Structural Estimation and Policy Evaluation in Developing Countries

Petra E. Todd   Kenneth I. Wolpin[1]

[1]University of Pennsylvania

March, 2010

Overview

- Ex post vs. ex ante approaches to policy evaluation
- Use of behavioral models for ex-ante evaluation
- Parametric assumptions not necessarily required
- Static vs. dynamic frameworks
- Applications
- Model validation

Ex Post Evaluation

- Goal is to evaluate impacts of an existing program
- Data on a treated group and on a comparison group
- Alternative approaches
    - Randomization
    - Difference-in-difference
    - Matching
    - Regression-discontinuity
    - Control function methods
    - IV methods, MTE, LATE
    - Estimation of a behavioral model

Ex Ante Evaluation

- Evaluate effects of changing parameters of an existing program
- Evaluate the impact of a new program prior to its implementation
  - Needed for optimal program design and placement, which requires simulating program effects and costs
- Evaluate effects of longer terms of exposure to an existing program than are observed in the data

Some Examples of Ex Ante Evaluations Using Static Models

- Forecast demand for a new good introduced into the choice set
    - e.g. McFadden (1977) - BART subway
- Forecast effect of changing the characteristics of a good on consumer demand
    - Berry,Levensohn and Pakes (1985) - changing car characteristics (e.g. price, fuel efficiency)

Some Examples of Ex Ante Evaluations Using Dynamic Models

- Wise (1985): Predict the effect of housing subsidy on housing demand
- Lumsdaine, Stock and Wise (1992): Predict the effect of retirement bonus on retirement patterns
- Lise, Seitz and Smith (2003) - Predict effects of welfare bonus program on job search
- Todd and Wolpin (2006) - Predict effects of school subsidy program on school attendance and work behaviors

The Importance of Economic Models in Ex Ante Policy Evaluation

- Koopmans (1947), Marschak (1953), Hurwicz (1962)
    - Recognize that an economic model provides a way of extrapolating from historical experience
    - Observe that it is not necessary to know the entire structure of the problem to answer certain policy questions (e.g. tax changes)

Recent Efforts at Nonparametric Ex Ante Evaluation

- Ichimura and Taber (1998, 2002)
  - Present general set of conditions under which nonparametric policy evaluation is possible
  - Estimate effects of a tuition subsidy using tuition variation in the data
- Heckman (2000,2001)
  - Discusses "Marshak's Maxim," provides new examples of where nonparametric assessment of new policies is feasible
- Blomquist and Newey (2002)
  - Nonparametric estimation of labor supply responses with nonlinear budget sets.
- Todd and Wolpin (2009)
  - Nonparametric estimation of effects of school and income subsidies on school attendance

Evaluate Effects of School Attendance Subsidy When Child Wage Offers are Observed (Todd and Wolpin (2009))

- Household makes a single period decision about whether to send a child to school or work
- Utility depends on consumption (c) and on whether the child attends school (indicator s).
- A child not attending school works at wage w.
- y denotes household income, net of the child's earnings
- The household solves the problem:

$$\begin{aligned} & \max_{\{s\}} \; U(c, s, \mu) \\ & s.t. \\ c &= y + w(1-s). \end{aligned}$$

The optimal choice $s^* = \varphi(y, w, \mu)$, where μ denotes unobservable preference heterogeneity.

- Consider a policy that provides a subsidy τ for school attendance. The problem becomes:

$$\begin{aligned} & \max_{(s)} \; U(c, s, \mu) \\ & s.t. \\ c &= y + w(1-s) + \tau s. \end{aligned}$$

The constraint can be rewritten as

$$c = (y+\tau) + (w-\tau)(1-s),$$

which shows that the optimal choice of s in the presence of the subsidy is $s^{**} = \varphi(\tilde{y}, \tilde{w}, \mu)$, where $\tilde{y} = y + \tau$ and $\tilde{w} = w - \tau$.

Estimation

- Under the assumption that :

$$f(\mu|y,w) = f(\mu|\tilde{y},\tilde{w}),$$

  Can estimate the effect of the subsidy program on the proportion of children attending school by comparing children from families with income $\tilde{y}$ and child wage offers $\tilde{w}$ to children from families with income y and child wages w.
- Clearly a stringent condition.
- To make more plausible, could condition on a vector of family characteristics, x, and assume:

$$f(\mu|y,w,x) = f(\mu|\tilde{y},\tilde{w},x).$$

Estimation

- A matching estimator of average program effects for those offered the program (the "intent-to-treat" estimator):

$$\frac{1}{n} \sum_{\substack{j=1 \\ j,i \in S_P}}^{n} \{E(s_i | w_i = w_j - \tau, y_i = y_j + \tau) - s_j(w_j, y_j)\},$$

  where $s_j(w_j, A_j)$ denotes the school attendance decision for a child of family j with characteristics (w_j, y_j).
- The average can only be taken over the region of overlapping support S_P, which in this case is over the set of families j for which the values $w_j - \tau$ and $y_j + \tau$ lie within the observed support of w_i and y_i.
- $E(s_i | w_i = w_j - \tau, y_i = y_j + \tau)$ can be estimated by nonparametric regression.

- We can evaluate the effects of a range of school subsidy programs.
- Nonparametric ex ante policy evaluation is feasible even when there is no variation in the data in the policy instrument (here, the price of schooling).

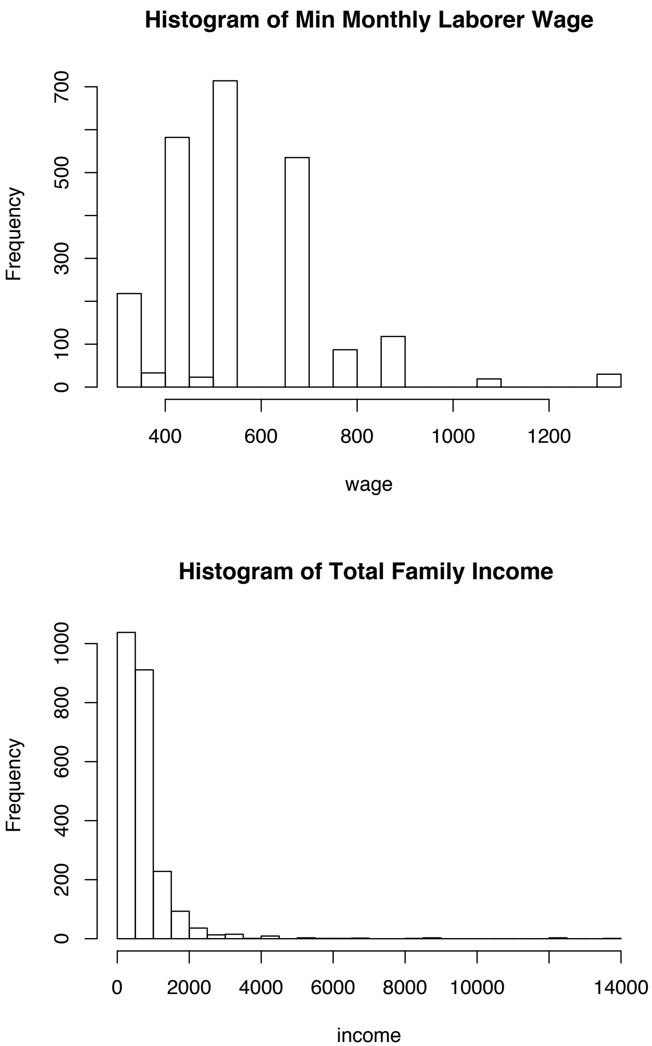
Application: The PROGRESA Program

- Large scale anti-poverty program
  - Begun in 1997, now has budget of about 1 billion US Dollars
  - About 20% of Mexican families participating
- Provides educational grants to mothers to encourage children's school attendance
- Benefit levels increase with grade level, higher for girls
- Subsidies amount to about 20% of average annual income
- Data from the initial rural evaluation that randomized 506 villages in or out of the program.

Table 1
Monthly Transfers for School Attendance

School Level	Grade	Gender	
		Female	Male
Primary	3	70	70
	4	80	80
	5	105	105
	6	135	135
Secondary	7	210	200
	8	235	210
	9	235	225

Figure 1



Histogram of Min Monthly Laborer Wage
Frequency
0
100
300
500
700
400
600
800
1000
1200
wage
Histogram of Total Family Income
Frequency
0
200
400
600
800
1000
0
2000
4000
6000
8000
10000
14000
income

Table 2(a)
Comparison of Ex-Ante Predictions and Experimental Impacts
Multiple-child model (Bootstrap standard errors in parentheses) †

	Boys			
Ages	Experimental	Predicted	Sample-Sizes‡	% overlapping support
12-13	0.05** (0.02)	0.05 (0.03)	374, 610	68%
14-15	0.02 (0.03)	0.09* (0.05)	309, 569	61%
12-15	0.03 (0.02)	0.06** (0.03)	683, 1179	64%
	Girls			
	Experimental	Predicted	Sample-Sizes‡	% overlapping support
12-13	0.07 (0.07)	0.04 (0.04)	361, 589	67%
14-15	0.11** (0.04)	0.11* (0.06)	361, 591	68%
12-15	0.09 ** (0.02)	0.07** (0.04)	677, 1180	68%
	Boys and Girls			
	Experimental	Predicted	Sample-Sizes‡	% overlapping support
12-13	0.06** (0.02)	0.04 (0.03)	735, 1199	67%
14-15	0.07** (0.03)	0.10** (0.04)	625, 1160	64%
12-15	0.06** (0.02)	0.07** (0.02)	1360, 2359	66%

†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.

Table 2(b)
Effects of Counterfactual Subsidy Levels
Multiple-child model (% in overlap region in parentheses)

	Boys		
Ages	2* Original	Original	0.75*Original
12-13	0.01 (50%)	0.05 (68%)	0.01 (92%)
14-15	0.16 (43%)	0.09 (61%)	0.04 (93%)
12-15	0.08 (47%)	0.06 (64%)	0.02 (93%)
	Girls		
	2* Original	Original	0.75*Original
12-13	0.04 (48%)	0.04 (67%)	0.04 (93%)
14-15	0.15 (52%)	0.11 (68%)	0.04 (93%)
12-15	0.09 (50%)	0.07 (68%)	0.04 (93%)
	Boys and Girls		
	2* Original	Original	0.75*Original
12-13	0.03 (49%)	0.04 (67%)	0.02 (93%)
14-15	0.15 (48%)	0.10** (64%)	0.04 (93%)
12-15	0.08 (49%)	0.07** (66%)	0.03 (93%)

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

Limitations of Nonparametric Policy Evaluation

- Not possible when there is an alternative use of children's time, such as leisure.
- Imposes strong assumptions on the distribution of unobservables.
- Child wage offers usually not observed when children are not working.

Parametric Static Model with Unobserved Heterogeneity and Partial Wage Observability

- A couple chooses between sending their child to work ($d_{it} = 1$) or school ($d_{it} = 0$)
- Utility is

$$U_{it} = C_{it} + \alpha_{it} (1 - d_{it}),$$

  where C_{it} is household i's consumption at period t.
- The utility the couple attaches to the child's school attendance, α_{it}, is time-varying:

$$\alpha_{it} = x_{it}\beta + \varepsilon_{it}$$

- x_{it} ($\subseteq X_{it}$) include, perhaps, parents' schooling or the child's gender.
- ε_{it} is an iid random preference shock to the utility of the child's school attendance (iid assumption can be relaxed)

- The child receives a wage offer of w_{it} and the household otherwise generates income y_{it}.
- The budget constraint is

$$C_{it} = y_{it} + w_{it} d_{it} ,$$

where there are assumed to be no costs associated with attending school.

- Wage offers only observed for children who work (partial observability), so we also need a wage offer equation:

$$w_{it} = z_{it}\gamma + \eta_{it},$$

- z_{it} ($\subseteq Z_{it}$) would contain, for example, the child's age, gender, or factors affecting the demand for child labor, such as distance to a city.
- η_{it} is an iid wage shock
- We do not include the child's current educational attainment in z to maintain the static nature of the model.

- Alternative-specific utilities, U_{it}^{1} if the child works and U_{it}^{0} if the child attends school as

$$\begin{aligned} U_{it}^{1} &= y_{it} + w_{it} \ , \\ U_{it}^{0} &= y_{it} + x_{it}\beta + \varepsilon_{it}. \end{aligned}$$

- Substituting the wage equation yields U_{it}^{1} - U_{it}^{0}

$$\begin{aligned} v_{it}^{*}(x_{it}, z_{it}, \varepsilon_{it}, \ \eta_{it}) &= z_{it}\gamma - x_{it}\beta + \eta_{it} - \varepsilon_{it} \\ &= \xi_{it}^{*}(\Omega_{it}^{-}) + \xi_{it}, \end{aligned}$$

where $\xi_{it} = \eta_{it} - \varepsilon_{it}$, $\xi_{it}^{*}(\Omega_{it}^{-}) = z_{it}\gamma - x_{it}\beta$ and Ω_{it}^{-} consists of z_{it} and x_{it}.

Estimation: Likelihood Function

- The likelihood function, incorporating the wage information, is

$$L(\theta; x_{it}, z_{it}) = \Pi_{i=1}^{I} Pr(d_{it} = 1, w_{it} | \Omega_{it}^{-})^{d_{it}} Pr(d_{it} = 0 | \Omega_{it}^{-})^{1-d_{it}}$$

Ex Ante Evaluation: Predict Effects of a Subsidy

- Assume that f(ε, η) is joint normal with variance-covariance matrix, $\Lambda = \begin{pmatrix} \sigma_\varepsilon^2 & \cdot \\ \sigma_{\varepsilon\eta} & \sigma_\eta^2 \end{pmatrix}$.
- Parameters to be estimated include β, γ, π, σ_ε^2, σ_η^2, and $\sigma_{\varepsilon\eta}$.
- Joint normality is sufficient to identify the wage parameters (γ and σ_η^2) as well as (σ_η^2 - $\sigma_{\varepsilon\eta)}$)/$\sigma_\xi$ (Heckman 1979).

- The probability that the child works is

$$pr(d_{it} = 1 | z_{it}, x_{it}) = \Phi(z_{it}(\gamma/\sigma_\xi) - x_{it}(\beta/\sigma_\xi))$$

  where Φ is the standard normal cumulative distribution.
- Data on work choices identify γ/σ_ξ and β/σ_ξ.
- To identify σ_ξ, there are three types of variables: - variables only in z (in the wage function), - variables only in x (in the utility function), and - variables in both x and z.
- Having identified the $\gamma' s$, the identification of σ_ξ (and thus also $\sigma_{\varepsilon\eta}$) requires at least one variable only in the wage equation.
- For example, a variable that affects the demand for labor but does not affect the utility value the couple places on the child's school attendance.

Predict effects of a subsidy

- Suppose the government wants to predict the effects of a schooling subsidy
- With the subsidy τ

$$pr(d_{it} = 1 | z_{it}, x_{it}) = \Phi(z_{it}(\gamma/\sigma_{\xi}) - x_{it}(\beta/\sigma_{\xi}) - (\tau/\sigma_{\xi}))$$

- It is necessary to have identified σ_{ξ} to predict the effects of the subsidy
- Government outlays on the program equal the number of children times the probability of attending school.
- Can study effects of a range of subsidies.
- Exogenous variation in the wage (independent of utility) is crucial for identification.

Ex Ante Evaluating Using Dynamic Models

- In the static model, there was no connection between the current period decision and future utility.
- Suppose that child's wage increases with work experience

$$w_{it} = z_{it}\gamma_1 + \gamma_2 h_{it} + \eta_{it},$$

where $h_{it} = \sum_{\tau=1}^{\tau=t-1} d_{i\tau}$ is work experience at the start of period t.

- Alternatively, parents' utility could depend on the number of school years completed, so that current attendance affects future utility.

Dynamic Model continued

- The couple maximizes the PDV of remaining lifetime utility starting from t=1 and ending at T.
- $V_t(\Omega_{it})$ denotes the maximum expected present discounted value of remaining lifetime utility at t given the state space and discount factor δ,
- The state space at t consists of all factors, known to the individual at t, that affect current utility or the probability distribution of future utilities.

$$V_t(\Omega_{it}) = max_{d_{it}} E(\sum_{\tau=t}^{T} \delta^{\tau-t}[U^1_{i\tau} d_{i\tau} + U^0_{i\tau}(1 - d_{i\tau}]|\Omega_{it}).$$

- With the wage equation, h_{it} becomes part of the state space and evolves according to $h_{it} = h_{i,t-1} + d_{i,t-1}$

- The value function can be written as the maximum over the two alternative-specific value functions, $V_t^k(\Omega_{it})$, k $\in\{0, 1\}$

$$V_t(\Omega_{it}) = max(V_t^0(\Omega_{it}), V_t^1(\Omega_{it}))$$

each of which obeys the Bellman equation

$$\begin{aligned} V_t^k(\Omega_{it}) &= U_{it}^k + \delta E[V_{t+1}(\Omega_{i,t+1})|\Omega_{it}, d_{it} = k] \text{ for } t < T, \\ &= U_{iT}^k, \text{ for } t = T. \end{aligned}$$

- The expectation is taken over the distribution of the random components of the state space at t+1 conditional on the state space elements (here the shocks are mutually serially independent.)

- The latent variable in the dynamic case is $V_t^1(\Omega_{it}) - V_t^0(\Omega_{it})$:

$$
\begin{aligned}
v_t^*(\Omega_{it}) &= z_{it}\gamma_1 + \gamma_2 h_{it} - x_{it}\beta - \varepsilon_{it} + \eta_{it} \\
&+ \delta([E[V_{t+1}(\Omega_{i,t+1})|\Omega_{it}, d_{it} = 1] \\
&- [E[V_{t+1}(\Omega_{i,t+1})|\Omega_{it}, d_{it} = 0]) \\
&= \xi_{it}^*(\Omega_{it}^-) + \xi_{it}.
\end{aligned}
$$

- A full solution of the dynamic programming problem consists of finding E[max($V_t^0(\Omega_{it})$, $V_t^1(\Omega_{it})$)] at all values of Ω_{it}^-, denoted by Emax(Ω_{it}^-), for all t=1,...,T.
- Same as static case, except now includes the difference in the future component of the expected value functions under the two alternatives.

Estimation: Likelihood function

- Assume researcher has data from t_{1i} to t_{Li}.

$$L(\theta; x_{it}) = \Pi_{i=1}^{I} \Pi_{\tau=t_{1i}}^{t_{Li}} Pr(d_{i\tau} = 1, w_{i\tau} | \Omega_{i\tau}^{-})^{d_{i\tau}} Pr(d_{i\tau} = 0 | \Omega_{i\tau}^{-})^{1-d_{i\tau}}$$

- where Pr(d$_{i\tau}$ = 1, w$_{i\tau}$) = Pr($\xi_{i\tau} \geq -\xi_{i\tau}^{*}(\Omega_{i\tau}^{-})$, $\eta_{i\tau}$ = w$_{i\tau}$ - z$_{i\tau}\gamma_1$ - γ_2 h$_{it}$) and Pr(d$_{i\tau}$ = 0) = 1 - Pr($\xi_{i\tau} \geq -\xi_{i\tau}^{*}(\Omega_{i\tau}^{-})$).
- If the error is not additive, then calculating the joint regions of the error that determine the probabilities that enter the likelihood can be done numerically.

Extension to Multinomial Choice

- If there are K>2 mutually exclusive alternatives, there will be K-1 latent variable functions (relative to one of the alternatives, arbitrarily chosen).
- Having to solve the dynamic multinomial choice problem, that is, for the $E[\max(V_t^0(\Omega_{it}), V_t^1(\Omega_{it}),......, V_t^K(\Omega_{it}))]$ function at all values of Ω_{it}^- and at all t, is computationally more intensive.
- Defining d_{it}^n as the discrete {0,1} choice variable corresponding to the nth choice (n =1,..,N) and $\widetilde{d}_{it}$ as the N element vector of those choices, there would be at most $K = 2^N$ mutually exclusive choices.

Allowing for Permanent Unobserved Heterogeneity

- In the example, unobservables were iid, but serial dependence is feasible.
- A standard specification assumes that agents can be distinguished, in terms of preferences and opportunities, by a fixed number of types. (Similar to approach of Heckman and Singer, 1981, in duration analysis)
- If a family was of type j, the preference for school attendance might be specified as $\alpha_{ijt} = \alpha_{oj} + x_{it}\beta + \varepsilon_{it}$ and the child's wage offer as $w_{ijt} = \gamma_{oj} + z_{it}\gamma_1 + \gamma_2 h_{it} + \eta_{it}$.
- The dynamic program must then be solved for each type and the likelihood function is a weighted average over each type in the sample.
- Type proportions are estimated along with the other parameters.

Applications in Development Economics

- What policies are effective in increasing educational attainment and improving school quality?
  - *Mexico*: Todd and Wolpin (2006), Attanasio, Meghir and Santiago (2005)
  - *India*: Duflo, Hanna and Ryan (2008)
  - *Chile*:Bravo, Mukhopadyay and Todd (2009)
- How do government pension programs affect household labor supply?
  - *Indonesia*: McKee (2006)
  - *Chile*: Velez-Grajales (2009), Joubert (2010)
- What policies are effective in increasing business investments by households?
  - *India*: Rosenzweig and Wolpin (1993)
  - *Thailand*: Kaboski and Townsend (2007)
- How does immigration policy affect the flow of migrants between developing and developed countries?
  - *Mexico-U.S.*: Colussi (2006)

Applications in Development Economics

- What policies are effective in increasing educational attainment and improving school quality?
  - *Mexico*: Todd and Wolpin (2006), Attanasio, Meghir and Santiago (2005)
  - *India*: Duflo, Hanna and Ryan (2008)
  - *Chile*:Bravo, Mukhopadyay and Todd (2009)
- How do government pension programs affect household labor supply?
  - *Indonesia*: McKee (2006)
  - *Chile*: Velez-Grajales (2009), Joubert (2010)
- What policies are effective in increasing business investments by households?
  - *India*: Rosenzweig and Wolpin (1993)
  - *Thailand*: Kaboski and Townsend (2007)
- How does immigration policy affect the flow of migrants between developing and developed countries?
  - *Mexico-U.S.*: Colussi (2006)

Applications in Development Economics

- What policies are effective in increasing educational attainment and improving school quality?
  - *Mexico*: Todd and Wolpin (2006), Attanasio, Meghir and Santiago (2005)
  - *India*: Duflo, Hanna and Ryan (2008)
  - *Chile*:Bravo, Mukhopadyay and Todd (2009)
- How do government pension programs affect household labor supply?
  - *Indonesia*: McKee (2006)
  - *Chile*: Velez-Grajales (2009), Joubert (2010)
- What policies are effective in increasing business investments by households?
  - *India*: Rosenzweig and Wolpin (1993)
  - *Thailand*: Kaboski and Townsend (2007)
- How does immigration policy affect the flow of migrants between developing and developed countries?
  - *Mexico-U.S.*: Colussi (2006)

Applications in Development Economics

- What policies are effective in increasing educational attainment and improving school quality?
  - *Mexico*: Todd and Wolpin (2006), Attanasio, Meghir and Santiago (2005)
  - *India*: Duflo, Hanna and Ryan (2008)
  - *Chile*:Bravo, Mukhopadyay and Todd (2009)
- How do government pension programs affect household labor supply?
  - *Indonesia*: McKee (2006)
  - *Chile*: Velez-Grajales (2009), Joubert (2010)
- What policies are effective in increasing business investments by households?
  - *India*: Rosenzweig and Wolpin (1993)
  - *Thailand*: Kaboski and Townsend (2007)
- How does immigration policy affect the flow of migrants between developing and developed countries?
  - *Mexico-U.S.*: Colussi (2006)

How Universal School Vouchers Affect Educational and Labor Market Outcomes: Evidence from Chile (Bravo, Mukhopadhyay, and Todd)

- Chile adopted a nationwide school voucher program in 1981
  - Part of broader market-oriented reforms led by military government
  - Chicago economists (incl Milton Friedman) played important role in voucher program design
- Since then, Chile has a relatively unregulated, decentralized, competitive market for private and public schooling
  - Three types of schools (public, subsidized private, nonsubsidized private
  - Free entry into private sector
  - "Funds follow child" voucher design"
  - Government oversight of teacher licensing, publicized standardized test results

How Universal School Vouchers Affect Educational and Labor Market Outcomes: Evidence from Chile (Bravo, Mukhopadhyay, and Todd)

- Chile adopted a nationwide school voucher program in 1981
  - Part of broader market-oriented reforms led by military government
  - Chicago economists (incl Milton Friedman) played important role in voucher program design
- Since then, Chile has a relatively unregulated, decentralized, competitive market for private and public schooling
  - Three types of schools (public, subsidized private, nonsubsidized private
  - Free entry into private sector
  - "Funds follow child" voucher design"
  - Government oversight of teacher licensing, publicized standardized test results

Goals of Paper

- Examine evidence for whether public or private schools improved after the introduction of the voucher reforms
- Study effects of school vouchers on school choice, educational attainment and on longer term earnings and labor market outcomes
- Explore effects of the reforms on inequality
- Approach:
  - Develop and estimate a dynamic model of schooling and working decisions using data from the *Enquesta Proteccion Social* (EPS) collected in 2002 and 2004
  - No exogenous variation in the timing of the voucher reform but exploit the fact that Chileans were differentially exposed to the program, depending on their age at the time of introduction

Model

- Builds on long labor literature on schooling and occupational choice models
  - Roy, 1951, Heckman and Honore, 1990, Heckman and Sedlacek, 1985 Ben-Porath, 1967, Rosen and Willis, 1979, Keane and Wolpin, 1997
- Wage offers represent a price paid to human capital and better schools augment human capital

  Behrman and Birdsall, 1983, Card and Krueger, 1992, Heckman, Layne-Farrar and Todd, 1996

Model Structure

- At ages 6-15, decide sequentially whether to continue in school or stay at home
- Choose whether to attend public, private subsidized or private nonsubsidized primary and secondary schools
- Can attend college for up to 5 years
- At age 16, start receiving wage offers and decide whether to work (schooling and work mutually exclusive)
- Wage offers depend on type of schooling attended, years attended, and whether attended during pre or post voucher regime, and accumulated work experience
- After leaving school, decide whether to work up to age 65.

Table 13
Simulated effect of voucher program on education outcomes
by family background status

	Complete sample†			Poor Subsample††			NonPoor Subsample‡		
	With Program	Without Program	Diff	With Program	Without Program	Diff	With Program	Without Program	Diff
% Attending private subsidized primary	26.1	17.3	8.8	25.3	16.7	8.6	26.5	17.6	8.9
% Attending private nonsubsidized primary	6.7	9.4	-2.7	6.4	8.9	-2.5	6.9	9.6	-2.7
% Attending private subsidized secondary	22.4	13.0	9.4	21.6	12.3	9.3	22.8	13.2	9.6
% Attending private nonsubsidized secondary	5.7	5.5	0.2	5.3	5.0	0.3	5.7	5.6	0.1
% Attending college	30.1	27.0	3.1	29.1	25.8	3.3	30.9	27.6	3.3
25% quantile years of education	10	10	0	10	10	0	11	10	1
Median years of education	12	12	0	12	12	0	12	12	0
75% years of education	13	13	0	13	13	0	13	13	0

†Refers to sample of individuals exposed to voucher program at any point in their schooling careers.
†† Refers to subsample that reported family background as indigent or poor.
‡Refers to subsample that reported family background as good or very good.

Table 15a
Voucher Program Impact on Labor Market Outcomes
(Earnings and Labor Force Participation)

	Complete sample†		Poor Subsample††		NonPoor Subsample‡	
	With Program	Without Program	With Program	Without Program	With Program	Without Program
Earnings of Workers ages 16-25	3153	3168	3040	3054	3211	3227
ages 26-35	4672	4733	4565	4619	4727	4791
ages 36-45	5258	5263	5129	5129	5324	5331
ages 16-45	4361	4388	4245	4267	4421	4550
Percent of time participate in the labor force ages 16-25	58.3	60.2	59.6	61.5	57.6	59.5
ages 26-35	92.8	92.7	93.0	93.0	92.7	92.6
ages 36-45	93.8	93.5	94.0	93.7	93.7	93.4
ages 16-45	81.6	82.1	82.2	82.7	81.3	81.8

†Refers to sample of individuals exposed to voucher program at any point in their schooling careers, over Ages 16-45.
†† Refers to subsample that reported family background as indigent or poor.
‡Refers to subsample that reported family background as good or very good.

Table 16
Voucher Reform Impact on Present Discounted Lifetime Earnings and Utility

	Discounted lifetime Earnings (from age 16 to age 45)		Discounted Lifetime Utility (from age 6 to age 45)	
Percentile	With Reform	Without reform	With reform	Without reform
1	11138	10980	9625	8309
5	11797	11663	10741	9382
10	12231	12122	11430	10048
50	13760	13542	13049	11640
90	17844	18015	15870	14675
95	18397	18568	16507	15271
99	19381	19689	17625	16322
Mean	14679	14646	13510	12217
S.D	2223	2360	1766	1851
90-10 ratio	1.46	1.49	1.39	1.46
50-10 ratio	1.12	1.12	1.14	1.16

"Migrant's Networks: An Estimable Model of Mexican Migration" Colussi (2006)

- Mexican illegal immigration has been and continues to be a central US policy concern
- Bracero program: 1942 negotiated treaty with Mexico to import agricultural workers on temporary visas
  - Peaked at 400,000 annually between 1955-1960
  - Program ended in 1965
- Immigration Reform and Control Act of 1986
  - Employer sanctions
  - Border patrol resources increased
  - Amnesty for those continuously in US since 1982
- Immigration Acts of 1990, 1996 - increased resources for border patrol

Model Structure

- Choice Set
  - In each period, either work in home village in Mexico or work in the US (allows return migration)
- Salient features
  - Probability of receiving a job offer in the US depends on the village network in the US, measured by the percent of villagers currently in US, and on tenure in the US
  - Non-pecuniary cost of residing in US that depends on current tenure and is heterogenous
  - One period cost of crossing the border
  - The US wage rate is exogenous and follows and autoregressive process

Model Structure

- Salient features:
  - Mexican village wage endogenously determined
    - Stochastic demand for agricultural works in village
    - Supply of workers = potential workers - migrants
  - Beginning from zero migrants (end of Bracero program), model solves for equilibrium path of number of migrants up to the steady state.

Data and Estimation

- Data: Mexican Migration Project
  - Retrospective information on households within a set of three villages surveyed in 1988
  - Migration histories of household heads starting in 1965, wages on last trip to US
  - Between 60% and 85% of crossings are undocumented
- Estimation method - Simulated Method of Moments

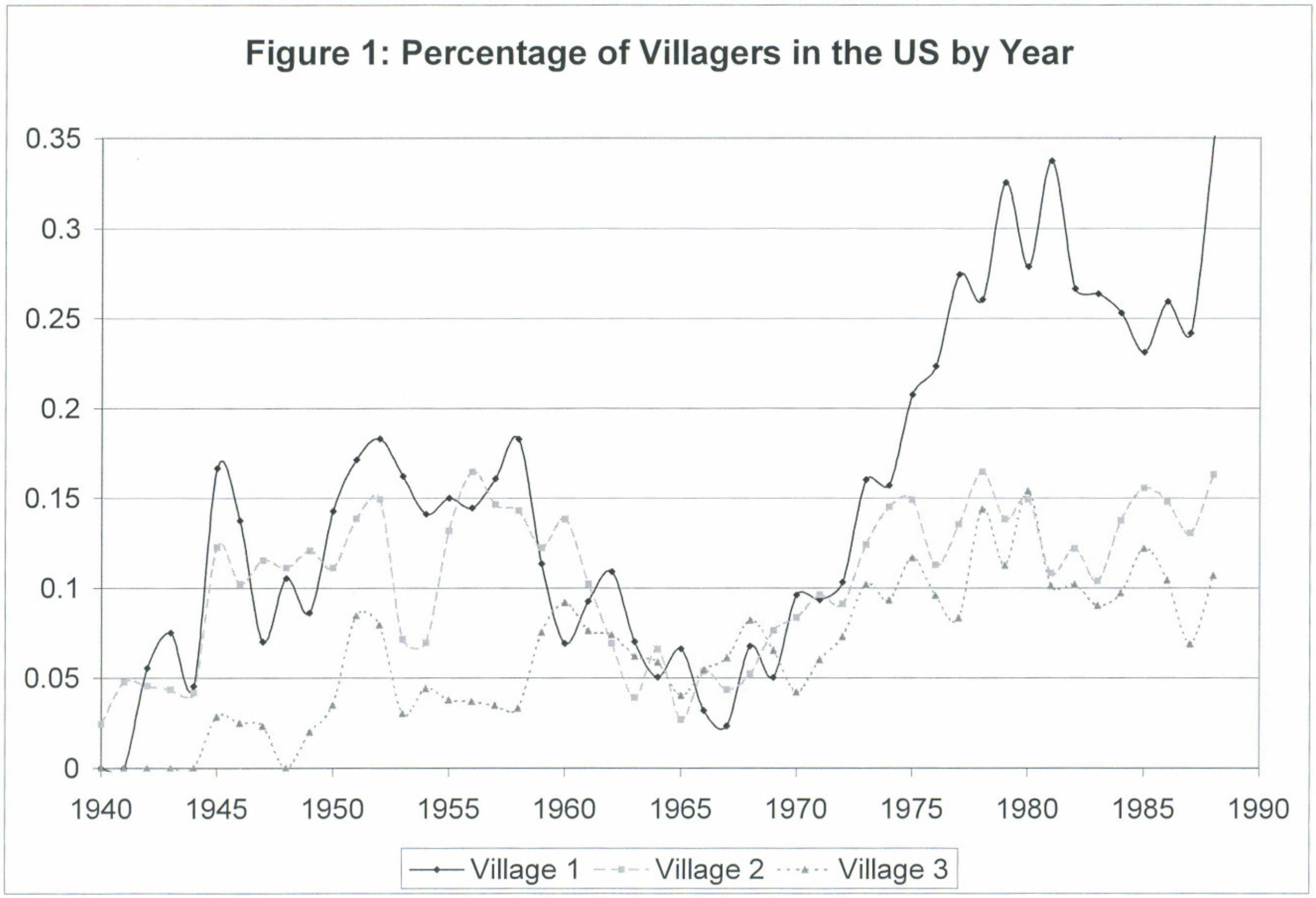


Figure 1: Percentage of Villagers in the US by Year
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1940
1945
1950
1955
1960
1965
1970
1975
1980
1985
1990
Village 1
Village 2
Village 3

Table 10: The Effect of Increased Enforcement on Illegal Immigration (Village 1)

	Baseline[a]	Cost of Border Crossing		U.S. Wage Reduction	
		One Month U.S. Wages	Six Months U.S. Wages	10%	25%
Steady State Fraction in U.S.	.32	.35	.28	.27	.02
Ave. Duration in U.S.(yrs.)	5	7	19	4	2

a. Cost of border crossing = one week of U.S. wages.

Conclusions

- A major benefit of the structural modeling approach is that it allows for ex ante evaluation of policy interventions
- However, models rely on extra-theoretic modeling and distributional assumptions, so model validation is an important concern
- Different approaches to Validation
    - Check robustness to alternative modeling assumptions
    - Examine within sample fit
    - Examine out of sample fit to data that were not used in estimation

Conclusions

- Randomized social experiments provide special opportunities for model validation
  - Can estimate the behavioral model on the control group and predict the behavior of the treatment group (or vice versa)
  - Randomization ensures that unobserved heterogeneity distribution same across groups that differ in the treatment
  - Recent applications: Todd and Wolpin, 2006, Duflo, Hanna, and Ryan, 2008, Kaboski and Townsend, 2007
- Exploiting the natural synergy between field experiments and structural approaches requires collecting rich data under the experiment of the type needed to estimate a behavioral model.