Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA

Orazio P. Attanasio, Costas Meghir & Ana Santiago

December 9, 2015

Central Questions

- Analyze the impact of monetary incentives on education choices in rural Mexico.
- Discuss the design of effective policies.

Methodological Point

- Structural estimation methods can be combined with experimental data to give better answers.
- Experimental data provides more variation than observational data
- Structural methods help
  - interpret experimental results (mechanisms) $\rightarrow$ external validity
  - General Equilibrium effects
  - Evaluation of counterfactuals/ policy experiments
- Todd and Wolpin (2006)

Outline

Overview of PROGRESA

- Started in 1997.
- Key provision: 'Eligible' mothers were given grants to keep their children in school.
- A household is eligible for the grant if it is deemed 'poor'.
- 506 low-income rural localities in Mexico were selected.
- 186 of these were randomized out of the program forming the control group to receive program later.
- Grants increasing with grades; higher for girls than for boys in secondary.

Grant Structure

TABLE 1
The PROGRESA grants

PROGRESA bimonthly monetary benefits				
Type of benefit	1998 1st semester	1998 2nd semester	1999 1st semester	1999 2nd semester
Nutrition support	190	200	230	250
Primary school				
3	130	140	150	160
4	150	160	180	190
5	190	200	230	250
6	260	270	300	330
Secondary school				
First year				
Boys	380	400	440	480
Girls	400	410	470	500
Second year				
Boys	400	400	470	500
Girls	440	470	520	560
Third year				
Boys	420	440	490	530
Girls	480	510	570	610
Maximum support	1170	1250	1390	1500

Data

- Baseline survey for all households in 1997.
- Program implemented first for school year 1998/99.
- Surveys conducted in October 1998, March 1999 and November 1999 for all households.
- This paper mainly uses data from October 1998 - and also from baseline survey.
- Only look at boys.

Average impact: Treatment - Control

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.

Average impact: Diff-in-diff

TABLE A1

Difference in difference estimates—August 1997 *to October* 1998

The impact of PROGRESA on boys school enrolment

Age group	Impact on Poor 97	Impact on Poor 97–98	Impact on non-eligible
10	0·0291	–0·0007	0·0723
	(0·027)	(0·026)	(0·060)
11	0·0240	0·0176	–0·0111
	(0·016)	(0·014)	(0·019)
12	0·0478	0·0420	0·0194
	(0·021)	(0·020)	(0·043)
13	0·0391	0·0396	0·0411
	(0·028)	(0·025)	(0·049)
14	0·0838	0·0731	–0·0460
	(0·032)	(0·027)	(0·051)
15	0·0963	0·0816	0·0617
	(0·035)	(0·033)	(0·062)
16	0·0350	0·0472	0·0517
	(0·036)	(0·030)	(0·059)
12–15	0·056	0·043	0·015
	(0·019)	(0·012)	(0·023)
10–16	0·048	0·053	0·006
	(0·013)	(0·012)	(0·032)

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

The ideal experiment

- Conditional grant can have differential benefit from other subsidies.
- The available variation only identifies effect of existence of grant.
- No variation in grant, conditional on class - BUT variation in ages within class.
- Model (assume something) about behavior (conditional on age) in order to use this variation.

Overview

- Dynamic model
  - Opportunity cost of work
  - Habit formation/ State dependence
- Estimated using cross sectional data

Overview

- Child chooses to go to school *or* work $\rightarrow$ (random) instantaneous utility .
- If work, get a wage.
- Habits $\rightarrow$ initial condition will matter.
- Possibility of going to school until age 17 $\rightarrow$ terminal value.
- Time invariant heterogeneity between children.
- Uncertainty about graduating a grade.
- Anticipation effects - no evidence.

Instantaneous Utility

Utility from schooling:

$$
\begin{aligned}
u_{it}^s &= Y_{it}^s + \alpha g_{it} \\
Y_{it}^s &= \mu_i^s + a^{s\prime} z_{it} + b^s \mathrm{ed}_{it} + \mathbb{I}(p_{it} = 1)\beta^p x_{it}^p + \mathbb{I}(s_{it} = 1)\beta^s x_{it}^s + \epsilon_{it}^s
\end{aligned}
$$

- z_{it} : household taste shifter - subsumes income
- mu_i^s : heterogeneity
- ed_{it} : 'habit' variable
- ϵ_{it}^s : extreme value, iid across i and t

Instantaneous Utility

Utility from work:

$$
\begin{aligned}
u_{it}^{w} &= Y_{it}^{w} + \delta w_{it} \\
Y_{it}^{w} &= \mu_{i}^{w} + a^{w\prime} z_{it} + b^{w}\text{ed}_{it} + \epsilon_{it}^{w}
\end{aligned}
$$

- δ : differential impact of g_{it}
- w_{it}: (potential) earnings

Instantaneous Utility

Summarizing,

$$
\begin{aligned}
u^s_{it} &= \alpha g_{it} + \mu^s_i + a^{s\prime} z_{it} + b^s \mathrm{ed}_{it} + \mathbb{I}(p_{it} = 1)\beta^p x^p_{it} + \mathbb{I}(s_{it} = 1)\beta^s x^s_{it} + \epsilon^s_{it} \\
u^w_{it} &= \delta w_{it} + \mu^w_i + a^{w\prime} z_{it} + b^w \mathrm{ed}_{it} + \epsilon^w_{it}
\end{aligned}
$$

Instantaneous Utility

Summarizing,

$$
\begin{aligned}
u_{it}^{s} &= \alpha g_{it} + \mu_i^{s} + a^{s\prime} z_{it} + b^{s}\mathrm{ed}_{it} + \mathbb{I}(p_{it} = 1)\beta^{p} x_{it}^{p} + \mathbb{I}(s_{it} = 1)\beta^{s} x_{it}^{s} + \epsilon_{it}^{s} \\
u_{it}^{w} &= \delta w_{it} + \mu_i^{w} + a^{w\prime} z_{it} + b^{w}\mathrm{ed}_{it} + \epsilon_{it}^{w}
\end{aligned}
$$

- Observe schooling choice $\implies$ only the difference between parameters is identified.

Instantaneous Utility

$$u_{it}^{s} = \alpha g_{it} + \mu_i^{s} + a^{s\prime} z_{it} + b^{s}\mathrm{ed}_{it} + \mathbb{I}(p_{it} = 1)\beta^{p} x_{it}^{p} + \mathbb{I}(s_{it} = 1)\beta^{s} x_{it}^{s} + \epsilon_{it}^{s}$$
$$u_{it}^{w} = \delta w_{it} + \mu_i^{w} + a^{w\prime} z_{it} + b^{w}\mathrm{ed}_{it} + \epsilon_{it}^{w}$$

- Observe schooling choice $\implies$ only the difference between parameters is identified.

Therefore, WLOG

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

Instantaneous utility

$$u_{it}^{s} = \gamma\delta g_{it} + \mu_i + a' z_{it} + bed_{it} + \mathbb{I}(p_{it} = 1)\beta^{p} x_{it}^{p} + \mathbb{I}(s_{it} = 1)\beta^{s} x_{it}^{s} + \epsilon_{it}$$
$$u_{it}^{w} = \delta w_{it}$$

- parameters are differences
- ϵ_{it} : logistic distribution (difference between two extreme value)
- γ : impact of grant proportional to impact of wage
- μ_i : discrete random variable, estimated

Instantaneous utility

$$u^s_{it} = \gamma\delta g_{it} + \mu_i + a' z_{it} + bed_{it} + \mathbb{I}(p_{it} = 1)\beta^p x^p_{it} + \mathbb{I}(s_{it} = 1)\beta^s x^s_{it} + \epsilon_{it}$$
$$u^w_{it} = \delta w_{it}$$

- parameters are differences
- ϵ_{it} : logistic distribution (difference between two extreme value)
- γ : impact of grant proportional to impact of wage
- μ_i : discrete random variable, estimated
- Add Dummy for eligibility

Instantaneous utility

$$u_{it}^{s} = \gamma\delta g_{it} + \mu_i + a' z_{it} + bed_{it} + \mathbb{I}(p_{it} = 1)\beta^{p} x_{it}^{p} + \mathbb{I}(s_{it} = 1)\beta^{s} x_{it}^{s} + \epsilon_{it}$$

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

- parameters are differences
- ϵ_{it} : logistic distribution (difference between two extreme value)
- γ : impact of grant proportional to impact of wage
- μ_i : discrete random variable, estimated
- Add Dummy for eligibility
- Add Dummy for treatment+ineligible

Aside: Income pooling

- Consider static model

$$U^s = \beta^s Y + \theta^s g$$
$$U^w = \beta^w Y + \theta^w w - \alpha$$

- $\beta^s = \theta^s$ and $\beta^w = \theta^w$: Income pooling
- Alternative estimation strategies:
  - Assume $\beta^s = \theta^s$ and $\beta^w = \theta^w$ - no variation in g
  - Assume $\theta^s = \theta^w$ - no variation in g
  - Use variation in g

Terminal Value function

- At age 18, could measure returns to education - but limited data.
- So,
$$V(\mathrm{ed}_{i,18}) = \frac{\alpha_1}{1+\exp(-\alpha_2 \mathrm{ed}_{i,18})}$$
where $\alpha_1, \alpha_2 > 0$.
- Assumption: Only education determines returns, everything else affects education.

Value functions

Value from schooling (V^s) and work (V^w)

$$V^s_{it}(\mathrm{ed}_{it}|\Gamma_{it}) = u^s_{it} + \beta\left\{p^s_t(\mathrm{ed}_{it})E\max\left[V^s_{it+1}(\mathrm{ed}_{it}+1), V^w_{it+1}(\mathrm{ed}_{it}+1)\right]\right.$$
$$\left. + (1-p^s_t(\mathrm{ed}_{it}))E\max\left[V^s_{it+1}(\mathrm{ed}_{it}), V^w_{it+1}(\mathrm{ed}_{it})\right]\right\}$$

$$V^w_{it}(\mathrm{ed}_{it}|\Gamma_{it}) = u^w_{it} + \beta E\max\left[V^s_{it+1}(\mathrm{ed}_{it}), V^w_{it+1}(\mathrm{ed}_{it})\right]$$

- $p^s_t(\mathrm{ed}_{it})$ - calculated from data
- Given this, E over ϵ_{it}

Wages

- Need to estimate a child wage equation
  - wage unobserved for school-goers
  - predict future wage for dynamic programming problem
  - GE effects (local)
- Estimate an equation of the following form

$$\ln w_{ij} = q_j + a_1 \mathrm{age}_i + a_2 \mathrm{educ}_i + \omega_{ij}$$

- Will use point estimates for prediction

Wages I - endogeneity

Education as a regressor

- Selection issue: more able children, higher wage
  - Tobit selection model
- Education choice is a function of wage
  - almost flat relationship between education and wages within village
  - Child wage + returns to education mostly through migration

Wages II - Modeling q_j

- Consider the village economy

$$[\text{labor supply}]:\ H_k = L_k w_k^{\gamma k}$$
$$[\text{production}]:\ Q = [\delta H_{child}^{\sigma} + (1-\delta) H_{adult}^{\sigma}]^{\frac{1}{\sigma}}$$
$$\sigma = \frac{\rho - 1}{\rho}, \rho > 0$$

Wages II - Modeling q_j

- In eqilibrium,

$$\ln w_{child} = \frac{\rho + \gamma_{adult}}{\rho + \gamma_{child}} \ln w_{adult} - \left[\frac{1}{\rho + \gamma_{child}} \ln \left(\frac{L_{child}}{L_{adult}} \right) + \kappa \right]$$

- PROGRESA $\implies \left(\frac{L_{child}}{L_{adult}} \right)$ goes down.
- Identification: assume all parameters are constant across localities, conditional on treatment.
- w_{adult} is a sufficient statistic for w_{child}.
- Also, w_{adult} is excluded from the selection equation.

Estimated Wage Equation

$$\ln w_{ij} = \underset{(0.384)}{-0.983} + \underset{(0.028)}{0.0605} P_j + \underset{(0.049)}{0.883} \ln w_j^{ag} + \underset{(0.027)}{0.066} \mathrm{age}_i + \underset{(0.0065)}{0.0116} \mathrm{educ}_i - \underset{(0.053)}{0.056} \mathrm{Mills}_i + \bar{\omega}_{ijt}$$

- *Mills* insignificant
- Small coefficient on *educ*
- At individual level: *age*
- At village level: adult wage ($\ln w_j^{ag}$) and P_j

Initial Conditions problem

- The issue: ed_{it} term in utility function, but only observe cross section.
- Specify a reduced form

$$\text{ed}_{it} = h_i'\lambda + \xi\mu_i$$

- μ_i captures the initial conditions (endogeneity) problem.
- h_i - instrument - distance to school changes between the two time periods.
- Ordered probit with age dependent thresholds.

Estimation

- Estimate 3 specifications:
  - Without dummy for treatment + ineligible
  - With dummy
  - Only control group - without variation in grants
- Wage equation estimated separately.
- MLE for the main model

Estimation

To obtain the log-likelihood function:

- Guess parameters and use data to
- Solve the dynamic programming problem for full state space
  - Initial conditions equation enters value function
- Calculate log-likelihood
  - Value functions enter as latent variables
  - 3 point support for μ_i distribution
  - Logistic distribution allows for parametrization

Maximize likelihood iteratively.

Results - μ_i

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.

TABLE 4
Equation for initial conditions

	A	B	C
Poor	−0·275	−0·243	−0·280
	0·030	*0·046*	*0·051*
Ineligible individual in a PROGRESA village	—	0·057	—
	—	*0·055*	—
Father's education			
Primary	0·180	0·181	0·218
	0·025	*0·025*	*0·04262*
Secondary	0·262	0·264	0·281
	0·030	*0·030*	*0·05302*
Preparatoria	0·559	0·558	0·499
	0·0160	*0·057*	*0·09107*
Mother's education			
Primary	0·159	0·158	0·231
	0·026	*0·026*	*0·04446*
Secondary	0·316	0·314	0·398
	0·030	*0·030*	*0·05139*
Preparatoria	0·301	0·301	0·334
	0·061	*0·061*	*0·09740*
Indigenous	−0·005	0·006	0·133
	0·036	*0·026*	*0·0461*
Availability of Primary 1997	0·373	0·372	0·691
	0·073	*0·073*	*0·19003*
Availability of Secondary 1997	0·808	0·804	−0·568
	0·188	*0·188*	*0·349*
Kilometer to closest secondary school 97	0·00004	0·00004	−0·0002
	0·00024	*0·00003*	*0·00007*
Availability of Primary 1998	−0·261	−0·264	−0·449
	0·127	*0·126*	*0·235*
Availability of Secondary 1998	−0·845	−0·841	0·516
	0·187	*0·187*	*0·348*
Kilometer to closest secondary school 98	−0·0001	−0·0001	0·00015
	0·00003	*0·00003*	*0·00007*
Cost of attending secondary	0·00006	0·0001	−0·00019
	0·00024	*0·00024*	*0·00037*

Notes: As in Table 3. State dummies included. Availability means school in the village.

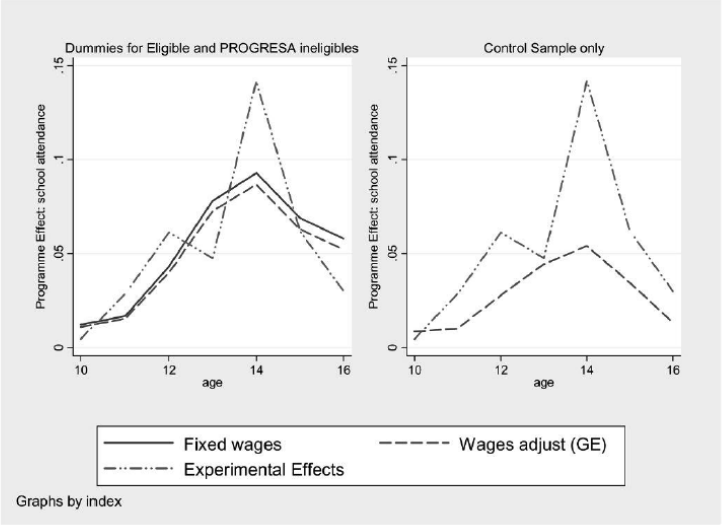
TABLE 5
Parameter estimates for the education choice model

	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	
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
Log-Likelihood	−23,403·98	−23,395·31	−8862·34
State dummies included *Notes* as in Table 3.			
LR test: equal effect of grant and wage χ^2_1	18·66	16·14	
p-value	*0·0*	*0·0*	

Simulations - I

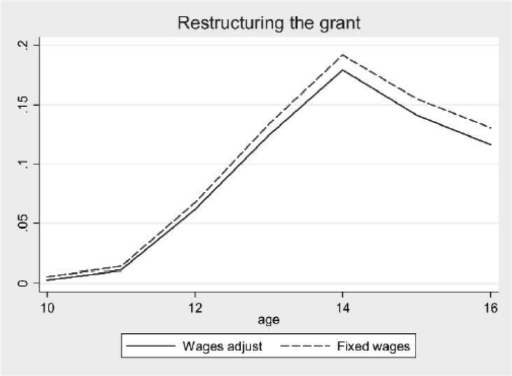
- Two scenarios -
  - grant (actual)
  - no grant (counterfactual)
    - GE wage
    - unchanged wage
- For full sample and only for control

Simulations - I



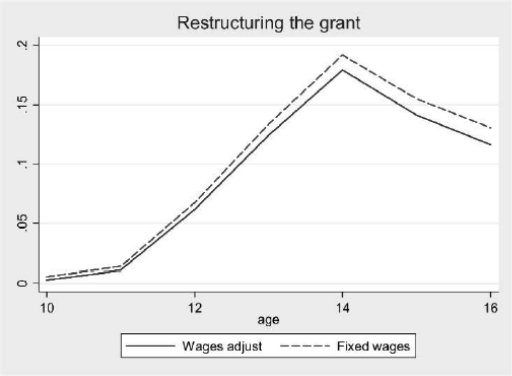
- Age pattern, GE effect
- Only control sample

Simulations - II



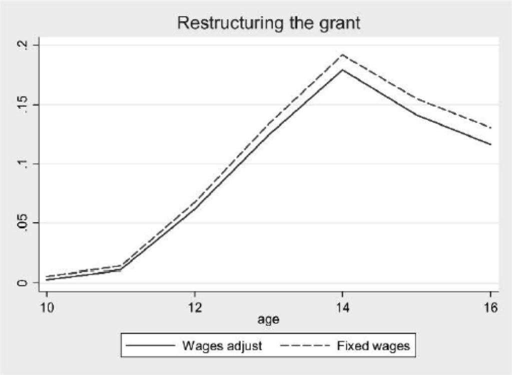
Experiment: Revenue neutral restructuring of grant profile

Simulations - II



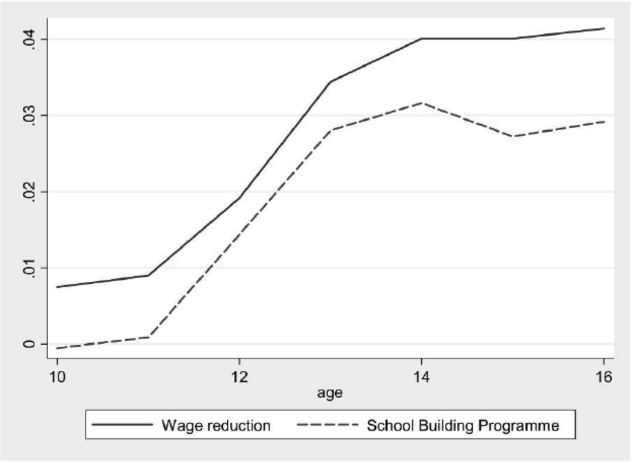
Impact: Almost double of original scheme

Simulations - II



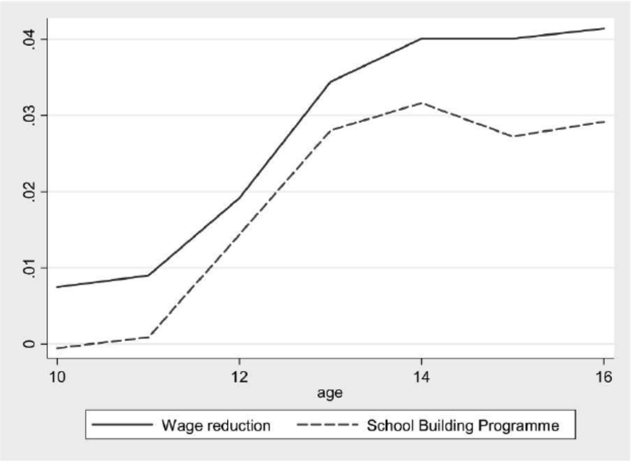
But: Not appropriate if credit constrained or early age health investments

Simulations - III



Experiment: Reduce wage by average grant size

Simulations - III



Experiment: Limit max school distance to 3 km

Conclusion

- Study the effect of PROGRESA on school choice.
- Structural modeling helps
  - Grants vs wages
  - General equilibrium
  - Anticipation effect
  - Counterfactuals
- Usefulness of using full variation in data.