Autonomy and Improved Performance:
Lessons from an NHS Policy Reform

Paul Anand, The Open University and
Health Economics Research Centre, Oxford University
Mark Exworthy, Royal Holloway University
Francesca Frosini, The Kings Fund
Lorelei Jones, London School of Tropical Medicine and Hygiene

June 2010

Acknowledgements
The authors are particularly grateful to the NHS Service Delivery Organisation for funding this research (http://www.sdo.nihr.ac.uk/projdetails.php?ref=08-1618-125) and to those who took part in the interviews.
Autonomy and Improved Performance: Lessons from an NHS Policy Reform

Autonomy is currently seen by policy-makers as a possible mechanism for enhancing public sector performance and in this paper we examine a health-service reform (in England) in which more autonomy is given to better performing hospitals. Drawing on data from interviews with senior managers in the NHS, our research suggests that despite being enmeshed in a politicised culture of regulations and guidance, autonomy is increasingly perceived positively and appears to depend on the extent to which organisations have the incentives and the capacity to respond to increased autonomy.

1. Introduction

Recent political interest in organisational freedom and privatisation follows a long-standing debate about decentralisation within public sector around the world which has started to give way to policies and language that emphasise the development of autonomous organisations within the public sector. Indeed a key aspect of the New Public Management (NPM) paradigm, Barzelay (2001) argues, which has been particularly influential in the OECD is the establishment of appropriate relations between the centre and subordinate agencies, even if central governments find the balance between political control and local managerial autonomy difficult to get right (Laegreid et al (2008)). In this paper, we examine an example of autonomy in the National Health Service in England with a view to developing a theoretically informed description of a significant policy experiment. Our main aim is to develop findings of value particularly to policy-makers.

In what follows, we discuss aspects of theories and frameworks that have been used by social scientists to understand the relationship between autonomy and performance. We then discuss in more detail these issues as they play out in the NHS before moving on to presentation of our evidence and interpretation. In a concluding discussion, we emphasise the distinction between formal and practical autonomy and highlight two intervening factors,

---

1 More recent, related discussions can be found in Peckham et al (2005), Mannion et al (2005) and Yamamoto (2006). See also a general literature on performance measurement in health e.g. Smith et al. (2010).

2 For our purposes, we take hospital autonomy to be enhanced by the creation of foundation status and the wider range of actions the status makes possible. At the same time, we note that some observers (e.g. Bruce and McConnell (1995)) have argued that the NHS has been subject to a ‘hollowing out’ process by which local accountability has declined whilst governance by central authorities, has effectively been on the rise.
Incentives and capacity, that appear to mediate the relationship between autonomy and performance.

2. Autonomy and its Rationale

*Autonomy and Decentralisation in Organisational Analysis*

In the UK, public sector management has experienced waves of change that have focussed on internal reform, privatisation, regulation and now an emphasis on autonomy for both providers and patients or citizens (Ferlie et al, 2005). It could be claimed that autonomy is just the latest fashionable concept in a long public sector history of decentralisation (Exworthy and Frosini, 2008). Furthermore, some might argue that decentralisation, though a perennial theme in public sector reform has recently been revived because of the inappropriateness of private sector solutions to public sector management problems (Saltman et al, 2007). For those public services not amenable politically or technically to privatisation, decentralisation could be a solution. Alternatively, others see privatisation as one form of decentralisation – i.e. a transfer of ownership and control not just to lower organisational levels but to independent providers (Rondinelli, 1983). Autonomy, combined with some form of regulatory oversight enables governments to pursue efficiency in various forms and represents a significant movement from the earlier approaches to public sector reform, inspired by neo-classical economics, which tended to emphasise the extent rather than the quality of government action. In the NHS policy experiment that we study here, autonomy has been linked, possibly for the first time, *conditionally*, and therefore as incentive, to good organisational performance.

*Justifications for Autonomy*

*Economic* justifications for autonomy within the public sector can be read as a relatively new application of older principles. Relevant theories tend to focus on the rationality of actor behaviour and the efficiency of the resulting outcomes. The *neo-classical theory* of production and markets, for example, specifies conditions under which inputs are used and outputs produced in ways that are optimal with respect to preferences and it is assumed, somewhat loosely though with some justification from the historical and international experience of central planning, that government or other centralised interventions impose constraints which prevent agencies from achieving optimal outcomes (such as efficiency or innovation). A complementary, and widely used (Propper (1995)), approach is that provided by the *principal-agent analysis* in which there is an emphasis on the reasons why bureaucrats might not do exactly what citizens or politicians would wish. This analysis has encouraged the development of proposals that often amount to the sharpening of economic incentives, for
example by the introduction of performance related pay or the promotion of inter-organisational competition.

*Political* justifications for autonomy are not un-related. A key trade-off for political actors in institutional design is the desire to steer agents of the state towards the production of *politically desired outcomes* whilst at the same time stepping back from responsibility (particularly blame) for operational issues which they cannot reasonably control in much detail and which might be largely instrumental in any case. This may be a tacit recognition that the centre never did or could ‘command and control’ but rather ‘exhorts and encourages’ (Exworthy et al, 1999). So from a political and managerial perspective, autonomy is often aimed at inputs and processes whilst offering the prospect of some outcome control. Moreover, decentralisation linked to autonomy offers the possibility of making more explicit tensions within policy: for example, it may enable centralisation of credit and decentralisation of blame (divide and rule) thereby diffusing blame of systemic issues.

Not all explanations of autonomy are based on rational or efficiency considerations – for example some *organisational and management* researchers working on public sector reform emphasise the flow of ideas and the fact it is possible for ideas to be adopted simply because they are fashionable (see, for instance, Marmor (2004) and Pollitt et al (2004) for discussions of different strands of theorising).

Within functionally oriented management research, the divide between strategic and operational functions is often cited as a reason for decentralisation – the centre retains control of strategic issues whilst newly decentralised units focus more on operational issues. Put another way, the centre determines the ends whilst the locality focuses on the means - an idea aligned with the loose-tight organisational model and steering-rowing analogy (Osbourne and Gaebler, 1992). The strategic-operational distinction is often hard to maintain in public services: see for instance the experience of the UK’s Prison Service in the 1990s and the role of its political/strategic leader Michael Howard and the managerial/operational leader Derek Lewis in target setting, or more theoretically, the street level bureaucracy (Lipsky, 1980) whereby practitioners’ daily decisions become the *de facto* policy of the organisation *de facto* autonomy has long/always existed in professional-dominated organisations by virtue of their professional (here, clinical) ‘autonomy’ (about whether/how to diagnose, treat and/or refer). Local autonomy thus pre-exists government initiatives to decentralise and offers *de jure* autonomy and so, *de jure* autonomy is a chimera.
3. Decentralisation and Autonomy in the NHS

Within the UK’s NHS, the internal (quasi-) market for health (first introduced in 1991) was an important experiment in which many decisions in the health service were supposed to be given over to local decision-makers. With the re-invention of the internal market (since 1997), there has, nonetheless, been a recognition that the NHS is over-centralised and over-politicised (Kings Fund 2002), a view that has encouraged the development of policies aimed at ‘increasing freedom from central control, patient empowerment and clinical empowerment’ (Peckham et al, 2008 p.572). Exworthy and Greener (2008) note that two incidents prompted a change in policy around 2000/2001 towards decentralisation – Bedford hospital mortuary scandal and experiences of first term in office (hence, the call for “delivery, delivery, delivery”). Many of the market concepts to do with freedom and choice have, as a result, carried over into the agenda of subsequent political regimes even if the language has changed. Within this new policy regime where autonomy and choice are central, both for providers and patients, the NHS set in train a set of policies (NHS 2000) that would lead to the creation of Foundation Trust hospitals, organisations whose performance justified operation on looser reins.

The policy initially began (DOH 2000) in terms of earned autonomy for high performing Trusts but has expanded into Foundation Trust policy which, in time, is expected to cover all NHS Trusts (DOH, 2002). However, if all Trusts become FT, the issue of performance still remains (see above) unless the government is satisfied that overall performance is above a threshold / minimum.

The implementation of autonomy, conditional upon centrally-determined performance measures, in the NHS adds a twist to the justifications of autonomy discussed in section 2. By offering hospitals (and other health-care providers such as Mental Health Trusts) that perform “well” the opportunity to become more autonomous, central government effectively gives scope for the logic of growth to apply within the public sector as autonomous hospitals are allowed to expand their activities more freely. The policy of autonomy also benefits from any in-built selection bias. If good performance is not random year on year but tends to persist, then those hospitals that take up the offer of greater of autonomy will be better than average and so any simple comparisons between autonomous and non-autonomous hospitals will tend to favour the policy of autonomy. (This issue is very similar to the policy of GP-fund holding according to which better performing group practices were in effect selected to be fund-holders thereby encouraging a conclusion that funding-holding was itself desirable (Robinson and Le Grand, 1994)). There may also be a political benefit in that the earning of autonomy
could place more responsibility for autonomous status on hospitals themselves, thereby giving rise to fewer objections from the service itself.

The alignment of autonomy with performance is significant because it is premised on an apparent clear definition and expression of performance. Performance is problematised and thus becomes a ‘problem’ from which escape is impossible – there will always be a distribution of performance where half of all Trusts are below the modal average! There is also an ever present danger that performance will be equated with easily measured dimensions neglecting other (qualitative) measures. The balanced scorecard (such as the Healthcare Commission’s annual health check) might partly address that but the ease of measurability issue, interpretation and relative priorities of the components remain (Bevan and Hood, 2006).

4. Autonomy in Health? A Case-Study
To explore empirically some of these theoretical issues about autonomy, performance and the connection between the two, we draw on a qualitative study of two local health economies in England. These case studies were carried out at a time when the Foundation Trust programme was, for some institutions, beginning to bed down and is based on a series of 52 interviews with senior clinicians and managers conducted over the period 2006-9. In this section, we explore some of the perceptions around autonomy and its implications for performance in a manner that is consistent with the autonomy-performance matrix discussed above whilst recognising that some interpretation is inevitable given that the language used by managers differs from the analytical categories used by academics. Whilst the study encompassed interviews with a wide range of role-holders, what follows highlights comments particularly relevant to autonomy as embodied in the Foundation Trust hospital status.

Overall, and despite questions about a high degree of vertical accountability to the centre, it has to be said that our interviews indicate that Foundation Trust hospitals are in some important respects more autonomous than counterparts without FT status and that this does provide the opportunity to develop and improve health service delivery for those who want it. As one Manager (in an FT) said:

“...I’m quite excited about the stuff we do with the other... Trusts and we’re already starting to do with the PCT – well some of the GPs and – there’s quite a lot of manoeuvre if you want to redesign the service – if you’ve the energy and the leadership and a bit of business…”
In addition, we found evidence that managers in hospitals without FT status feel quite constrained and see the Foundation Trust hospital as something that would help improve their situation:

“… we don’t actually feel that we have space at the moment as an organisation, we’re still very pushed, very drained…but we feel that one way out of this is to become a Foundation Trust...” (Trust Director in a non-FT)

Without necessarily making clear distinctions between different aspects of autonomy, these quotes reflect a view that the increased autonomy given to hospitals is tangible. However, the existence of increased autonomy does not imply that it will be exercised, particularly if Foundation Trust managers feel their that their exposure means that risks are borne more directly by individual managers and organisations rather than being diffused within a hierarchy of command and control. For example one Chief Executive summed the situation up as follows:

…Well we’re on our own, so if we get it wrong, there’s no one to bale us out...

The same CE went on to suggest that people felt rewarded if they were able to share in any surplus made but it seemed likely that this was a matter of pride or benefit to the system rather than any personal gain given the small size of the personal income bonuses being used.3

Waiting times and targets for their reduction have become something of an intractable issue and they illustrate how service quality can be affected by autonomy (or otherwise). In some, there was a recognition that national level policies could be enabling – one manager was able to introduce a stroke unit, something he had wanted to do ‘for years’ because there was a national framework that required it. This framework does, however, imply a value in centralism (rather than autonomy). Similarly another General Manager said to us:

“...I don’t think we would have probably ever cracked the A&E problem if we hadn’t had the four hour target around it for example but those targets bring along problems of their own...”

But there is also evidence of a perennial concern about the extent to which targets gain traction within the organisation and/or are gameable (Propper et al, 2007). At least one manager suggested that A and E targets could be achieved by careful, but honest, work on the

3 We encountered one FT hospital offering income bonuses for organisational performance and for most employees these were no more than £100 per year.
data, a point that highlights the need to recognise the status of performance data when based on self-reporting. Similarly, there was some questioning about the introduction of legally binding contracts for FTs: given the ever present risk of political intervention if things are not going well within the NHS, what does legally binding mean? In one case, an interviewee offered opinions about guidelines and their connection to performance in a manner that makes such views hard to misinterpret:

“…Inundated with pieces of paper, guidelines, questionnaires, you know, can’t remember all…frameworks, you know, have you got posters up telling people what to do if they want alcohol rehabilitation,…and it takes up vast amounts of my time…and I really think somebody in NICE ought to – they ought to do a NICE guidance on NICE guidance…I don’t think it translates into good patient care..” (attribute anon)

A number of respondents discussed the way in which their organisations had dealt with the prospect of large financial deficits and these interviews paint a slightly more favourable attitude to the autonomy of Foundation Trust status than emerges from the Mannion et al (2005) survey of senior managers. We have little evidence that sheds light on the difference though it could be that the more favourable attitudes of our managers reflected learning and adaptation over time as the policy bedded in.

Limits on the Decision Space
A significant finding is that notwithstanding the economic justifications for autonomy there are two forces that contrive to constrain the extent to which autonomy leads to behavioural change and therefore the possibility of performance improvements. Firstly, on the agency side, even where hospitals have foundation status, there may be weak incentives to make use of the autonomy this implies (a similar issue was found in an evaluation of the 1990s internal market (Le Grand et al, 1998)). The increasingly exposed position that some hospital managers find themselves in appears to have encouraged behaviour that is sometimes risk-averse – counter to the general policy aim of promoting innovation. The key distinction is between FTs’ willingness, and ability, to exercise autonomy. They have the legal ability to exercise discretion but often do not because of on-going centralisation, years of having been inured to centralisation, limited entrepreneurial skills and capacity, fear of destabilizing good local working relations with other agencies etc (Health Select Committee, 2009, Exworthy et al, 2008). Secondly, political actors appear to find it hard to manipulate the way in which

---

4 The key point is that innovation is stymied by a combination of factors: past behaviour, uncertain ‘rules of the game’, greater exposure to financial risk and the desire to accumulate reserves for ‘a rainy day’ or possible capital projects.
they are held accountable. In this case, even if an autonomous hospital is responsible for not providing a service, the inequalities of provision across geographical areas are blamed on central government. This is not unlike the case of the poll-tax in which the prime-minister Thatcher’s desire to attribute local taxes to local authorities was scuppered as people attributed changes in the tax-system to her policies. In a similar vain, recent attempts to reconfigure services in London, Sussex, Surrey and elsewhere have had a high political profile and were subject to decisions that not taken locally (Oborn 2008). The Darzi Review is, in effect, an admission that these are centralised plans for the organisation of services which might otherwise be taken locally.

*Accountability Dynamics*

The second issue also confirms a point made by Hoque et al (2004) namely that the goals externally set by central government are sufficiently numerous that autonomy may have limited bite. Even if politicians in government would like to distance themselves from operational responsibilities, the pressure from their counterparts in opposition and the media is such that they find this difficult to do. It is as if there is a web of accountability from which no one in the system can escape. Indeed if one recalls the taxonomy of organisational autonomies developed by Verhoest et al (2004), it is clear that foundation trust hospitals are still subject to limits on their managerial autonomy, outputs and objectives, finance and legal status. What they call structural autonomy (the appointment of a head) and interventional autonomy (e.g., when performance dips below a certain level), it is clear that in most categories, autonomy is low or minimal. In many cases, central government has a direct role or could constrain the use of a hospital’s autonomous powers. One of the case-studies FT commented to us on the inimical effect that a government requirement for the Trust to employ a certain number of matrons had upon the Trust. Furthermore, a retained surplus might also be subject to raids by governments in time of financial stringency. That said, the establishment of Foundation Trusts has removed many powers previously held by the Secretary of State (cf. self-governing Trust in 1990s).

These latter points raise an important question about the material significance of the theory of incentive alignment which has been used extensively in the new management literature (Le Grand (2003)). Principal agent analysis provides useful guidance when the agent’s objectives diverge from those of the principal but in the case of health-care provision, whether one takes the principal to be citizens or their political representatives, objectives tend to be partial and

---

5 This is not a new concern as the distinction between formal autonomy and freedom on the ground in Tupper and Doern (1981) indicates.
fuzzy (whereas principal-agent analysis assumes these are completely defined\(^6\)). In addition, there are multiple objectives which have to be weighed and traded-off and it is very unlikely that all stakeholders in health provision will weigh and trade off these objectives in an identical fashion. Moreover, one could even argue that the principal-agent analysis breaks down because in reality it is not clear who the principal is: and at least two sets of possible principals (government and citizens) do not have well-defined objectives.

That said, where there are a number of similar organisations, as in the case of hospitals and schools, *benchmarking* competition might be possible (even if direct monitoring of effort is not). For benchmarking to work, however, there needs to be some mechanism for facilitating comparison and in multi-product organisations, such as hospitals, this is difficult as prices are part of the negotiating mix (under the Payment-by-Results policy). In such cases, quantitative, comparative performance measurement (such as star ratings) seems unavoidable.

A somewhat negative finding that might apply to the NHS concerns what Majone (2001) refers to as the logic of delegation. A rationale for delegating decision-making is to reduce its costs and this is consistent with the move to lighter touch regulation for those hospitals that achieve autonomous status. However, so long as some hospitals are not autonomous, central government still needs to retain the capacity to make detailed regulatory decisions at the centre with the result that any decision-cost savings might not be that extensive. But Majone also discusses the need to commit as a rationale for policy and it seems that in this case, government finds it difficult to commit to not directing the actions of ‘autonomous’ hospitals, perhaps because it has not made sufficiently difficult its own ability to direct the actions of apparently autonomous hospitals. Health in the UK at least, is not an area where the ‘hollowing out’ thesis seems to apply (a conclusion that chimes with Jessop (1994)). Despite this process of autonomisation, the state seeks to take ‘more control over less’\(^7\).

It is also worth noting that there is little evidence that, within the local health economies observed, the move to decouple vertical relations from hospitals to the centre has been matched by policies that would enhance horizontal forms of accountability (i.e. with clients, stakeholders or peers (Schilleman 2008)). Foundation Trusts are also significant because they are designed as ‘mutual organisations’, with members taking a greater involvement in their

---

\(^6\) Williamson’s (1975) ‘classical’ minimal role for government included objective setting, evaluation, appointing and rewarding the CEO and providing resources. Our findings confirm his view about the rarity of objective setting. Similarly, Kenneth Arrow (1985) argues that some methods of monitoring bureaucratic activity cannot be analysed with standard principal agent analyses that emphasise hidden information and action.

\(^7\) So perhaps it is fallacious to think that more autonomy will reduce monitoring costs – particularly if the central regulatory function becomes fractionalised.
affairs. Limited evidence suggests that Foundation Trusts have not made great strides in this area, often focusing on those with whom they were already engaged (see Health Select Committee, 2009).

One possible sign of future de-coupling might be greater private sector involvement in commissioning and/or different decentralisation to PCTs (e.g. links with local government). However, so long as the NHS remains centrally funded, it is likely that the government will take a keen interest in local NHS matters. This seems to be a widespread feature of reforms dealing with similar issues in other countries – the price of greater autonomy is often stronger accountability (Yamomoto 2006 p35; Schick, 1996)).

Performance Measurement
Whilst our evidence in this area is limited, comments about the way in which ‘contracting’ parties in the local health economy might contrive to help each other out, particularly when it comes to the 6% rate of return (ROR) target for hospitals (inter alia), suggest a simple but important point. If players see autonomy as new constraint on the way in which they engage over time, then it could be serving as a limit on collaborative activity rather than a minimum threshold on any surplus to be generated. Limiting activity is inevitable and so not undesirable but it suggests that ROR targets may not contribute to the efficiencies (appropriate input mix, lowest cost and adaptive dynamics) as one would expect them to in private sector settings. This in turn suggests that league tables based on objective criteria should be seen as an important component of the performance management regime. However, drawing again on the evidence of our interviews, it seems important that performance management regimes should depend on the objective measurement of these objective criteria. The problem is by no means limited to health but suggests a need to beef up auditability requirements of such data, giving more attention to the standardisation of what is measured, and/or relying on data from alternative sources (such as patient surveys).

Principal Agent Analysis and Preference Matching
When it comes to preference matching, the first kind of efficiency discussed in section 2, there are a variety of processes that contribute to this – ranging from regulatory decisions by NICE at national level through to choices of health service provider made by the patient or GP on the patient’s behalf. It does seem that an important part of the patient’s concern is with non-medical, quality aspects of health-care - where they are treated, how far from home, at what time of day, the flexibility of appointments and so on. For this, if for no other reason, patient choice may play a crucial role in signalling information about personal preferences that would be difficult to capture and analyse usefully at a centralised level. So perhaps, even
if patient choice does not match up to the ideal of fully informed individuals being offered a wide range of alternatives, it is useful to have some choice in the system as a way of giving providers at local level some information about the preferences for the trade-offs that people actually have.

5. Concluding and Summary Discussion

Drawing on the management aspects of this research, in sum, the Foundation Trust experiment in organisational autonomy does seem to have given senior management in hospitals room for manoeuvre, despite being embedded in an increasingly dense mesh of (on the one hand) laws, directives, guidance and policies and (on the other) a local network of alliances and collaborations (Exworthy and Frosini, 2008). Within this dense mesh, there is still room for managers to deliver performance that varies widely. The logic of conditional autonomy means, however, that any simple comparison of performance would be confounded by other factors so it may be quite difficult to produce a quantitative evaluation of the reform. Nonetheless, and combined with a modicum of patient choice, the use of performance league tables and a close performance management regime for hospitals lower down the league tables, autonomy appears capable of making a contribution to the enhancement of performance. As we have seen from the manager statements, this has tended to cash out in personal values that range from excitement through to trepidation.

Adding to this an economics perspective, our empirical evidence suggests that policy-makers persuaded by the logic of autonomy need to realise that incentives and capacity to make use of autonomy need to be present if organisational freedom is to generate changed behaviours. Furthermore, we believe this conclusion is likely to be of importance in many areas of public sector reform around the world where autonomy or decentralisation are seen as possible pathways to enhanced performance. In effect, and given the regulatory environment in which large scale public sector organisations operate, increased autonomy is unlikely to be a sufficient condition for improved performance. Unlike the private sector, where the consequences of autonomy are amplified by the ability of firms to grow and fail, such forces are relatively weak in the public sector where different systems of accountability prevail. Our argument is consistent with, and may help explain recent quantitative findings that find autonomy having limited impacts on performance (e.g. Manion et al (2005) and Allen (2010)). In our incentives-capacity model of the autonomy-performance relationship in the public sector, autonomy needs to be accompanied by suitable rewards, remits, skills development and the genuine granting of freedom from policy directives and other governance constraints, if performance is to be impacted.
References


