

A new economics of public services: Lessons from using targets and quasi- markets in the UK and New Zealand

Gwyn Bevan (Professor, London School of Economics)

Rodney Scott (Associate Professor, University of New South Wales)

WORKING PAPER JUNE 2019

In 2018, both authors were Visiting Fellows at the Blavatnik School of Government at the University of Oxford. This paper is the summary of several discussions had over this period, in which public management experiences in the UK and New Zealand were compared and contrasted.

Abstract

Since the 1980s, in New Zealand and the countries of the UK, there have been various attempts to deploy elements of New Public Management to address the traditional absence of externally driven incentives for those who work in public services. These attempts have included quasi-markets and various versions of accountability for performance against targets. Quasi-market reforms proved to be problematic, while well-designed governance by targets proved to be effective. We argue that an effective system of governance generates sanctions and rewards for poor and high performance that are not financial but reputational. The rationale for this is based on concepts of identity economics and reciprocal altruism.

Keywords

Quasi Markets, Targets, Delivery Units, Reputation, Reciprocal Altruism

Introduction

Public administration is made more complicated by the traditional absence of external incentives for good performance. One of the main features of New Public Management ('NPM') is introducing such incentives through various means (Hood, 1995; Scott, 2001), as tried in New Zealand and England since the 1980s. Lodge and Gill (2011) describe New Zealand as '*the* primary example of an extensive and intellectually coherent attempt at public sector reform that

followed the key characteristics associated with New Public Management' (p142, emphasis in original). Hood (1995) pointed out that the UK offered fertile territory for NPM, because it combined motive, for a government looking for efficiency, and opportunity, given the absence of checks on constitutional reform.

In this paper we consider three basic governance models in the context of New Zealand and the UK over the past thirty years. One, 'trust and altruism', is the traditional model of trusting public servants to behave altruistically and lacks external incentives. The other two aim to introduce rewards and sanctions for good and poor performance: 'quasi-markets', based on money following choice that determines the funding of providers; and 'reciprocal altruism' based on reputation effects from publishing measures of provider performance by 'naming and shaming' and 'naming and faming' that which is unacceptably poor and exceptionally good. We develop conceptual arguments and present empirical evidence to explain why 'trust and altruism' is inadequate for addressing poor performance and why 'reciprocal altruism' offers more powerful sanctions and rewards than quasi-markets for local services. Reflecting on our experiences in being involved in the design and implementation of 'naming and shaming' in England and 'naming and faming' in New Zealand, we argue each proved effective because they were appropriate for their respective contexts.

The inadequacy of 'trust and altruism'

In the 1980s, in both the UK and New Zealand, governments were dissatisfied with performance of their public services. These Governments had trusted public

servants to do their best, and did not have effective means for sanctioning those who did not (Schick, 1996; LeGrand, 2003). We examine this model using the simple dichotomy of 'knights' and 'knaves', proposed by Smullyan (1978) and LeGrand (2003), in which public servants may be virtuous and intrinsically motivated (knights) or unscrupulous and self-interested (knaves). The 'trust and altruism' model assumes that all public servants are 'knights', who see their jobs as a calling (Perry and Hondeghem, 2008), and are driven by intrinsic motivation with a mission to deliver services of high quality. Therefore, all that is needed to improve performance of public services is supplying these 'knights' with comparative information (LeGrand, 2003); and there is no need for external incentives in terms of sanctions for failure or rewards for success. This model has four dysfunctional consequences.

First, a corollary of the assumption that all are assumed to be 'knights' doing the best they possibly could, is that failure entails a need for extra resources (Bevan and Fasolo, 2013). So 'knaves' learn that their lack of effort will be rewarded. Second, sanctions are a much stronger motivator than rewards, because people feel losses much more keenly than gains of equivalent magnitude (Tversky and Kahneman, 1991). Smith (1759, p103) pointed out that 'A man of sensibility is apt to be more humiliated by just censure than he is ever elevated by just applause'. The absence of any sanctions for failure removes the high-powered incentives likely to be necessary to overcome organisational inertia for improving performance. Third, this model can undermine the motivations of the 'knights' who see their efforts ignored whilst 'knaves' are rewarded. Fourth, by lacking enforcement of what constitutes acceptable standards of performance for public

servants this model has potentially corrosive effects in defining their identity. Nevertheless, governance by 'trust and altruism' is the default model for many countries because of its popularity with professionals and its minimal costs of monitoring.

Tackling 'knaves' through performance management

While the model of 'trust and altruism' treats all public servants as 'knights', Hume (Miller and Hume, 1994) argued the contrary:

Political writers have established it as a maxim, that in contriving any system of government, and fixing the several checks and controls of the constitution, *every man ought to be supposed a knave* and to have no other end, in all his actions, than private interest. By this interest, we must govern him and, by means of it, notwithstanding his insatiable avarice and ambition, co-operate to the public good (Miller and Hume, 1994:42–43, italics in original).

As Hume recognises that 'this is false in fact' (Miller and Hume, 1994, p43), his argument is that an effective model of governance must be designed to try to eliminate 'knavish' behaviour. This principle was used in England from 2000 to 2005 by the Blair Government for departments accountable to the Prime Minister's Delivery Unit ('PMDU') for improving performance of schools and hospitals (Barber, 2007). These departments published annual rankings of performance of both schools (league tables of examination results) and hospitals

(star ratings) – see below. These were designed to satisfy what Hibbard *et al.* (2003) argued are the three requirements for publishing information on performance to inflict reputational damage on those performing poorly; that this information: is published and widely disseminated, ranks performance in a way that can be easily understood by the public, and is followed up by future reports. Such systems subject poor performers to a loss of reputational esteem through ‘naming and shaming’ and so being publicly ‘unfrocked’ as ‘knaves’. We now consider a natural experiment comparing performance under ‘naming and shaming’ in England, and ‘trust and altruism’ in Wales.

A natural experiment between England and Wales

England and Wales offer a good ‘natural experiment’, because each used to have common governance models for the NHS and schools prior to devolution in 1999, but afterwards their governments implemented different systems of governance (Bevan and Wilson, 2013). (Before 1999, legislation for public services in the UK was common for England and Wales, but differed for Scotland and Northern Ireland.)

For schools in England and Wales, from 1988, their government departments annually published the outcomes of inspections, and league tables of their pupils’ performance in examinations at 16 and 18. The stated objective was to inform parental choice in quasi-markets (Burgess *et al.*, 2013). Following devolution, in response to calls from the National Union of Teachers (BBC news, 1999), the Welsh government stopped publishing school league tables and, in effect, moved

to governance by the model of 'trust and altruism'. Burgess *et al.* (2013) found that this resulted in a dramatic relative worsening of performance, as compared with England. This is shown in Figure 1, which gives the percentages achieving the target of five 'good' grades (A* to C) in the General Certificate of Secondary Education (GCSE) in England and Wales. These grades are based on assessments across a number of subjects over two or three years to the end of compulsory education at 16. Five 'good' grades are the minimum requirement for continuing with further education or leaving school with good employment prospects. Burgess *et al.* (2013) estimated that the adverse impact of not publishing league tables would be equivalent to worsening the pupil teacher ratio in Wales by increasing the typical class size by over 25 per cent. They also found that the relatively poor performance of pupils in Wales was largely due to worse outcomes for the lower 75% of schools, measured by student prior attainment (measured by Key Stage 3 exams that aim to test learning from ages 12 to 14), and poverty (measured by eligibility for Free School Meals that are defined in relation to income and welfare payments). Hence, 'naming and shaming' by publication of school league tables seems to have narrowed social inequalities, whereas as we argue below, quasi-markets tend to widen them.

The NHS in the UK experienced one of its recurrent 'crises' in the winter of 1999/2000, after devolution. The principal finding of a Chapter on the health sector in an OECD (2000) report on the UK was summarised by Clive Smee, Chief Economist in the Department of Health, as: having 'highlighted poor cancer survival rates in the UK, suggested that other disease-specific outcomes were also poor, and noted the limited progress on waiting times and the apparent under-

investment in both doctors and buildings'. Smee also points out that: 'In private the authors went further and indicated that they had been unable to identify any features of the NHS that were particularly commendable' (Smee, 2008, p. 92). The government in England recognised its failings and described the NHS then as 'a 1940s system operating in a 21st century world' (Secretary of State for Health, 2000, p26). This was the justification for changing governance from 'trust and altruism' (from 1997 to 2000) to 'naming and shaming' (from 2000 to 2005) in the regime of an annual (star) performance ratings, which were dominated by targets for reducing waiting times. In 2000/01, twelve acute hospitals were 'zero rated' as 'failing'. Six of their chief executives were sacked by the centre (in a regime of 'targets and terror'), because they were deemed to be incapable of delivering performance to meet the new standards (Bevan and Hood, 2006). But the government in Wales persisted with governance by 'trust and altruism'. This 'natural experiment' has been the subject of a series of evaluations (Bevan and Hood, 2006; Bevan and Hamblin, 2009; Bevan and Wilson, 2013, Bevan *et al.*, 2014). Here we highlight differences in performance for hospital waiting times and response times by ambulances to potentially life-threatening emergency ('Category A') calls; each was the subject of damning reports on failures in governance to tackle unacceptably poor performance by the Auditor General for Wales (2005 and 2006).

Figure 2 gives the targets for hospital waiting times from referral by a general practitioner to hospital admission for elective care in England and Wales. In England, for 2001/02 (the second year of the 'star rating' regime) the targets were 104 weeks; for 2005/06 (the end of that regime), 39 weeks; for 2008, 18 weeks.

Over this period, hospitals in England and Wales were allocated substantial extra funding and those in England managed to achieve these increasingly demanding targets, but those in Wales failed to hit the target of 156 weeks in 2005/06 (Bevan et al, 2014). There is a similar story for performance of ambulance services against the target (set in both countries for 2001) that 75 per cent of potentially life-threatening emergencies (Category A calls) be met within 8 minutes (see Figure 3). This target was not achieved in England until the 'star rating' regime was applied (for 2001/02), with 65 per cent as the threshold set for a service that had been zero-rated as 'failing'. The continued failure by the Welsh Ambulance Service to meet the 75 per cent target resulted, not in public censure, but in less demanding 'milestone' targets being set: 65 per cent from 2004 to 2005; 60 per cent from 2005 to 2008, and 65 per cent from 2008 (Bevan et al, 2014).

Gaming and regulation

Any account of improvements from linking measures of performance to high-powered incentives has to recognize the substantial body of literature identifying its many dysfunctional consequences (see e.g., Holmstrom and Milgrom 1991; Power, 1999; Bevan and Hood, 2006; Arnaboldi et al., 2015). Without adequate checks from independent auditors, gaming can have corrosive consequences on what is deemed to be acceptable behaviour in which those performing poorly learn that the most promising route to avoiding sanctions is to game the system rather than try to achieve genuine improvements. Kelman and Friedman (2009, p. 26, emphasis in original) point out that:

Simply to note that a performance measurement regime produces some level of dysfunctional response does not by itself imply that such a regime fails on balance to improve organizational performance.... The appropriate comparison is between an organization's performance level with performance measurement and the dysfunctional responses, and the counterfactual performance level *with no measurement*.

Bevan and Hamblin (2009) argue that the absence of any attempts at gaming in response to measures of performance suggests that providers do not take these measures seriously. What matters is the scale of gaming. A series of careful studies over the period for which we have compared performance in England and Wales found that gaming in England was limited in its scale: Burgess et al (2013) for schools, Propper et al (2010) for hospital waiting times, and Bevan and Hamblin (2009) for ambulance response times to Category A calls. A plausible explanation for these findings is that regulation of quality in schools, hospitals and ambulance services in England was by inspections based on visits (Bevan and Cornwell, 2006). Arnaboldi et al. (2015) argue that the scandalously poor care provided by Mid-Staffordshire NHS Foundation Trust from 2006, which followed draconian cuts in numbers of nurses, was caused by targets for waiting times and finance. Bevan (2011), argues, however, that this was a failure of regulation that allowed such flagrant gaming to go on for four years before it was detected as a consequence of seeking supply-side flexibility from the independent sector in implementing market-based reforms from 2006. As a consequence, the government changed in scope of the independent auditor of quality in the English NHS. From 2000, the Commission for Health Improvement (CHI) was organised

to inspect the NHS only through week-long visits to NHS Trusts. To implement the market reforms, the government abolished CHI and required its successor, the Health Care Commission to design a regulatory 'level playing field' for the hundreds of NHS hospitals and thousands of independent hospitals and clinics. To do this, the Health Care Commission abandoned inspection based on visits for relying on routinely-available data and, for NHS Trusts, their own assessments of quality of care. This is why, in 2006/07, when a later investigation found that the previously poor quality of care at Mid-Staffordshire was deteriorating to be become scandalously poor, the Health Care Commission's annual Health Check highlighted that Trust as being as one of the four 'most improved acute and specialist trusts'.

Limitations in application

Both Barber (2007) and Le Grand (2007) recognised the successes of 'naming and shaming' in improving services in England from 'awful' to 'adequate'; but argued that such improvements would still leave the public dissatisfied, and a different approach was required to further improve services from 'adequate' to 'good' or 'great'. Hood and Dixon (2010) found that the outcomes of the 'natural experiment' had brought neither electoral benefits for the government in England nor losses in Wales. Blair, having observed that 'I bear the scars on my back' from reforms of public services (BBC News, 2007), sought an alternative to 'keeping flogging those who deliver public services'. We now consider the model that was

then tried to achieve further improvements in performance; namely, the use of quasi-markets.

The disappointment of quasi-markets

The rationale for quasi-markets is that providers compete for 'customers', and 'money follows choice' (Le Grand, 2007) so that providers' funding depends on their performance. In a quasi-market with effective demand and flexibility on the supply side, 'knights' who provide high quality services would be rewarded with increased market share, gains in income and expansion; and knaves, who provide poor quality services would be penalised with reduced market share, losses of incomes and, eventually, having to exit the market. Thus the 'invisible hand' of the market promises to remedy the failings of the 'trust and altruism' model and the limitations of 'naming and shaming' with a system of rewards and sanctions that are commensurate with the scale and degree of high and poor performance.

Implementing and abandoning quasi-markets

Quasi-markets were introduced for schools in England and Wales from 1988, and in New Zealand from 1989. This change: allowed parental choice, instead of schools being determined by zones; introduced money following the pupil instead of incremental budgeting; and was supported by performance information published by the respective governments (Burgess *et al.*, 2013; Hughes *et al.*, 1996). Quasi-markets in which public and private providers competed for service contracts from public purchasers were introduced in health care in England,

Wales, Scotland and Northern Ireland from 1991; and in New Zealand from 1993 (Bartlett et al, 1998; Borren and Maynard, 1993). Of these eight quasi-markets only one continues: the quasi-market for schools in England (Walford, 2014). We now describe how the other seven were abandoned.

Following devolution in 1999, Wales moved from quasi-markets for schools to a model based on collaboration between educational partners (Reynolds, 2008). In New Zealand, the settings that enabled quasi-markets for schools introduced between 1989 and 1991 were gradually reversed from 1995 to 2001, with progressive reinstatement of: equity funding, prohibition of merit-based selection processes for students, the partial reintroduction of zoning, and the gradual reintroduction of a national curriculum (Gordon, 2003; Stubbs and Strathdee 2012).

All UK and New Zealand quasi-markets in healthcare have been abandoned. The Blair government ended the first quasi-market for health care in the UK in 1997. Following devolution, England gradually reintroduced a second quasi-market from 2006, but none of the other UK countries did so (Bevan, 2014). The Coalition government proposed changes to a third attempt at a quasi-market in England in 2010 (Timmins, 2012) but by 2019, after trying the quasi market model for over 30 years this has been abandoned for a policy of development of Integrated Care Systems [*The NHS Long Term Plan*, 2019], which requires a different model of governance. In New Zealand, the quasi-market for health care was abandoned in 2001 because it was unable to demonstrate performance improvements (Cumming and Salmond, 1998; Easton, 2002; Ashton, 2002; Ashton et al, 2005)

and was incompatible with the prevailing group identity of the health system (Hornblow, 1997; Ashton, 2005). Hornblow (1997, pp. 1894) in particular, characterised the failures of the health reforms as due to a 'clash of cultures' between clinical and commercial interests.

An absence of evidence

Le Grand (2007) considers the argument that 'consumers' of hospitals want a good local hospital or school rather than choice. His counterargument is that choice in quasi-markets offers the best way of achieving those outcomes. But successive attempts at quasi-markets for schools and hospitals shows there are intrinsic problems on the supply and demand sides.

On the supply side, because Ministers are required to ensure good local access, 'failing' providers cannot simply exit the market through the workings of its 'invisible hand'. This results in what Williamson (1985, p 95) describes as site asset specificity, as one of the causes of market failure from small numbers of competitors. This severely limits supply-side flexibility and so impairs the potential dynamism of these markets. For schools in England, Chowdry et al., (2008) identified the consequence of the lack of supply side flexibility that entry and exit of schools was 'extremely rare' even for poorly-performing schools, which suffered reductions in funding and so provided worse services for their pupils. The Coalition government, elected in 2010, sought to create more supply-side flexibility by enabling private funding for 'free' schools (Walford, 2014). We know outcomes from another 'natural experiment': between Sweden, which introduced

'free' schools (*friskolor* in Swedish) in 1992, and Finland, which focused on making the teaching profession an elite group. These are adjacent countries with similar cultures. Finland transformed its previously failing school system (before the 1960s: 'comparable to that of Malaysia or Peru') by making the teaching profession attractive. This was not achieved through high salaries but by raising it up as an elite group and defining its identity through 'high social prestige, professional autonomy in schools, and the ethos of teaching as a service to society and the public good' (Sahlberg, 2010, pp 1 and 2). Now Finland is the top-ranked European nation in the Programme for International Student Assessment ('PISA' – OECD 2015); while Sweden's achievement has fallen further than any other nation (Wigmore, 2016). For hospitals, because exit is so problematic for Ministers (Tuohy, 1997), the first NHS internal market did not develop a competition with winners and losers but rather, as Royce (1995) observed, operated as an Alice in Wonderland version in which 'all must have prizes'.

On the demand side, studies of the effects of school choice in New Zealand produced extensive evidence of increased socio-economic segregation without improvements in performance (e.g. Lauder et al, 1999; Gordon, 1994; Fiske and Ladd, 2000; Nash, 1999; Waslander and Thrupp, 2006). Such findings have been replicated in other countries that have introduced choice (Musset 2012). Education (unlike health care) has the distinctive characteristic among public services of being what Hirsch (2005) described as a 'positional good', the value of which depends on its relative scarcity. So, if a high-performing school were to expand supply, this would diminish the value of its positional good. In New Zealand, there is evidence that declaring a school 'full' increased its popularity,

which provided an incentive for high performing schools to limit enrolment (Gordon 2003). In England, one of the consequences of the lack of supply side flexibility is that, instead of parents choosing good schools, good schools were choosing parents through 'cream skimming' (Whitty et al, 1998; West et al 2006). For hospital care in England, as in other countries (Ham, 2008; Maarse, et al, 2016), effective selective contracting failed to develop (Audit Commission for Local Authorities and the National Health Service in England and Wales, 1996; Le Grand et al, 1998; Wyke et al, 2003; Smith and Curry, 2011). Lavery et al (2012) found that the attempt to remedy this abiding weakness by introducing patient choice also failed to have an impact on hospitals' market shares, even following scandals that attracted local and national publicity (including that at Mid-Staffordshire NHS Foundation Trust). They concluded that 'reporting designed to affect provider reputation may ultimately be more effective (in driving improvements in quality) than reliance on patients' choosing their providers' (p 599). This followed similar findings from studies on hospitals in the US (Hibbard, 2008).

The argument for quasi-markets as an effective form of governance of public services comes, not from evidence of their success, but rather from an appeal to their potential. Their attractions are those that Smith (1759, Kindle p 275) recognised as a tendency to value a 'beautiful orderly system' regardless of the ends it delivers. In the 1980s, even though the various market mechanisms that were tried violated the requirements for the perfect markets of micro economic theories (Besley and Ghatak, 2003), we also knew that for various other goods and services imperfect markets had proved effective. In the following thirty years,

however, we have learned about problems of lack of flexibility on the supply side, of inequities in demand for schools and the ineffectiveness in demand for hospitals. These problems have undermined quasi-markets as the means of creating sanctions and rewards for poor and good performance. So whilst it made sense in the 1980s to try quasi-markets to explore how their promise as a novel way to foster intrinsic motivation in those who deliver public services, the evidence is that these have not fulfilled that promise.

Rediscovering knights

Governance by the 'trust and altruism' model is flawed because it assumes that all are 'knights'. 'Naming and shaming' is limited as it is designed to tackle failure only. Although quasi-markets seem to be designed both to reward success and punish failures, experience and evidence shows that, as failure does not result in exit, they fail to correct 'knavish' behaviour. And markets can crowd out 'knightly' motives, as shown in the classic study comparing the systems of voluntary blood donation in England with its marketisation in the US (Titmuss et al, 1997). Bowles (2016) argues, based on a substantial body of carefully designed experiments, that it is quite wrong to assume that using market mechanisms to tackle 'knavish' behaviour will not undermine 'knightly' behaviour. Hence quasi-markets in practice threaten to do the opposite of the idealised version by undermining 'knightly' and encouraging 'knavish' behaviour for which they lack effective sanctions. This section describes an example of the second element of reciprocal altruism designed to support an aspiration to excellence, rather than mere adequacy, where knightly behaviours were celebrated by 'naming and faming'.

'Naming and faming' in New Zealand.

In 2011, the New Zealand government sought to improve the performance of public services, and commissioned an independent Better Public Services Advisory Group (State Services Commission, 2011) to report on how this might be achieved. The group found high performance except where responsibility fell across or between multiple agencies (State Services Commission, 2011), an example of the 'silo mentality' that frequently accompanies NPM (Lodge and Gill, 2011). The government's response was the introduction, in 2012, of the BPS programme, based on governance by targets. The government selected ten objectives that spanned the responsibility of multiple departments from the government's top priorities that had been resistant to past attempts at change, and set challenging targets for each (State Services Commission, 2012) (see Table 1). For each of the ten objectives, the government described the intended result, a target for improving performance over five years, and the method by which progress would be measured. Because these objectives spanned department boundaries, small groups of departments were held collectively responsible for progress made (and successes were attributed collectively to that group). Reporting consisted of a line graph showing historic and current performance, the further improvement required to reach the target, a qualitative assessment of progress, and descriptions of some of the best and most important changes as experienced by New Zealanders. There was collegial competition among public servants to be 'named and famed' as examples of success and for choice of powerful stories to capture the public's imagination (Jensen et al. 2014). These

stories were about the consumers of services and the changes they experienced, in which the contributions of named public servants were mentioned in support of the narrative. This indirect faming, where public servants were famed for helping New Zealanders without being the subject of direct praise, may have been instrumental; direct awards can be demotivational and create resentment and backlash (Mackenzie, 2007; Rosenblatt, 2011). Public servants described this as a strong motivator (Scott and Boyd, 2016). Detailed examples of how performance was reported, and how this supported positive competition, are shown in Scott and Boyd (2017). (Ed. note: The Better Public Service Results programme is explored in more detail in Scott and Bardach 2019; Scott and Boyd 2017, 2020, 2022; and Scott and Merton, 2020.)

The ten BPS targets ran for five years from 2012-2017. For each of the ten objectives, departments began from different levels of readiness to take up the challenge. Some made rapid progress, visible from the first six months, whereas others took two full years to stop downward trends and exceed even their baseline measure. Nonetheless, all ten eventually made substantial progress when measured against their baselines and historic trends (Scott and Boyd, 2017). We illustrate the impact of the Better Public Services Results with two examples, where many years of effort prior to 2012 had produced little improvement. Figure 4 shows the increase in the percentage of children aged 2 or 3 enrolled in early childhood education after the programme. Figure 5 shows the increase in the percentage of infants receiving their full vaccination schedule by 8 months, from 86% to 94%.

In the absence of a 'natural experiment' for the BPS programme, we consider six counterfactual explanations for the reported improvements. From these we conclude that some portion of the improvements were associated with the programme itself, as other explanations could only explain part of the observed changes in some of the objectives.

1: Improvements represent some sort of natural progression. But, after 2012, performance for each objective was better than the trend.

2: Improvements were as a result of increased attention. But the objectives were chosen because each had been found to be intractable despite significant prior attention and action.

3: Problems were going to improve as a result of the delayed effects of prior actions. Discussion with the managers and practitioners involved indicates that this may be a factor in one problem only, that of increasing infant immunisation rates (from 86 per cent), but they acknowledged that this was likely to explain only a small change in the initial stages of the five year programme in the measured outcome (of 94 per cent).

4: Measures are imperfect, and may not be good reflections of the underlying problem. Seven initiatives rely on a full count of the actual incidence of the problem. Two (workforce qualifications and ease of doing business with government) are assessed through a survey based on random samples and not subject to bias. The tenth (assaults against

children) relies on substantiated cases, which reflect a combination of the underlying incidence of violence and the likelihood of those assaults being reported. The latter is likely to increase by the greater attention and effort to address family violence.

5: Measures were ‘gamed’, or intentionally manipulated, such that the measure improved without an underlying improvement in the problem. This is possible for two of the problems only. Reducing the measured crime rate could be achieved by diverting minor offences away from the criminal justice system (without formal charges). Similarly, the criminal recidivism rate could be reduced by focussing resources on preventing reoffending among those who have committed minor offences (where progress is easier), and away from more serious offences that are associated with greater social harm. A previous study could not find evidence of gaming at a systematic level (Scott and Boyd 2016).

6: Measures improved due to additional resources provided. But the programme was put in place when budgets that were declining in real terms, due to the after effects of the global financial crisis.

In the earlier discussion of performance measurement, we identified three elements of a successful system for tackling ‘knavish’ tendencies: shaming, exit (the sack), and an independent auditor specifically tasked with countering gaming. The BPS programme suggests a fourth element: ‘naming and faming’ may be used to encourage ‘knights’ to achieve high performance. Doing so generates

reputation effects from competition for awards (Frey, 2013), and is consistent with pedagogical theories of performance suggesting that recognition of success is a strong motivator in experimentation and innovation (Stipek 1993). (Ed. note: New Zealand's approach to public service motivation is explored in more detail in Scott 2019; Scott and Macaulay 2020.)

It may be that 'naming and faming' in this way applies best at a small scale, as it relies on a credible chance of being highlighted as a top performer. Reputational rewards have proven to be effective in New Zealand and Finland, but may need to be applied at a sub-national level to be effective in larger countries.

Discussion

We have organised this paper around 'knights' and 'knaves' because this has proved to be a useful construct in discussion of governance models. However, most public servants, as humans, are neither pure 'knights' nor knaves, but display elements of both extremes, in different contexts and in response to the behaviours of others. We aim to respond to the calls from Arnaboldi et al. (2015) for a system of governance of requisite complexity by deepening the discussion of governance models using two other constructs: reciprocal altruism, and social identity.

Reciprocal altruism (Rothstein, 2011; Oliver 2017) suggests that people behave altruistically toward each other because this increases the likelihood that this kindness will be returned. Axelrod (1981) describes this as a 'tit-for-tat' evolutionary strategy. Wilson (2015, p 23) summarises the advantages of altruism

as 'Selfishness beat altruism within groups. Altruistic groups beat selfish groups. The rest is commentary'. This suggests that if we want high performing public services we should be aiming to encourage altruistic behaviours among public servants. Reciprocal altruism is a kind of 'interactive rationality' (Kahan 1974), where individuals adapt their behaviour on the basis of their expectations of the behaviour of others. From this, we infer that knightly tendencies are encouraged by observations of knightly behaviours in others and, conversely, by evidence that 'knavish' behaviours will not be tolerated. By not correcting knaves, 'trust and altruism' redefines identity as self-interest and discourages altruism by diminishing reciprocity. By only correcting knaves, 'naming and shaming' provides an identity defined negatively, as not being terrible, and doesn't provide a pathway to further improvement

In acknowledging that the behaviours of others influence our actions, the psychology literature has progressively argued for a more fluid, contingent, and socially constructed conception of identity (Rosenau, 1991; Huddy, 2001). 'Knights' or 'knaves' do not arrive in the public service as fixed identities, but are influenced by the social norms of the groups with which they identify. The military group exemplifies 'trust and altruism' by the ways they risk their lives and assume complete trust in other members. But Akerlof and Kranton (2010) use identity economics to show the importance of designing systems of selection, initiation and discipline, which is why for the military, the decision on whether or not to break a disciplinary rule is based on the fear of falling short in the eyes of one's buddies; and not, as in standard economics on assessing the expected costs and benefits (p 47). Hence 'trust and altruism' ought to be sought as an objective of a

model of governance rather than assumed. The fundamental flaw in allowing, or even rewarding, the worst and most flagrant misbehaviour, is that this can come to define the identity of the group (see Powell, 1985). This is why it is so vital to set standards of acceptable behaviour for a group's membership. For example, a headteacher recognised that he had succeeded in reforming the 'shocking' state of Baldwin Elementary school (in a blighted area in New Haven, Connecticut) when he saw a student stop a fight with the words: 'We don't do that in this school' (Akerlof and Kranton, p 72). From this perspective, the shaming of 'knaves' should not only be thought of as a deterrent to other potential knaves, but also as something that strengthens the knightly identity of the remaining members of the group.

An effective model of governance that aims to match the complexity of public services must learn from the failings of simple systems, which, as a consequence of using measures of performance, result in loss of morale and gaming. Indeed, the working assumption of our model based on reciprocal altruism is that most of those who choose to work in public services are 'knights': they have an intrinsic motivation or a mission (Besley and Ghatak, 2003); they see their jobs as a calling (Perry and Hondeghem, 2008). The core of their identity is providing a good service and being esteemed for this by those they serve and their professional peers. The economics of identity suggest a contingent and contextual use for performance incentives, where performance is influenced by the way that good and bad performance are recognised. We describe performance measurement as a contextual solution because of the relative importance of rewards and sanctions based on existing levels of performance. In a poor performing system, where

'knavish' behaviour has not been corrected, reputational sanctions ('naming and shaming') drive the poorest performing public servants to move closer to their peers (through 'loss aversion'). At its most extreme, this reframes the meaning of threats of the sack from a severe kind of sanction for poor performance, to a system of defining identity through minimally acceptable standards for providers of public services. However, as Oliver (2017) argues, 'naming and shaming' ought to be reserved for only the minority of 'knaves' who lack the intrinsic motivation to deliver acceptable levels of performance. Furthermore, it should be combined with an independent auditor tasked with countering the 'knavish' responses of gaming that objective. In higher performing systems, where risk may be more tolerable if it is associated with the prospect of innovation, it may be appropriate to emphasise reputational rewards ('naming and faming'). In either case, that 'knightly' behaviour can be encouraged by creating an identity for public servants where high performance is desirable and poor performance is unacceptable.

References

- Akerlof, G. A. & Kranton, R. E. (2010). *Identity Economics: How Identities Shape Our Work, Wages, and Well Being*. Woodstock: Princeton University Press.
- Anonymous (2017) The National Health Service. Policy Transplant. *Economist*, 4 November, 56-57.
- Arnaboldi, M., Lapsley, I. & Steccolini, I. (2015). Performance management in the public sector: The ultimate challenge. *Financial Accountability & Management*, 31(1), 1-22.
- Ashton, T. (2002) Running on the spot: lessons from a decade of health reform in New Zealand. *Applied Health Economics and Health Policy*, 1, 97-106.
- Ashton, T. (2005) Recent developments in the funding and organisation of the New Zealand health system. *Australia and New Zealand Health Policy* 2, 9. doi: 10.1186/1743-8462-2-9
- Ashton, T., Mays, N. & Devlin, N. (2005). Continuity through change: the rhetoric and reality of health reform in New Zealand, *Social Science & Medicine*, 61(2), 253-262.
- Audit Commission for Local Authorities and the National Health Service in England and Wales. (1996). *What the doctor ordered: a study of GP fundholders in England and Wales*. London: HMSO.
- Auditor General for Wales. (2005). NHS waiting times in Wales. Volume 1 – The Scale of the problem and Volume 2 - Tackling the problem. Cardiff: The

Stationery Office.

<http://www.wao.gov.uk/reportsandpublications/2005.asp>.

Auditor General for Wales. (2006). *Ambulance Services in Wales*. Cardiff: The

Stationery Office.

http://www.wao.gov.uk/assets/englishdocuments/Ambulance_Inquiry.pdf.

Axelrod, R. (1981). *The evolution of cooperation*. New York: Basic Books.

[http://www.sfs.uni-tuebingen.de/~roland/Literature/Axelrod\(81\)_the_evolution_of_cooperation.pdf](http://www.sfs.uni-tuebingen.de/~roland/Literature/Axelrod(81)_the_evolution_of_cooperation.pdf)

Barber, M. (2007). *Instruction to deliver*. London: Methuen.

Bartlett, W., Roberts, J. & Le Grand, J. (1998). *A Revolution in Social Policy:*

Lessons from Developments of Quasi-markets in the 1990s. Bristol: Policy Press.

BBC News. (1999). NUT's call for new education policy. Retrieved from

<http://news.bbc.co.uk/1/hi/wales/579734.stm>

BBC News. (2007). Blair: In his own words. Retrieved from

http://news.bbc.co.uk/1/hi/uk_politics/3750847.stm

Besley, T. & Ghatak, M. (2003). Incentives, choice, and accountability in the provision of public services. *Oxford Review of Economic Policy*, 19(2), 235-249.

Bevan G. (2014). The impacts of asymmetric devolution on health care in the four countries of the UK.

https://www.researchgate.net/profile/Gwyn_Bevan/publication/267864997_The_impacts_of_asymmetric_devolution_on_health_care_in_the_four_countries_of_the_UK_Research_report/links/545b841d0cf249070a7a714f.pdf

- Bevan, G. (2011). Regulation and system management. *Understanding new labour's market reforms of the English NHS*. London: The King's Fund.
- Bevan, G. & Cornwell, J. (2006). Structure and logic of regulation and governance of quality of health care: was OFSTED a model for the Commission for Health Improvement?. *Health Economics, Policy and Law*, 1(4), 343-370.
- Bevan, G. & Fasolo, B. (2013). 'Models of Governance of Public Services: Empirical and Behavioural Analysis of Econs and Humans'. In: Oliver, A. (Ed.) *Behavioural Public Policy*. Cambridge: Cambridge University Press.
- Bevan, G. & Hamblin, R. (2009). Hitting and missing targets by ambulance services for emergency calls: effects of different systems of performance measurement within the UK, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(1), 161-190.
- Bevan, G. & Hood, C. (2006). What's measured is what matters: targets and gaming in the English public health care system. *Public Administration*, 84(3), 517-538.
- 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 & Management*, 33(4). 245-252.

Bevan, G., Karanikolos, M., Exley, J., Nolte, E., Connolly, S., & Mays, N. (2014). The four health systems of the United Kingdom: how do they compare?.

London: Nuffield Trust.

Borren, P. & Maynard, A. (1993). Searching for the Holy Grail in the Antipodes: the market reform of the New Zealand health care system. University of York: Centre for Health Economics.

Bowles S. (2016). The Moral Economy: Why Good Incentives Are No Substitute for Good Citizens. London: Yale University Press.

Burgess, S., Wilson, D. & 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.

Chowdry, H., Muriel, A. & Sibieta, L. (2008). Level playing field? The implications of school funding. CfBT Education Trust.

<http://s3.amazonaws.com/zanran_storage/www.cfbt.com/ContentPages/43889325.pdf>

Cumming, J. & Salmond, G. (1998). Reforming New Zealand health care. In: Ranade W (Ed.) *Markets and Health Care*. London and New York: Longman, 122-146.

Easton, B. (2002). The New Zealand health reforms of the 1990s in context. *Applied Health Economics and Health Policy*, 1, 107-112.

Fiske, E. & Ladd, H. (2000) *When schools compete: a cautionary tale*. Washington: Brookings Institute.

- Frey, B. (2013). How should people be rewarded for their work? In: Oliver, A. (Ed.) Behavioural Public Policy. Cambridge: Cambridge University Press.
- Gordon, L. (1994). Rich and poor schools in Aotearoa/New Zealand. *New Zealand Journal of Educational Studies*, 29(2), 113-126.
- Gordon, L. (2003). School choice and the social market in New Zealand: Education reform in an era of increasing inequality. *International Studies in Sociology of Education*, 13(1). 17-34.
- Ham, C. (2008). World class commissioning: a health policy chimera?, *Journal of Health Services Research & Policy*, 13(2), 116-121.
- Hibbard, J. H. (2008). What Can We Say about the Impact of Public Reporting? Inconsistent Execution Yields Variable Results. *Annals of Internal Medicine*, 148(2). 160-161.
- Hibbard, J.H., Stockard, J., & Tusler, M. (2003). Does publicizing hospital performance stimulate quality improvement efforts?, *Health Affairs*, 22(2), 84-94.
- Hirsch, F. (2005). *Social limits to growth*. London: Routledge.
- Holmstrom, B. & Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, 7, 24-52.
- Hood, C. (1995). The "New Public Management" in the 1980s: variations on a theme. *Accounting, organizations and society*, 20(2-3), 93-109.

- Hood, C. & Dixon, R. (2010). The political payoff from performance target systems: no-brainer or no-gainer?. *Journal of Public Administration Research and Theory*, 20(2), 281-298.
- Hornblow, A. (1997). New Zealand's health reforms: a clash of cultures. *BMJ*, 314, 1892-1894.
- Huddy, L. (2001). From social to political identity: A critical examination of social identity theory. *Political psychology*, 22(1), 127-156.
- Hughes, D., Lauder, H., Watson, S., Hamlin, J. & Simiyu, I. (1996) Markets in Education: testing the polarisation thesis, The Smithfield Project Phase Two: Fourth Report to the Ministry of Education (New Zealand).
- Jensen, K., Scott, R., Slocombe, L., Boyd, R., & Cowey, L. (2014). The management and organisational challenges of more joined-up government: New Zealand's Better Public Services reforms. *Wellington, New Zealand Government*.
- Kahan, J. P. (1974). Rationality, the prisoner's dilemma, and population. *Journal of Social Issues*, 30(4), 189-210.
- Kelman, S. & Friedman, J. N. (2009). Performance improvement and performance dysfunction: an empirical examination of distortionary impacts of the emergency room wait-time target in the English National Health Service. *Journal of Public Administration Research and Theory*, 19(4), 917-946....
- Lauder, H., Hughes, D. & Watson, S. (1999). The introduction of educational markets in New Zealand: Questions and consequences. *New Zealand Journal of Educational Studies*, 34(1), 86-98.

- Laverty, A. A., Smith, P. C., Pape, U. J., Mears, A., Wachter, R. M. & Millett, C. (2012). High-profile investigations into hospital safety problems in England did not prompt patients to switch providers. *Health Affairs*, 31(3), 593-601.
- Le Grand J., Mays, N. & Mulligan, J. A. (1998). *Learning from the NHS internal market*. London: King's Fund.
- Le Grand, J. (2003). *Motivation, agency and public policy: Of 'knights' and knaves, pawns and queens*. Oxford: Oxford University Press.
- Le Grand, J. (2007). *The other invisible hand: Delivering public services through choice and competition*. Woodstock: Princeton University Press.
- Lodge, M. & Gill, D. (2011). Toward a New Era of Administrative Reform? The Myth of Post-NPM in New Zealand. *Governance*, 24(1). 141-166.
- Maarse, H., Jeurissen, P. & Ruwaard, D. (2016). Results of the market-oriented reform in the Netherlands: a review, *Health Economics, Policy and Law*, 11(2), 161-78.
- Mackenzie, N. (2007). Teaching excellence awards: an apple for the teacher?. *Australian Journal of Education*, 51(2), 190-204.
- Miller, E. F. & Hume, D. (1994). *Essays: Moral, Political, and Literary*. Indianapolis: Liberty Fund.
- Musset, P. (2012). *School Choice and Equity: Current Policies in OECD Countries and a Literature Review*, OECD Education Working Papers, No. 66, OECD Publishing, Paris. <http://dx.doi.org/10.1787/5k9fq23507vc-en>

Nash, R. (1999). Social capital, class identity, and progress at school: case studies.

New Zealand Journal of Educational Studies, 34(2), 267-280.

The NHS Long Term Plan <https://www.longtermplan.nhs.uk/wp>

content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf

OECD. (2015). Programme for International Pupil Assessment (PISA). Country

Note: United Kingdom [http://www.oecd.org/pisa/PISA-2015-United-](http://www.oecd.org/pisa/PISA-2015-United-Kingdom.pdf)

Kingdom.pdf

OECD. (2000). OECD Economic Surveys: United Kingdom, pp 123-146.

<https://www.keepeek.com//Digital-Asset->

Management/oecd/economics/oecd-economic-surveys-united-kingdom-

2000_eco_surveys-gbr-2000-en#page121 OECD (2015) Programme for

International Pupil Assessment (PISA). Country Note: United Kingdom

<http://www.oecd.org/pisa/PISA-2015-United-Kingdom.pdf>

Oliver, A. (2017). Do Unto Others: On the Importance of Reciprocity in Public

Administration, *American Review of Public Administration*.

<http://journals.sagepub.com/doi/10.1177/0275074016686826>

Perry, J. L. & Hondeghem, A. (Eds.). (2008). *Motivation in public management:*

The call of public service. Oxford: Oxford University Press.

Powell, A. G. (1985). The Shopping Mall High School. *Winners and Losers in the*

Educational Marketplace. *NASSP Bulletin*, 69(483), 40-51.

Power, M. (1999). *The Audit Society: Rituals of Verification.* Oxford: Oxford

University Press.

- Propper, C., Sutton, M., Whitnall, C. & Windmeijer, F. (2010). Incentives and targets in hospital care: evidence from a natural experiment. *Journal of Public Economics*, 94(3-4), 318-335.
- Reynolds, D. (2008). New Labour, education and Wales: the devolution decade, *Oxford Review of Education*, 34, 6, 753-765.
- Rosenau, P. M. (1991). *Post-modernism and the social sciences: Insights, inroads, and intrusions*. Princeton University Press.
- Rosenblatt, M. (2011). The use of innovation awards in the public sector: Individual and organizational perspectives. *Innovation*, 13(2), 207-219.
- Royce, R. G. (1995). Observations on the NHS internal market: will the dodo get the last laugh?, *BMJ*, 311: 431.
- Sahlberg, P. (2010). The secret to Finland's success: Educating teachers. Stanford Center for Opportunity Policy in Education, 2, 1-8.
<<http://www.nnstoy.org/download/preparation/Secret%20to%20Finland's%20Success%20-%20Education%20Teachers.pdf>>
- Schick, A. (1996). The Spirit of reform: managing the New Zealand State Sector in a time of change. State Services Commission of New Zealand, Wellington.
- Scott, G. (2001). Public Sector Management in New Zealand: lessons and challenges. Wellington: Australian National University.
- Scott, R. J. (2019) *Public service motivation and social identity*. Wellington: International Research Society of Public Management.
- Scott, R. J. & Boyd, R. (2016). Results, targets and measures to drive collaboration: Lessons from the New Zealand Better Public Services

reforms. In: Butcher, J. R., & Gilchrist D. J., (Eds.) *The three sector solution: Delivering public policy in collaboration with not-for-profits and business.* Canberra: Australian National University Press.

Scott, R. J. & Boyd, R. (2017a) *Interagency performance targets: a case study of the New Zealand Results programme.* Washington: IBM Business of Government.

Scott, R. J., & Boyd, R. (2017b). *Joined-up for what? Response to Carey and Harris on adaptive collaboration. Australian Journal of Public Administration, 76(1), 138-144.*

Scott, R. J., & Boyd, R. (2020). *Determined to succeed: Can goal commitment sustain interagency collaboration?. Public Policy and Administration, 0952076720905002.*

Scott, R., & Boyd, R. (2022). *Targeting Commitment: Interagency Performance in New Zealand.* Brookings Institution Press.

Scott, R. J., & Bardach, E. (2019). *A comparison of management adaptations for joined-up government: Lessons from New Zealand. Australian Journal of Public Administration, 78(2), 191-212.*

Scott, R. J., & Macaulay, M. (2020). *Making sense of New Zealand's 'spirit of service': social identity and the civil service. Public Money & Management, 40(8), 579-588.*

Scott, R. J., & Merton, E. R. (2021). *When the going gets tough, the goal-committed get going: overcoming the transaction costs of inter-agency collaborative governance. Public Management Review, 1-24.*

- Smee, C. (2008) 'Speaking truth to power', in Timmins N ed (2008) *Rejuvenate or Retire? Views of the NHS at 60*. Nuffield Trust.
- Smith, A. (1759). *Theory of Moral Sentiments*. Strand & Edinburgh: A. Millar; A. Kincaid & J. Bell.
- Smith, J. & Curry, N. (2011). 'Commissioning'. In Dixon A., Mays N. and Jones L. (Eds) *Understanding New Labour's market reforms of the English NHS*. London: King's Fund: 30-51.
<http://www.kingsfund.org.uk/sites/files/kf/chapter-3-commissioning-new-labours-market-reforms-sept11.pdf>
- Smullyan, R. (1978). *What is the Name of this Book?*. Prentice-Hall.
- State Services Commission, (2011). *Better Public Services Advisory Group Report*. Wellington: New Zealand Government. Accessed at:
http://www.ssc.govt.nz/sites/all/files/bps-report-nov2011_0.pdf
- State Services Commission, (2012). *Better Public Services Results: Targets and Public Communication*. CAB (12) 315. Wellington: New Zealand Government.
- Stipek, D. J. (1993). *Motivation to learn: from theory to practice*. Boston: Allyn and Bacon.
- Stubbs, T. & Strathdee R. (2012) *Markets in education: the impact of school choice policies in one market context in New Zealand*. *International Studies in Sociology of Education*, 22(2). 97-124
- Timmins, N. (2012). *Never Again?: The Story of the Health and Social Care Act 2012: a Study in Coalition Government and Policy Making*. London:

- Institute for Government and the King's Fund.
https://www.kingsfund.org.uk/sites/default/files/field/field_publication_file/never-again-story-health-social-care-nicholas-timmins-jul12.pdf
- Tversky, A. & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model, *The quarterly journal of economics*, 106(4), 1039-1061.
- Walford, G. (2014). From city technology colleges to free schools: sponsoring new schools in England. *Research papers in education*, 29(3), 315-329.
- Waslander, S. & Thrupp, M. (1995). Choice, competition and segregation: an empirical analysis of a New Zealand secondary school market, 1990-93. *Journal of Education Policy*, 10(1), 1-26.
- West, A. (2010). High stakes testing, accountability, incentives and consequences in English schools. *Policy & politics*, 38(1), 23-39.
- West, A., Ingram, D. & Hind, A. (2006). "Skimming the Cream" Admissions to Charter Schools in the United States and to Autonomous Schools in England. *Educational Policy*, 20(4), 615-639.
- Whitty, G., Halpin, D. & Power, S. (1998). Self-managing schools in the marketplace: the experience of England, the USA and New Zealand. In Legrand, J., Bartlett, W., & Roberts, J. (eds) (1998). *A revolution in social policy: quasi-market reforms in the 1990s*. Bristol: Policy Press.
- Wigmore, T. (2016). Why Sweden's free schools are failing. *New Statesman* 18 June 2016...
- Williamson, O. E. (1985). *The Economic Institutions of Capitalism: firms, markets, relational contracting*. London: Macmillan

Wilson, D. S. (2015). Does altruism exist?: culture, genes, and the welfare of others. London: Yale University Press.

Wyke, S., Mays, N., Street, A., Bevan, G., McLeod, H. & Goodwin, N. (2003). Should general practitioners purchase health care for their patients? The total purchasing experiment in Britain. *Health policy*, 65(3), 243-259.

Tables and Figures

1. Reduce the number of people who have been on a working age benefit for more than 12 months.
2. Increase participation in early childhood education.
3. Increase infant immunisation rates and reduce the incidence of rheumatic fever.
4. Reduce the number of assaults on children.
5. Increase the proportion of 18-year-olds with National Certificate of Educational Achievement (NCEA) level 2 or equivalent qualification.
6. Increase the proportion of 25 to 34-year-olds with advanced trade qualifications, diplomas and degrees (at level 4 or above).
7. Reduce the rates of total crime, violent crime and youth crime.
8. Reduce reoffending.
9. New Zealand businesses have a one-stop online shop for all government advice and support they need to run and grow their business.
10. New Zealanders can complete their transactions with the Government easily in a digital environment.

Table 1: The ten BPSR objectives

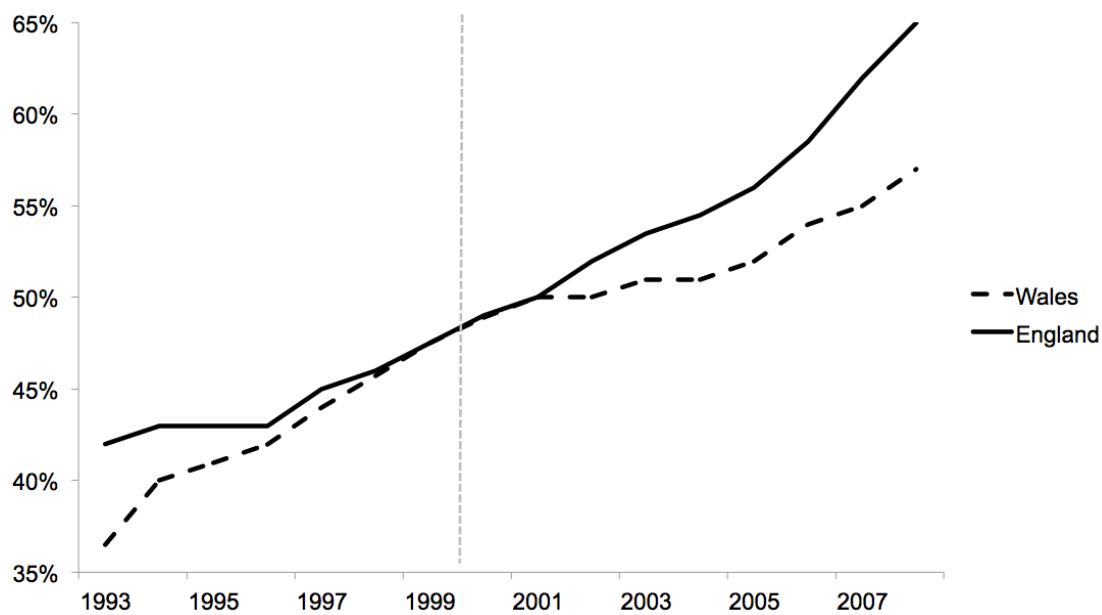


Figure 1: Education results in England vs. Wales
 (% > 5 good grades at age 16; note publication of school league tables stops in Wales in 2000, indicated by vertical line. Source Office for National Statistics, *Regional Trends for each year*)

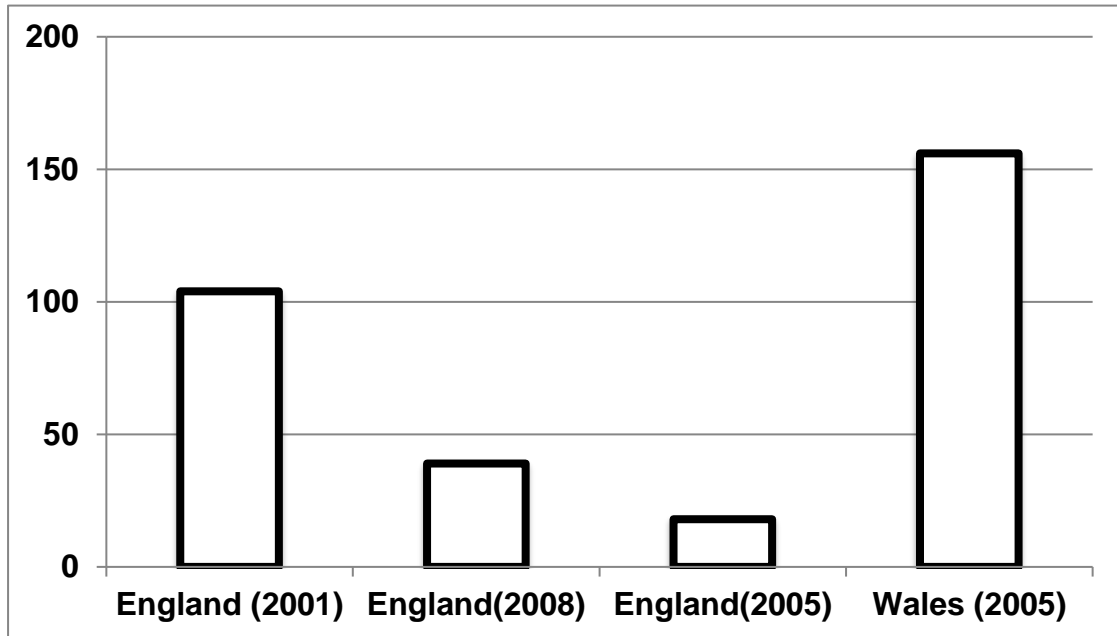


Figure 2: Health results in England vs. Wales
(weeks hospital waiting times; source Auditor General for Wales (2005))

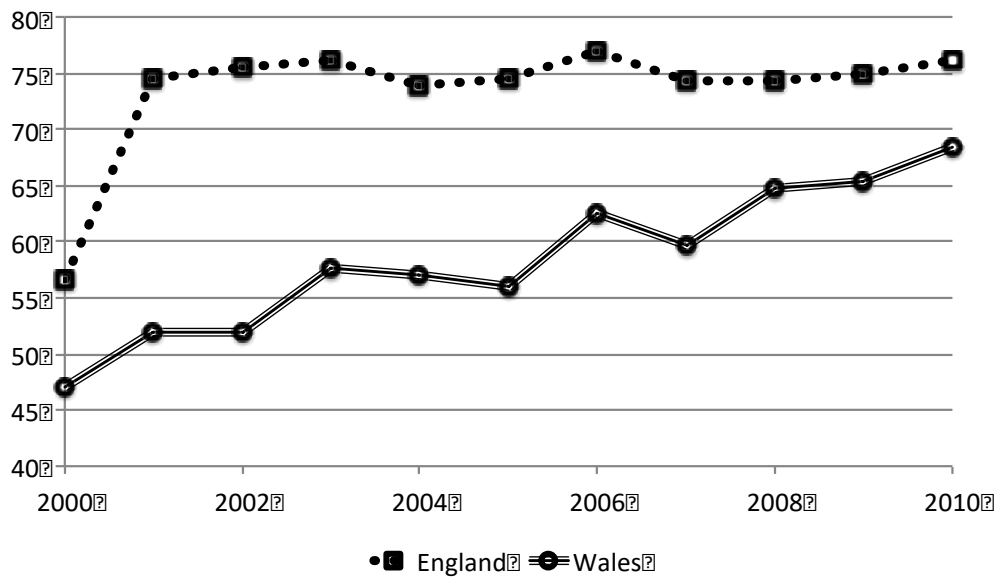


Figure 3: Health results in England vs. Wales
(percentages of ambulance response times to potentially immediately life threatening emergencies (Category A calls) within eight minutes; sources NHS Digital for England and Health Statistics Wales)

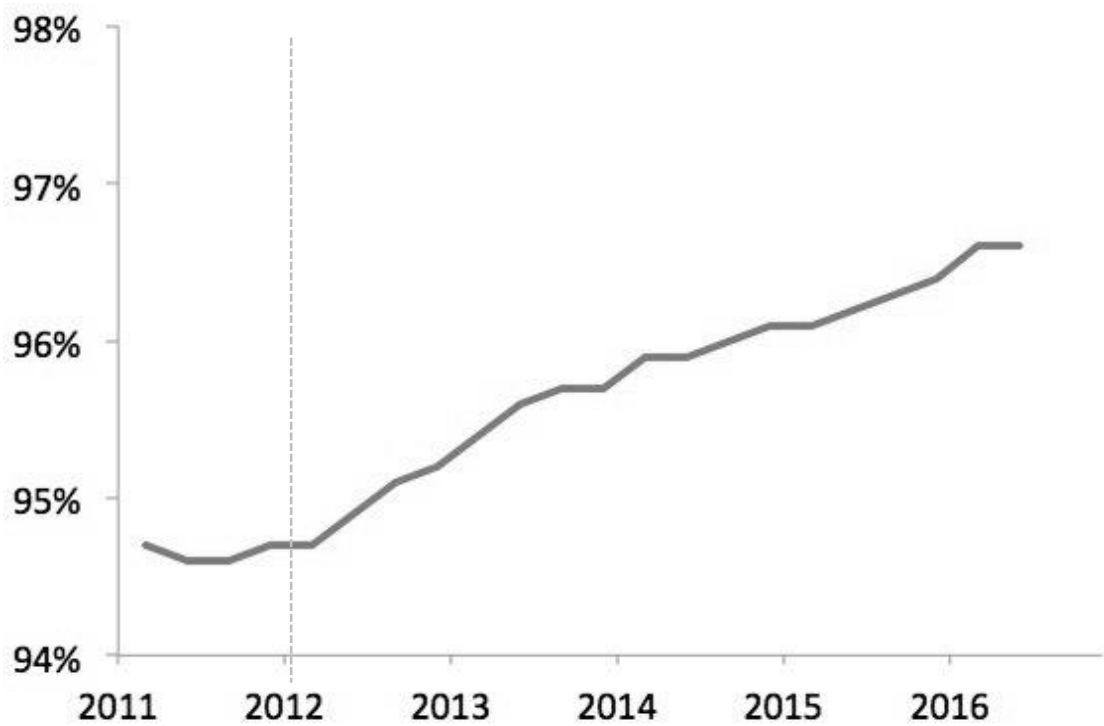


Figure 4: Education in New Zealand
(percentage of children aged 2 or 3 enrolled in early childhood education; note education target introduced in 2012, indicated by vertical line)

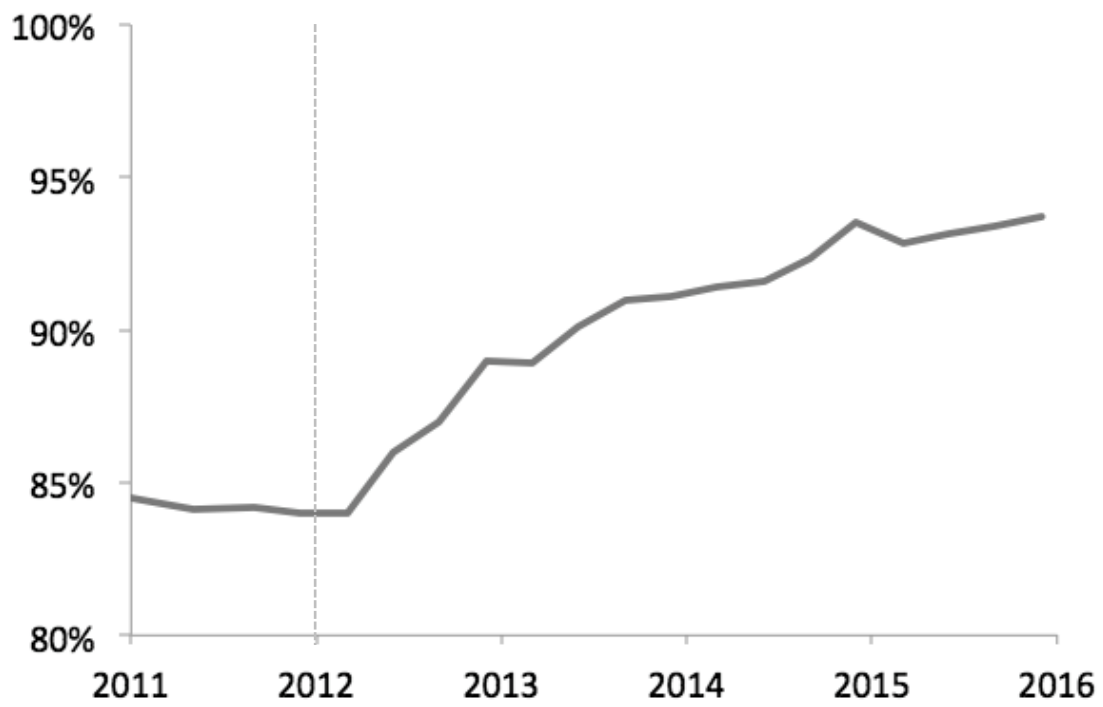


Figure 5: Health in New Zealand
(percentage of infants receiving full vaccination schedule by 8 months; note immunization target introduced in 2012, indicated by vertical line)