

Schooling and labor market consequences of school
construction in Indonesia : Evidence from an unusual policy
experiment

Esther Duflo

Main objectives

- Evaluating the outcomes of a large-scale education policy
 - Impact of infrastructure on schooling
 - cost-benefit analysis
- Estimating the return of schooling on :
 - education level
 - wages

The sekolah Dasar INPRES program

- favorable economic context (oil boom)
- 1973-74 , 1978-79 : 61,807 schools
- The fastest primary school construction program ever undertaken (1990) : 1 school per 500 children aged 5 to 14 in 1971.
- Teachers training,

Methodology

- Identification assumption :
 - Control group: individuals aged 12 or older in 1974 (not exposed) – assuming that in the absence of the program the increase in educational attainment would not have been systematically different in low and high program regions
 - Region of birth = region of education (for 91,5% according to IFLS)
- Controlling exogenous changes by controlling time-varying (date of birth) and region specific (region of birth) effects.

Methodology

- Controlling through difference in

TABLE 3—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) |
| <i>Panel A: Experiment of Interest</i> | | | | | | |
| 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) |
| <i>Panel B: Control Experiment</i> | | | | | | |
| 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) |

Methodology

Incremental effect one more school :

- On schooling

$$S_{ijk} = c_1 + \alpha_{1j} + \beta_{1k} + (P_j T_i) \gamma_1 \\ + (C_j T_i) \delta_1 + \varepsilon_{ijk}$$

$$y_{ijk} = c_1 + \alpha_{1j} + \beta_{1k} \\ + (P_j T_i) \gamma_1 + (C_j T_i) \delta_1 + \varepsilon_{ijk}$$

$$Y_{ijk} = \alpha_0 + \alpha_1 P_j + \alpha_2 T_k + \alpha_3 (P_j T_k) + \varepsilon_{ijk}$$

- ▶ Y_{ijk} = the outcome for student in i in region j in birth cohort k
- ▶ P_j = a dummy variable for high-intensity areas
- ▶ T_k = a dummy variable representing the younger cohort
- ▶ α_3 is the DD estimate for the effect of the program

TABLE 4—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

| | Observations | Dependent variable | | | | | |
|---|--------------|--------------------|--------------------|--------------------|---------------------|----------------------|---------------------|
| | | Years of education | | | Log(hourly wage) | | |
| | | (1) | (2) | (3) | (4) | (5) | (6) |
| <i>Panel A: Experiment of Interest: Individuals Aged 2 to 6 or 12 to 17 in 1974</i> | | | | | | | |
| <i>(Youngest cohort: Individuals ages 2 to 6 in 1974)</i> | | | | | | | |
| 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) |
| <i>Panel B: Control Experiment: Individuals Aged 12 to 24 in 1974</i> | | | | | | | |
| <i>(Youngest cohort: Individuals ages 12 to 17 in 1974)</i> | | | | | | | |
| 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) |
| <i>Control variables:</i> | | | | | | | |
| 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.

Interaction term analysis



FIGURE 3. COEFFICIENTS OF THE INTERACTIONS AGE IN 1974* PROGRAM INTENSITY IN THE REGION OF BIRTH IN THE WAGE AND EDUCATION EQUATIONS

Instrumental variables : estimating the impact of schooling on wages

- Assuming that the program had an effect on wages only through increasing years of schooling.
- Instrument variables estimates a regression of wages on the level of schooling predicted by the number of schools.

Methodology

Estimate regression

- OLS on :

$$y_{ijk} = d + \alpha_j + \beta_k + S_{ijk}b + \eta_{ijk}$$

- 2SLS on :

$$y_{ijk} = d + a_j + b_k + S_{ijk}b$$

$$+ \sum_{l=2}^{12} (C_j d_{il}) \pi_l + \eta_{ijk}.$$

Controlling the possible bias on 2SLS

- if the program affected both quantity and quality of school : 2SLS may be biased.
- Through pupil/teacher ratio and the in-existent effect of the program on children completing 9 or more years of school she deducts that the program had no significant effect on the quality.
- Thus increase of wage may be attributed to education

Controlling a possible sample selection problem

- The problem of the wage sample : only 45 % of the whole sample are working for a wage.
(working for a wage is affected by schooling, thus the sample selection is likely to induce a correlation between the instruments and the error term.)
- Solution :
 - introducing correction for sample selection
(slight change of coefficient from 8.1 to 9.2%)
 - Imputing an income for self-employed individuals and examine whether the results change.

Results : at what level the program was effective

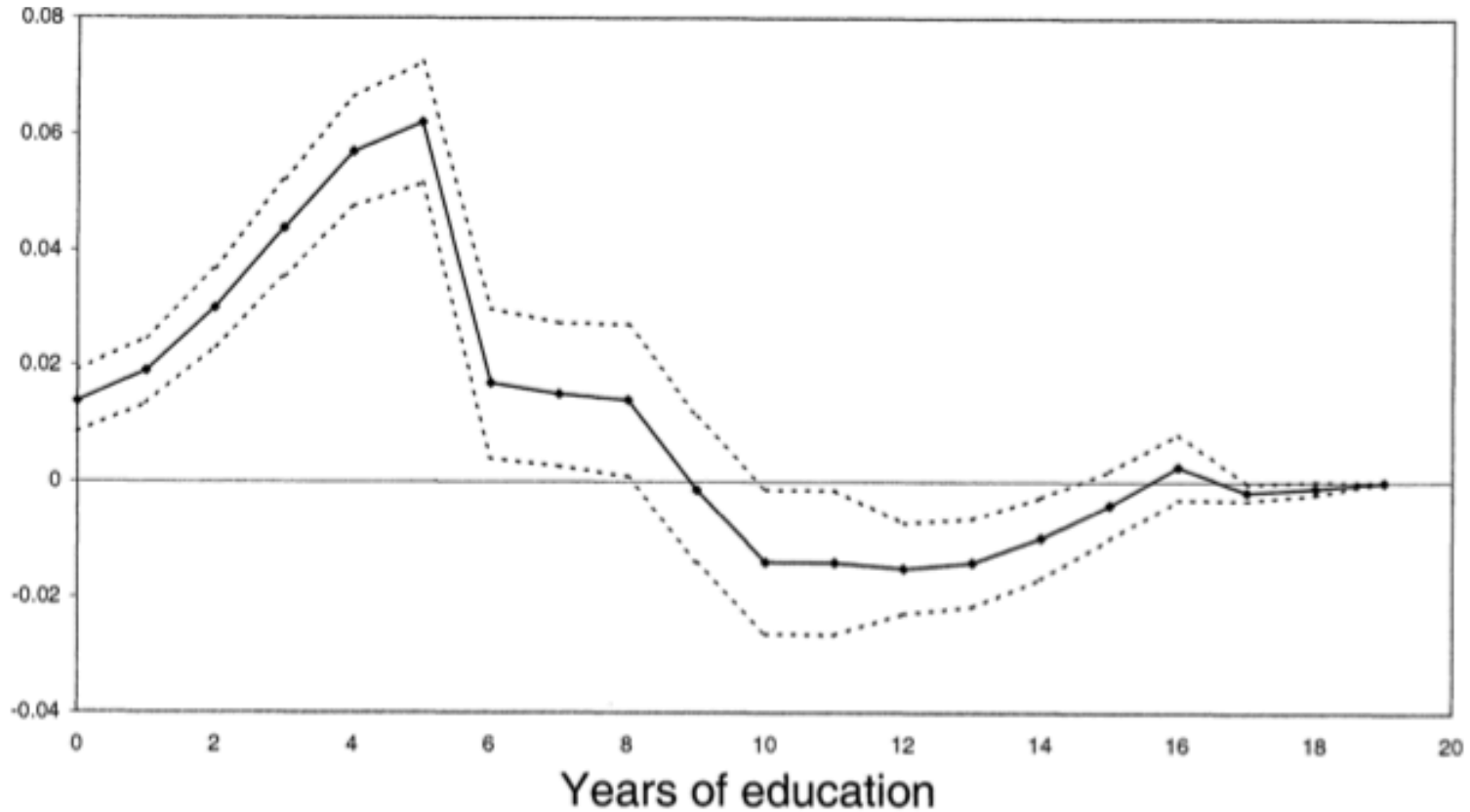


FIGURE 2. DIFFERENCE IN DIFFERENCES IN CDF (ESTIMATED FROM LINEAR PROBABILITY MODEL) WITH 95-PERCENT CONFIDENCE INTERVAL

Results : Cost – benefit analysis

- Under many assumptions :

Cost : $C(t) = rK + rTC + W(t)1.25$

Benefit . $B(c, t) = \alpha GDP(t)S(c, t)E(c)$

Results (table 8) :

- School construction program takes a very long time to generate positive returns
- The returns generated are large (8.8% – 12%)

Results : education

- Reducing the distance children traveled to school was the most important effect of the program (strong effect in low density region, Table 6, though there might be other characteristics correlated with density)
- The program effect is higher for wage earners than it is in the whole sample (Table 5)
- The program increased average schooling through increasing primary schooling essentially (figure

Results : wage

- Constructing one school per children increased the 1995 wages of individuals aged 2 in 1974 by 1.6% to 4.0%

- Program where educational wage



Results : estimation of return

- Return to education : 6.8% to 10.6% per year of additional schooling
- Variance across region : returns are higher in sparsely populated region, lower in regions where initial education was high.

Even after the program, returns were still higher in regions that received more schools.

Conclusions

- The INPRES program was effective in increasing the quantity of education and wages in Indonesia.
- Since 2SLS and OLS estimates are not significantly different, we can assume that OLS estimates of returns to education in developing countries are not **particularly biased** (as a result of omitted families, community background variables)

Comments

A methodology model to follow : OLS estimates can be used in similar context, development studies should focus on scientific ways of testing policies efficiency.

Relying on many assumptions (cost-benefit analysis) and really particular to one policy, no generalization possible