The Best of Both Worlds

Huan Deng

April 29, 2025

Outline

Introduction

Motivation

- Empirical policy analysis in economics traditionally takes one of two broad paths:
  1. **Program Evaluation Approach (Reduced Form):** Focuses on direct "treatment effects" (e.g. ATE, LATE) using methods like randomized controlled trials, matching, or IV.
  2. **Structural Approach:** Builds explicit economic models (preferences, technologies, constraints) to forecast policy impacts in new settings.
- **More recent view:**
  - Each approach has strengths and weaknesses.
  - We need *both* if we want credible causal identification and the ability to forecast novel policies.
- I follow Heckman (2010) and Todd and Wolpin (2023) in today's discussion.

The Program Evaluation Approach

Program Evaluation (Reduced Form)

- Emphasizes experimental analogies (RCTs) or quasi-experimental methods (IV, matching) to isolate "treatment effects."
- Typically good for:
  - Internal validity: Estimating causal impacts in settings close to an experiment.
  - Simple policy questions: "What is the average effect of X on Y?"
  - Complicated or "non-econ" settings: it would be too hard to specify a good model or no model as a guide at all.
- However:
  - Often struggles with *external validity*, i.e. transporting results to different contexts or policies that haven't been implemented.
  - Due to ethical or monetary constraints, we can't implement RCTs at "desired" scale or intensity.

The Structural Approach

Structural Approach

- Builds explicit models capturing key economic mechanisms (preferences, constraints).
- Can handle:
  - *Forecasting novel policies* or large-scale environmental changes (thanks to deeper microfoundations).
  - *Counterfactuals* that go well beyond the data observed, good external valifity.
- However:
  - More complex, requires stronger modeling assumptions (risk of misspecification).
  - More costly for researchers to produce and communicate

Heckman's Three Broad Classes of Policy Evaluation

(P1) Evaluating Impacts of Implemented Interventions

- How did a *current* or *historical* policy affect outcomes?
- Example: Assessing the ex-post returns of a job-training program that already ran.

(P2) Forecasting Impacts from One Environment to Another

- External validity challenge: applying results from one setting to a different setting.
- Example: If job-training worked in City A, how might it perform in City B?

(P3) Forecasting Impacts of Never-Before-Seen Interventions

- Most ambitious: predicting effects of *entirely new* or *unprecedented* policies.
- Example: A new form of education subsidy with no close historical analog.

Policy Invariance

- **Definition:** A *policy-invariant* parameter remains stable across a class of policy or environmental shifts.
- Often, the structural approach would claim that it estimates some deep parameters.
- **Key Role in P2 and P3:**
  - If preferences, technologies, or distributional parameters do *not* change when policies change, we can *transport* those estimates to new settings or design new interventions.
  - Essential for accurate forecasting of policies not observed in the past.
- **Potential Limitations:**
  - Large policy shocks can alter fundamental behaviors (violating invariance).
  - Requires careful *theoretical* and *empirical* justification.

The Best of Both Worlds

The Best of Both Worlds

The Idea

Introduction

- There is a longstanding debate on whether economic theory (structural models) or purely statistical methods (reduced-form or “experimental” methods) should dominate policy evaluation.
- Can we combine the strengths of **Randomized Controlled Trials (RCTs)** and **structural econometric modeling**?
- Todd and Wolpin (2023) argue that there is a **natural synergy** between structural approaches and RCT-based reduced-form approaches.

New Emerging View: Synergy

- RCT evidence can **validate or discipline** structural models:
  - Use **holdout samples** for out-of-sample checks.
  - Randomized design can help **identify parameters** that might be otherwise underidentified.
- Structural models can **extend RCT findings**:
  - Investigate **mechanisms** behind treatment effects.
  - Evaluate policy designs **not tested** in the RCT.
  - Offer **longer-term or general equilibrium** analyses.
- Todd and Wolpin (2023) emphasize **two main uses** of RCT in structural models:
  1. Parameter estimation: RCT variation as an extra source of identification.
  2. Model validation/selection: Using control/treatment groups as a holdout sample.

Ex Ante Evaluation and Counterfactuals

- One major goal of structural models is to predict outcomes under **new policies** (ex ante).
- RCT-based evidence is typically **ex post** for a specific treatment.
- By embedding RCT data in a behavioral model, researchers can:
  - Check how well the model **replicates RCT outcomes**.
  - Once validated, **simulate alternative** designs (e.g., different subsidy amounts, different target populations).
- **Crucial assumption**: The structural parameters are *invariant* to policy or environment changes relevant to the new designs (Lucas critique / Marschak critique).

Model Validation: Holdout Samples

- A validation approach:
  - Estimate structural model parameters using *either* the treatment group *or* the control group.
  - Predict outcomes in the *holdout* sample (the other group).
  - Compare predictions with **actual data**.
- If the model systematically fits *both* arms of the experiment, it gains credibility for broader policy counterfactuals.
- Sometimes, the RCT provides crucial identification variation. In this case, we can't afford to use one treatment arm as a hold-out sample.

The Best of Both Worlds

An Illustrative Example

Motivation

- We consider a static model where individuals choose whether to **work or not** in the presence of a possible **welfare program**.
- With this example, we:
  - illustrate **nonparametric** and **parametric** ways of doing *ex ante* policy evaluation using a structural model, a task that can hardly be done by the program evaluation approach.
  - show how **RCT data** can be leveraged to estimate or validate structural models.
- Framework is adapted from Todd & Wolpin (2023), demonstrating how *experimental* variation and *structural* approaches can be combined.

Basic Setup

Single Decision: $L_i = 1$ (not working) or $L_i = 0$ (working).

- Let y_i = *non-earned* income (e.g. child support).
- Let w_i = *wage* if working.
- Let $b(y_i, n_i) \geq 0$ = *welfare benefit* for nonworkers, possibly depending on number of children n_i.

Consumption C_i:

$$C_i = \begin{cases} y_i + w_i, & \text{if working } (L_i = 0), \\ y_i + b(y_i, n_i), & \text{if not working } (L_i = 1 \text{ and eligible}). \end{cases}$$

Utility Function:

$$U_i = U(C_i, L_i; \varepsilon_i),$$

where ε_i is an unobserved preference shock.

Nonparametric Ex Ante Evaluation

- Suppose we observe data from a **no-program** environment ($b = 0$).
- We want to *predict* labor supply if a program **were** introduced with benefit $b(y_i, n_i)$.
- **Insight:** If the new environment's budget set can be "re-labeled" to look like the old one, we may do a *matching* approach.

$$\tilde{y}_i = y_i + b(y_i, n_i), \quad \tilde{w}_i = w_i - b(y_i, n_i).$$

- Then deciding work vs. not work under $(\tilde{y}_i, \tilde{w}_i)$ in the new setting is analogous to deciding work vs. not work under (y_i, w_i) in the old setting.

Matching Estimator

- Let $H_i = 1 - L_i$ = indicator of *working* in original data.
- For each person j, define matched group as those with:

$$(y_i, w_i) \approx \big(y_j + b(y_j, n_j),\ w_j - b(y_j, n_j)\big).$$

- Todd and Wolpin (2008) develop a matching estimator of the policy impact on the employment rate:

$$\tilde{\Delta} = \frac{1}{J} \sum_{\substack{j=1 \\ j,i \in S_p}}^{J} [\tilde{\mathbb{E}}[H_i \,|\, y_i = y_j + b(y_j, n_j),\ w_i = w_j - b(y_j, n_j)] - H_j(y_j, w_j, n_j)], \quad (1)$$

  where S_p is the region of overlapping support.
- This difference approximates *the change in employment* from introducing $b(\cdot)$.

Why This Nonparametric Strategy Works

- Requires that the **unobserved preference** distribution (ε_i) is *independent of* (y_i, w_i).
- We exploit direct observation of how people behave at different (y, w) points.
- If $(\tilde{y}, \tilde{w})$ from the new policy *lies in the support* of the original (y, w) data, we can **match** to it and read off the typical labor supply choice.
- **In practice**, sample size constraints can complicate matching, and many real policies (like an EITC with phase-outs) break the one-to-one re-labeling.

Parametric Ex Ante Evaluation

- Many real-world programs have more complex structures:
  - Earnings-based benefit reduction rates,
  - Phase-in, phase-out segments,
  - Different rules for different subgroups.
- Nonparametric approach may fail if we cannot neatly "relabel" to old data.
- **Solution**: Specify a functional form for:
  - *Utility*, e.g. $U(C_i, L_i; \varepsilon_i)$,
  - *Wage offers*, e.g. $w_i = z_i\gamma + \eta_i$,
  - *Distribution* of unobservables (ε_i, η_i).

A Parametric Example

Utility Model:

$$U_i = C_i + \alpha_i L_i + \lambda\, C_i\, L_i, \quad \alpha_i = x_i \beta + \varepsilon_i.$$

Budget Constraint:

$$C_i = \begin{cases} y_i + w_i, & \text{if working} \\ y_i + b(y_i, n_i), & \text{if not working} \end{cases}$$

Labor Supply Decision:

$$v_i^* = U(\text{work}) - U(\text{not work}) = y_i + w_i - \Big[(1+\lambda)\big(y_i + b(y_i, n_i)\big) + x_i\beta + \varepsilon_i\Big].$$

- Person i works if $v_i^* \geq 0$.
- Distributions of (ε_i, η_i) must be specified for maximum likelihood or method of simulated moments.

Identification and Estimation

Key Points:

- Identification of the above model follows our discussion of Roy model.
- Exclusion restrictions help identify preference vs. wage parameters. E.g. some z_i that affects w_i but not preferences.
- Typically assume joint normal (ε_i, η_i) or rely on semiparametric methods (e.g. selection correction).

Once Estimated:

- With $(\hat{\beta}, \hat{\lambda}, \hat{\gamma}, \dots)$ in hand, the policymaker can *simulate* labor supply choices under new welfare or tax policies .
- This is the hallmark advantage of a structural model: **ex ante policy evaluation**.

Value of RCT Data

- If an RCT with random assignment to *treatment* (welfare) vs. *control* (no welfare) is available:
  - We see outcomes in both policy states.
  - That direct variation can identify parameters—especially if wages or preferences might differ under the new policy.
- **Holdout Validation**:
  - Estimate the model using the *treatment* group data only.
  - Predict behavior of the *control* group out-of-sample.
  - Compare predictions to observed outcomes.
  - If accurate, the model is more credible for further counterfactuals.

Stigma Effects (Illustration)

- Suppose there is a **stigma cost** ϕ_i for taking welfare:

$$U_i = C_i + \alpha_i L_i + \lambda\, C_i\, L_i - \phi_i\, P_i,$$

  where $P_i = 1$ if welfare is claimed (i.e. if $L_i = 1$ and the program is taken up).
- **Identification:** We need to observe how *eligible* individuals behave. If some choose not to claim, that reveals ϕ_i.
- Without actual *treatment* data, ϕ_i is unidentifiable. Observing *eligible non-takers* in the RCT is crucial.

Applications

Introduction

- Now we see some examples that illustrate the approach we just discussed.
- There are many excellent papers out there, so we can only cover a very small proportion of the current literature.
- This unavoidably reflects my limited knowledge and idiosyncratic preference.
- We discuss Todd and Wolpin (2006) and Allende et al. (2019): the former exploits the Conditional Cash Transfer (CCT) program in Mexico (namely, PROGRESA); the latter is based on an information RCT in Chile.

Applications

Use RCT for Model Validation: Todd and Wolpin (2006)

Context of PROGRESA

- **Origins:** Introduced by the Mexican government in the late 1990s (originally named PROGRESA, later Oportunidades).
- **Motivation:** Many rural families kept children out of school to work or help at home, leading to lower education levels and perpetuating poverty.
- **Program Mechanism:**
  - Conditional Cash Transfers (CCTs): Parents receive monthly cash payments only if children attend school regularly (typically 85% attendance).
  - Subsidy amounts increase with grade level (and are slightly higher for girls).
  - Program also supports health/nutrition.
- **Coverage:** Targeted low-income rural areas – eventually expanded nationwide.

Experimental Design of PROGRESA

- **Randomized Control Trial (RCT):**
  - 506 rural villages selected, then randomly assigned to *treatment* (receive subsidies) or *control* (no subsidies).
  - Baseline data collected prior to subsidy rollout. Follow-up surveys tracked changes over time.
- **Data Content:**
  - Household composition, income, children's schooling, labor activities, wages, distance to schools, etc.
  - Rich data allowing for both *reduced-form* impact analysis (comparing T vs. C) and *structural* modeling.
- **Previously Found Impacts:**
  - Direct *experimental* comparisons: 5–15 percentage point increase in school attendance in treatment villages (depending on child age/grade).
  - Gains particularly notable at junior secondary level, where dropout is common.

Descriptive Statistics

- **High enrollment at young ages:** Most 6–10-year-olds attend primary school, but attendance drops off after primary (i.e. after grade 6).
- **Child labor:** By ages 13–15, noticeable fraction (especially boys) work for pay; girls more likely to do home chores if not in school.
- **Fertility Patterns:** Typically, families have 3–4 children; childbearing often starts soon after marriage.
- **Distance factors:**
  - If no secondary school in the village, the child has to travel (or move) to attend. This raises the cost of schooling, especially for older girls.
  - Wage offers in local labor markets vary with distance to a city (urban wage premium).

Core Idea of the Model

- Todd and Wolpin (2006), TW06 hereafter, build a **dynamic decision model** where:
  - Household decides each period:
    1. Whether to have a new child (fertility choice),
    2. How each school-age child allocates time: attend school, work in the labor market, or stay at home.
  - The model extends over the mother's fertile years (up to mid-40s) and covers children's early years until 15 (or so), after which they "become independent."
- This structure captures:
  - *Utility from children:* Parents value having children, the children's schooling levels, and possibly having them at home.
  - *Budget constraints:* Parents have exogenous income plus any child's earnings if working; no savings/borrowing.
  - *Policy intervention (subsidies):* If a child attends school, the family may receive PROGRESA's monthly transfer.

Utility Function and States (Simplified)

- **Utility in each period t:** $U_t = U\big(C_t, \text{children's schooling}, \text{number/ages of children at home}, \dots\big)$
- **Income constraint:**

$$C_t = Y_{parents}(t) + \sum_{\substack{\text{children } i \\ \text{who work}}} w_i(t) - \text{possible tuition costs} + \text{subsidy if child attends school.} \quad (2)$$

- **State Variables:**
  - Ages of parents, of children, children's ages & schooling levels,
  - Distance to school/city (affects cost of schooling or child wages),
  - Unobserved household "type" capturing permanent preference/income differences,
  - Period-specific random shocks to preferences, wages, etc.
- **Decision Set:**
  - *Fertility:* become pregnant or not (if still feasible).
  - *Child i's activity:* school, work, or home (with some simplifying assumptions on ordering).

Household Heterogeneity

- **Permanent "Types":**
  - Households differ by unobserved, time-invariant parameters:
    1. Preference for more (or fewer) children,
    2. Strength of preference for schooling or for child's leisure,
    3. Average earning power (parental income, child wage offers),
    4. Probability a child fails a school grade, etc.
  - Todd & Wolpin assume a discrete distribution of these types (e.g., 3–4 types).
- **Shocks Each Period:**
  - Realization of parental income and child wage offers,
  - Random taste shocks to having a new child, or cost of schooling if child is "behind grade," etc.
- **Why This Matters:**
  - Households with high preference for schooling or high parent income $\rightarrow$ more likely to keep kids in school,
  - Households facing higher child wages $\rightarrow$ more likely to have kids work, especially at older ages,
  - Program evaluation must account for these differences to accurately forecast.

Estimation and Computational Strategy

- **Two main challenges**:
  1. Large decision space (fertility plus schooling/work choices for multiple children each period).
  2. Large state space (differing ages, schooling histories, repeated over many periods).
- **Strategies**:
  - *Approximate* the value function (the expected future utility for each possible choice).
  - Use *Monte Carlo integration* to handle random shocks.
  - Restrict the choice set in certain ways (e.g., older kids can only attend school if younger kids do too, etc.) to reduce dimensionality.
- The final result is an estimated dynamic model that captures the interplay of preferences, constraints, and random shocks in these households.

Validation (Table 12 From TW06)

TABLE 12—ACTUAL AND PREDICTED SCHOOL ATTENDANCE RATES BY CHILD AGE, SEX, AND SCHOOL ATTAINMENT: CONTROL AND TREATMENT GROUPS BY YEAR[a]

	Girls				Boys			
	Control group		Treatment group		Control group		Treatment group	
	1997	1998	1997	1998	1997	1998	1997	1998
Age 6–11								
Actual	96.9	96.5	97.6	98.5	96.6	96.7	97.6	98.7
Predicted	96.1	96.2	96.4	97.1	96.4	96.4	96.3	97.1
No. obs.	449	431	632	600	471	460	671	678
Age 12–15								
Actual	65.3	66.5	62.9	74.4	68.8	72.5	69.5	76.3
Predicted	61.6	61.8	61.8	74.9	68.8	68.8	68.0	77.1
No. obs.	190	176	205	223	189	182	279	262
Age 12–15 behind in school								
Actual	58.3	58.7	56.9	71.4	64.0	67.4	64.2	71.6
Predicted	54.2	55.5	55.6	72.3	63.9	65.3	62.7	72.9
No. obs.	127	121	144	161	139	135	204	190
Age 13–15 HGC $\geq$ 6 behind in school								
Actual	40.9	44.4	30.3	51.5	59.0	57.1	52.6	58.3
Predicted	40.2	45.3	37.3	58.7	55.0	53.0	51.7	66.7
No. obs.	66	72	66	66	61	56	95	96

[a] Based on 200 simulation draws per family.

Validation Continued (Table 13 From TW06)

TABLE 13—ACTUAL VERSES PREDICTED SUBSIDY EFFECTS ON PERCENT ATTENDING SCHOOL

	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)	

Counterfactual Analyses

- **Motivation:** The real RCT only tested one schedule of subsidies. Could we do *better* with a different structure?
- **Counterfactual Experiments**:
  - *Vary the amount of the subsidy by grade* (e.g. reduce or eliminate subsidies in lower grades—where attendance is already high—and shift more resources to junior secondary).
  - *Pure Income Transfer* (unconditional cash) vs. *conditional* on attendance.
  - *Infrastructure investment*: building more schools to reduce travel distance/time for older students.
  - *Child labor laws* enforcement.
- **Model Simulation Results:**
  - *Re-targeting older grades* can yield higher overall schooling attainment at roughly the same cost as standard PROGRESA.
  - *Unconditional transfer* yields significantly smaller schooling gains (parents do not face the attendance requirement).

Policy Implications

- The validated structural model shows that:
  - **Conditional** incentives (CCT) are more effective at boosting schooling than just handing out money.
  - Shifting subsidy amounts to higher grades (where dropout is more likely) can increase total years of schooling *more* than current baseline schedules.
  - Cost-effectiveness can be improved by fine-tuning the structure of payments.
- **Broader Lesson for Policy:**
  - RCT alone tells us the *impact of the tested policy*, but not about *untried* designs.
  - Structural models validated by experiments can explore these *alternative worlds* without needing a new field experiment each time.

Applications

Program Evaluation With Spillover/GE Effects: Allende et al. (2019)

Why Informed School Choice?

- **Key Issue:** Parents often lack perfect information about school quality, which can lead to suboptimal choices in educational markets.
- **Potential Consequences:**
  - Low-SES families may be especially misinformed, contributing to under-investment in quality schools.
  - If enough parents lack information, schools' incentives to improve quality may be muted.
- **Policy Relevance:**
  - Government or NGOs can provide better information (e.g., test scores, "report cards") to families.
  - How do such interventions *scale up*? Are there equilibrium changes in school behavior?

Setting: The Chilean School Market

- **Context**: Chile's education system has a widespread voucher system.
  - Families choose among public (municipal) schools, private voucher schools, and private non-voucher schools.
  - A majority of schools are "voucher-funded," with possible top-up fees.
- **Information Issues**: Despite standardized tests, parents may not know or may undervalue differences in school quality.
- **Socioeconomic Gaps**: Low-SES families tend to attend lower-quality schools. Differences persist even though voucher expansions have tried to level resources.
- This paper: *Intervenes* with a personalized info policy for Pre-K families and uses a structural model to capture *equilibrium* supply reactions.

Policy Intervention and RCT Design

Goal: Provide parents of Pre-K students with personalized information about local schools.

- **Sample:** The network of Integra preschools that provide Pre-K education to 25,229 students in the cohort of 3 to 4 years
- **Treatment:**
  - *Video* emphasizing the importance of school quality and future returns to education.
  - *Report Card* showing test scores and relevant data on nearby schools.
- **Random Assignment:** Some schools assigned to Treatment, others to Control.
- **Timeline:** Intervention delivered as families begin deciding on children's primary school enrollment.

Data Collection

- **Administrative Data**:
  - Records of all Integra preschools (locations, enrollment, SES).
  - Ministry of Education data on *all* primary schools (addresses, fees, test scores, enrollment).
  - Student-level data: test scores, maternal education, birth records, health info.
- **RCT** :
  - Gathered background on parents, whether child already enrolled/decided on a primary school.
  - Contact details for follow-up.
  - Survey after choice is made.
  - Linked to actual school matriculation data in subsequent years.
- **Longer-Term Outcomes**:
  - Standardized tests (SIMCE) in later grades (up to 4th or 8th).
  - Observed whether families switch or remain in initially chosen school.

Core Experimental Results

- **School Choices:**
  - Treated parents more likely to choose higher test-score schools and (on average) slightly higher-priced schools.
  - They also end up traveling a bit farther to reach these higher-quality schools.
- **Academic Achievement:**
  - After 4–5 years, the children in treated families score $\approx 0.2\sigma$ higher on standardized tests.
  - Suggests real improvements, not just short-run re-sorting.
- **Interpretation**: The targeted information changes families' awareness of (and willingness to pursue) better schools, leading to nontrivial test-score gains.

Limitations of Direct RCT Evidence

- The RCT is small-scale:
  - It is infeasible to conduct a large-scale information RCT.
  - The supply side (schools) sees relatively small changes in demand, $\Rightarrow$ minimal or no direct price/quality adjustments.
- **Scaling Up Concern**:
  - If the policy is implemented at a larger scale, *capacity constraints* in high-quality schools may bind.
  - Schools may respond by altering prices or quality if enough new demand arises.
- This paper use a *structural model of demand and supply* to approximate how equilibrium might change if the policy is expanded to all families.

Demand Side: School Choice Under Incomplete Information

- **Family Utility:**
  $U_{i,j} = \beta_k \cdot q_j - \alpha_k \cdot \text{price}_{i,j} + \lambda_k \cdot \text{distance}_{i,j} + \dots$
  - q_j: school quality (value-added/test-score-based measure).
  - $\text{price}_{i,j}$: out-of-pocket cost for family i.
  - $\text{distance}_{i,j}$ depends on family's location vs. school location.
  - Heterogeneous coefficients $\beta_k, \alpha_k, \lambda_k$ for different SES types k.
- **Imperfect Information:**
  - Families observe noisy signals of q_j (and possibly of price_j, distance_j).
  - $\Rightarrow$ The perceived quality might underweight actual q_j, especially for low-SES families with less info.
- **Information Treatment Shifts** the weight on quality (less noise).
- **Logit/Discrete Choice:** Aggregates over families' random tastes $\varepsilon_{i,j}$.

Supply Side: School Pricing and Quality Choices

- **Schools as Differentiated Firms:**
  - Each school chooses (p_j, q_j) to maximize profit, subject to capacity constraints.
  - Profit depends on government voucher, top-up fee, cost function.
- **Cost of Quality**:

$$\mathrm{MC}_j(q_j) = \gamma_0 + \gamma_1 \cdot q_j + \omega_j,$$

  plus any fixed cost.
- **Capacity Constraints**:
  - If many new families want a high-quality school, it may *fill up*.
  - Some families are "crowded out," or the school can choose to expand if profitable.
- **Equilibrium Concept**:
  - Demand side: families pick best school given beliefs.
  - Supply side: schools set (p_j, q_j) to maximize objective, or stay within constraints.
  - Market clears, ensuring no school is over its capacity (short-run) or can expand (medium-run scenario).

Estimation Strategy

- **1. Demand Estimation**:
  - Use administrative microdata on actual school choices in the entire market.
  - Incorporate random coefficients for SES types, location dummies, etc.
  - The RCT informs how the "treatment" changes perceived weights on school quality/price/distance.
- **2. Supply Estimation**:
  - Use variations in voucher policy + observed distribution of (p_j, q_j).
  - Identify cost structure parameters γ.
- **3. Counterfactual Simulation**:
  - Impose *all* families are "treated" with more accurate info signals.
  - Solve for new equilibrium: (p_j, q_j) with capacity constraints, schools re-optimizing.

Short-Run vs. Medium-Run Policy Effects

- **Short-Run (Fixed p_j, q_j):**
  - If all families receive better info, **demand** for high-quality schools jumps.
  - $\Longrightarrow$ Potential crowd-out at top schools if capacity binds.
  - Net effect on average achievement: *positive* but smaller than RCT partial equilibrium if many families are "pushed out."
- **Medium-Run (Schools Adjust (p_j, q_j)):**
  - If demand for better schools grows, some schools *raise quality* or expand capacity to capture new enrollment or set higher prices.
  - Authors find net effect is even more *positive* on average because:
    - Price may rise, but so does quality.
    - Low-SES students who see big returns to quality still can benefit from improved voucher policy.
  - *Heterogeneity:* some markets or families might face bigger crowd-out, but overall equilibrium effect is beneficial.

Main Quantitative Findings

- **Partial Equilibrium Gains**: $+0.20$ SD test scores for the typical low-SES child (seen in RCT).
- **Short-Run Equilibrium Effects** (capacity constraints only):
  - Gains shrink but remain positive, as highest-quality schools fill up quickly.
  - On average, effect is around half of the partial equilibrium estimate, e.g. $+0.10$ SD.
- **Medium-Run w/School Response**:
  - Some schools respond by raising q_j, expanding or adjusting price.
  - Net effect: Gains can exceed $+0.20$ SD if schools respond strongly to the new demand.
  - Policy can reduce socioeconomic achievement gaps.
- **Overall:** Affirmative evidence that an information policy at large scale can yield meaningful achievement improvements, despite capacity constraints.

Implications for Policy

- **Informing Families Works:**
  - Even a modest, low-cost intervention can shift demand toward better schools.
  - Gains in test scores persist, especially for poorer households.
- **Crowd Out vs. Quality Upgrading:**
  - Capacity constraints partly limit short-run gains if everyone gets the info at once.
  - However, supply-side responses (raising quality) can offset or amplify the net effect.
- **Methodological Insight**:
  - RCT data alone cannot fully predict scaled-up equilibrium outcomes.
  - A structural model can approximate equilibrium constraints and supply reactions.