Assessing the Impact of School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility

Petra Todd and Kenneth I. Wolpin

2003

Ex post evaluation

- Policy evaluation based on a randomized experiment.
- The results obtained in the experiments are limited. Usually the data that can be obtained from the experiments are over short time horizon, and the set of policy instrument that can be changed as well as the magnitude of the policy changes are small.
- The program specific treatment effect can be obtained with miminal functional form assumptions. Interpretation of the result is more difficult.

Ex ante evaluation

:

- Estimate parameters of a structural behavioral model of decision makers, which are assumed to be policy invariant.
- After the estimation of the structural parameters, wide range of policy experiments can be easily conducted. (Rust: structural models: virtual crash dummies).
- Estimation of structural models often require strong functional form assumptions. In complex models, identification often rely on functional form assumptions.

Validation of Structural Models Using Randomized Experiments

- Estimate structural models using the data from the control group.
- Based on the estimated strcutural parameters, solve the model given the policy change imposed on the treatment group.
- If the simulated data is close to the data obtained from the treatment group, then the model specification has good predictive performance for policy experiments.
- Once the estimated model is validated by the treatment data, then one can use the model and the parameter estimates for other policy experiments.

The PROGRESA Program

- Treatment group: 320 villages, control group: 186 villages.
- baseline survey: Oct. 1997, March 1998, followup survey: Oct. 1998, May 1999, November 1999.
- Household demographics, income, school attendance, employment, wages of children of all households in villages.
- local data: distance to nearest secondary school, nearest city.

- household in treatment villages get subsidies if the benefit is one fourth of average family income.
- parents receive subsidies for each grade eligible child that attends school 85 % of the time.
- subsidies increase with the grade level, up to grade 9.
- Data: landless nuclear household, spouse less than 50 years old.
- 1,316 households in control villages, 1,885 households in treatment villages.
- Both eligible and ineligible households were used for estimation.

- School attendance almost universal from age 7 to age 11.
- Attendance rate decline rapidly after 12, more for girls.
- Once children leave school, they never return.
- Fertility occurs rapidly after marriage.

The Model

Parents maximize the discounted present value of lifetime utility:

$$V(\Omega, t) = Max_{d_k(t)} E\left[\sum_{t=1}^{T} \delta^{\tau - t} U(t) | \Omega(t)\right]$$

T: terminal period: woman's age 59.
Bellman equation:

$$\begin{aligned} V(\Omega, t) &= Max_{k \in K(t)} \left[V^k(\Omega, t)\right] \\ V^k(\Omega, t) &= U^k(t, \Omega(t)) + \delta E(V(\Omega(t+1), t+1 | d_k(t) = 1, \Omega(t))) \\ &\quad , t < T \\ V^k(\Omega, t) &= U^k(T, \Omega(T)), t = T \end{aligned}$$

Parents' per period utility function:

$$\begin{aligned} U(t) &= U(C(t), p(t), n(t), s_b(t), s_g(t), S_b(t), S_g(t) \\ &\quad , l_b(t), l_g(t), z_s; \epsilon(t), \mu) \end{aligned}$$

C: household consumption.
p: pregnancy
n: history of birth.
s: school attendance.
S: cumulative schooling.
l: children at home (homework).
μ: household unobserved heterogeneity.
ϵ: i.i.d. shock.
z_s: distance to the nearest village with the secondary school.

Other

- utility loss of child lagging in grades completed,
- utility loss of child attending grade 10.
- Value of having older girl home depends on whether young children at home.

Family consumption

$$C(t) = y_p(t) + \sum_n y_o(t, \tau_n) h(t, \tau_n)$$

$y_p(t)$: parents' income in time t.
$y_o(t, \tau_n)$ child's income born in year τ_n

Parents' Income

$$logy_p(t) = y_p(a_p(t), z_c, \epsilon_{y_p}; \mu_{y_p})$$

a_p:husband's age
z_c:distance to the nearest city.
ϵ_{y_p}:income shock.
μ_{y_p}: parents specific unobserved heterogeneity.

Other Income

$$log(y_o(t, \tau_n)) = y_o(t - \tau_n, I(b(\tau_n) = 1), z_c, \epsilon_{y_o}(t); \mu_{y_o})$$

$b(\tau_n)$: gender
$t - \tau_n$: age

Grade completion probability

$$\pi_c(t, \tau_n) = \pi(t - \tau_n, S(t, \tau_n) | s(t, \tau_n) = 1, \mu_c)$$

$S(t, \tau_n)$: years of schooling
ϵ: seriall uncorrelated, jointly normally distributed.
μ: discrete k types.

Parents' choice variable

- pregnancy
- which children to send to school
- which children to work in the market
- which children to stay home

Control Sample and Treatment Sample Fit

In general, the model fits well in both control sample and treatment sample.

Table 7
Predicted Selected Characteristics by Unobserved Type

	Type 1		Type 2		Type 3	
	Girls	Boys	Girls	Boys	Girls	Boys
% of children age 6-11 in school	98.5	99.4	97.6	99.9	78.7	64.2
% of children age 12-15 in school	37.3	50.2	84.6	86.9	44.5	36.8
% of children age 12-15 at home	55.9	31.0	11.3	7.0	33.5	30.9
% of children age 12-15 at work	6.8	18.8	4.1	6.1	21.9	32.3
Mean wage of children age12-15	2675	3599	2599	3499	2738	3665
Mean parental income	9916		11927		10124	
Percent becoming pregnant	15.3		5.7		15.0	
Percent of Sample	36.0		55.9		8.1	

Table 8
Actual and Predicted Choice Distribution
by Child Age and Sex

	Actual			Predicted			
Age	School	Work	Home	School	Work	Home	χ^2
Boys							
6	0.934	-	0.067	0.923	-	0.077	0.58
7	0.982	-	0.019	0.980	-	0.020	0.02
8	0.987	-	0.013	0.980	-	0.020	0.99
9	0.994	-	0.006	0.980	-	0.020	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	0.521	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	0.987	-	0.013	0.976	-	0.024	1.96
9	0.991	-	0.009	0.976	-	0.024	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	0.723	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

$\chi^2(.05,1)=3.84$, $\chi^2(.05,2)=5.99$

Table 9
Actual and Predicted School Attendance Rates by Number of Years Lagging Behind in School: Age 13-15

Age	Boys Actual	Boys Predicted	Boys χ^2	Girls Actual	Girls Predicted	Girls χ^2
Not behind	88.3	82.1	8.50	83.8	78.2	6.02
Behind one year	79.8	76.4	1.56	75.4	74.5	0.09
Behind two years	65.8	62.5	0.91	52.9	51.0	0.20
Behind three years or more	49.1	51.7	0.62	44.7	42.7	0.39

$\chi^2(.05,1)=3.84$

Table 10
Actual and Predicted Annual Wage if working by Child Age and Sex[a]
(number of observations in parentheses)

Age	Boys Actual	Boys Predicted	Girls Actual	Girls Predicted
12	6233 (6)	9298	3720 (2)	7301
13	7064 (21)	7618	5460 (6)	6907
14	7643 (34)	10218	8726 (19)	9306
15	10189 (53)	10313	6386 (22)	9848

a. in real 1997 pesos

Table 11
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[b]	96.6	96.7	97.6	98.7[b]
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[b,c,d]	68.8	72.5	69.5	76.3[c]
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[b,c,d]	64.0	67.4	64.2	71.6[c]
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≥6 Behind in School								
Actual	40.9	44.4	30.3	51.5[c,d]	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
b. cross-section treatment effect (T98-C98) p-value ≤ .10
c. longitudinal treatment effect (T98-T97) p-value ≤ .10
d. difference-in-difference treatment effect ((T98-T97) – (C98-C97)) p-value ≤ .10

Table 12
Actual verses Predicted Subsidy Effects on Percent Attending School

Subsample	Girls, Age 12-15 (1) Actual Attendance Rate	Girls, Age 12-15 (2) Pred. with Subsidy	Girls, Age 12-15 (2)-(1)	Girls, Age 12-15, Behind in school (1) Actual Attendance Rate	Girls, Age 12-15, Behind in school (2) Pred. with Subsidy	Girls, Age 12-15, Behind in school (2)-(1)	Girls, Age 13-15, HGC≥6, Behind in school (1) Actual Attendance Rate	Girls, Age 13-15, HGC≥6, Behind in school (2) Pred. with Subsidy	Girls, Age 13-15, HGC≥6, Behind in school (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: X-Section, Longitudinal, Difference-in-Difference	7.9*, 11.5*, 10.3*			12.7*, 14.1*, 14.5*			7.1, 21.2*, 17.7*		
	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: X-Section, Longitudinal, Difference-in-Difference	3.8, 6.8*, 3.1			4.2, 7.4*, 4.0			0.8, 5.7, 2.7		

* p-value of treatment effects ≤ .10

Table 13
Actual and Predicted Choice Distribution by Child
Age, Sex and School Attainment: Post-Subsidy Treatment

	Girls						Boys					
	In School		Home		Work		In School		Home		Work	
	Act[a]	Pred[b]	Act.	Pred.	Act.	Pred.	Act.	Pred.	Act.	Pred.	Act.	Pred.
Age 12-15[c,d]												
All	74.4	74.9	21.2	22.3	4.1	2.8	76.3	77.1	14.9	15.0	8.8	7.9
Not Behind	82.3	82.1	14.5	14.8	3.2	3.1	88.9	88.2	9.7	9.7	1.4	2.1
Behind	71.9	72.3	23.7	25.0	4.4	2.7	71.6	72.9	16.8	16.9	11.6	10.2
Not Behind and HGC ≥ 6	52.3	58.7	41.5	37.7	6.2	3.6	58.3	66.7	25.0	20.9	16.7	12.4

a. Based on observations in which neither the school nor work choice is missing.
b. Based in all observations including those missing school or work.
c. Numbers of observations for each of the four rows are 222, 62, 160 and 65 for girls, and 262, 72, 190 and 96 for boys.
d. Based on 200 simulation draws per family.

Table 14
Comparison of Actual and Predicted Attendance and Fertility
Based on N-Year Predictions Using Initial Conditions

	Controls, 1997		Controls, 1998		Treatments, 1997	
	Actual	Predicted	Actual	Predicted	Actual	Predicted
Percent Attending School						
Age 6-11						
Girls	96.9	95.3	96.5	95.4	97.6	95.3
Boys	96.6	93.3	96.7	93.5	97.6	93.2
Age 12-15						
Girls	65.3	58.2	66.5	58.5	62.9	56.6
Boys	68.8	62.5	72.5	62.7	69.5	61.2
Age 12-15, Behind in School						
Girls	58.3	52.4	58.7	52.6	56.9	51.1
Boys	64.0	56.4	67.4	56.8	64.2	55.2
Age 12-15, HGC≥6, Behind in School						
Girls	40.9	41.3	44.0	41.0	30.3	39.6
Boys	59.0	51.1	57.1	50.1	52.6	48.9
Percent Pregnant						
Age 20-24	17.9	21.2	17.0	19.6	17.3	20.8
Age 25-29	16.7	20.0	14.6	19.4	16.4	19.8
Age 30-34	13.1	10.8	9.3	10.8	12.8	11.0
Age 35+	5.2	7.8	6.7	8.1	6.3	7.7

Table 15
Short-run and Long-run Effects of the Subsidy on the
Percent of 12-15 Year-olds Attending School

	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

a. predicted value with subsidy minus predicted value without subsidy, conditional on current state space

b. predicted value with subsidy minus predicted value without subsidy, based on initial conditions

Table 16
Predicted Effect of the Subsidy on Completed Schooling of Children by Age 16:
All Children Ever Born[a]

	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	25.8	22.8	28.1

a. completed schooling at grade 10

Table 17
Predicted Effect of Subsidy on Completed Fertility:
All Children Ever Born

	Without subsidy	With subsidy
Mean Number of Children Ever Born	4.24	4.28
Percent of Families with		
Zero Children	0.05	0.04
One Child	1.16	1.13
Two Children	9.23	8.74
Three Children	22.97	22.48
Four Children	24.43	24.60
Five Children	21.54	21.46
Six Children	14.78	15.30
Seven Children	5.05	5.36

Table 18
The Effectiveness and Cost of Alternative Programs

	Baseline[a]	Compulsory School Attendance through Age 15	Original Subsidy	2x Subsidy	.5x Subsidy	Restricted Subsidy[b]	1.43x Restricted Subsidy
Mean Completed Schooling							
Girls	6.29	8.37	6.83	7.30	6.56	6.67	6.97
Boys	6.42	8.29	6.96	7.44	6.68	6.79	7.07
Percent Completed Grade 6 or more							
Girls	75.8	95.1	82.3	86.9	79.3	77.4	82.0
Boys	78.8	93.7	83.3	86.7	81.1	79.6	82.8
Percent Completed Grade 9 or more							
Girls	19.2	55.5	25.8	31.6	23.1	26.2	29.3
Boys	22.8	54.7	28.0	34.6	25.5	29.2	31.8
Cost per Family	0	-	26,096	59,956	12,318	15,691	25,193
Mean Number of Children	4.24	4.21	4.28	4.32	4.27	4.25	4.27

a. Predicted: control and treatment families.
b. Subsidy for attending school in grades 6-9 only.

Table 18 continued
The Effectiveness and Cost of Alternative Programs

	Bonus for Completing 9th Grade[c]	Junior Secondary School in Each Village	Unconditional Income Transfer 5,000 Pesos /Yr	No Child Labor through Age 15	Original Subsidy and 25% Wage Increase
Mean Completed Schooling					
Girls	6.50	6.39	6.41	6.30	6.75
Boys	6.58	6.55	6.53	6.52	6.79
Percent Completed Grade 6 or more					
Girls	74.9	76.0	77.6	76.1	81.5
Boys	76.9	79.0	80.0	79.9	81.8
Percent Completed Grade 9 or more					
Girls	28.8	21.2	20.8	19.7	25.3
Boys	32.7	24.1	23.6	23.5	26.5
Cost per Family	36,996	-	237,000	-	25,262
Mean Number Children	4.20	4.25	4.23	4.25	4.29

c. The bonus is set at 30,000 pesos for girls and boys.

Long Term Impacts of Education Subsidy

- The existence of subsidy from the start of marriage would increase years of completed education at age 16 by 0.54 years for both boys and girls.
- Fertility without policy: 4.24 children
  Fertility with policy: 4.28 children
  Fertility does not change much.

Counterfactual Experiments

- Compulsory school attendance between ages 6 to 15. Because of failure rates, mean completed schooling is 8.37 (girls), 8.29 (boys) where the baseline without policy is (6.29, 6.42) and the original data (6.83, 6.96)
- Restrict subsidy to attendance in the 6th grade or higher. Zero subsidy for 3 to 5 grades.
  Fall in completed schooling : 30 % (girls), 33 % (boys), even though attendance of 3 to 5 years old children are almost universal.
  No subsidy for young childrens' attendance reduce income to parents, thus induce more work and less school attendance for older children.

- Enforcing child labor law: prohibit children under age 16 to work. Little effect on schooling since they would do home production.
- Build a junior secondary school in every village: increases mean schooling years by 0.1 (girls) , 0.13 (boys)