

Institution: London School of Economics and Political Science
Unit of Assessment: 19: Business and Management Studies
Title of case study: Using targets and incentives to improve the quality of public services
<p>1. Summary of the impact (indicative maximum 100 words)</p> <p>Gwyn Bevan's research used the 'natural experiment' whereby each of the four countries of the UK applied a different model of governance to its 'National' Health Service. From 2000, each devolved government allocated unprecedented increases in health spending, each set similar goals for improved performance, and yet performance was transformed only in the NHS in England. Bevan's research explained why and changed the understanding of key policy actors about the use of targets to achieve performance goals. The evidence base it created also influenced the governance of health services, notably a general shift to the use of targets for all the UK National Health Services.</p>
<p>2. Underpinning research (indicative maximum 500 words)</p> <p>RESEARCH INSIGHTS AND OUTPUTS. In the late 1990s, across the four UK NHSs, there was a crisis of poor quality indicated by long hospital waiting times. From 2000, the Blair government sought to transform the quality of the NHS in England through two policies that were radically different from the past. The first policy in England was sustained and generous increases in NHS funding each year. This policy was followed by the other governments of the devolved countries. The second policy was to abandon the old model of governance, which rewarded failure to deliver targets by NHS organisations with extra resources. In its place a new model was proposed, which set ambitious targets ('P45 targets') with sanctions for failure. The sanctions that applied from 2000 to 2005 in the regime of annual 'star ratings' were 'naming and shaming' and 'targets and terror' (i.e. Chief Executives were at risk of being sacked). This second policy was not followed by the non-English governments, which continued with the old model of governance. These differences in governance between the NHSs in England and the devolved countries are the basis of the 'natural experiment'.</p> <p>The underpinning research began when Bevan was on secondment (from 2000 to 2003) from LSE as a Director at the Commission for Health Improvement (CHI), the regulator of NHS quality in England and Wales. He had the lead responsibility for CHI's contribution to the development and implementation of the regime of annual 'star ratings' in the English NHS. He learned that, consistent with its policy of not publishing school league tables, the Welsh Government had decided not to introduce a similar regime of 'naming and shaming'. At a CHI seminar in 2002, Bevan presented comparisons that showed that hospital waiting times were reduced in England but increased in Wales. Bevan also discovered, through work by colleagues at CHI, that there was 'gaming' by hospitals and ambulance services in response to targets in 'star ratings'.</p> <p>On his return to LSE Bevan evaluated the 'natural experiment' in a number of studies:</p> <ul style="list-style-type: none"> • At a national level across the four NHSs for hospitals and ambulance services [1,2, 3]; • In comparisons of England and Wales for ambulance services [4], and hospital waiting times [5], and the effects of the presence or absence of 'naming and shaming' [6]. <p>The research provided three principal findings. First, that the regime of 'star ratings' transformed reported performance of the NHS in England. Second, that outcomes in England were not matched by the performance of health services in the devolved countries. Third, that there was also gaming in the English NHS.</p> <p>These findings suggested two strategic lessons for policy design.</p> <ul style="list-style-type: none"> • Policies which link high-powered incentives to targets can work for public services where targets accurately measure key dimensions of performance and gaming is of secondary importance (e.g. hospital waiting times, ambulance response times and school league

Impact case study (REF3b)

tables);

- However, where the targets can only be proxies (and hence not an accurate measure) for important dimensions of performance, the dysfunctional effects of gaming may make their use problematic (e.g. policing).

Governments seek to improve public services and like to claim that their policies are evidence-based. But, the nature of policy making means that there is typically little good experimental evidence; hence the importance of Bevan's research across the four NHSs of the UK, which were so similar except for different models of governance.

KEY RESEARCHERS. Gwyn Bevan has been at LSE since 1997 (seconded to CHI 2000-2003).

3. References to the research (indicative maximum of six references)

1. Alvarez-Rosete A, Bevan G, Mays N, Dixon J. 2006 'Diverging policy across the UK NHS: what is the impact?' *British Medical Journal*, 331: 946-50. DOI: 10.1136/bmj.331.7522.946
2. Bevan G, Hood C. 2006 'What's measured is what matters: targets and gaming in the English public health care system'. *Public Administration* 6, 84(3): 517-38. Available on the World Bank website at <http://siteresources.worldbank.org/EASTASIAPACIFICEXT/Images/226299-1251872399239/bevanhoodpubadmin%5B1%5D.pdf> (Google Scholar: 447 citations) DOI: 0.1111/j.1467-9299.2006.00600.x
3. Connolly S, Bevan G, Mays N. 2011 *Funding and Performance of Healthcare Systems in the Four Countries of the UK before and after Devolution*. London: the Nuffield Trust, (revised edition). http://www.nuffieldtrust.org.uk/sites/files/nuffield/funding_and_performance_of_healthcare_systems_in_the_four_countries_report_full.pdf
4. Bevan G, Hamblin R. 2009 'Hitting and missing targets by ambulance services for emergency calls: impacts of different systems of performance measurement within the UK'. *Journal of the Royal Statistical Society (A)*, 172(1): 1-30. Available at <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-985X.2008.00557.x/full>
5. Besley T, Bevan G, Burchadi K. 2009 *Naming & Shaming: The impacts of different regimes on hospital waiting times in England and Wales*. DP7306. <http://www.cepr.org/pubs/dps/DP7306>
6. Bevan G, Wilson D. 2013 'Does 'naming and shaming' work for schools and hospitals? Lessons from natural experiments following devolution in England and Wales'. *Public Money and Management*, 33(4): 245-252. DOI: 10.1080/09540962.2013.799801

Evidence of quality. Papers 1, 2, 4 and 6 are in good peer-reviewed journals. Paper 3 was submitted to RAE 2008.

4. Details of the impact (indicative maximum 750 words)

IMPACTS. The research in section 2 had impacts through public engagement, influence over the terms of the debate about targets, and downstream policy changes, especially in devolved governments. Bevan's research has featured in the media and led to personal invitations to policy debate. It also provided a model for other research enquiries into health services and school systems, and into emerging policies for policing.

Public engagement

- 2009: Bevan's research featured in a BBC Radio 4 programme, Analysis - In Defence of Targets [7].
- 2009: Bevan was invited by the editor of the *British Medical Journal* to participate in a debate on 'Have targets done more harm than good in the English NHS?' [8].
- 2011: Bevan was invited by an editor of the *British Medical Journal* to review the evidence

Impact case study (REF3b)

that competition between hospitals improves clinical quality [9].

- 2013 Bevan and Wilson were invited to a House of Lords/Parliamentary Seminar on Benchmarking Public Services for Excellent Performance, to present their paper on models of governance of the NHSs and schools in England and Wales.

Public engagement also had an international dimension:

- 2010: Bevan was interviewed by and quoted in the *WHO's* report on assessment of the performance of health systems [10].
- 2012: Bevan was invited to contribute to seminars at the World Bank on performance measurement in developing countries – the papers by Bevan and Hood [2] and Bevan and Hamblin [3] are posted on the World Bank website.

Influencing the terms of the debate and policymakers

- 2010: The Report by Connolly et al [3] was covered extensively in the media and by the Scottish Parliament, with defensive responses in Wales and Scotland, as described by Timmins [11].
- 2011: Bevan's research was cited by Burgess et al [12], who analysed the 'natural experiment' between England and Wales in terms of performance in GCSEs at age 16. This study produced strong evidence that devolution resulted in worse grades in Welsh schools relative to those in England.
- 2011: Peter Preston in *The Guardian* [13] cited Bevan's research as evidence of the effectiveness of governance by targets in commenting on the shift to a quasi-market model for the English NHS by the then Secretary of State for Health.
- 2013: The Nuffield Trust Report on public reporting of performance, written for Jeremy Hunt, the new Secretary of State for Health in England, extensively cited Bevan's research into 'star ratings' [14].
- 2013: Guilfoyle [15], a serving Police Inspector, cites the research by Bevan and Hood [2] which identified scope for gaming in response to targets in his account of how he improved performance in his police force. Guilfoyle argued that the inflexible nature of targets in policing led to such dysfunctional consequences that they should not be used in attempts to improve performance of the police.

Policy changes

In addition to its influence on the terms of debate, Bevan's research is also strongly associated with policy changes in the health services of the four devolved nations through its contribution to an evidence base about targets, which the devolved governments could not ignore. Timmins, who was public policy editor of the *Financial Times* between 1996 and 2011, and is now a senior fellow at both the King's Fund and the Institute for Government, is the author of Report published by the Kings Fund, which is the most influential health policy think tank in the UK. His Report [11] draws on research by experts on each country's NHS and provides evidence of the impact of Bevan's research. His Report states, "In terms of day-to-day management of the NHS ... one of the most obvious examples of lessons learnt is over the use of targets" (p 13). He made three points to justify that statement.

First, *the strength of the research evidence*: "there is no doubt that 'targets and terror', as the approach was originally dubbed by Gwyn Bevan and Christopher Hood, worked in England by focusing management effort on reducing waiting times as a key priority (Bevan and Hood 2006; Proper et al 2007). In England they duly fell, sharply and steadily. By 2005, even despite the difficulty of directly comparing waiting times *between* the countries, it was clear that waiting times *within* countries were coming down faster in England than in Scotland, Wales or Northern Ireland" (p 13).

Second, *the shift by governments* in the devolved countries from initial rejection of, to following, the approach taken in England: "Initially, Wales rejected, and Scotland was at best lukewarm, about the target-driven approach to waiting times that the English NHS adopted, ahead of the re-

introduction of choice and competition” (p 13); but, given the evidence, by 2005, decided to follow suit the English approach, “adopting such targets, and in each case waiting times fell” (p 13).

Third, “the coalition government in England, having first given the impression that it was abandoning waiting time targets, has now reaffirmed them” [11, p 14].

WHY DOES THE IMPACT MATTER? Improved governance leads to improved performance of public services in terms of the objectives that matter to the public and hence to governments, such as reducing long hospital waiting times and good performance in GCSEs at age 16. The ‘natural experiment’ following devolution shows which types of governance are and are not effective in achieving highly salient measures of performance of public services.

5. Sources to corroborate the impact (indicative maximum of 10 references)

All Sources listed below can also be seen at: https://apps.lse.ac.uk/impact/case_study/view/21

7. Analysis - *In Defence of Targets*, BBC, 27th September 2009, http://www.bbc.co.uk/iplayer/episode/b00mr16g/Analysis_In_Defence_of_Targets
8. Bevan, G. 2009 ‘Have targets done more harm than good in the English NHS?’ No. *BMJ*; 338: a3129 <http://www.bmj.com/content/338/bmj.a3129>
9. Bevan G; Skellern M. 2011 ‘Does competition between hospitals improve clinical quality? A review of evidence from two eras of competition in the English NHS’. *BMJ*; 343:d6470 <http://www.bmj.com/content/343/bmj.d6470>
10. World Health Organisation. *Health systems performance assessment. A tool for health governance in the 21st century*. Copenhagen: WHO Regional Office for Europe. http://www.euro.who.int/_data/assets/pdf_file/0019/160813/HSPA_A-tool-for-health-governance-in-the-21st-century.pdf
11. Timmins N. 2013 *The Four UK Health Systems*. London: King’s Fund. http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_summary/four-uk-health-systems-jun13.pdf
12. Burgess S, Wilson D, and Worth J. 2013 ‘A natural experiment in school accountability: The impact of school performance information on pupil progress’. *Journal of Public Economics* 106: 57-67. <http://www.sciencedirect.com/science/article/pii/S0047272713001291>
13. Preston P. ‘Andrew Lansley’s reality test’, *The Guardian*. 21 November 2011. <http://www.guardian.co.uk/commentisfree/2011/nov/20/andrew-lansley-targets-nhs>
14. The Nuffield Trust. 2013 *Rating providers for quality: A policy worth pursuing?* London: The Nuffield Trust,. <http://www.nuffieldtrust.org.uk/publications/rating-providers-quality>
15. Simon Guilfoyle, 2013 *Intelligent Policing: How Systems Thinking Methods Eclipse Conventional Management Practice*. Axminster: Triarchy press. <http://www.triarchypress.com/pages/Intelligent-Policing.htm>