14.771 Development Economics: Microeconomic issues and Policy Models

Fall 2008

Private and Social Returns to Education

Esther Duflo

14.771

Education and Development

- Tremendous correlation between education and level of income across countries.
- The R2 of the regression in ▸ figure 1 is 0.65. Human capital is given a weight of two thirds in Cobb Douglas models.
- Cross Countries studies often regress GDP *growth* on level of education, and also find large coefficient (one extra year of average schooling is associated with 0.3 percent extra growth every year in GDP, between 1960 and 1990).
- This raises a number of questions:
  - Sources of this strong correlation (in level and in the growth regressions)
  - If education is so important, need to understand the determinants of its provision, who should pay for it, the optimal way to pay for it, etc.

Mincerian Returns to Education

- Mincer hypothesizes that each extra year of education raise income by b%.

$$y_i = a + bS_i + cE_i + \epsilon_i$$

- Where S is schooling and E is experience.
- Why call this returns to education?
- Social returns may differ from private returns:
  - Costs
  - Externalities

Estimating Returns to Education

- This question can be estimated from micro data.
- Concerns:
  - Functional form (why log linear? Convex? Concave?)
  - Omitted variables
- Randomly identifying education is not easy, and convincing control strategies are difficult to come by.
- Therefore a large literature in labor searches for *instruments*: something that affects educational achievement but does not affect income directly.

Instrumental Variables

- Let Z_i be an *instrument*, which affects the probability that an individual is treated
- Let $W_i(1)$ be the treatment status for individual i if $Z = 1$, and $W_i(0)$ the treatment status of the same individual if $Z_i = 0$.
- The observed treatment is : $W_i = Z_i W_i(1) + (1 - Z_i) W_i(0)$
- As before, $Y_i(1)$ is potential outcome of treated (if $W_i = 1$) and $Y_i(0)$ is potential outcome if non-treated.
- Identification assumptions (Imbens and Angrist):
  1. *All* Potential outcomes are independent of the Instrument

     $$(Y_i(1), Y_i(0), W_i(1), W_i(0)) \perp Z_i$$

  2. What does this imply?
     - Treatment assignment is randomly assigned (or can be treated as such)
     - Treatment has no direct impact on the outcome (that is not implied by randomization of the instrument and has to be argued on a case by case basis!)
  3. Monotonicity: $W_i(1) \geq W_i(0)$ for everyone

More on Monotonicity

Three groups of people :

1. The Compliers: $Y_i(1) = 1$ and $Y_i(0) = 0$.
2. The Never-Takers: $Y_i(1) = 0$ and $Y_i(0) = 0$
3. The Always-Takers: $Y_i(1) = 1$ and $Y_i(0) = 1$
4. The Defiers: $Y_i(1) = 0$ and $Y_i(0) = 1$

The monotonicity assumption means that there are no defiers.
This is not a testable assumption, and needs to be assessed on a case by case basis.

Wald estimate and its interpretation

Wald estimate: Ratio of *Reduced form* and *First stage*.

$$\hat{\beta}_{IV} = \frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[W_i|Z_i = 1] - E[W_i|Z_i = 0]}$$

Case of constant treatment effect:

$$Y_i = a + bW_i + \epsilon_i$$

$$W_i = \alpha + \gamma Z_i + \upsilon_i$$

Substituting:

$$Y_i = a + b(\alpha + \gamma Z_i + \upsilon_i) + \epsilon_i$$

$$Y_i = a + \pi Z_i + \omega_i$$

Independence assumption insures that $\omega_i = \epsilon_i + b\upsilon_i$

$$b = \frac{\pi}{\gamma}$$

Two stage least squares

- Regress W on Z
- Regress Y on predicted W
- (in practice this is done in one step by the "two stage least square" procedure)
- Can be generalize to multiple instruments (and multiple treatment):
  1. Project (regress) X onto the vector of instruments Z
  2. Regress Y on the predicted value of X

$$\beta_{2SLS} = (W'Z(Z'Z)^{-1}Z'W)W'Z(Z'Z)^{-1}Z'Y$$

- Intuition: we are only using the part of the variance in the X for which we believe the identification assumptions.

Heterogenous treatment Effect

$$E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]$$

$$= E[W_i(1)Y_i(1) + (1 - W_i(1))Y_i(0)|Z_i = 1]$$
$$-E[W_i(0)Y_i(1) + (1 - W_i(0))Y_i(0)|Z_i = 0]$$

$= E[(W_i(1)-W_i(0))(Y_i(1)-Y_i(0))]+E[Y_i(0)|Z_i = 1]-E[Y_i(0)|Z_i = 0]$

$= E[(W_i(1) - W_i(0))(Y_i(1) - Y_i(0))]$ (by independence)
$= E[-(Y_i(1)-Y_i(0))|W_i(1)-W_i(0) = -1]P(W_i(1)-W_i(0) = -1)$
$+E[0 * (Y_i(1) - Y_i(0))|W_i(1) - W_i(0) = 0]P(W_i(1) - W_i(0) = 0)$
$+E[(Y_i(1) - Y_i(0))|W_i(1) - W_i(0) = 1]P(W_i(1) - W_i(0) = 1)$
$= E[Y_i(1) - Y_i(0)|W_i(1) - W_i(0) = 1] * P(W_i(1) - W_i(0) = 1)$
(by monotonicity)
$= E[Y_i(1) - Y_i(0)|W_i(1) - W_i(0) = 1] * (E[W_i(1)] - E[W_i(0)]$

Wald Estimate is treatment effect on the compliers

$$\hat{\beta}_{IV} = \frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[W_i|Z_i = 1] - E[W_i|Z_i = 0]}$$
$$= E[Y_i(1) - Y_i(0)|W_i(1) - W_i(0) = 1]$$

Who are the compliers?

- Special case: Treatment on the Treated:
  - When $W_i(0) = 0$ (e.g. randomized evaluation: all the control stays control)
- General case: Those are compelled by the instrument to get the treatement: external validity?
- While we cannot know who the compliers are, we can describe their characteristics

The INPRES Experiment: First Stage and Reduced form

- The Set up is a DID set-up similar to Bleakley's: Cohorts and Region
  - School construction campaign started in 1973: affect cohort age 12 or younger in 1973
  - More schools were built in regions that were initially lagging behind in term of education
- Results: Impacts of the program on Education and on log(wages)
  - ▸ Basic DID
  - ▸ Placebo experiment
  - Use all the regional variation (keep 2 cohorts) ▸ Table
  - Use all the regional variation, and all cohorts
- Check identification assumptions by estimating effects for all the cohorts ▸ Graph
- Force the earlier cohort to have an 0 effect: more precision ▸ Table

Instrumental variable

- What can we use as instruments?
  - If we wanted to use just one instrument
  - If we wanted to use many instruments?
- What are the identification assumptions? Do we believe in them?
- ▸ Results Did the IV make a big difference?
- What is the interpretation of the estimate? What are the years of education we are estimating the returns for?
- Interpretation of IV when the treatment takes more than one value: weighted average of marginal effects (going from 0 to 1, 1 to 2, etc..), where the weights are the fraction of people who are moved from one value of the instrument to another.
- See ▸ impact of programs by year of education

Reconciling Macro and Micro pictures

- Returns to education estimated from Mincerian specification ranges from 2.7% to 15.4%. Mean of 9%, stdv 2.2%. Generally at individual level IV is roughly equal to OLS.
- Puzzle 1: levels
- Countries in top decile of education distribution have about 8 more years of education than those in the bottom. They should have GDP no more than twice the size if private returns were the only part of the story. In fact they are about 15 times richer.
- Puzzle 2: Does the effect of *level* of education on *Growth* of GDP follows from the Mincer Framework? What can explain an effect of *level* of education on *growth* of income.
- Potential solution to both puzzle: Externalities.

Estimating Externalities

- The same experiment can be used to estimate the "social returns" to education
- Do we expect externalities to be positive or negative? (why?)
- We are looking to estimate:

$$y_i = \alpha + \beta S_i + \beta \overline{S_i} + \epsilon_i$$

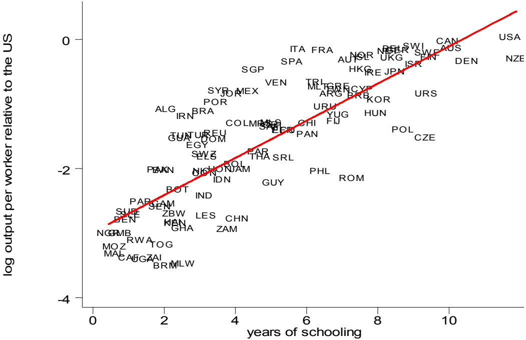
- Two estimation problem: we need an instrument for S_i and an instrument for $\overline{S_i}$ (Acemoglu and Angrist).
- Consider a cohort who was 12 or older in 1973, and is thus not exposed by the program
- Until 1979, no-one in the labor market is educated in the new schools.
- Starting in 1979, slow influx of the graduate of the new schools Graph

Empirical Strategy

- Fix the cohort, let the years vary.
- Survey Year*Region are instrument for $\overline{S_i}$. Are they correlated with S_i?
- Results (Graph, Table): Mushy, but if anything, equilibrium effects are negative.

Reconciling Macro and Micro picture (2)

- Externalities are not doing the trick...
- Other potential explanations:
  - Omitted variable
  - Endogeneity: Future growth in income motivates people to invest in education (Bills and Klenow)
- Micro-evidence of this channel
  - Foster+Rosenzweig HYV revolution in India (AER, 1995)
  - Jensen, Nguyen: Young people sensitive to *perceived* returns to education.


log output per worker relative to the US
years of schooling
0
-2
-4
0
2
4
6
8
10
USA
CAN
NZE
DEN
URS
POL
CZE
HUN
ROM
PHL
GUY
CHN
ZAM
IND
SGP
ITA
FRA
SPA
VEN
ALG
BRA
POR
MEX
KOR
JPN

Duflo (2001)

Means of Education and Log(Wage) by Cohort and Level of Program Cells

	Years of education			*Log(wages)*		
	Level of program in region of birth			*Level of program in region of birth*		
	High (1)	*Low (2)*	*Difference (3)*	*High (4)*	*Low (5)*	*Difference (6)*
Panel A: Experiment of interest						
Aged 2 to 6 in 1974	8.49 (0.043)	9.76 (0.037)	-1.27 (0.057)	6.61 (0.0078)	6.73 (0.0064)	-0.12 (0.010)
Aged 12 to 17 in 1974	8.02 (0.053)	9.40 (0.042)	-1.39 (0.067)	6.87 (0.0085)	7.02 (0.0069)	-0.15 (0.011)
Difference	0.47 (0.070)	0.36 (0.038)	0.12 (0.089)	-0.26 (0.011)	-0.29 (0.0096)	0.026 (0.015)
Panel B: Control experiment						
Aged 12 to 17 in 1974	8.02 (0.053)	9.40 (0.042)	-1.39 (0.067)	6.87 (0.0085)	7.02 (0.0069)	-0.15 (0.011)
Aged 18 to 24 in 1974	7.70 (0.059)	9.12 (0.044)	-1.42 (0.072)	6.92 (0.0097)	7.08 (0.0076)	-0.16 (0.012)
Difference	0.32 (0.080)	0.28 (0.061)	0.034 (0.098)	0.056 (0.013)	0.063 (0.010)	0.0070 (0.016)

Notes: The sample is made of the individuals who earn a wage. Standard errors are in parentheses.

Figure by MIT OpenCourseWare.

Duflo (2001)

Effect of the Program on Education and Wages: Coefficients of the Interactions Between Cohort Dummies and the Number of Schools Constructed per 1,000 Children in the Region of Birth

		Dependent variable					
		Years of education			*Log(hourly wage)*		
	Observations	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Panel A: Experiment of interest: Individuals aged 2 to 6 or 12 to 17 in 1974 (Youngest cohort: Individuals ages 2 to 6 in 1974)							
Whole sample	78,470	0.124 (0.0250)	0.15 (0.0260)	0.188 (0.0289)			
Sample of wage earners	31,061	0.196 (0.0424)	0.199 (0.0429)	0.259 (0.0499)	0.0147 (0.00729)	0.0172 (0.00737)	0.0270 (0.00850)
Panel B: Control Experiment: Individuals aged 12 to 24 in 1974 (Youngest cohort: Individuals ages 12 to 17 in 1974)							
Whole sample	78,488	0.0093 (0.0260)	0.0176 (0.0271)	0.0075 (0.0297)			
Sample of wage earners	30,225	0.012 (0.0474)	0.024 (0.0481)	0.079 (0.0555)	0.0031 (0.00798)	0.00399 (0.00809)	0.0144 (0.00915)
Control variables:							
Year of birth*enrollment rate in 1971		No	Yes	Yes	No	Yes	Yes
Year of birth*water and sanitation program		No	No	Yes	No	No	Yes

Notes: All specifications include region of birth dummies, year of birth dummies, and interactions between the year of birth dummies and the number of children in the region of birth (in 1971). The number of observations listed applies to the specification in columns (1) and (4). Standard errors are in parentheses.

Duflo (2001)

Effect of the Program on Education and Wages: Coefficients of the Interactions Between Dummies Indicating Age in 1974 and the Number of Schools Constructed per 1,000 Children in Region of Birth

Age in 1974	*Dependent variable: years of education*						*Dependent variable: Log(hourly wage)*		
	Whole sample			*Sample of wage earners*					
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*
12	-0.035	-0.025	0.002	-0.040	-0.010	0.009	0.016	0.019	0.027
	(0.047)	(0.048)	(0.054)	(0.077)	(0.078)	(0.091)	(0.013)	(0.013)	(0.015)
11	0.011	0.025	0.018	0.008	0.014	-0.003	-0.014	-0.013	-0.009
	(0.046)	(0.047)	(0.051)	(0.073)	(0.074)	(0.083)	(0.012)	(0.013)	(0.014)
10	0.059	0.049	0.078	0.10	0.092	0.13	0.0036	0.0042	0.0059
	(0.047)	(0.049)	(0.054)	(0.075)	(0.076)	(0.090)	(0.013)	(0.013)	(0.015)
9	0.14	0.14	0.15	0.067	0.063	0.17	0.0095	0.010	0.018
	(0.039)	(0.041)	(0.044)	(0.065)	(0.066)	(0.077)	(0.011)	(0.011)	(0.013)
8	0.088	0.11	0.11	0.19	0.20	0.28	0.019	0.021	0.027
	(0.049)	(0.050)	(0.054)	(0.078)	(0.079)	(0.089)	(0.013)	(0.013)	(0.015)
7	0.12	0.14	0.16	0.11	0.13	0.16	-0.0095	-0.0049	0.0066
	(0.044)	(0.046)	(0.051)	(0.072)	(0.073)	(0.084)	(0.012)	(0.012)	(0.014)
6	0.14	0.17	0.26	0.23	0.23	0.32	0.011	0.013	0.018
	(0.042)	(0.044)	(0.049)	(0.070)	(0.070)	(0.084)	(0.012)	(0.012)	(0.014)
5	0.10	0.13	0.13	0.14	0.16	0.27	0.021	0.023	0.052
	(0.043)	(0.045)	(0.050)	(0.075)	(0.075)	(0.088)	(0.013)	(0.013)	(0.015)
4	0.11	0.12	0.18	0.19	0.19	0.29	0.019	0.020	0.038
	(0.039)	(0.041)	(0.046)	(0.069)	(0.069)	(0.082)	(0.012)	(0.012)	(0.014)
3	0.11	0.14	0.20	0.15	0.17	0.30	0.0079	0.013	0.027
	(0.044)	(0.046)	(0.053)	(0.079)	(0.080)	(0.097)	(0.013)	(0.014)	(0.016)
2	0.14	0.19	0.19	0.20	0.22	0.25	0.016	0.023	0.040
	(0.041)	(0.043)	(0.049)	(0.073)	(0.074)	(0.088)	(0.012)	(0.013)	(0.015)
Control variables:									
Year of birth*enrollment rate in 1971	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Year of birth*water and sanitation program	No	No	Yes	No	No	Yes	No	No	Yes
F-statistic	4.03	5.18	6.15	2.70	2.74	4.38	1.13	1.29	2.05
R^2	0.19	0.19	0.17	0.14	0.14	0.13	0.14	0.15	0.13
Number of observations	152,989	152,495	143,107	60,633	60,466	55,144	60,633	60,466	55,144

Figure by MIT OpenCourseWare.

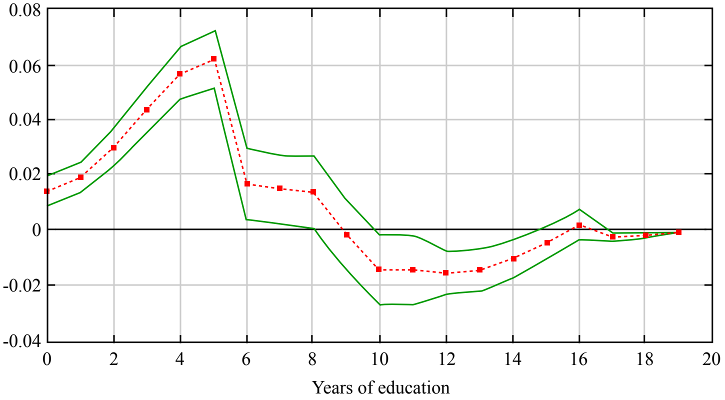
Duflo (2001)

Effect of Education on Labor Market Outcomes: OLS and 2SLS Estimates

Method	*Instrument*	*(1)*	*(2)*	*(3)*
Panel A: Sample of wage earners				
Panel A1: Dependent variable: log(hourly wage)				
OLS		0.0776 (0.000620)	0.0777 (0.000621)	0.0767 (0.000646)
2SLS	Year of birth dummies*program intensity in region of birth	0.0675 (0.0280) [0.96]	0.0809 (0.0272) [0.9]	0.106 (0.0222) [0.93]
2SLS	(Aged 2-6 in 1974)*program intensity in region of birth	0.0752 (0.0338) (0.0338)	0.0862 (0.0336) (0.0336)	0.104 (0.0304) (0.0304)
Panel A2: Dependent variable: log(monthly earnings)				
OLS		0.0698 (0.000601)	0.0698 (0.000602)	0.0689 (0.000628)
2SLS	Year of birth dummies*program intensity in region of birth	0.0756 (0.0280) [0.73]	0.0925 (0.0278) [0.63]	0.0913 (0.0219) [0.58]
Panel B: Whole sample				
Panel B1: Dependent variable: participation in the wage sector				
OLS		0.0328 (0.00311)	0.0327 (0.000311)	0.0337 (0.000319)
2SLS	Year of birth dummies*program intensity in region of birth	0.101 (0.0210) [0.66]	0.118 (0.0197) [0.93]	0.0892 (0.0162) [1.12]
Panel B2: Dependent variable: log(monthly earnings), imputed for self-employed individuals				
OLS		0.0539 (0.000354)	0.0539 (0.000354)	0.0539 (0.000355)
2SLS	Year of birth dummies*program intensity in region of birth	0.0509 (0.0157) [0.68]	0.0745 (0.0136) [0.58]	0.0346 (0.0138) [0.16]

Figure by MIT OpenCourseWare.

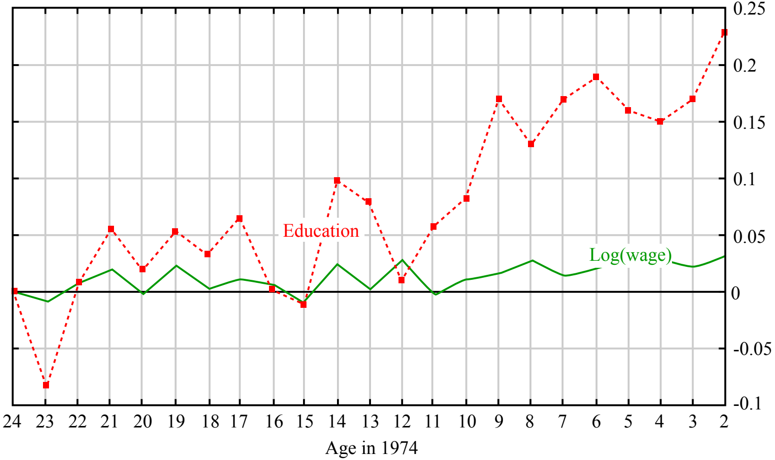
Duflo (2001)



Difference in differences in CDF (estimated from linear probability model) with 95-percent confidence interval

Figure by MIT OpenCourseWare.

Duflo (2001)



Coefficients of the interactions (age in 1974) * (program intensity in the region of birth) in the wage and education equations

Figure by MIT OpenCourseWare.

Duflo (2004)

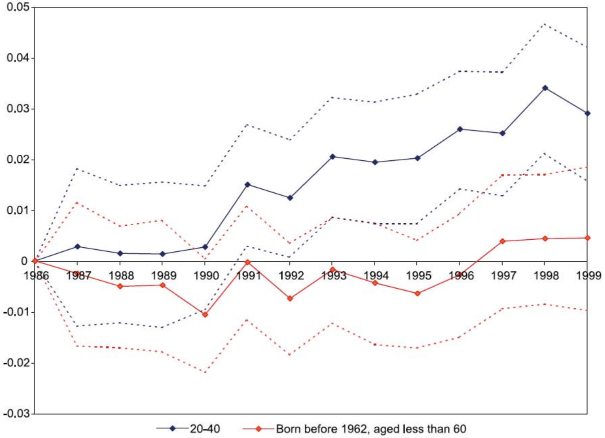


Fig. 2. Coefficients of the interactions of program intensity and survey year dummies. Dependent variable: % of primary school graduates.

Duflo (2004)

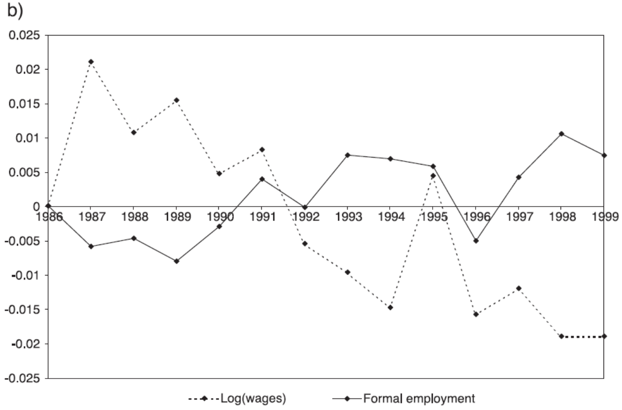


Fig. 4. (a) Coefficients of the interactions of program intensity and survey year dummies. Dependent variables: log(wage) and formal sector employment (individuals born before 1962 and aged less than 60). Sample: urban and rural regions. (b) Coefficients of the interactions of program intensity and survey year dummies. Dependent variables: average log(wage) and average formal sector employment among individuals born before 1962 and aged less than 60. Sample: rural regions.

Duflo (2004)

Table 6
2SLS estimates of the impact of average education on individual wages

	Independent variable: % of primary school graduates in the 20–40 sample		Independent variable: % of primary school graduates in the 20–60 sample	
	Sample: rural and urban areas	Sample: rural areas only	Sample: rural and urban areas	Sample: rural areas only
	(1)	(2)	(3)	(4)
Panel A: years 1986–1999				
Log (wage)	−0.204 (0.443)	−0.834 (0.701)	−0.208 (0.615)	−0.871 (0.837)
Log (wage) residual	−0.292 (0.355)	−0.633 (0.431)	−0.379 (0.512)	−0.994 (0.556)
Skill premium	−0.434 (0.916)	−0.982 (1.408)	−0.596 (1.197)	−0.636 (1.645)
Formal employment	0.441 (0.159)	0.454 (0.203)	0.661 (0.238)	0.745 (0.352)
Formal employment among educated workers	0.432 (0.197)	0.501 (0.259)	0.543 (0.264)	0.713 (0.406)
Formal employment among uneducated workers	0.379 (0.203)	0.409 (0.232)	0.510 (0.354)	0.318 (0.318)
Panel B: years 1986–1997				
Log (wage)	−0.358 (0.493)	−0.710 (0.821)	−0.451 (0.716)	−0.480 (0.801)
Log (wage) residual	−0.330 (0.412)	−0.588 (0.529)	−0.437 (0.618)	−0.902 (0.602)
Skill premium	−0.225 (1.033)	−0.635 (1.461)	−0.291 (1.488)	0.536 (1.576)
Formal employment	0.463 (0.183)	0.442 (0.233)	0.716 (0.282)	0.694 (0.379)
Formal employment among educated workers	0.428 (0.229)	0.473 (0.301)	0.530 (0.317)	0.622 (0.479)
Formal employment among uneducated workers	0.478 (0.249)	0.449 (0.277)	0.624 (0.415)	0.263 (0.319)

Men aged 20–60 and born before 1962.
1. Survey year dummies, region dummies, interactions between survey year dummies and the enrollment rate in 1971, and interactions between survey year dummies and the number of children are included in the regressions. 2. Regression run using kabupaten-year averages, weighted by the number of observations in each kabupaten-year cell. 3. The instruments are interactions between survey year dummies and the program intensity. 4. The standard errors are corrected for auto-correlation within kabupaten.