing what they like allows inferring preferences
- Can then evaluate how they value different policy environments

Examples

1. How do individuals value the social security program? How would they value different (τ, b) combinations?
2. What is the willingness to pay for a lump-sum payment b financed by proportional taxes τ?

To Make Sense of Otherwise Puzzling Data

The Fundamental Question

- What is the effect of price on quantities?

- A silly question — but only because supply and demand theory is fundamental to our worldview

Without Theory

- Observe prices and quantities in a market over time
- All you see is a cloud of points with no apparent pattern

With Theory

- Each point is the simultaneous equilibrium outcome of two structural relationships
- Wright proposed **Instrumental Variables** to disentangle supply from demand
- Perhaps the earliest estimated structural model

When Variation Does Not Identify the Counterfactual of Interest

- In examples above, we defined "treatment" very specifically
- Models articulate what *can* and *cannot* be identified by observed variation
- They outline assumptions under which related counterfactuals can be forecast
- Often framed as *internal vs external validity* — but external validity deserves more depth

Key Insight (Heckman and Vytlacil 2005)

- Estimands from simple causal inference methods (DD, IV, RD) are rarely parameters of exact policy interest

---

Social Security: What DD Doesn't Tell Us

Additional Assumptions for Concrete Analysis

1. Each individual faces a *known* income sequence $\{y_{n,t}\}_{t=1}^{T}$
2. Agents face a *natural borrowing constraint*:

$$\sum_{s=t}^{T} q_{s-t} c_s \leq a_t + \sum_{s=t}^{T} q_{s-t}(1-\tau) y_{n,s} + \sum_{s \geq 65}^{T} q_{s-t} b$$

   where $q_\tau = 1/(1+r)^\tau$ is the price of consumption τ periods ahead
3. $\beta(1+r) = 1$ and $\psi = 0$: agents perfectly smooth consumption to NPV of net income

Consumption Effect

- Mean effect for cohort t^* at any $t \geq t^*$:

$$\Delta C_{t^*,t} = \sum_{s=60}^{64} q_{s-t^*} b + \sum_{s=t^*}^{T} (\tau - \tau') \overline{y}_{t^*,t}$$

  where $\overline{y}_{t^*,t}$ is average income for cohort t^* at age t

What We Learn

1. Parallel trends holds: cross-cohort differences are constant with age
2. DD robustly identifies $\alpha_{t^*,t}$
3. Model implies $\alpha_{t^*,t}$ is constant in t
4. Each effect depends on **how each cohort expects financing** $(\tau' - \tau)$ — not specified

What DD Does NOT Identify

1. Total effect of social security (not just the expansion)
2. Effect of additional changes in eligibility age
3. Effects of changes in b and τ at different ages and anticipation horizons
4. Effects on cohort t^* if they learned at any age other than t^* (one-to-one cohort-treatment mapping)
5. Effects for cohorts who anticipated the policy their whole life
6. Effects under alternative financing arrangements

Tax Reform: What Experiments Don't Tell Us

Concrete Case

- Effect of reform A when $\beta(1+r) = 1$ (stationary consumption):

$$\Delta H_{n,t} = \psi \log(1-\tau) - \psi\sigma\Delta \log(C_n^*)$$

- ΔC_n^* embodies the income effect; solved via optimal labor supply in budget constraint

Key Points

1. ATE depends on income effects, which depend on *perceived persistence* of the reform
2. Substantial heterogeneity in treatment effects
3. Different wage/work cost distributions imply different impacts

What Experiments Do NOT Identify

1. Effects of reforms with different persistence (real or perceived)
2. Effect of reform A on population B, C, etc. (and vice versa)
3. Distribution of individual treatment effects
4. Effects of *partly anticipated* reforms
5. Effects at scale with general equilibrium wage responses
6. Model can invert out latent distributions and estimate ATEs along latent dimensions

---

Entry-Exit: What the Event Study Doesn't Tell Us

The model-free approach does **not** identify:

1. Effect when the minimum wage introduction is *anticipated*
2. Effect on markets that do *not* adopt

3. Effect of *repealing* the minimum wage
4. Effect of *nominal changes* to the minimum wage

- Each of these is arguably a more useful policy calculation

To Interpolate Existing Variation in the Data

The Problem

- Each tax reform is related: each comes as a function $\mathscr{Y}(W, H)$ of wages and hours
- Model-free estimation of labor supply as a nonparametric function of $\mathscr{Y}$ is infeasible
- Would require implausible quantities of policy variation

The Solution

- Economic models provide a way to interpolate related — but not functionally identical — variation
- In the labor supply model, any $\mathscr{Y}$ enters through the budget constraint
- Different reforms are comparable without additional parameters

Articulated Variation

- **Definition:** when a policy's effect can be modeled without additional parameters
- Typically involves *a priori* known changes to prices and endowments
- With structural parameters in hand, any $\mathscr{Y}$ can be studied

Counterexample

- Minimum wage in the entry-exit model is embodied as parameter ϕ_4
- Within-period production not modeled (reduced form)
- Policy is *not* well-articulated: further nominal wage changes cannot be studied

---

Marschak's Maxim

The Principle

> Researchers should specify the minimal set of model ingredients in order to answer the research question of interest.

- May seem obvious; hard part is figuring out the minimal set
- Surprisingly easy to forget once deep inside your model
- Repeatedly ask: *"Is this essential to my question of interest?"*
- Your research question is the mast you tie yourself to
- Decide what it is as early as possible

Marschak's Original Example (1953)

Setup

- Monopolist chooses quantities to maximize profit; government taxes quantities
- Demand: $p = \alpha_0 - \alpha_1 q$
- Profits (constant marginal cost c):

$$\Pi(q) = (\alpha_0 - c - \tau)q - \alpha_1 q^2$$

- Optimal quantity:

$$q^* = \frac{\alpha_0 - c - \tau}{2\alpha_1}$$

- Tax revenue:

$$R = q\tau = \frac{\alpha_0 - c - \tau}{2\alpha_1}\tau$$

Three Problems

Suppose q varies exogenously across markets. Three agents, three information requirements:

1. **Firm maximizing profit:** Observe $\Pi = aq - bq^2$, set $q^* = a/(2b)$
2. **Government forecasting revenue:** Only needs τ and observed quantities
3. **Government optimizing taxes:** Extract $a = \alpha_0 - c - \tau$ from reduced form, set $\tau^* = (\alpha_0 - c)/2$

Lessons

1. Each question requires only certain *combinations* of parameters, observable in reduced-form relationships
2. Once agents optimize, reduced-form relationships no longer suffice for forecasting policy effects

- Point (2) is the Lucas critique, articulated two decades earlier
- Point (1): answers to some questions are invariant to particular model ingredients
- Motivates seeking a minimal set of ingredients

Application: Two-Stage Budgeting

Setup

- Utility depends on K commodities x with price vector p:

$$V_t(p, a, y) = \max_{x, a'} \left\{ v(x) + \beta \mathbb{E}_{p', y' | p, y} V_{t+1}(p', a', y') \right\}$$

subject to:

$$p \cdot x + \frac{1}{1+r} a' \leq y_t, \qquad a \geq \underline{a}$$

The Question

- Suppose research question concerns only savings accumulation
- Policy counterfactuals affect intertemporal allocation of resources

The Result (Gorman 1959)

- If v is homothetic, there exists price index $P(p)$ and indirect utility u:

$$V(p,X) = \max_x v(x) \text{ s.t. } p \cdot x \leq X = u(X/P(p))$$

- Consumers first allocate aggregate expenditure across periods
- With access to price index, can disregard intratemporal allocations

Application of Marschak's Maxim

- If question concerns intertemporal allocation and counterfactuals don't affect intratemporal margins:
- Under regularity conditions on demand, no need to model intratemporal consumption
- All you need is a reasonable price index

Application: Sufficient Statistics

Setup

- Static labor supply for individuals $i \in [0,1]$:

$$H_i = \arg\max_h C - v(h)$$

  where $C = b + (1-\tau)hW_i$
- Transfer financed through taxes: $b = \int \tau H_i W_i \, di$
- Utilitarian welfare objective:

$$V(\tau) = \int (C_i - v(H_i))di$$

Question

- What is the welfare effect of a marginal expansion in b, financed through τ?

Derivation

- Taking $dV/d\tau$ and applying envelope theorem:

$$\frac{dV}{d\tau} = -\int W_i H_i di + \int W_i H_i di + \tau \int d(W_i H_i)/d\tau \, di$$

- With constant elasticity of taxable income $Z_i = W_i H_i$ w.r.t. τ:

$$\frac{dV}{d\tau} = \bar{Z} \cdot \varepsilon_{Z,\tau}$$

Interpretation

- Willingness to pay for marginal policy expansion **exactly cancels** mechanical cost
- Only the **distortionary cost** of behavioral adjustments remains (Harberger 1954)

Sufficient Statistics Literature

- Derives approximations to local policy effects
- Formulas depend on empirical quantities and elasticities, not specific functional forms
- As models get more elaborate, required elasticities can become intractable
- Still a useful guide to understanding policy fundamentals
- Great example of Marschak's Maxim

Summary

Key Takeaways

- Quasi-experimental methods are powerful when observed variation directly identifies the counterfactual
- Models become essential for:
  - Welfare calculations
  - Optimal policy design
  - Out-of-sample counterfactuals
- Models interpolate related variation through articulated policy channels
- **Marschak's Maxim:** always aim for the minimal set of ingredients your question requires