Development Econometrics
Combining structural & experimental methods to evaluate programs

Måns Söderbom, University of Gothenburg

16 September 2011

1 Introduction

Reference: Attanasio, Orazio, Costas Meghir, and Ana Santiago (2011). "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA," Review of Economic Studies, published online August 2, 2011 doi:10.1093/restud/rdr015.

- PROGRESA: one of the first conditional cash transfer programs, implemented in order to improve human capital formation (as well as reducing short term poverty).

- Main component: education - eligible mothers were given grants to keep their children in school.

- At the outset of the program, randomization was used to enable researchers to study the impact of the program.
  - 320 villages got the program
  - 186 villages were randomized out of the program - program implementation postponed by about 2 years.
- Given the randomization, program impacts can be estimated by comparing mean outcomes between treatment & control groups.
- Schultz (2004; lecture 2): At junior high level, the estimated program effect is to increase enrollment rates by 7-9 percentage points for girls and by 5-6 percentage points for boys.

- The premise of the AMS paper is that conventional, reduced-form, evaluations of the impact of programs on outcomes sometimes tell us little about how and why the program works. They also are not suitable for generalization (this point is closely linked to the 'how and why' argument; unless we know how/why it would seem inappropriate to extrapolate to other contexts).

- Aim of the paper:
  - Analyze the impact of monetary incentives on education choices in rural Mexico
  - Discuss effective design of interventions aimed at increasing school enrolment

  - Illustrate the benefits of combining randomized experiments with structural models.

- Approach: Estimate a **structural model** of education choices using the data from the PROGRESA randomized experiment. Use the model to simulate the effect of changes to some of the parameters of the program. Sample: boys aged between 9 and 17.

- Structural estimation: Write down a fully parameterized theoretical model of individual behaviour (e.g. modeling the decision to go to school vs. work); and estimate the parameters of the model using a suitable dataset

- General point: Experimental variation can help identify economic effects under more general conditions than the observational data (because of exogeneity), while the structural model can help provide an interpretation of the experimental results and broaden the usefulness of the experiment (because the behavioral model is richer).

- Todd and Wolpin (2006; AER) use a structural model to identify the effect of the PROGRESA program **without** using data from the experiment; that is, using their methodology, it would be possible (of course, subject to assumptions) to identify the effect of the program *even before its implementation*.

- This is one illustration of how a structural approach may be useful.

- Too good to be true?

- Inferring the effect of the program from analysis based on the sensitivity of schooling decisions to changes in wages (the opportunity cost of schooling) is fine, provided the marginal utility of the grant does not differ from the marginal utility of any other source of income (e.g. wage).

- But it may well be that money received through the grant program affects behavior differently compared to money received from other sources. Why?

- Makes ex ante evaluation doubtful? Maybe.

- Key contributions of AMS:

  - Test for separability of income earned by the child (in school as a scholarship or in work as a wage) from the activity that generated it.

That is, is a peso of child income from work the same as a peso of child income from a school grants programme?

– Builds general equilibrium effects of the grant on village child wages into the model

– Presents results from simulations of how schooling decisions would change if the structure of the PROGRESA program changed.

2 Recap: The PROGRESA Program and the Impact

References: Sections 2-3 AMS; Schultz (2004); notes for lecture 2.

- Started in 1997; 506 low-income rural localities in Mexico were selected; 186 of these were randomized out of the program forming the control group.

- A household is eligible for the grant if it is deemed "poor" (proxy means testing).

- Main component is education (other: nutrition, health).

- Grants in grades 3-6 primary; 1-3 secondary; increasing with grades; higher for girls than for boys in secondary. See Table 1 in AMS.

- Random assignment means it's straightforward to estimate the overall impact of the conditional cash transfer on enrollment (Schultz, 2004). This approach is silent on the mechanics through which the program operates (how? why?).

TABLE 2
Experimental results October 1998

Difference estimates of the impact of PROGRESA on boys school enrolment				
Age group	Enrolment rates in control villages (eligible)	Impact on Poor 97	Impact on Poor 97–98	Impact on. non-eligible
10	0·951	0·0047	0·0026	0·0213
		(0·013)	(0·011)	(0·021)
11	0·926	0·0287	0·0217	−0·0195
		(0·016)	(0·015)	(0·019)
12	0·826	0·0613	0·0572	0·0353
		(0·024)	(0·022)	(0·043)
13	0·780	0·0476	0·0447	0·0588
		(0·030)	(0·027)	(0·060)
14	0·584	0·1416	0·1330	0·0672
		(0·039)	(0·035)	(0·061)
15	0·455	0·0620	0·0484	0·1347
		(0·042)	(0·039)	(0·063)
16	0·292	0·0304	0·0355	0·1063
		(0·038)	(0·036)	(0·067)
12−15	0·629	0·0655	0·0720	0·0668
		(0·027)	(0·024)	(0·022)
10−16	0·708	0·0502	0·0456	0·0810
		(0·018)	(0·015)	(0·026)

Note: standard errors in parentheses are clustered at the locality level.

Overview:
Dynamic Model of School Participation

- Each child decide whether to attend school or work (child labor).
  - Going to school: child incurs a utility (cost) and becomes eligible for next grade.
  - Working: child receives a wage(educ,village,age)
- PROGRESA grant: additional monetary reward to schooling.
- Children can attend school up to and including age 17; but not later in life.

Outline of behavioral model

Age 18:

- Individuals have now made all the decisions; they enter the adult labor market.
- Earnings depend on education. The 'terminal value function', which represents the present discounted value of current & future earnings at age 18, is written

$$V(\mathrm{ed}_{i,18}) = \frac{\alpha_1}{1 + \exp(-\alpha_2 * \mathrm{ed}_{i,18})}$$

where α_1 and α_2 are parameters determining the return to education (how?).

Outline of behavioral model

Age 17:

- Taking into account the effect of education on the terminal value, individuals choose between:
  - One more year of schooling
  - Child labor

Age 17	Chooses schooling		Chooses schooling	
Current grade (ed)	Current utility	Future utility	Current utility	Future utility
0	$U^{s}_{i,17}$(ed=0)	Terminal value(ed=1)	$u^{w}_{i,17}$ (ed=0)	Terminal value(ed=0)
1	$u^{s}_{i,17}$ (ed=1)	Terminal value(ed=2)	$u^{w}_{i,17}$ (ed=1)	Terminal value(ed=1)
2	$u^{s}_{i,17}$ (ed=2)	Terminal value(ed=3)	$u^{w}_{i,17}$ (ed=2)	Terminal value(ed=2)
3	$u^{s}_{i,17}$ (ed=3)	Terminal value(ed=4)	$u^{w}_{i,17}$ (ed=3)	Terminal value(ed=3)
(…)				
8	$u^{s}_{i,17}$ (ed=8)	Terminal value(ed=9)	$u^{w}_{i,17}$ (ed=8)	Terminal value(ed=8)

Net utility of going to school instead of working:

$\{u^{s}_{it}(ed) + \beta \text{ x Terminal value}(ed+1)\} – \{u^{w}_{it}(ed) + \beta \text{ x Terminal value}(ed)\}$

where β=1/(1+r) is the discount factor.

- If net utility positive, optimal choice is ”school”; if negative, ”work”.

- Total utility, ***given the optimal decision,*** is given by the <u>value function</u>.

Age 17

Current grade (ed)

Value function

Current grade (ed)	Value function
0	$V_{i,17}(ed=0)=\max [\ u^{s}_{17,t}(ed=0)+\beta*terminalvalue(ed=1)\ ,\ u^{w}_{i,17}(ed=0)+\beta*terminalvalue(ed=0)\ ]$
1	$V_{i,17}(ed=1)=\max [\ u^{s}_{17,t}(ed=1)+\beta*terminalvalue(ed=2)\ ,\ u^{w}_{i,17}(ed=1)+\beta*terminalvalue(ed=1)\ ]$
2	$V_{i,17}(ed=2)=\max [\ u^{s}_{17,t}(ed=2)+\beta*terminalvalue(ed=3)\ ,\ u^{w}_{i,17}(ed=2)+\beta*terminalvalue(ed=2)\ ]$
3	$V_{i,17}(ed=3)=\max [\ u^{s}_{17,t}(ed=3)+\beta*terminalvalue(ed=4)\ ,\ u^{w}_{i,17}(ed=3)+\beta*terminalvalue(ed=3)\ ]$
(…)	
8	$V_{i,17}(ed=8)=\max [\ u^{s}_{17,t}(ed=8)+\beta*terminalvalue(ed=9)\ ,\ u^{w}_{i,17}(ed=8)+\beta*terminalvalue(ed=8)\ ]$

Now we can solve for the optimal school/work decision for **16-year** olds – i.e. we take one step back in time… (backward induction).

Age 16 Current grade (ed)	Chooses schooling: Current utility	Chooses schooling: Future utility	Chooses child labor (work): Current utility	Chooses child labor (work): Future utility
0	$U^{s}_{i,16}(ed=0)$	$V_{i,17}(ed=1)$	$u^{w}_{i,16}(ed=0)$	$V_{i,17}(ed=0)$
1	$u^{s}_{i,16}(ed=1)$	$V_{i,17}(ed=2)$	$u^{w}_{i,16}(ed=1)$	$V_{i,17}(ed=0)$
2	$u^{s}_{i,16}(ed=2)$	$V_{i,17}(ed=3)$	$u^{w}_{i,16}(ed=2)$	$V_{i,17}(ed=0)$
3	$u^{s}_{i,16}(ed=3)$	$V_{i,17}(ed=4)$	$u^{w}_{i,16}(ed=3)$	$V_{i,17}(ed=0)$
(…)				
8	$u^{s}_{i,16}(ed=8)$	$V_{i,17}(ed=9)$	$u^{w}_{i,16}(ed=8)$	$V_{i,17}(ed=0)$

Net utility of going to school instead of working:

$$\{u^{s}_{i,16}(ed) + \beta\, V_{i,17}(ed+1)\} - \{u^{w}_{i,16}(ed) + \beta\, V_{i,17}(ed)\}$$

- If positive, the optimal decision is ”school”; if negative, ”work”.
- Use optimal decision to figure out value at age 16… and then consider decisions for children aged 15, and then 14,13,…6.

Taking stock

- Above we have an illustration of how individuals' schooling decisions are determined in the model.
- For the model to be useful, the utility functions associated with schooling and work have to be specified.
- As we shall see, these functions will depend on observable variables, and unknown parameters to be estimated.
- Once we have estimated the parameters of the model, we can simulate optimal schooling decisions, and do a range of policy experiments (e.g. alter the structure of the grant).

Principle of estimation

1. Specify start values for the parameters to be estimated (utility function parameters, terminal value parameters, discount rate etc.)
2. Use the dynamic model outlined above to solve for optimal schooling for each individual.
3. Match the model predictions to the actual decisions observed in the data, and compute the log likelihood for the sample.
4. Go back to step 1 and change the parameter values; step 2 to obtain new optimal decisions; step 3 new log likelihood value. Continue until no further improvements in the log likelihood can be made; at this point you have ML estimates of the structural parameters.

Note that the theoretical model is embedded in the estimation algorithm!

Key ingredients of theoretical model

1. Instantaneous utility of schooling
2. Instantaneous utility of child labor (work)
3. Uncertainty
4. Terminal value of schooling

Nonstandard features of the econometrics

1. Unobserved heterogeneity in utility functions.
   - Assume a discrete distribution of the unobservables with M points of support (Heckman & Singer, 1984). (Similar to random effects in probits & logits, except not assumed noramally distributed – more below.)
2. Initial conditions. Because the data is cross-sectional, the initial condition (grade in the beginning of the period) may be endogenous, i.e. correlated with unobservables.
   - Use a flexible ordered probit to model initial education, which depends on inter alia the unobservables modeled in (1) above. Similar to IV.

Model Details

Instantaneous Utility

- Linear utility of going to school:

$$u_{it}^{s} = Y_{it}^{s} + \alpha g_{it},$$

$$Y_{it}^{s} = \mu_{i}^{s} + a^{s\prime} z_{it} + b^{s} \mathrm{ed}_{it} + 1(p_{it} = 1)\beta^{p} x_{it}^{p} + 1(s_{it} = 1)\beta^{s} x_{it}^{s} + \varepsilon_{it}^{s},$$

- Possible entitlement to PROGRESA grant (g). Effect on utility: α.
- Other determinants: unobserved heterogeneity; taste shifters; acquired education; costs specific to primary (p) & secondary (s) school; error term.

Instantaneous Utility

- Linear utility of working (not attending school):

$$u_{it}^{\mathrm{W}} = Y_{it}^{\mathrm{W}} + \delta w_{it},$$

$$Y_{it}^{\mathrm{W}} = \mu_i^{\mathrm{W}} + a^{w\prime} z_{it} + b^{\mathrm{W}} \mathrm{ed}_{it} + \varepsilon_{it}^{\mathrm{W}},$$

- Effect of wage on utility: δ.
- Other determinants: unobserved heterogeneity; taste shifters; acquired education; error term.
- Note: δ=α would imply that income earned by the child is separable from the activity that generated it. This potentially strong assumption is **not** imposed in the analysis (i.e. can test for separability – why interesting?).

Instantaneous Utility

Subtract Y^{w}_{it} from the RHS of both utility functions:

$$u^{s}_{it} = \gamma\,\delta g_{it} + \mu_i + a' z_{it} + b\mathrm{ed}_{it} + 1(p_{it}=1)\beta^{p} x^{p}_{it} + 1(s_{it}=1)\beta^{s} x^{s}_{it} + \varepsilon_{it},$$

$$u^{w}_{it} = \delta w_{it}$$

where $a = a^{s} - a^{w}$, $b = b^{s} - b^{w}$, $\gamma = \alpha/\delta$, $\mu_i = \mu^{s}_{i} - \mu^{w}_{i}$, and $\varepsilon_{it} = \varepsilon^{s}_{it} - \varepsilon^{w}_{it}$.

- The residual ε follows a logistic distribution
- The coefficient γ measures the impact of the grant as a proportion of the impact of the wage on the education decision. If γ=1, then lowering the wage by some percentage would have the same effect on schooling as increasing the grant by the same percentage. What if γ>1? Why might this be? Discuss.

Wages

- Wages are a key determinant of schooling in the model, since they affect the opportunity cost of schooling.
- That is, in the present model, an increase in wages will reduce school participation.
- Wages are determined in the local labor market.
- *General equilibrium effects*: Wages may be **affected** by the PROGRESA program: e.g. grant => more schooling => less child labor supply => wages increase => less schooling....
- GE effects could thus **mitigate** the effect of the program on schooling.

Wages

- Estimated wage regression for boy *i* living in village *j*:

$$\ln w_{ij} = \underset{(0.384)}{-0.983} + \underset{(0.028)}{0.0605}\, P_j + \underset{(0.049)}{0.883}\, \ln w_j^{\text{ag}} + \underset{(0.027)}{0.066}\text{age}_i$$

$$+ \underset{(0.0065)}{0.0116}\text{educ} - \underset{(0.053)}{0.056}\,\text{Mills}_i + \varpi_{ijt},$$

  - Keep in mind: This is for child labor earnings only.
  - Positive & significant effect of PROGRESA grant (P)
  - Positive & significant effect of male agricultural wage
  - Positive & significant effect of age
  - Small effect of education
  - No evidence of sample selection bias (Mills = inverse Mill's ratio; Heckit)

Using the wage regression estimates

- Recall:

$$u_{it}^{\mathrm{s}} = \gamma\,\delta g_{it} + \mu_i + a' z_{it} + b\,\mathrm{ed}_{it} + 1(p_{it} = 1)\beta^{\mathrm{p}} x_{it}^{\mathrm{p}} + 1(s_{it} = 1)\beta^{\mathrm{s}} x_{it}^{\mathrm{s}} + \varepsilon_{it},$$

$$u_{it}^{\mathrm{w}} = \delta w_{it}$$

- Thus the wage is a key determinant of schooling. Based on the wage regression above, the authors compute predicted wages as a function of age, education, PROGRESA and the village agricultural wage.
- These predictions get plugged into the model based on which we solve for optimal schooling decisions.
- *Hard*: What's the exclusion restriction? Why do we need this?

Uncertainty

Two sources of uncertainty in the theoretical model:

- i.i.d. logistic shock to schooling costs (ε; see previous slide). The individual knows ε in the current period but not its value in the future.
- Risk of failing to complete a grade. Calculated as the ratio of individuals who are in the same grade as the year before (details in appendix).

Return to education & terminal value of schooling

- An important incentive for going to school is that it gets you higher future earnings. This needs to be modelled somehow.
- One option might be to estimate the returns to education using wage data for the adult population (c.f. Schultz, 2004); however the authors prefer another approach.
- They write the terminal value function as

$$V(\mathrm{ed}_{i,18}) = \frac{\alpha_1}{1 + \exp(-\alpha_2 * \mathrm{ed}_{i,18})}$$

which implies the only thing that matters for lifetime (expected) earnings is education. The parameters α_1 and α_2 (>0) are estimated 'structurally'. Note that returns to education can then be computed, even though we don't have wage data! How is this possible?

The Value Functions

$$V_{it}^{\mathrm{s}}(\mathrm{ed}_{it}|\Upsilon_{it}) = u_{it}^{\mathrm{s}} + \beta\{p_t^{\mathrm{s}}(\mathrm{ed}_{it}+1)\,E\max[V_{it+1}^{\mathrm{s}}(\mathrm{ed}_{it}+1),\, V_{it+1}^{\mathrm{w}}(\mathrm{ed}_{it}+1)]$$

$$+(1-p_t^{\mathrm{s}}(\mathrm{ed}_{it}+1))\,E\max[V_{it+1}^{\mathrm{s}}(\mathrm{ed}_{it}),\, V_{it+1}^{\mathrm{w}}(\mathrm{ed}_{it})]\},$$

$$V_{it}^{\mathrm{w}}(\mathrm{ed}_{it}|\Upsilon_{it}) = u_{it}^{\mathrm{w}} + \beta\,E\max\{V_{it+1}^{\mathrm{s}}(\mathrm{ed}_{it}),\, V_{it+1}^{\mathrm{w}}(\mathrm{ed}_{it})\}.$$

- Interpret!

Habits & initial conditions

- Recall that the utility of attending school depends on years of education:

$$u_{it}^{\mathrm{s}} = \gamma\,\delta g_{it} + \mu_i + a' z_{it} + b\mathrm{ed}_{it} + 1(p_{it} = 1)\beta^{\mathrm{p}} x_{it}^{\mathrm{p}} + 1(s_{it} = 1)\beta^{\mathrm{s}} x_{it}^{\mathrm{s}} + \varepsilon_{it},$$

Such **state dependence** (= current outcome depends on previous decisions) is potentially important – e.g. may be a mechanism that reinforces the effect of the grant.

However state dependence creates an econometric problem: ed_{it} cannot be assumed uncorrelated with the unobserved heterogeneity term μ_i (why not?).

Habits & initial conditions

- To tackle this problem, education already attained is modeled as a function of observables h_i and the unobserved term μ_i .
- This is done by means of **ordered probit**, where the index function is specified as $h'_i\zeta + \xi\mu_i$ where ξ is a coefficient ('factor loading') to be estimated, and h_i contains variables reflecting **past** schooling costs (cf. IV approach).
- Probability of attending & eduction=ed:

$$P(\mathrm{ed}_{it} = e, \mathrm{Attend}_{it} = 1 | z_{it}, x^{\mathrm{p}}_{it}, x^{\mathrm{s}}_{it}, h_i, \mathrm{wage}_{it}, \mu_i)$$

$$= P(\mathrm{Attend}_{it} = 1 | z_{it}, x^{\mathrm{p}}_{it}, x^{\mathrm{s}}_{it}, \mathrm{wage}_{it}, \mathrm{ed}_{it}, \mu_i) \times P(\mathrm{ed}_{it} = e | z_{it}, x^{\mathrm{p}}_{it}, x^{\mathrm{s}}_{it}, h_i, \mathrm{wage}, \mu_i)$$

[NOTE: THE FIRST PROBABILITY SHOULD DEPEND ON THE GRANT (g_{it}) ALSO.]

The likelihood function

Contribution to the sample log likelihood of indvidual *i* aged *t*:

$$\begin{aligned}\log L_{it} &= \log[\sum_{m=1}^{M} p_m \Lambda^{attend}\left(g_{it}, z_{it}, x_{it}^{s}, x_{it}^{s}, wage_{it}, ed_{it}, \mu_i = s_m\right) \\ &\times oprobit^{ed_{it}}\left(z_{it}, x_{it}^{s}, x_{it}^{s}, h_i, wage_{it}, \xi\mu_i = \xi s_m\right)]\end{aligned}$$

where Λ is the logistic distribution (remember: wage is predicted).

Note: joint estimation of oprobit and logit. Interpretation of ξ?

Note that p_m and s_m are parameters to be estimated. That is, the entire distribution of the unobserved heterogeneity term is estimated. The number of support points (M) is pre-specified (usually a small number; in the present paper M=3).

Results

Three sets of results

- Estimates of the heterogeneity distribution (Table 3)
- Estimates of the parameters in the ordered probit, modeling initial education (Table 4)
- Estimates of the utility parameters and the discount rate (Table 5)
- Three different specifications:
  - (A) Model doesn't do (B)…
  - (B) Model allows for differences in pre-program enrollment btw treatment & control villages
  - (C) Control sample only – i.e. experiment not used.

TABLE 3
The distribution of unobserved heterogeneity

	A	B	C
Point of Support 1	−9·706	−8·327	−4·290
	1·041	*1·101*	*2·46*
Point of Support 2	−14·466	−13·287	−17·62
	1·173	*1·208*	*3·144*
Point of Support 3	−5·933	−4·301	−0·267
	0·850	*0·941*	*2·45*
Probability of 1	0·513	0·518	0·490
	0·024	*0·023*	*0·032*
Probability of 2	0·342	0·335	0·270
	0·022	*0·021*	*0·017*
Probability of 3	0·145	0·147	0·240
Load factor for initial condition	0·108	0·102	0·068
	0·016	*0·014*	*0·013*

Notes: Column A: eligible dummy only; B: eligible dummy and non-eligible in treatment village dummy. C: model estimated on control sample only. Asymptotic standard errors in italics.

- Three 'types' of children; prob 0.51, 0.34, 0.15
- Point of support indicates the value of the heterogeneity term for the particular type
- Load factor is the coefficient on mu in the initial conditions equation.
- Heterogeneity term is expressed as a determinant of the cost of schooling – hence large negative means high likelihood of schooling.

TABLE 4

Equation for initial conditions

	A
Poor	−0·275
	0·030
Ineligible individual in a PROGRESA village	—
	—
Father's education	
Primary	0·180
	0·025
Secondary	0·262
	0·030
Preparatoria	0·559
	0·0160
Mother's education	
Primary	0·159
	0·026
Secondary	0·316
	0·030
Preparatoria	0·301
	0·061
Indigenous	−0·005
	0·036
Availability of Primary 1997	0·373
	0·073
Availability of Secondary 1997	0·808
	0·188
Kilometer to closest secondary school 97	0·00004
	0·00024
Availability of Primary 1998	−0·261
	0·127
Availability of Secondary 1998	−0·845
	0·187
Kilometer to closest secondary school 98	−0·0001
	0·00003
Cost of attending secondary	0·00006
	0·00024

- Ordered probit
- Contains age-specific cut-off points (not shown)
- Contains discrete random term mu with factor loading (see Table 3)
- Variables proxying for cost of initial schooling (availability 1997) significant; cf. IV approach.
- Parental education has expected effects
- Poverty => less schooling

Education choice model

- Have a look at the results in Table 5.
- All variables except the grant & the wage are expressed as determinants of the cost of schooling. Hence: negative sign => increase means lower cost & more schooling.
- Wage is expressed as a determinant of utility of work (so a positive coef means higher wage => more work, less schooling)
- Grant is expressed as a determinant of utility of schooling (so a positive coef means more schooling, less work)
- Coefficient on grant: ratio to the coefficient on wage, hence a unity coefficient means wage and grant have same effects on schooling & work.

Selected results Table 5

	A	B	C
Wage	0·134	0·168	0·357
	0·043	*0·045*	*0·100*
PROGRESA grant	3·334	2·794	—
	1·124	*0·796*	—
Parameter in terminal function $\ln(\alpha_1)$	5·876	5·886	6·59
	0·115	*0·113*	*0·175*
Parameter in terminal function $\ln(\alpha_2)$	−1·276	−1·286	−1·62
	0·025	*0·024*	*0·089*
Poor	0·676	0·105	0·431
	0·154	*0·215*	*0·274*
Ineligible individual in a PROGRESA village		−1·079	
		0·261	

- PROGRESA coef approx. 3, implying larger positive effect of grant than wage cut on schooling. H_0: =1 rejected.
- Partial effects:
  - Reducing wage by 44% (~grant) raises likelihood of attending by 2.1%
  - This effect is higher if we don't use the experiment (C)
  - The grant raises likelihood of attending by 4.7%.
- Poverty raises the cost of schooling, and thus lowers attendance.
- Estimates of terminal function parameters imply an average return on education of about 5%.

Additional results Table 5

	A	B	C
Father's Education - Default is less than primary			
Primary	−0·462	−0·509	−0·486
	0·120	*0·123*	*0·217*
Secondary	−0·746	−0·803	−0·959
	0·147	*0·150*	*0·261*
Preparatoria	−1·794	−1·819	−2·176
	0·323	*0·328*	*0·558*
Mother's Education - Default is less than primary			
Primary	−0·488	−0·488	−0·870
	0·123	*0·126*	*0·233*
Secondary	−0·624	−0·613	−1·119
	0·143	*0·145*	*0·254*
Preparatoria	−1·576	−1·681	−2·158
	0·351	*0·355*	*0·645*
Indigenous	−0·783	−0·777	−1·018
	0·132	*0·135*	*0·241*
Availability of Primary 1998	3·600	3·765	3·092
	0·285	*0·295*	*0·499*
Availability of Secondary 1998	−0·030	−0·074	0·789
	0·193	*0·197*	*0·425*
Kilometer to closest secondary school 98	0·0003	0·0003	0·00078
	0·00005	*0·00005*	*0·00014*
Cost of attending secondary	0·007	0·007	0·013
	0·001	*0·001*	*0·0033*
Age	2·291	2·249	2·903
	0·160	*0·157*	*0·354*
Prior years of education	−2·785	−2·896	−3·621
	0·256	*0·261*	*0·621*
Discount rate	0·95	0·96	0·975

State dependence!
(why important?)

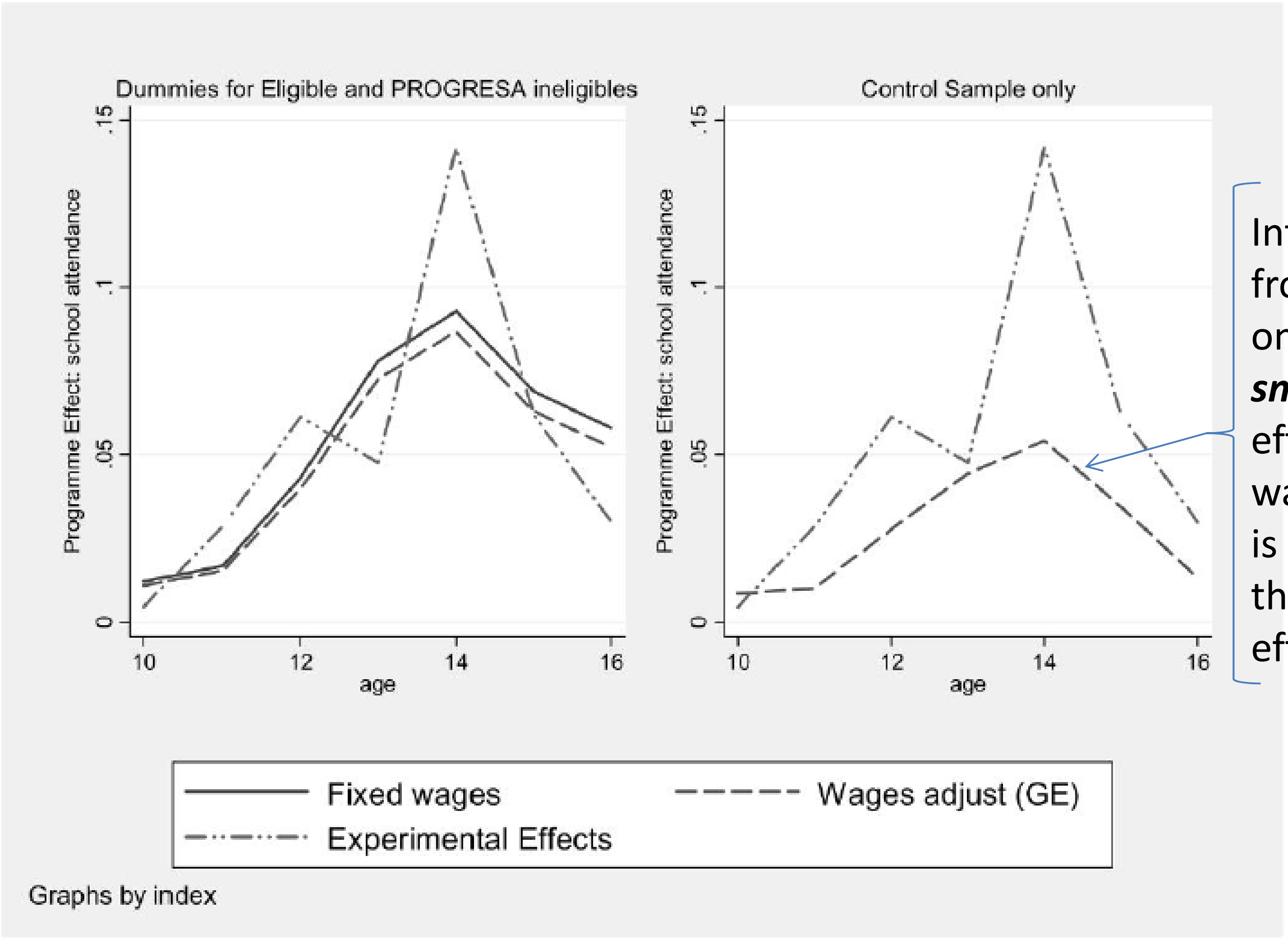
Simulated treatment effects

- Baseline scenario : Actual situation (with grants)
- Counterfactual scenario (i): Set grants to zero, no effect of grant on wages (partial eq:m)
- Counterfactual scenario (ii): Set grants to zero, allow for effect of grant on wages (general eq:m)

➢ Simulate enrollment decisions, and average across children of differing characteristics

➢ Results in a 'simulated average treatment effect' of the program.

Simulated impact of program



Inferred from coef. on wage; ***smaller*** effect since wage effect is smaller than grant effect

Policy simulations

Now use the model to simulate school participation under different scenarios.

1. Compare effect of current program with that of a similar program (same total cost) that differs with respect to how the grant varies with grades attended, targeting those most responsive.
2. Decrease the wage by an amount equivalent to the grant
3. Reduce the distnace to school to a maximum of 3 kms.

Policy reform: set the grant to zero for grades below 6, and increase the grant for grades above 6, retaining a balanced budget.

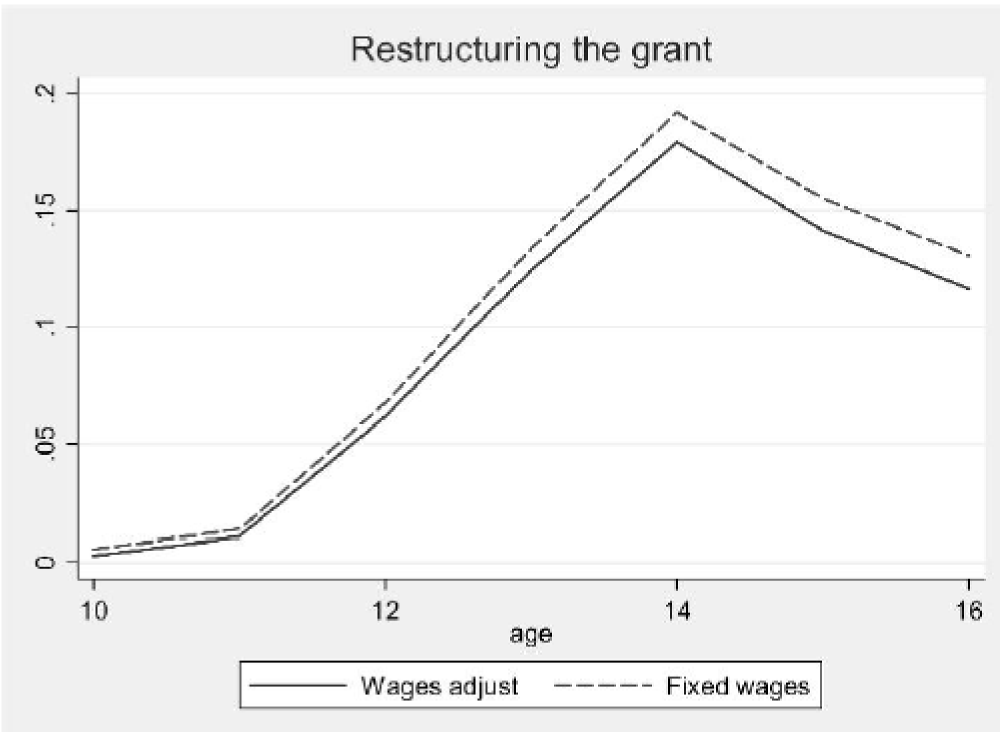
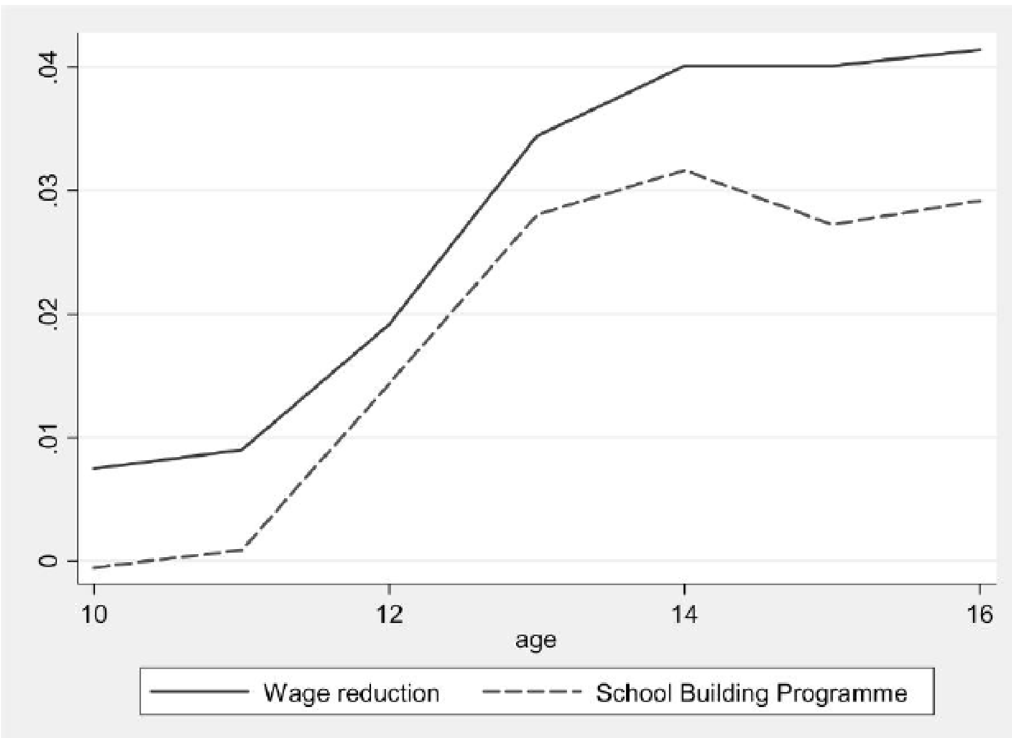


FIGURE 2

Redistributing the grant to those above grade 6 only—revenue neutral

- The impact of this program is nearly twice as high as the actual program!
- This is because removing the grant for lower grades has very small effects – almost all children attend anyway! So grants at low levels are basically unconditional!
- Suggests structure of grants should change – but…
- Credit constraints?
- Effects operating through nutrition?

Additional experiments: Cut the wage; build schools



- Quite modest effects!

Conclusions

- Simple treatment/control comparisons leave many questions unanswered.
- Combining a structural & experimental approach, we can learn new things; e.g.
  - Investigate separability (grant vs. Wage)
  - Investigate effects of changes to the program structure
  - Consider effects of alternative policies (e.g. building schools)
  - General equilibrium effects

Conclusions

- How much faith should one have in predictions based on the structural model?
- The present model is silent on program effects on other aspects of child development
- Doesn’t take into account liquidity constraints