Empirical Methods for Policy Evaluation

Matteo Bobba

Toulouse School of Economics (TSE)

TSE PhD Program (MRes)
Fall 2025

Motivation/Background

- Debate on economic theory $\Leftrightarrow$ econometric modeling and estimation

  $\Rightarrow$ Design-based vs. model-based approaches to policy evaluation

- Natural synergy for a better characterization of policy impacts
  - Middle ground approach dates back to Marschak (1953)
  - Many advocates since then....

$\Rightarrow$ We will focus on both methods and applications (Labor, Devo, Public,..)

Overview of the Course – Design-based Methods

1. Randomized Control Trials
2. Regression Discontinuity Designs
3. Difference-in-Differences and Event Studies
4. Shift-Share Instrumental Variables

Overview of the Course – Model-based Methods

- RCT1 Risk sharing models
- RCT2 Dynamic factor models

- RD1 Discrete choice models
- RD2 GE model of education and the labor market

- DD Job search models

- IV1 Firm dynamics
- IV2 Spatial equilibrium models

Overview of the Course – Applications

$\Rightarrow$ Seasonal migration in Indonesia

$\Rightarrow$ Human capital accumulation in Colombia

$\Rightarrow$ Teacher sorting and student outcomes in Peru

$\Rightarrow$ The equilibrium effect of a schooling expansion in India

$\Rightarrow$ Informal labor markets and schooling investments in Mexico

$\Rightarrow$ Rural-urban migration in Brazil

$\Rightarrow$ Urban public works in India

Course Evaluation

1. Problem sets [55% of the grade, you can work in pairs]

   ⇒ Three exercises based on class material (I will provide the datasets)

2. Referee report [30% of the grade, individual assignment]

   ⇒ Pick one paper from a list that I will circulate soon

3. Class participation [15% of the grade]

   ⇒ Take a look at required readings (* in the syllabus) before class

Causal inference meets structural models (and viceversa)

1. Ex-ante policy evaluation
   - ⇒ Chapter 2 in Wolpin's book (MIT press, 2013)
2. Principles for combining descriptive and model-based analysis
   - ⇒ Chapter 2 in Wolpin's book (MIT press, 2013)
   - ⇒ Mahoney (JEP, 2022)

Ex-ante policy evaluation

The Ex-ante Approach for Evaluating Public Policies

- Economic models allow predicting the effects of public policies
  - $\Rightarrow$ Before they are implemented and/or variants of existing policies
- The ex-ante approach is particularly suitable for informing policy
  - $\Rightarrow$ Improve program design to maximize impacts given costs
  - $\Rightarrow$ Inform targeting by identifying sub-populations for which impacts are highest
  - $\Rightarrow$ Analyze program impacts over a longer time horizon than variation in data
  - $\Rightarrow$ Study the effect of programs in the presence of spillover or GE effects

An Example

- Many governments have adopted conditional cash transfer (CCT) programs
  - Provide cash transfers to HHs conditional on school attendance of children

  $\Rightarrow$ Alleviate poverty and stimulate human capital investments

- Can we evaluate those programs before they are implemented?

  $\Rightarrow$ A model of schooling decisions where transfer decreases schooling costs

Economic Model

- Consider the following (static) optimization problem for the household

$$\max_{s\in\{0,1\}} U(c,s) \text{ s.t. } \begin{cases} c = y + w(1-s) \\ c = y + w(1-s) + \tau s \end{cases}$$

- Optimal schooling choices without and with the subsidy

$$\begin{aligned} s^{\star} &= g(y,w) \\ s^{\star\star} &= g(\tilde{y},\tilde{w}), \text{ where } \tilde{y} = y + \tau \text{ and } \tilde{w} = w - \tau \end{aligned}$$

$\Rightarrow$ The subsidy acts as an income-compensated reduction in child wages

Bringing the Model to the Data

- Add observable and unobservable preference shifters

$$U(c, s; X, \epsilon)$$

- Unobserved heterogeneity is not systematically related to wages and income

$$f(\epsilon|y, w, X) = f(\epsilon|X) \tag{CIA}$$

$\Rightarrow$ Given CIA, variations in wages and income identify the impact of the program

Estimation

- Ex-ante average treatment effect is:

$$\text{ExATE}_{np} = \frac{1}{N}\sum_{j=1}^{N}\left[\underbrace{\hat{\mathbb{E}}(s_i \mid w_i = w_j - \tau_j, y_i = y_j + \tau_j, X_i)}_{\text{Predicted schooling under the program}} - \underbrace{s_j(w_j, y_j, X_j)}_{\text{Observed schooling}}\right]$$

- $\mathbb{E}(s_i \mid \tilde{w}_i, \tilde{y}_i, X_i)$ can be estimated non-parametrically (kernel, LLR, or series)
  - $\Rightarrow$ Need common support in the data: $(w, y) \in R^2$ such that $f(w, y) > 0$

Counterfactual Subsidy Levels

	Boys		
Ages	2* Original	Original	0.75*Original
12-13	0.04 (59%)	0.01 (87%)	0.003 (98%)
14-15	0.24 (45%)	0.01 (83%)	0.05 (98%)
12-15	0.12 (53%)	0.06 (86%)	0.02 (98%)
	Girls		
	2* Original	Original	0.75*Original
12-13	0.06 (48%)	0.06 (91%)	0.05 (98%)
14-15	0.23 (51%)	0.07 (89%)	0.03 (98%)
12-15	0.14 (50%)	0.06 (90%)	0.05 (98%)
	Boys and Girls		
	2* Original	Original	0.75*Original
12-13	0.05 (54%)	0.04* (89%)	0.03 (98%)
14-15	0.23 (48%)	0.09 (86%)	0.04 (98%)
12-15	0.13 (52%)	0.06 (88%)	0.03 (98%)

† Bandwidth equals 200 pesos. Trimming implemented
using the 2% quantile of positive density values as the cut-off point.

Unconditional Income Grant

	Boys		
Ages	Predicted	Sample-Sizes‡	% overlapping support
12-13	-0.02 (0.03)	374, 610	89%
14-15	-0.06 (0.05)	309, 569	90%
12-15	-0.04 (0.03)	683, 1179	89%
	Girls		
	Predicted	Sample-Sizes‡	% overlapping support
12-13	-0.03 (0.04)	361, 589	88%
14-15	0.00 (0.05)	316, 591	88%
12-15	-0.02 (0.03)	677, 1180	88%
	Boys and Girls		
	Predicted	Sample-Sizes‡	% overlapping support
12-13	-0.03 (0.03)	735, 1199	88%
14-15	-0.03 (0.03)	625, 1160	89%
12-15	-0.03 (0.02)	1360, 2359	89%

†Standard errors based on 500 bootstrap replications. Bandwidth equals 200 pesos. Trimming implemented using the 2% quantile of positive density values as the cut-off point.

‡The first number refers to the total control sample and the second to the subset of controls that satisfy the PROGRESA eligibility criteria.

Adding Home Production

- Allow for an alternative use of children's time, home production $l \in \{0, 1\}$

$$\max_{(s,l)} U(c, l, s) \text{ s.t. } \begin{cases} c = y + w(1 - s - l) \\ c = y + w(1 - s - l) + \tau s \end{cases}$$

- Optimal schooling choices without and with the subsidy are different

$$s^{\star} = g(y, w)$$
$$s^{\star\star} = h(\tilde{y}, \tilde{w}, \tau)$$

$\Rightarrow$ Non-parametric ex-ante approach is not feasible in this case

Parametric Approach

- Consider the following functional form (no child leisure and no X)

$$U(C, s; \epsilon) = C + \alpha s + \beta C s + \epsilon s,\ \epsilon \sim N(0, \sigma_\epsilon^2)$$

- The probability of school attendance under the subsidy is

$$P(s = 1) = 1 - \Phi\left(\frac{(w - \tau) - \alpha - \beta(y + \tau)}{\sigma_\epsilon}\right)$$

$\Rightarrow$ Model parameters can be estimated by ML from data with no subsidy ($\tau = 0$)

Parametric Approach

- Given parameter estimates, the effect of τ on the attendance rate is

$$\text{ExATE}_p = \Phi\left(\frac{(w-\tau)-\hat{\alpha}-\hat{\beta}(y+\tau)}{\hat{\sigma}_\epsilon}\right) - \Phi\left(\frac{w-\hat{\alpha}-\hat{\beta}y}{\hat{\sigma}_\epsilon}\right)$$

$\Rightarrow$ There is no condition on the support of y and w

Wrapping Up on Ex-Ante Policy Evaluation

- Estimating ex-ante policy impact does not require specifying a full model
- Nonparametric approaches may be feasible
  - $\Rightarrow$ Even when there is no variation in the data on the policy (price of schooling)
- If not feasible, extra-assumptions on the distribution of unobs. heterogeneity

Principles for combining descriptive and model-based analysis

The Causal Inference Approach

- Binary random variable $D_i = \{0, 1\}$ and potential outcomes (Y_i^1, Y_i^0)

$$\begin{aligned} \text{ATE} &= \mathbb{E}(Y_i^1 - Y_i^0) \\ &= \mathbb{E}(Y_i^1 \mid D_i = 1) - \mathbb{E}(Y_i^0 \mid D_i = 0) \\ &= \underbrace{\mathbb{E}(Y_i^1 - Y_i^0 \mid D_i = 1)}_{\text{ATT}} + \underbrace{\mathbb{E}(Y_i^0 \mid D_i = 1) - \mathbb{E}(Y_i^0 \mid D_i = 0)}_{\text{Selection bias}} \end{aligned}$$

$\Rightarrow$ Research designs attempt to eliminate bias by means of assumptions:

RCT Random assignment

RD Continuity of potential outcomes around cutoff

DD Same trend in potential outcomes over time

IV(LATE) First-stage + exclusion restriction (+ Monotonicity)

The (Old School) Structural Approach

1. Lay out an economic model of the phenomenon being studied
2. Addition of a stochastic structure if the model itself does not possess one
3. Identification of the parameters given the data and model

parametric: Two sets of parameters yielding the same likelihood are necessarily equal

non-param: Observables pick only one set of parameters irrespectively of unobservables

4. Estimation technique given model, identification, and data
5. Empirical comparative statics, welfare metrics, and/or policy experiments

Example Continued: The *Progresa* Program in Mexico

- Large scale anti-poverty program
  - Began in 1997 in rural areas and rapidly expanded throughout the country
  - About 20% of Mexican families participating
- Educational grants to mothers to encourage children's school attendance
  - Benefits levels increase with grades attained, higher for girls
  - Subsidies amount to about 20% of average annual income
- Data from the initial rural evaluation of the program

  $\Rightarrow$ Randomized phase-in design at the village level
  - Within villages, both eligible and non-eligible HHs (wealth index)

Design-based ⇒ Model-based: Out of Sample Validation

- Plausibility of the assumptions determines the credibility of predictions
  - ⇒ Within-sample goodness-of-fit tests are necessary but not sufficient
- The credibility of the model is better assessed in terms of out-of-sample fit
  - ⇒ Estimate a model by holding out the treatment/control group
  - ⇒ Validate its predictions about program impacts

Todd and Wolpin (AER 2006)

- Dynamic discrete choice of children's time allocation and family fertility

  $\Rightarrow$ Ex-ante evaluation (control group): MU of subsidy=MU of income

- Each year a married couple decides on whether

  $\Rightarrow$ Each of their children attend school/stay-at-home/work for wage

  $\Rightarrow$ The wife becomes pregnant

- Parental and children earnings are subject to idiosyn. time-varying shocks
  - No parental labor supply decisions
  - No saving or borrowing
  - No equilibrium effects of the subsidy on children's wages

Model Validation: Within-Sample Fit (Control Group)

Boys							
	Actual			Predicted			
Age	School	Work	Home	School	Work	Home	χ^2
6	0.933	—	0.066	0.923	—	0.077	0.58
7	0.981	—	0.019	0.980	—	0.020	0.02
8	0.987	—	0.013	0.980	—	0.020	0.99
9	0.994	—	0.006	0.979	—	0.021	3.49
10	0.982	—	0.018	0.974	—	0.026	0.86
11	0.977	—	0.023	0.964	—	0.036	1.45
12	0.885	0.021	0.094	0.846	0.039	0.115	3.99
13	0.780	0.084	0.136	0.736	0.078	0.186	4.51
14	0.677	0.157	0.166	0.619	0.191	0.190	3.41
15	0.490	0.276	0.235	0.520	0.251	0.229	0.88
Girls							
6	0.965	—	0.035	0.942	—	0.058	3.84
7	0.976	—	0.024	0.968	—	0.032	0.77
8	0.989	—	0.011	0.976	—	0.024	1.96
9	0.991	—	0.009	0.975	—	0.025	3.26
10	0.979	—	0.021	0.970	—	0.030	0.93
11	0.969	—	0.031	0.948	—	0.052	2.97
12	0.896	0.007	0.097	0.854	0.020	0.126	4.61
13	0.726	0.028	0.245	0.676	0.025	0.299	2.85
14	0.582	0.089	0.329	0.566	0.092	0.342	0.22
15	0.419	0.123	0.458	0.402	0.157	0.442	1.68

Note: χ^2 (0.05, 1) = 3.84, χ^2 (0.05, 2) = 5.99.

Out-of-Sample Model Validation Using the Experiment

	Girls age 12–15			Girls age 12–15, behind in school			Girls age 13–15, HGC ≥ 6, behind in school		
	(1) Actual	(2) Pred. with Subsidy	(2)–(1)	(1) Actual	(2) Pred. with Subsidy	(2)–(1)	(1) Actual	(2) Pred. with Subsidy	(2)–(1)
97 Control	65.3	72.7	7.4	58.3	67.0	8.7	40.9	58.6	17.7
98 Control	66.5	72.9	6.4	58.7	66.9	8.2	44.4	60.6	16.2
97 Treatment	62.9	73.0	10.1	56.9	67.6	10.7	30.3	56.2	25.9
Experimental treatment effect:									
Cross section		8.0 (4.6)			12.8 (5.7)			7.1 (8.6)	
Difference-in-difference		10.3 (6.7)			14.1 (8.3)			17.7 (12.0)	

	Boys age 12–15			Boys age 12–15, behind in school			Boys age 13–15, HGC ≥ 6, behind in school		
	(1)	(2)	(2)–(1)	(1)	(2)	(2)–(1)	(1)	(2)	(2)–(1)
97 Control	68.8	79.6	10.8	64.0	75.8	11.8	59.0	72.7	13.7
98 Control	72.5	80.2	7.7	67.4	78.0	10.6	57.1	72.8	15.7
97 Treatment	69.5	79.4	9.9	64.2	75.8	11.6	52.6	71.6	19.0
Experimental treatment effect:									
Cross section		3.8 (4.2)			4.2 (5.2)			1.2 (8.4)	
Difference-in-difference		3.1 (6.1)			4.0 (7.4)			3.8 (11.7)	

Design-based $\Rightarrow$ Model-based: Identification/Estimation

- Use policy variation for estimation of the model parameters
  - $\Rightarrow$ Allows researcher to relax some behavioral/distributional assumptions
- In the Progresa case, MU of subsidy$\neq$MU of income
  - $\Rightarrow$ Transfers to mothers, and who receives the money likely matters

Attanasio, Meghir and Santiago (ReStud, 2012)

- Similar dynamic discrete choice structure with some differences
  - ⇒ No fertility decision
  - ⇒ Binary choice: school vs. work
  - ⇒ Each child's utility is independent of that of the parents/other children
  - ⇒ Allow for MU of the subsidy to differ from MU of other sources of income
  - ⇒ Allow for equilibrium effects of the program on children's wages

Identification and Estimation

$\Rightarrow$ Treatment-control + eligible-ineligible identify the effect of the subsidy

$\Rightarrow$ Within-grade variation in age identifies the effect of the subsidy amount

- Parameters of the model are estimated by simulated maximum likelihood
  - (Discrete) distribution of unobservables assumed independent of all observables
  - Distance to school as an IV to solve initial condition problem

Model-based$\Rightarrow$Design-based: Long-term Effect of *Progresa*

- Randomized phase-in design allows to study only short-term effect of CCT
- Todd-Wolpin forecast the impact of the program on completed schooling
  - Beginning at marriage and extending over the entire lifetimes of HHs

$\Rightarrow$ Fertility outcomes are essentially invariant to the subsidy

$\Rightarrow$ Small long run effects (compared to short-run) on attendance and attainment

Predictions of TW Model on Attendance and Attainment

	Girls		Boys	
	Short-run effect[a]	Long-run effect[b]	Short-run effect	Long-run effect
Control group				
1997	10.9	11.9	10.7	12.0
1998	11.2	12.3	11.4	12.7
Treatment group				
1997	11.2	12.3	11.3	12.4
1998	11.7	12.7	12.1	12.4

	Girls		Boys	
	No subsidy	Subsidy	No subsidy	Subsidy
Mean schooling	6.29	6.83	6.42	6.96
Percent completing grade six or more	75.8	82.2	78.8	83.3
Percent completing grade nine or more	19.8	25.9	22.8	28.0

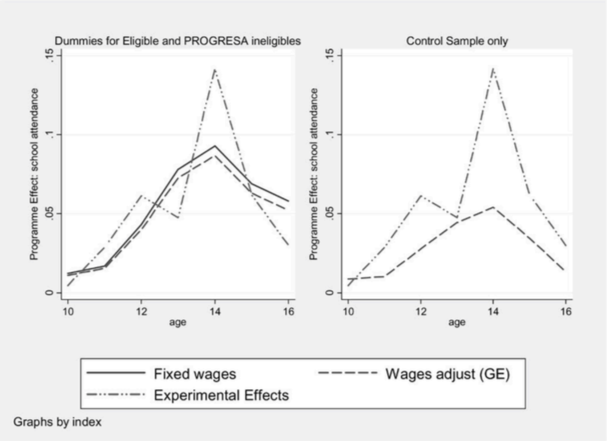
Model-based $\Rightarrow$ Design-based: GE Effect of *Progresa*

- An increase in child wages will reduce schooling
- Wages may be affected by the subsidy as it reduces children's labor supply
- This channel is important as it attenuates the program's impact on schooling
  - $\Rightarrow$ Depends on the elasticity of substitution between child and adult labor

$\Rightarrow$ Attanasio et al. use Model+RCT to quantify this GE effect

- Document an increase in the wage rate by 6%

In-sample Fit of the Attanasio et al. Model



Validation vs. Identification?

- Counterfactuals from Todd-Wolpin may be more credible
- But Attanasio et al. is more parsimonious, and yet more general

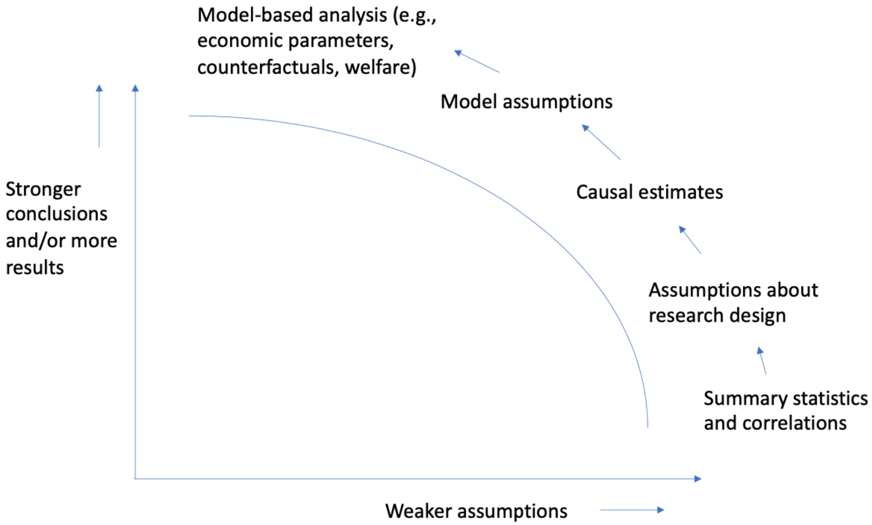
$\Rightarrow$ Difficult to account for behavioral responses using only the control group

- One should use (at least some) policy variation for identification/estimation
- Then, if possible, use extra sources of variation for out-of-sample validation

Toward a Synthesis Between the Two Approaches

1. Show your data/variation with descriptive analysis
2. Use the design-based analysis to provide preliminary evidence
3. Clearly articulate the value-added of the model
4. Use the design-based analysis to guide modeling choices and identification
5. Choose parameters and counterfactuals that are linked to your variation

Strength of assumptions and economically relevant results



Data-then-Model or Model-then-Data?

- Highly heterogenous preferences about types of model-based assumptions
- Readers accept model-based assumptions if they get something in return
- The Data-then-Model structure has some advantages
  - Allows readers to situate themselves at their favorite point on the frontier
  - "Get off the train" when they are no longer comfortable with the tradeoff