

Evaluations of Educational Interventions

Stephen Machin

November 15 2005

Issues

- A key challenge for policy makers is to identify when a government intervention can improve the allocation of resources.
- In general policy makers will want to know which interventions are the most attractive in terms of the net benefits they can deliver.
- Accurate assessment of the benefits and costs of each potential intervention are crucial, particularly when public funds, which have an obvious opportunity cost, need to be committed.

Aims

- First to establish what would be the precise impact on certain outcomes of the policy intervention .
- Second what would be the value of any improved outcome?

Approach

- Estimating the current or future impact of a new policy initiative requires assumptions to be made.
- This is regardless of whether a scheme is yet to be implemented, has been implemented nationwide, or has been piloted.
- The key is in estimating a counterfactual – namely what would have happened in the absence of, or alternatively the presence of, the policy intervention.

Concerns

- Often the parameter of most interest to the policy maker might not be precisely the same as the parameter that the evaluator can estimate. So it is vital to be careful to define what a particular evaluation approach is doing.
- In practice evaluations cannot be simply lifted off the shelf, but instead need to be carefully designed with respect to the specific problems that arise with a particular programme that is to be evaluated.

Definition of Evaluation Problem

- Simple general model with treatment and control group

$$\left\{ \begin{array}{l} Y_i^1 = a + b + \varepsilon_i + u_i \\ \\ Y_i^0 = a + u_i \end{array} \right.$$

- This shows the outcome in sector 1 (say with treatment) and sector 0 (no treatment). The return to treatment is

$$Y_i^1 - Y_i^0 = b + \varepsilon_i$$

The Policy Evaluation Problem

- The key evaluation issue is that the individual effect of a policy is not and cannot ever be observed, because the same individual cannot be in and out of treatment at the same point in time.
- The response of the evaluation literature has been to identify different parameters of interest that could perhaps be estimated under different evaluation designs and/or assumptions. These parameters of interest relate to different aspects of the distribution of the effect.

Parameters of Interest

- An obvious parameter of interest is the average treatment effect, which is b in the notation above (assuming $E(\varepsilon_i) = 0$). This is the expected effect of the policy for a random individual.
- If there is selection (on observables X) into the programme may be interested in the the effect of the treatment on the treated (TT): this is

$$TT = b + E(\varepsilon_i | D_i = 1)$$

- And we may be interested in the entire distribution of returns , which may be very difficult to obtain, or perhaps impacts for subgroups of the population defined by particular values of the vector X .

Methods

1) Random assignment

Where (otherwise identical) individuals are randomly assigned into treatment and control groups.

Resolves selection issues so can estimate average treatment effect.

Not seen very often in education programmes (e.g. JTPA programme in the US, Tennessee STAR).

Methods

2) Difference-in-differences

The key idea is that we can split up the population into two groups: one is exposed to a policy/treatment and the other one is not. Alternatively one is more likely to obtain treatment and the other is less likely.

The key assumption is that the *change* in the outcome variable is the same between the two groups in the absence of this exposure. The “effect” is then estimated as the difference in the growth of the outcome variables over time for the two groups, divided by the difference in the proportion of those exposed to the policy in the two groups.

Methods

3) Matching

An alternative approach to evaluation is that of matching.

In the matching framework it is recognised that selection of individuals into treatment is not random but that selection takes place based on observable variables plus random factors unrelated to outcomes.

Try to identify, using statistical methods a set of people in a control group who have similar observable characteristics.

Methods

4) Combining matching with difference-in-differences

The matching approach allows for an arbitrary dependence of the program effect on characteristics X . The simple difference-in-differences approach does not.

But they can be combined, to obtain a closer matched control group.

General Equilibrium Effects

- All methods considered so far make the implicit assumption that treatment does not affect the comparison group either directly or indirectly.
- Can be a problem: in the presence of peer group effects; or if there is displacement or diffusion.

Applications From the UK:

i) The Literacy Hour

ii) Education Maintenance Allowances

Literacy Hour - Motivation

- Literacy matters
- ‘One in five adults in the UK are not functionally literate’ - Moser Report A Fresh Start (DfEE, 1999).
- Key question - How to ensure literacy of next generation?

The Literacy Hour

- Constitutes a discrete, well-defined and systematic change in how literacy is taught within primary schools.
- Daily literacy hour: 10-15 minutes whole class reading or writing; 10-15 minute session on word level work (e.g. phonics, spelling); 25-30 minutes directed group activities; plenary session at the end.

The Literacy Hour

What happened before?

- General standard of reading was a concern, especially with wide variations across Local Education Authorities.
- John Stannard, Director of National Literacy ‘in some LEAs teaching of literacy had fallen apart’.
- Criticisms included ‘free reading, with little or no intervention by the teacher; too much time hearing individual pupils read; insufficient attention to the systematic teaching of an effective programme of phonic knowledge and skills’ (OFSTED, 1996)

How Were NLP Schools Selected?

First LEAs, then schools within LEAs:

- 1). LEAs – in the main part those perceived to be LEAs with lowest national performance, though somewhat mixed in actuality (and demographics and disadvantage considered). Bigger list picked than actually participated (possibly resource constraints).
- 2). Schools – advised to select schools most in need for programme. Did have school level performance data, own knowledge and OFSTED inspections information. Again not necessarily lowest english attainment, with a clear spread within LEAs.

Evaluation of the Literacy Hour

- The National Literacy Project was introduced in 400 schools in September 1996 and September 1997.
- Gives setting where some children exposed to up to two years of the literacy hour, when other children were not.

Differences in Reading Percentiles Between Pupils in NLP and non-NLP Schools

	No controls	All controls	Fixed effects	Matching	Matching Fixed effects
NLP*	2.143	2.423	2.631	1.715	1.791
Policy On NLP	(.741)	(.647)	(.650)	(.631)	(.628)
	-8.808	-1.075	--	-.962	--
	(.845)	(.883)		(.892)	
No. Pupils	104654	104654	104654	96083	96083
No. Schools	841	841	841	761	761

Pre-Policy Period Comparison

	No controls	All controls	Fixed effects
1. Above Estimates			
NLP*Policy On	2.710 (1.068)	3.182 (0.924)	3.182 (0.924)
2. School Level Estimates 1996-98 (school performance tables)			
NLP*Policy On	2.346 (1.063)	2.746 (0.936)	2.607 (1.139)
3. School Level Estimates 1995-96 (school performance tables)			
NLP*Policy On	0.198 (1.320)	0.244 (1.337)	0.304 (1.902)

Cost-Benefit Calculations

- Present discounted value of the cumulative effect of the literacy hour is estimated as £5476, £4302 and £2103.
- Costs: planned cost of the NLP was £12.5 million over 5 years. The main costs were 14 local centres (each costing about £25,000 per year) and literacy consultants in each participating Local Education Authority (about £27,000 per year for each consultant).
- The total cost per annum was thus £2.5 million (or about £2.8 million in 2001 prices). We observe the number of students affected from pupil numbers in the schools within Cohorts 1 and 2 in 1997 and 1998 (i.e. 222,261 pupils in aggregate). Hence cost per pupil is £25.52 per annum.
- Therefore highly cost effective.

Education Maintenance Allowances

- EMA pilots launched in September 1999 in ten local education authorities in England.
- Provide a means tested benefit to 16-18 year olds to encourage them to stay on in full-time education after the compulsory school leaving age (15/16 in Year 11).
- Researchers at Institute for Fiscal Studies have looked at impact on post-compulsory education participation in treatment-control setting.

Education Maintenance Allowances

Table 4.1. Impact of EMA on Year 12 destinations of eligibles

	Participation in Pilot Areas	Unmatched	OLS	Fully Interacted OLS	Matching Estimate
<i>Total:</i>					
FT Education (S.E)	69.2	3.9 (1.4)	3.8 (1.3)	4.5 (1.4)	4.5 (2.3)
Work/Training (S.E)	16.4	-0.4 (1.1)	-1.0 (1.1)	-1.7 (1.1)	-1.7 (2.4)
NEET (S.E)	14.5	-3.5 (1.1)	-2.8 (0.9)	-2.7 (1.0)	-2.7 (2.0)
Sample size:		5,315	5,315	5,299	5,299

Policy Evaluations in Education

– Some Conclusions

- Impact and cost/benefit evaluation of public policy programmes should take a central role in the design of public policy. This is the only way to ensure that scarce public funds are used effectively and for the purpose(s) that they were allocated.
- It is also an important means of ensuring that in the long run we will accurately learn what works and what does not, allowing governments to offer more beneficial and cost effective interventions.
- Evaluation also has the major advantage that it places the public policy debate on ‘scientific’ rather than ‘ideological’ grounds; this makes it more likely that effective policies will be preserved while ineffective ones will not continue to be a burden on the public purse.