

Evaluating public policy in the UK: history, politics & practice

Stephen Morris,
NatCen Social Research &
Policy Studies Institute, University of Westminster

Overview

- Market for evaluation of public policy in the UK & how it operates
- A brief history of evaluation since 1997
- Methods and approaches
 - Debate over causation
 - An example of randomised trial in public policy
 - Issues and challenges
- Concluding remarks

Market for evaluation in the UK

Producers and consumers I

Consumers of evaluation (demand):

- Central government
- Local government
- Pressure groups & Trade Unions (even the Church)
- Not-for-profits/charities and third sector
- Private providers of public services

Producers and consumers II

Producers (supply)

- Institutions demanding evaluation also supply some of their needs internally
- Suppliers general external to government:
 - Research institutes & universities
 - Not for profit research companies
 - Market research companies (TNS, Ipsos, etc.)
 - Management consultancy firms (Deloitte, PWC, Ecorys, etc.)

5

How does government commission evaluation?

- Social Researchers (professional group within UK Civil Service)
 - Located within each department
- Work with policy makers and administrators to develop ToRs & specifications
- Work is competitively tendered by individual departments
 - Framework (usually with domain specific Lots)
 - Non/framework
- EOI → Shortlist → Full Tender → Award of contract

Can be completely open or based on framework

Usually no more than five providers

Short listed providers invited to tender

Tenders assessed on quality/value

What are evaluations used for?

- Testing of new policy ideas – pilots & demonstrations (limited to specific areas/time periods)
- Identifying whether existing interventions need to be reformed or abandoned
- Improvement of implementation & management of interventions
- Development of new interventions – action research
- Pragmatic – politicians want evidence that their programmes/interventions work as a political defence

Recent history of evaluation in UK & what politicians think of evaluation

New Labour - 'What matters is what works'

Rise of evidence-based policy & evaluation

Political factors:

- Technocratic government (competence) – centre ground
- New Labour = Thatcherism (neo-liberalism) with social justice
- Sceptical of government but activist

Outside politics:

- Evidence-based healthcare
- Influence of North America (both Canada and US)
- Expansion of social sciences in 1960s/70s

Evaluation under New Labour

- Policy hyperactivity
 - Welfare reform/education/social inclusion
 - Crime & justice
- Pilots and trials
- Expansion of social scientists within government
- Budgets
- Consultancy firms & market research agencies enter the market
 - Existing suppliers expanded

Evidence & politics an uncomfortable relationship?

NatCen
Social Research

“Social science should be at the heart of policy making... we need social scientists to help determine what works and why”

David Blunkett, Secretary of State for Education

But.....

“No one with the slightest common-sense could take seriously suggestion by University researchers that homework is bad for you..” David Blunkett, Secretary of State for Education

From a presentation by Professor Huw Davies

The coalition - ‘Evidence-based policy’ to ‘values-based policy’?

NatCen
Social Research

- Surface rejection of New Labour
- But much policy continuity
- Less emphasis on piloting policies
- Austerity
- Sceptical of evaluation commissioned under New Labour – poor value for money
- Reduced budgets – big market adjustment

Healthy dose of realism...

“People don't elect a politician simply with a manifesto that says I will look at the evidence and then decide what to do - they elect politicians because they have a vision for the country”

"I certainly believe in evidence-based policy and the Prime Minister does and the cabinet are committed to it. the world is imperfect. Politics has more to it than simply sitting in a kind of permanent seminar looking at academic evidence”

David Willetts, Minister of State for Universities and Science

Methods and approaches

Evaluation design and approaches (since 1997)

- Growing sophistication in methods
 - Much greater focus on design and not just on data collection
 - Mixed methods within an overarching design framework
- Randomised control trials (social experiments/field trials)
- Quasi-experimental approaches
 - Control group not constructed at random
 - Application of econometric methods in analysis
- Greater understanding of complementarity between impact evaluation & process evaluation
- Improvement in use of qualitative research
 - More analytical less descriptive
 - Role in providing an explanation – how and why?

Data, technology & evaluation

- Data
 - A lot of money spent on primary data collection
 - Access to administrative data
 - Evaluations of welfare policies relied on benefit receipt data & tax records
 - Education policies – National Pupil Data Base
 - Crime & justice – Police National Computer
 - Increased concern over data protection and confidentiality
- Development of computing technology
 - Sophisticate analytical techniques
 - Collection of survey data via CAPI

Causal attribution debate

- We want to spend limited resources on interventions which offer best value for citizens – that have impact
- Aim: determine whether observed outcomes can be attributed to programmes activities
- UK debate: experimentalists v realists
 - Experimentalists – emphasise causal identification through experimentation, i.e. randomised trials/control groups, etc.
 - Realists – critical of experiments - ignore social contexts & causal mechanisms (e.g. black box problem)
 - Build theory – generate predictions – test theory
- An example of how experimentalists approach the problem.....

ERA Demonstration – the experimentalist ideal approach?

- Test new approaches to supporting low-income groups in work
- Help those leaving welfare retain work & advance
- But also a ‘showcase’ for experimental methods
- Build capacity in the UK to conduct such large-scale trials (outside health care/clinical research seldom done)
 - Learn from North American experience
- Launch-pad for many further such evaluations
 - Represents the ideal from an experimentalists perspective
 - A high point in evidence-based policy (welfare)

What did ERA test?

A voluntary in & out of work case-managed advice service & cash incentives, providing:

- Up to 9 months pre-employment support
- Two years of personal in-work job coaching
- Up to 6 tax-free payments of £400 for working more than 29 Hrs.p.w.
- Up to £2,000 for in-work job training

Target groups

Three target groups - considered to suffer from job retention and advancement problems

- Lone parents out of work (New Deal for Lone Parents)
- Long-term unemployed aged > 25 (New Deal 25 Plus)
- Lone parents working part-time (16-29 hours a week)

Programme sites:

- Six areas of the country: NE London, C Manchester, Tyneside & Gateshead, Derbyshire, SE Wales & Renfrewshire.
- Programme delivered by PES Offices - around 50 in total across six areas

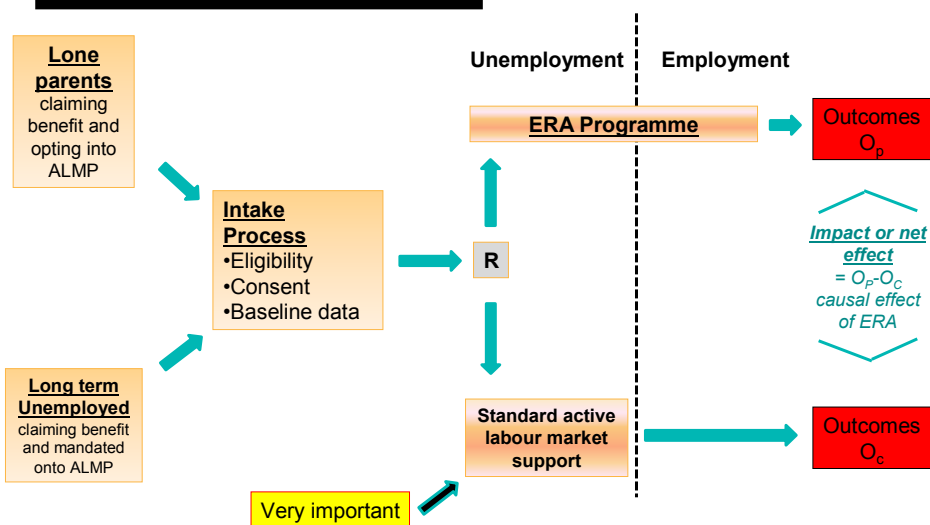
Timing

- Programme ended in October 2007 – evaluation ended 2011

Evaluation – mixed methods

- Programme of evaluation activities (almost a third of the intervention's budget)
- Mixed method:
 - **Design stage** – theoretical justification for intervention
 - Process study – mainly qualitative
 - **Impact study** – based on a randomised design - centre piece of the design & its standout feature
 - Cost and cost-benefit analysis
- Evaluation collected both survey and administrative data over a five year follow-up period
- Consortium of research institutes and firms led by MDRC (US consultancy) with UK research organisations PSI, IFS and NIESR.

Evaluation design



How did randomisation go?

- Across all target groups 16,834 individuals were assigned at random (Oct 2003 to Jan 2005)
- Programme group 8,206 control group 8,178
- Groups are well balanced on observables
 - Driving licence – 43% programmes/42% controls
 - Education (no qualifications) – 30% programmes/28% controls
 - Previous employment – 87% of programmes and 87% of control have had previous work

23

Age profile

	ERA Programme	Control
Less than 25	19%	18%
25 - < 30	15%	16%
30 - < 35	17%	18%
35 - < 40	19%	18%
40 - < 45	13%	12%
45 - < 50	8%	8%
Greater than 50	9%	10%
Base:	8,206	8,178

24

Advantages

- Two groups are statistically equivalent at baseline in all respects
- Subsequently only the programme group get the new intervention
- Control represents what would have happened to programme group had they not participated in ERA (counterfactual outcomes)
- Any differences in outcomes between the two groups which emerge are attributable solely to intervention
 - we can rule out alternative explanations
- Analysis and reporting straightforward & clear –compared to quasi-experimental alternatives

25

Results – lone parents

Outcome		Programme group	Control group	Impact
Ever employed (%)	Year 1	57.1	56.5	0.6
	Year 2	57.8	55.6	2.2*
	Year 3	53.7	53.8	-0.2
	Year 4	53.2	54.0	-0.8
	Year 5	52.9	53.9	-1.0
Earnings (£s) Tax years 2005-2009 (4 years)		17,280	16,742	538
Total amount of welfare benefits received (years 1-5, £s)		9,349	9,848	-499

Long term unemployed

Outcome		Programme group	Control group	Impact
Ever employed (%)	Year 1	37.3	35.4	1.9*
	Year 2	36.3	32.7	3.6***
	Year 3	34.6	32.5	2.1*
	Year 4	35.0	32.1	2.9***
	Year 5	32.8	30.9	1.9*
Total earnings (£s) Tax years 2005-2009 (4 years)		14,162	12,681	1,481**
Total amount of welfare benefits received (years 1-5, £s)		7,067	7,493	-426***

Results – part-time workers

Outcome		Programme group	Control group	Impact
Ever employed (%)	Year 1	77.0	76.6	0.4
	Year 2	74.1	73.4	0.8
	Year 3	71.2	69.9	1.3
	Year 4	71.3	70.1	1.2
	Year 5	68.6	68.2	0.3
Total earnings (£s) Tax years 2005-2009 (4 years)		30,615	29,695	921
Average total amount of welfare benefits received (years 1-5, £s)		1,190	1,149	40

Summarising the results

- Early gains in ave. earnings for lone parents and part-time workers due to increased hours which faded post-programme (mainly women) – effective for better qualified only
- Intervention not cost-effective for lone parents except better qualified
- Results much more encouraging for long-term unemployed (mainly men and very disadvantaged)
- Improved rates of employment and earnings – emerging after 1 year but persisting for 5 years
- Reduced dependence on welfare
- Cost-effective for this group from perspective of participants, exchequer and society more widely

Problems/issues

Set-up issues

- Hard to persuade administrators of the need for randomisation
- Long time to design and set-up
- Initially some ethical objections

Implementation

- Members of control group did receive in-work support
- Randomisation changing the composition of those who take-up the scheme (some evidence that eligibles had been diverted)
- Treatment diffusion

Non-response and timing of measurements

- Reliance on surveys for early results – year two findings suggested very different policy response

Costs of randomisation

- Randomisation of subjects can appear expensive (£7-10m)
- ERA - tracking & randomisation system
 - Generated a sample frame
 - Collected important baseline data
 - Administrators didn't like using the system
- Special algorithm designed to avoid subversion
- Technical advisers located at experimental sites
- Attempting to design-out the bias much more obvious cost than the cost of tackling it in analysis
- Possible to conduct much cheaper experiments

31

Looking back at ERA

Policy lessons

- PES can deliver in-work support effectively and efficiently for highly disadvantaged long-term unemployed
- Among lone parents – better educated did benefit
- Some conclusions drawn too soon

Evaluation lessons

- Randomised control trials feasible and deliverable
- From a 'raising our game' perspective ERA failed
- Randomisation became an unattainable ideal – sceptics point to its budget, timetable and uniqueness
- Some successes in terms of (US to UK) technology transfer

Looking to the future

- Experimentalists have won the debate for now but there won't be a lot of randomised experiments
 - Very few theory-based impact evaluations/lots of quasi-experiments
 - But maybe theory-based approaches better placed
- Challenges - evaluation set to get very much more difficult
 - Private and not-for-profit delivery (privatisation of public policy)
 - Payment by results & black box commissioning
 - Social impact bonds – private investors to invest in social/public programmes
 - Encouraging local solutions – 'a thousand flowers bloom'
 - These innovations are generating huge complexity and considerable challenges for evaluators

Finally.....

- Evaluation is hard
- We can be scientific
- Honesty is the best policy
- You learn a lot when policies don't work & you save money (if you know they don't work)!
- Timing is important

To end

“There is nothing a government hates more than to be well-informed; for it makes the process of arriving at decisions much more complicated and difficult.”

John Maynard Keynes, *Collected Writings*, vol. 21, p. 409.