The impacts of asymmetric devolution on health care in the four countries of the UK

Gwyn Bevan

September 2014
About this report

This report accompanies The Four Health Systems of the United Kingdom: How do they compare?, which compares the publicly financed health systems of England, Scotland, Wales and Northern Ireland in terms of funding, inputs and performance. This accompanying report looks at how the health care systems in the four countries of the UK have fared before and after devolution.

It is published after the outcome of the Referendum on Independence for Scotland on 18 September 2014, which means that Scotland remains in the UK. This has implications for greater devolution to Scotland and raises questions over constitutional arrangements for the other countries in the UK. In the debate before and after that Referendum, questions have been raised about the devolution settlement: the Barnett formula, which is the basis for the funding of public services; and what is known at the ‘West Lothian question’ – whether MPs from the devolved countries (Scotland, Wales and Northern Ireland) ought to be allowed to vote in the Westminster Parliament on policies that affect England only. The NHS also featured prominently in that debate, including the implications for funding public services in the devolved countries, were England to abolish the NHS.

This report is an account of current constitutional arrangements for devolution and how these have impacted on the health care systems of the four countries of the UK. Its title reflects the fundamental problem that the disproportionate scale of England, in relation to the devolved countries, means that while creating an English Parliament would provide a remedy to some constitutional problems, it would create other greater problems. But, without an English Parliament it is difficult to create a federal system of governance, which means that the UK lacks arrangements that operate in federal systems, for example Canada and Australia, over: what changes governments can make to public services (such as abolishing the NHS and NHS charges), developing a fair system of funding public services, and requiring governments to report data so that their performance can be compared.

The source report, summary report, data appendices and digital outputs are available to download free of charge at www.nuffieldtrust.org.uk/compare-uk-health and www.health.org.uk/compareUKhealth.
# Contents

Introduction 3

1 The historical background to devolution 4
   England and the British Isles 4
   Scotland 5
   Wales 6
   Northern Ireland 7

2 Governance before devolution 10
   The Offices of State of Scotland, Wales and Northern Ireland 10
   Funding of public services 11
   Four models of governance 12
   Governance of the NHS 14

3 The constitutional arrangements of asymmetric devolution 17
   The different parliaments and assemblies 17
   Funding and accountability 18

4 Governance of health care after devolution 20
   England 20
   Scotland 24
   Wales 27
   Northern Ireland 30

5 Comparing governance of health care in the four countries 33
   Models of governance 33
   Patient choice 34
   Integration 35
   The pressures of austerity 36

6 Policy outcomes of asymmetric devolution 38
   The absence of an English Parliament 38
   What is reserved for the UK and what is devolved? 38
   Charges 39
   Funding 40
   Accountability 42

References 44
Introduction

Devolution took effect when powers were transferred from Westminster to the Scottish Parliament and Welsh Assembly on 1 July 1999; and to the Northern Ireland Assembly on 2 December 1999. This report on the impact of the devolution settlement on health care in the four countries of the UK can be read separately from, but aims to give the policy context for, the companion report, *The Four Health Systems of the United Kingdom: How do they compare?* (Bevan and others, 2014). This report uses some material from that report.

There are two sets of guiding ideas in this account. The first is that of path dependency. As Tuohy (1999) has argued in her development and application of the concept to analyse the evolution of the systems of health care of the UK, the USA and Canada, the “accidental logics” of history “have shaped national systems at critical moments in time, and in the distinctive ‘logics’ of the systems thus created” (p. 6). She showed how governments responded to “windows of opportunity” to enact major reforms through political settlements and policies. These reforms would then tend to develop on paths that followed the underlying logic of those settlements.

The key political settlements that have shaped policies following devolution are of two kinds. The first kind is the unions with England of what are now the devolved countries of Wales, Scotland and Northern Ireland. Chapter 1 gives necessarily brief accounts of these long, complex histories, with the objective of illuminating the different influences on the capacities of their governments after the second kind of political settlement, namely, devolution itself. This report aims to bring out its problematic consequences, which is an example of what Bogdanor (1999, p. 233) describes as ‘asymmetric devolution’, in which there is a giant unitary state and pygmies: for example in Denmark with the Faroe Islands and Greenland; and in Portugal with Madeira and the Azores.

This account of the consequences of asymmetric devolution for health care is important in its own right, but it also illustrates the consequences of the lack of arrangements for constitutional governance that would normally be found in a federal system for the governance, funding and accountability of public services in the UK following devolution. The second set of ideas used to analyse policy paths for health care are models of governance. These are developed in this report.

The rest of this report is organised into six chapters. Chapters 1 and 2 outline the historical background to devolution and arrangements for governance before devolution. Chapter 3 outlines the devolution settlement and its arrangements for governance. Chapter 4 looks at how the different countries have developed and implemented different models of governance. Chapter 5 compares governance in the four countries. Finally, Chapter 6 considers the outcomes of asymmetric devolution.
The historical background to devolution

Devolution followed a long and complicated historical development of governance arrangements in the UK, which explains why our constitutional arrangements mean that there is now one set of policies for the NHS that serves the 50 million people who live in England, while the 10 million who live in Scotland, Wales and Northern Ireland are each governed by different policies.

England and the British Isles

Those from outside of the UK are likely to assume that there is still one UK NHS to which the English variant is assumed to apply, and may find it hard to believe that it is right to call, for example, Scotland a ‘country’ rather than a ‘region’. There is confusion between the various terms: England, Great Britain, the UK and the British Isles. One illustration of the confusion that applies is that those from Northern Ireland would play football for Northern Ireland, rugby for Ireland and the British and Irish Lions, cricket for England, and athletics for Great Britain in the Olympics. Often ‘England’ is taken to be synonymous with: England and Wales, Great Britain, and the UK. Great Britain ought to include Wales and Scotland, but not Northern Ireland; the UK ought to include Wales, Scotland and Northern Ireland; the ‘British Isles’ includes the Republic of Ireland, which as Davies (1999) points out ignores the vital historical fact that this country is independent of the UK (Davies, 1999, pp. xxiii–xli). Hence, the history by Davies of the ‘British Isles’ is titled simply ‘The Isles’. He found that neither the Bodleian library of Oxford University nor the Library of Congress in Washington had an entry for the history of the UK, and their entries for Great Britain assumed this to be identical to the history of England (pp. xxxiv–xxxvi). Wright (2000) comments on the puzzlement reported by Bill Bryson (Bryson, 1995) in his tour of Britain that Asquith’s headstone describes him as having been Prime Minister of England at the ‘highpoint of the British Empire’ (p. 8). Furthermore, this ‘Prime Minister of England’ represented, as an MP, Scottish constituencies (East Fife and Paisley).1

Greer (2004, p. 94) points out that Richard Rose entitled his text (Rose, 1989) on the politics of the UK, Politics in England, explaining (p. vi) that one can write about ‘British Government and English society’, which ignores the fact that Northern Ireland, although part of the UK, is not part of Great Britain. Davies (1999, p. 552) argues that although historical accounts of Scotland and Ireland highlight their unions with England in 1707, and in 1800, there is no reciprocal recognition in historical accounts of England, which suggests that “England never united with anyone”. Davies here overlooks England’s union with Wales in line with the early editions of the Encyclopaedia Britannica, which made no distinction between England and Wales; the index read: “For Wales – see England” (Robbins, 2001).

Bogdanor (1999, p. 5) describes the formation of the UK as coming about through the expansion of England via a process of “conquest, treaty and negotiation”
which came to be unified “sometime between 899 and 956... far earlier than any Continental State” so that “there was already a powerful English consciousness and sense of national identity before the Norman conquest” (p. 5). He argues that notions of Scottishness and Welshness were to be formed in opposition to Englishness, which had become so dominant. Similarly, in making sense of the NHS policies of Scotland and Wales after devolution, it is essential to know the policies in England, as these have defined a point of departure for the devolved governments in Scotland and Wales. Key developments in the formation of the current UK included historical defeats by the English of the Welsh, by the Scottish of the English; and, in Northern Ireland, the plantation of Ulster in the seventeenth century and the movement against Irish Home Rule in the 1880s (Davies, 1999). Their different histories mean that as Colley (1992, p. 14) points out, what are now the three devolved countries ought not to be assumed to be similar and labelled simply as the ‘Celtic fringe’.

Scotland

Bogdanor (1999) outlines key historical differences in the development of the Scottish nation from that of Wales or Ireland. Scotland emerged as a kingdom in about the same period as England (during the ninth and tenth centuries), but, unlike Wales and Ireland, was never permanently conquered by England. The failure of Edward I, the ‘hammer of the Scots’, to subjugate Scotland was followed by the Treaty of Northampton in 1328, which recognised Scottish independence (Bogdanor, 1999, p. 8). In medieval times, Scotland became a monarchy and united the Celtic and non-Celtic populations, unlike in Northern Ireland. Scotland remained independent with its own parliament until the merger of the Edinburgh and Westminster Parliaments in the Act of Union of 1707, in which Scotland secured greater autonomy than either Wales or Ireland in their earlier and later unions (Davies, 1999, p. 583). Bogdanor (1999) uses a Marxist distinction to highlight a crucial difference between Scotland and Wales following their unions with England: Scotland remained a ‘historic’ nation, “which succeeded in retaining the institutions of statehood” and Wales did not (p. 144). Colley (2014, p. 89) points out that “Scotland retained its own systems of Roman Law and local government, its own parish schools and excellent universities (more extensive in number in 1707 than their English counterparts)”. Greer (2004) argues that the geopolitics that made Scotland a separate successful Scottish state before 1707 also allowed it to assemble and develop an imposing and distinctive civil society, and that the legacies of the distinctiveness of Scottish institutions prior to devolution meant that of the three devolved countries, Scotland had “the most complicated, competitive and well-worked out policies” (p. 64) long before 1999.

Politics in Scotland has changed markedly since the 1955 general election, when the Conservative Party secured the majority of Scottish seats at Westminster and of the vote (Bogdanor, 1999, p. 132). This was followed by a decline in the popularity of the Conservative Party in Scotland, with Labour becoming the majority party between 1979 and 1997; and the rise and electoral success of the Scottish National Party (SNP) from the 1970s, which Bogdanor (1999, pp. 119–43) argues was important in putting devolution on the political agenda. He points out that nationalism in Scotland, unlike in Ireland, “has not displayed the emotional hostility to the British connection... and it has adhered strictly to constitutional methods” (p. 119). In the 1979 referendum on devolution in Scotland, there was
support from only 51.6 per cent of those voting with a 32.85 per cent turnout of the electorate (p. 191). This led to the Callaghan Government losing a ‘no confidence’ vote at Westminster and its defeat in the 1979 general election, with victory for the Conservatives under Margaret Thatcher, although as Bogdanor (1999) points out, in that election, Labour improved its share of the vote and number of Scottish MPs, and the SNP lost nine of its 11 seats (pp. 142–3). By the 1987 election, the Conservatives had less than one seventh of Scottish MPs (Bogdanor, 1999, p. 194). In 1989, the Thatcher Government imposed the deeply unpopular poll tax in Scotland, a year before its introduction in England and Wales. After the 1997 general election, there were no Conservative MPs in Scotland, and the two major parties were the Labour Party and the SNP (Norris, 1997). In the next referendum on devolution held in 1997, on a 60.2 per cent turnout, 74.3 per cent supported the establishment of a Scottish Parliament and 63.5 per cent were in favour of it having tax-varying powers (Bogdanor, 1999, p. 199).

**Wales**

The Statute of Rhuddlan (1284) subjugated Wales to English jurisdiction following its conquest by Edward I (Davies, 1999, p. 317). The Act of Union of England and Wales (1536) that followed the Tudor assault on Wales clearly defined the border for the first time, and renamed the Kingdom of England as the Kingdom of England and Wales (Davies, 1999, pp. 406–7). Colley (1992) observes that the union meant that Wales, in contrast to Scotland, “had lost its own legal system, its religious organisation was modelled on England’s own, and it had no universities or capital city like Edinburgh as a focus for cultural life” (p. 13).

Bogdanor emphasises that the union with England has “since the sixteenth century... been seen as a success”, in part because it allowed “the retention of Welsh cultural identity and a sense of Welshness” (p. 7). Later, English common law was imposed and English declared the sole language of administration. Bogdanor points out, however, that there was no objective by the Tudors, who were in origin a Welsh dynasty, to suppress the Welsh language. Bogdanor argues that the emergence of Welsh nationalism in the nineteenth century was in response to threats to a distinctive Welsh culture and way of life; and was not motivated “as in Ireland by a desire to break the British connection; nor as in Scotland, by a desire to refashion government so as to guarantee her distinctive institutions” (p. 7). Although Colley (2014, p. 83) points out that: “By the outbreak of the First World War, Wales possessed its own university system, its own National Library and Museum, its own recognised national anthem, ‘Hen Wlad Fy Nhadau’”, and that “Welsh self-government was openly canvassed by politicians in the late nineteenth century... [which] divided Welsh opinion”. Greer (2004, p. 131) contrasts the existence of a powerful medical elite that influences NHS and health policy in Scotland with its absence in Wales, which lacks its own Royal Colleges, and powerful and distinctive academic centres. Greer (2004, p. 130) observed that “Wales has historically been, and for many purposes still is, part of a unit called ‘England and Wales’”. And, as Colley (2014, p. 76) points out, “unlike Ireland, Scotland and England, Wales has never been emblematically represented on the Union Jack”.

Prior to devolution, for public services, Wales was essentially treated in terms of governance as an English region, although there were some differences in education policy, including in particular Welsh language education, which reflects
the focus of Welsh nationalism on preserving and developing the distinctive
Welsh culture rather than its own national institutions (Bogdanor, 1999, p. 157).
Legislation for the UK was on a common basis for England and Wales; whereas
Scotland and Northern Ireland each had different Acts before devolution (Greer
and Trench, 2008). Greer (2004, p. 131) observed that “Westminster passed no
Wales-specific legislation between the seventeenth century (when Wales lost its
separate structure of government) and a ban on Sunday pub openings in Wales
passed in 1881”. Longley (2004) noted that for Wales: “The courts and policing
arrangements are identical to those in England and many other aspects of public
life: higher education and schools; health care; social welfare policy and transport
have been characterized much more by their similarity with those of the much
larger neighbour than by their differences” (p. 2). McClelland (2002) states that
the NHS in Wales was perceived as “forming an adjunct to the English health
service”. In Wales, there was much less popular support for devolution than in
Scotland: in the 1979 referendum only 20.2 per cent voted for devolution. In the
1997 referendum, on a turnout of just over 50 per cent, 50.1 per cent supported the
establishment of a Welsh Assembly (Bogdanor, 1999, p. 199). Bogdanor (1999,
p. 199) points out that this was held “a week after the Scottish referendum, so that
Welsh opinion might be influenced by a favourable result in Scotland”, but the
majority favouring devolution was “less than 7,000 votes of over a million cast”.
This again illustrates, in contrast with Scotland, the relative lack of enthusiasm in
Wales for its own national institutions. The history of party politics in Wales “in
modern times has been characterised by its consistent and powerful support for
the Left” (Bogdanor, 1999, p. 145). Although that history is different from Scotland,
in both countries, after the 1997 general election, there were no Conservative MPs
in Wales.

Northern Ireland

Davies (1999) describes how at the end of the fifteenth century, the Dublin
Parliament had been obliged to pass Poyning’s Law, which invalidated all Irish
legislation not previously approved in England, and remained on the statute book
for nearly 300 years. King James I and VI signed the Act for the confiscation
of Ulster to achieve its subjugation by the plantation (principally of English
leaseholders and Scottish tenants). Ulster had previously been the most Irish,
Gaelic, Catholic and traditional province of Ireland (Davies, 1999, pp. 479–81)
with consequences that have shaped, and continue to shape, the policies
and politics of Northern Ireland. Bogdanor (1999, p. 8) describes how the
consequent developments in Ireland differed from those in Scotland and
Wales. In Ireland, the Catholic Celtic population was not assimilated with the
non-Celtic population, but “came to be treated as second class citizens, being
excluded from the Irish Parliament in 1692 and largely disenfranchised in 1727”.
He points out that, in Ireland, there were no dynastic links with the monarchs
ruling the union, unlike in Wales, through the Tudors, and for Scotland through
the Stuarts (p. 16).

In 1800, the British and Irish Parliaments were merged by two identical Acts of
Union passed simultaneously in Dublin and London, which followed the successful
defeat of the struggle for independence in 1789 by the United Irishmen led by
Theobold Wolfe Tone (with support from the French) (Davies, 1999, p. 532). This
union, however, in the view of the nationalists exacerbated Ireland’s economic
and social problems, as shown by the great famine of the 1840s (Bogdanor, 1999,
The impacts of asymmetric devolution on health care in the four countries of the UK

The British government’s response to the large number of nationalists elected (85 out of 103 Irish seats), in 1885, in the first general election in which the majority of Irish males could vote, was to introduce legislation for Irish Home Rule with the objective that Ireland could govern itself in her domestic affairs within the framework of the UK. Bogdanor (1999) describes this as “the most convulsive issue in British politics between 1886 and 1914” (p. 19). What is known in Ireland as the War of Independence of 1919 to 1921 ended with the partition of Ireland by the Government of Ireland Act (1920), which was confirmed by the Anglo–Irish Treaty (1921). This created two Parliaments, in Dublin, for the Irish Free State, and in Belfast, for the British Province of Northern Ireland (Davies, 1999, pp. 759–60). This means that Northern Ireland was the only part of the UK with practical experience of devolution prior to 1999 (from 1921 to 1971). As Bogdanor (1999, p. 55) points out, however, this is of limited relevance to devolution in Scotland and Wales for two reasons. First, the Northern Ireland Parliament “was not established, as devolved legislatures generally are, to meet a nationalist or separatist threat”. It was “pressed upon the Ulster Unionists by the British Government as a means of ending British rule in Northern Ireland and so resolving the Irish problem”. It was accepted by the Ulster Unionists not for the purpose of separation from, but rather to ensure that Northern Ireland remained part of, the UK. Second, Northern Ireland has its own political parties distinct from the major British parties (p. 101).

As Greer (2004, p. 163) observes, the politics of Northern Ireland is not one of Left versus Right, but of Unionist versus Nationalist. Bogdanor (1999) describes how the change in the voting system in Northern Ireland from proportional representation to first-past-the-post in local government in 1922 (together with the re-arrangement of boundaries) meant that “despite Catholic majorities in Fermanagh, Tyrone and Londonderry city, their local councils came to enjoy unionist majorities” (p. 76); and that change in 1929 “for elections to Stormont (the Parliament of Northern Ireland) consolidated Unionist strength... and anti-unionist parties at the expense of the Northern Ireland Labour party” (pp. 77–8; see also Bardon and Keogh, 2003, p. lxi). Greer (2004, p. 162) describes the Stormont regime in Northern Ireland, which governed Northern Ireland from 1922 to 1972, as a “Protestant ascendancy”, which “almost completely excluded Catholics from meaningful social, political, or economic roles, while restricting power to a narrow elite among Protestants” and acted independently of Whitehall or Westminster: a convention at Westminster ruled out parliamentary discussions of devolved (that is, Northern Irish) matters (Greer, 2004, p. 163). This regime collapsed in 1972 and was followed by Direct Rule, which lasted until devolution in 1999.

The Belfast Agreement (also known as the Good Friday Agreement), which led to the establishment of the Northern Ireland Assembly, was approved in the 22 May 1998 referendum in both Northern Ireland, by 71.1 per cent on an 81.1 per cent turnout, and the Irish Republic, by 94 per cent on a 56 per cent turnout (Bogdanor, 1999, p. 108). Greer (2004) argues that the consequences of the Belfast Agreement having been designed to “induce the Northern Irish parties to emerge from their conflict and participate in any sort of Northern Irish government” was that it was not designed to deliver effective government in developing policies or running services (p. 179). The constitutional arrangements for the Assembly, in which elections use a system of proportional representation (known as single transferable vote), have enshrined sectarian differences between the Nationalist and Unionist parties. Any party that wins 10 per cent of the vote wins ministerial positions with seats in the Executive. The two parties with the largest and second
largest percentages (Unionist and Nationalist) choose the First Minister and the Deputy First Minister, but each is empowered to appoint, shuffle or sack ministers within their own party only. Greer highlights two consequences for the governance of Northern Ireland: the absence of collective Cabinet responsibility and of the relationship between policy performance, ministerial accountability, and the party office. For example, “10 per cent of the population can keep an underperforming Minister in place whose mandate has nothing to do with his or her job” (p. 179). The continuing problems of governing Northern Ireland resulted in the Assembly being suspended on four occasions between its establishment and 2007 (O’Neill and others, 2012, p. 5).

Notes

1. I am grateful to Vernon Bogdanor for pointing this out.

2. Policing was not devolved, but the second report from the Silk Commission recommended that: “policing and related areas of community safety and crime prevention should be devolved” (Silk, 2014, p. 111). This is currently under discussion by the Silk Commission.
2
Governance before devolution

The Offices of State of Scotland, Wales and Northern Ireland

Within the UK, before devolution, there were Secretaries of State for Scotland, Wales and Northern Ireland, who were all members of the UK Cabinet. Greer (2004) describes the creation of the Scottish Office, in 1885, in Edinburgh, “as part of Scotland’s ongoing institutional adaptation to the British welfare state” (p. 65). After that, the development of the welfare state in Scotland, including the administration of the NHS after 1948, was supervised by the Scottish Office (pp. 65–6). Bogdanor (1999) argues that the convention of collective responsibility of the UK Cabinet meant that there was little scope for the Secretary of State for Scotland to pursue policies that diverged from those applied to England, except for matters “where English ministers did not particularly care what happened in Scotland, and where there seemed to be no implication for policy across the border” (p. 113); and that the wide range of responsibilities for the Secretary and Minister of State meant that, compared with England, “fewer decisions were taken by Ministers and more by civil servants” (p. 114). The relative weakness of the government offices of what are now devolved countries prior to devolution is illustrated by the way the Thatcher Government imposed the poll tax in Scotland, with the Scottish Office being “excluded from all forums engaged with the detailed work of the review until it became de facto policy” (Butler and others, 1994, p. 195). Even in the Blair Government, prior to devolution, Barber (2007, p. 82), head of the Prime Minister’s Delivery Unit, reported that officials, who were urging caution in response to ministers’ desires for bold reforms, would suggest: “Why not try it north of the border first?”.

Following the collapse of the Stormont regime in Northern Ireland in 1972, Direct Rule was imposed, which lasted until devolution in 1999, and the Northern Ireland Office was created in Whitehall. Bogdanor (1999, p. 101) argues that the nature of politics in Northern Ireland meant that “the Secretary of State and junior ministers were the executive and effective centres of power” and, unlike their counterparts in Scotland and Wales, lacked political accountability: Greer (2004, p. 167) describes this as “a type of vice-regal politics in which the separation of government and the vote was nearly total”.

The Welsh Office was created in 1964, but, as Greer (2004) observes: “Unlike its Scottish or Northern Ireland predecessors, it had to hew a distinctive Welsh policy arena out of the unified England and Wales organisations, policy regimes, and Whitehall departments responsible for policy until its creation” (p. 133); and Whitehall departments supervised the development of the NHS in England and Wales. McLean (2000) states that the Welsh Office has “always been the most junior of the three territorial departments” (p. 86). Bogdanor (1999) argues that, as in Scotland, the wide range of issues for which the Welsh Secretary of State was responsible meant that officials had a greater influence on policies than in England. And although, unlike in Northern Ireland, the political parties at Westminster did contest elections in England, he points out that: “The position of
Welsh Secretary became particularly difficult under Conservative governments since the Conservatives have never in modern times, gained a majority in Wales. Indeed, only one of six Conservative Secretaries of State for Wales... actually sat for a Welsh constituency” (p. 161).

**Funding of public services**

Before political devolution in 1999, the Secretaries of State for Scotland, Wales and Northern Ireland were accountable for expenditure on their public services within these countries. Decisions on spending on social security (although identifiable within each country), defence and foreign affairs were, and still are, made on a UK-wide basis. The three Secretaries of State were allocated a global sum for their public services and were free to allocate money to their chosen spending priorities.

The problems of funding of public services in different countries have long been fraught because of political lobbying and the mismatch between countries’ needs for these services and their tax base to pay for them (Bogdanor, 1999, pp. 235–54). These are ‘wicked problems’ to which there are no obvious good solutions. This is illustrated by the history of unsatisfactory attempts to try to develop workable and fair financial arrangements following Home Rule for Ireland for its ‘Imperial contribution’ “to cover the costs of its reserved services” which “led to difficulties as serious as that of representation” (Bogdanor, 1999, p. 35). McLean (2000, p. 81) points out that the problem faced by Goshen, the Chancellor of the Exchequer, in developing a mechanism for grants in aid for elected county councils in the 1880s, was that: “the areas on which most money needed to be spent had the lowest tax base” and so the proportions were based on tax receipts (and not populations). Following devolution to Northern Ireland in the 1920s, Bogdanor (1999) argues that there was a strong case that, as its citizens paid the same taxes as those in the rest of the UK, they were entitled to the same levels of public services. He describes how the second report of the Colwyn Committee, which was established to find a resolution to “long and irritating controversies between the Northern Ireland Ministry of Finance and the Treasury” (Bogdanor, 1999, p. 83) recommended a principle, which foreshadowed that of the Barnett Formula: that spending per head on services in Northern Ireland ought to increase at the same rate as in the rest of the UK. He described the process of allocating spending to services for Northern Ireland in the 1960s and 1970s as violating “the canons of efficiency and democracy”, and “drastically undermining her financial autonomy and destroying the connection between expenditure and revenue” (p. 89).

The allocation to Scotland benefited from what McLean (2000) calls the ‘Johnston gambit’, after Tom Johnston, Secretary of State for Scotland (1941-45), who, according to a Cabinet colleague, Lord Morrison of Lambeth (1960), would impress on Cabinet Committees that: “there was a strong nationalist movement in Scotland and it would be a danger if it grew through lack of attention to Scottish interests”. McLean (2000, p. 86) argues that the reason why Wales did not do as well in the funding of its public services was because of its history of never having had a voluntary treaty of union with England and of being effectively wholly incorporated into England until 1964.

Since 1979, funding of public services in the three devolved countries has been determined by the Barnett Formula, which began to operate in Scotland and Northern Ireland in 1979, and in Wales in 1980 (when political devolution was first being considered). McLean (2000, p. 82) states that the Barnett Formula
was supposed to counter the ‘Johnston gambit’. Its principle is that ‘growth’ in resources for ‘public services’ should be allocated to each country in proportion to its share of the UK’s population, with annual per-head spending increases derived from the percentage increase granted to the English baseline. Although the Barnett Formula was seen, at its introduction, as a short-term measure, it continued after devolution was enacted 20 years later, and has remained in place, largely unaltered, for 30 years.

A formula designed for the long term ought to take account of the relative needs of countries’ populations, but the Treasury’s two studies of relative needs in 1979 and 1993 were disregarded. If implemented, adjustment for relative needs would have meant reductions in Northern Ireland’s and Scotland’s allocations relative to England in both 1979 and 1993; and an increase in spending in Wales in 1979, but not in 1993 (Select Committee on the Barnett Formula, 2009, pp. 20–1). McLean (2000, p. 85) argues that the evidence showed that: Scotland benefited and the “conspicuous losers” were Northern England, the East Midlands, East Anglia and the South West. The House of Lords Select Committee on the Barnett Formula (2009) reported that, before the formula was used, England had had the lowest per-head spending; in 1976/77 per head spending on ‘public services’ was much higher than England in Northern Ireland (by 35 per cent) and Scotland (by 22 per cent), and a little higher in Wales (by 6 per cent) (p. 21). The Select Committee (pp. 21–2) also found that, although in principle the design of the Barnett Formula implies gradual convergence in spending per head, this did not happen for two reasons. First, relative populations were not updated until the 1990s, despite significant changes (for example, Scotland’s share of the UK population declined from 9.3 per cent in 1976 to 8.7 per cent in 1995). Second, the formula did not determine all allocations of devolved public spending: there were extra allocations negotiated bilaterally with the Treasury outside the formula, in particular, to cover public sector wage increases (which appear to have benefited Scotland and Northern Ireland). At the time of devolution to Scotland, Wales and Northern Ireland, Scottish public spending per head was around 25 per cent higher than in England, which McLean (2000, p. 82) argues was explained by the ‘Johnson gambit’.

**Four models of governance**

This report uses four models to analyse the way each NHS has been governed before and after devolution: ‘trust and altruism’; ‘choice and competition’; ‘targets and terror’; and ‘naming and shaming’, which are similar to those described in books written, after serving in the Blair Government, by Julian Le Grand (Le Grand, 2007), as senior policy adviser to the Prime Minister (2003 to 2005), and Sir Michael Barber (Barber, 2007), as head of the Prime Minister’s Delivery Unit (2001 to 2005).

The model of ‘trust and altruism’ assumes that providers of public services are ‘knights’ (Le Grand, 2003; 2007) and, as they are driven by altruism, they can be trusted to do the best they can for those they serve within the available resources, without any need for external incentives and, indeed, poor performance should be taken to indicate a need for extra resources. It is difficult to find any theoretical justification for this model, which can be criticised on two grounds. First, as Le Grand (2003) argued, those who deliver public services are not all purely driven by ‘knightly’ motives of altruism, but also by ‘knavish’ motives of self-interest. This mix cannot be guaranteed to provide the incentives necessary to overcome organisational inertia to implement changes for quality improvement. This
The impacts of asymmetric devolution on health care in the four countries of the UK

process requires using comparative information on performance to understand how performance needs to change to deliver better outcomes and to win support for implementing the necessary changes (Berwick and others, 2003). Second, prospect theory, developed by Tversky and Kahneman (Tversky and Kahneman, 1991; Kahneman, 2011), shows that people feel losses much more keenly than gains of equivalent magnitude. Thus, sanctions for failure can generate the high-powered incentives necessary to overcome organisational inertia to improve quality. As the ‘trust and altruism’ model eschews the use of sanctions, it removes these high-powered incentives, and rewarding failure, of course, creates perverse incentives (Bevan and Fasolo, 2013). Nevertheless, this model has low monitoring costs, is popular with professionals, is common in public services (Le Grand, 2003; 2007) and is assumed to have been the default model for the governance of health services in the countries of the UK for most of the life of the NHS.

The model of ‘choice and competition’ is based on competition between providers and ‘money follows choice’, so that providers respond to the consequences of choices on their market shares and incomes. In quasi-markets for public services, such as schools and hospitals, services remain free at the point of consumption but ‘money follows the pupil or patient’ (Le Grand, 2007). It has proved, however, difficult to design effective quasi-markets, as they require good information, supply-side flexibility and freedom to manage. Quasi-markets have high transaction costs, but are increasingly popular with governments, because pressure on poor performance is perceived to come from the ‘invisible hand’ of the market (Le Grand, 2007). Both Le Grand (2007) and Barber (2007) favour this model as having the greatest potential to deliver high performance in public services on the grounds of greater potential to respond to users’ needs than the two centrally driven models, below.

The model of ‘targets and terror’ holds providers to account against a limited set of targets that clearly signal priorities to those responsible for running organisations, with clear threats of sanctions for failure and rewards for success. Le Grand (2007) and Barber (2007) describe this model as ‘command and control’. This model also assumes that providers respond to clear economic incentives. Prospect theory tells us that, of these incentives, sanctions will have a stronger impact than rewards. The ‘targets and terror’ model imposes external incentives by strong performance management, has monitoring costs and is unpopular with professionals.

The model of ‘naming and shaming’ assumes that providers respond to threats to their reputation and is a system of performance measurement. Hibbard and others (2003) show that it requires an ability to rank providers’ performance so that the public, managers and professionals can easily see which providers are performing well and which are performing poorly on a regular basis, enabling change to be monitored. The classic model of ‘naming and shaming’ is the publication of annual league tables of schools based on the performance of their pupils in examinations (Bevan and Wilson, 2013).

In relating these models to actual practices, three points need to be made. First, although these models are conceptually distinct in practice, governments tend to use a mix of these models. Second, however convincing each model appears to be a priori, if it lacks effective sanctions for failure then it is, in effect, one of ‘trust and altruism’. Third, models other than ‘trust and altruism’ are vulnerable to ‘gaming’ because they create high-powered incentives in relation to inevitably imperfect measures of performance (Holmstrom and Milgrom, 1991; Bevan and
Hood, 2006). Indeed, there is no perfect model and it is a profound mistake to change policies in the belief that such a model is there to be discovered by trial and error.

**Governance of the NHS**

The creation of the NHS in 1948 had established largely the same organisational forms and common policies across the whole of the UK: with access to the NHS being free at the point of delivery (except for the subsequent introduction of prescription charges, which were later abolished by the devolved governments – see below) and typically via a general practitioner (GP), who acts as gatekeeper to specialist services. Initial arrangements illustrate how Wales was essentially seen as an English region, with Scotland and Northern Ireland seen as distinct administratively: the regulation to control the distribution of GPs was by three medical practices’ committees: for England and Wales; for Scotland; and for Northern Ireland. A hospital management committee governed each hospital (with separate boards of governors for teaching hospitals), and was accountable to 15 regional hospital boards in England and Wales (14 in England and one in Wales), five in Scotland (Levitt and Wall, 1984), and a hospitals authority in Northern Ireland (Leathard, 2000, p. 294).

The first major reorganisation of the NHS after its inception (implemented in 1973 in Northern Ireland, and 1974 in England, Wales and Scotland) aimed to shift the NHS from an organisation based on hospitals to one based on serving the needs of populations. This resulted in a regional structure in England of 14 regional health authorities (RHAs); for the other countries, the government department of health fulfilled both roles of region and department of state. Within English RHAs and Wales there were area health authorities responsible for running hospital and community health services, and planning for populations within each area (in England and Wales these were later reorganised into district health authorities); and family practitioner committees, later family health service authorities, for family practitioner services. In Scotland, health boards were created with the joint responsibilities of what were two separate bodies in England and Wales: area health authorities and family health service authorities. In Northern Ireland, health and social services boards were created with responsibility for health and social services. The constitution of executive teams was similar in England, Wales and Scotland; but differed in Northern Ireland, where the health and social services boards included the Director of Social Services, but excluded the Treasurer (Levitt and Wall, 1984). Greer (2004, p. 165) points out that the principal objective for moving social services out of local government in 1973 was not to achieve the now popular goal of integration with health services, but because of “systematic discrimination by local governments, which was one of the factors causing the troubles. Northern Ireland’s social services, thoroughly politicised and sectarian, had to be changed – since they were a major cause of Catholic grievance”.

Before political devolution, each country was subject to common NHS policies based on those developed for the NHS in England and then applied throughout the UK with latitude for minor variations. Hunter and Wistow (1987) did identify some policy diversity in community-based care in the 1980s, rather than uniformity across England, Scotland and Wales, but they also highlighted two powerful constraints on Scotland and Wales developing distinctive policies: the relative lack of policy-making resources in Edinburgh and Cardiff; and the limited growth in public spending.
Klein (2010a) describes governance in the first three decades of the English NHS by the centre as “largely to exhort and hope” (p 294). He argues that key to the development of a system of command and control in the 1980s was the development of electronic data, but this was not a system of central ‘targets and terror’ because of ambiguity in the power of the centre to hold RHAs to account. This changed with the controversial policies of the White Paper Working for Patients (Secretaries of State for Health, Wales, Northern Ireland and Scotland, 1989; Le Grand, 2003; Bevan and Robinson, 2005). These policies introduced governance by ‘choice and competition’ in an ‘internal market’, which was implemented throughout the UK. This created a purchaser/provider split, based on the idea that purchasers would contract with separate NHS providers on grounds of price and quality, with ‘money following the patient’ (although in practice money followed choice through contracts). In the ‘purchaser/provider’ split, district health authorities in England and Wales, health boards in Scotland, and health and social services boards in Northern Ireland, became ‘purchasers’, and their hierarchical role in governing providers was replaced with contractual arrangements, as providers became more autonomous NHS trusts. Another innovation of the ‘internal market’ was the creation of new small-scale purchasing by GPs who opted to become fund holders, of which various forms emerged over time (Mays and Dixon, 1996). In England, the emphasis on ‘light touch regulation’ of the ‘internal market’ resulted in the abolition of RHAs. This regulatory tier became a monitoring arm of the Department of Health and has been reorganised every few years ever since. The Coalition Government has continued to develop the idea of competition with a further major structural reorganisation of the NHS in England after 2010 (see below).

As Greer (2004) argues, England is the odd one out in the countries of the UK in being so much larger, having a strong Conservative Party, significant independent sectors for health care (and schools), think tanks across the right-left political spectrum, advocates of market-based reforms and ideas of New Public Management, a large and critical press, and political debates that put “the operation and even the existence of the NHS in question” (Greer, 2004, p. 103). The massive scale of the NHS in England means that it is more complicated to run than in the devolved countries. Greer (2004) argues that the consequences are that its “civil service policy and administrative capacity… far exceeds that of the devolved administrations” (p. 96) and: “The decisions of the government can be turned into policy and implemented far more quickly and with less hazard” (p. 97). English political influences thus naturally shaped policies for the NHS in England, which, prior to devolution, were then exported to the other countries of the UK. Shock (1994) describes the imposition of the ‘internal market’ on the medical profession as being “struck by the blitzkrieg from the right… with little consultation, indeed not much discussion, even with civil servants let alone those outside Whitehall”; but warned then that: “A blitzkrieg can certainly achieve conquest but it cannot ensure effective occupation”. As is explained below, the influences that shaped policies in Scotland and Wales following devolution meant that there was much more hostility to the internal market in those countries, in part, because it was seen as a policy invented in England and imposed from without. Following devolution, the governments in Scotland and Wales have both rejected the policy of provider competition, abolished the purchaser/provider split and reorganised their health systems so they are more similar to the organisational model of the 1970s.
Notes

3. This was named after Joel Barnett, then Chief Secretary to the Treasury, but McLean (2000, p. 82) points out that “he has always denied authorship and attributes it to the mandarin Sir Leo Pliatzky”.

4. For example, although the regime of ‘star ratings’ combined the two models of ‘targets and terror’, and ‘naming and shaming’; in schools and US health care, the latter model is used alone and there is no centralised system of accountability of providers to a government department. Although naming and shaming can lead to those responsible for running services being sacked, this is the outcome of a local rather than a national decision.

5. For example, in the ‘choice and competition’ model it has proved difficult for ministers to let ‘failing hospitals’ or services within hospitals exit the market; and a system of targets without sanctions is one of ‘hope and exhortation’ rather than ‘targets and terror’ (or ‘command and control’).

6. The 14 RHAs were abolished in 1994 and replaced by eight regional offices; which were abolished in 2001 and replaced by four new regional directorates of health and social care; these were abolished and replaced by 28 strategic health authorities in 2003, which were abolished and replaced by 10 strategic health authorities in 2006. This organisational turbulence continued with the changes made by the Coalition Government. Timmins (2012, p. 3) describes Andrew Lansley as: “A man in a hurry who was part of a coalition government (that just weeks earlier had promised the country ‘no more top-down reorganisations’ of the National Health Service) launched arguably the biggest restructuring it had seen in its 63-year history”.
3

The constitutional arrangements of asymmetric devolution

The different parliaments and assemblies

The act of devolution of 1999 gave the Scottish Parliament a wide range of legislative powers with freedom to legislate on all matters except those reserved for Westminster, which for health and health care included regulation for almost all the health professions. The powers of the Assembly in Wales were initially more circumscribed and limited to executive matters, which were those of the Welsh Office prior to devolution. However, these powers were enhanced with provisions in the Government of Wales Act 2006 that enabled the Assembly to legislate in relation to 20 areas that cover the delivery of local services (including health services). The Assembly in Northern Ireland was also empowered to legislate except for reserved matters, but it was suspended after only 11 weeks, restored in May 2000, and then suspended again on 14 October 2002 until 8 May 2007 following the St Andrews Agreement (the result of negotiations held at St Andrews, Scotland, in October 2006). During the periods of suspension, Northern Ireland returned to Direct Rule from Westminster.

Davies (1999, pp. 870–1) argues that the nature of the union of Scotland with England (and Wales) in 1707 prevented the UK from developing either a federal structure (as in Germany, where each Land was established on an equal basis with none designed to dominate the others) or a wholly unitary structure (as in France when it had a highly centralised character in which the political nation developed within one territory). The outcome is that the UK is “essentially a dynastic conglomerate, which could never equalise the functions of its four constituent parts” and lacks a unified legal system, centralised educational system, common cultural policy or history (Davies, 1999, pp. 870–1). Hazel (2000, p. 29), who described the absence of an English Parliament as a “hole in the devolution settlement”, ruled out its creation on the grounds of its scale: that “as a rule of thumb, there is no successful federation in the world where one of the parts is greater than one third of the whole. England with four fifths of their population would be hugely dominant, even more dominant than Prussia in the old Germany” (p. 35). Bogdanor (1999, p. 267) also argues that this would be unbalanced because of England’s size and cites the observation made by the Report of the Royal Commission on the Constitution (Kilbrandon, 1973) that an English Parliament “would rival the United Kingdom Federal Parliament; and in the Federal Parliament itself, the representation of England could hardly be scaled down in such a way as to enable it to be outvoted by Scotland, Wales and Northern Ireland”. Hence, Bogdanor argued that: “Devolution in England, therefore, if it is to serve the same ends as devolution in Scotland, Wales and Northern Ireland must be devolution to English regions” (p. 268). But, as Bogdanor argued, for various reasons, the development of regional government in England is “highly problematic” (p. 275), the “fundamental difficulty” being “that there is so little
The impacts of asymmetric devolution on health care in the four countries of the UK

...demand for it”. And, indeed, when the option of elected regional government was piloted in 2004 in North East England, it was rejected heavily in the subsequent referendum (BBC News, 2004). Asymmetric devolution means that the Westminster Parliament acts, in non-devolved matters, for the UK, and, for devolved matters, for England, and (in principle) as the body with oversight across the subordinate legislatures in the devolved countries, which creates two ‘wicked problems’.

First, what has become known, after the former MP for West Lothian, Tam Dalyell, as the ‘West Lothian question’. This question is: why should MPs from non-English constituencies be able to vote on policies for public services in England, when English MPs cannot vote on these policies for each devolved country, even though their finance comes from the budget for the UK? Bogdanor (1999, pp. 29-35) points out that this constitutional problem has a long history: it was recognised as a serious problem to which Gladstone struggled and failed to find an answer in relation to Irish Home Rule in 1889; and was also an issue in 1965. Second, there is no constitutional basis for agreeing what should be the UK-wide elements of policy for public services. Greer (2004, p. 197) observed that the UK is “an international outlier because of its lack of frameworks constraining the regional governments (of the different countries). The various mechanisms by which the central government constrain decentralised jurisdictions by law in other countries is simply lacking in the UK’s formal constitution”. He states that: “the UK, almost uniquely in the world does not constrain substantive policy divergence in health” (p. 179). Thus, for health care, there is no analogy to the Canada Health Act, which sets out nine requirements that provincial governments must meet through their public health care insurance plan in order to qualify for the full federal cash contribution under the Canada Health Transfer (Moore, 2005; Flood and Choudhry, 2004). One of these requirements is that, to qualify for federal contributions, provincial plans must prohibit ‘user charges’. In principle, the legislation that created the Scottish Parliament allows the Westminster Parliament to assert its supremacy over that body, but Bogdanor (1999, p. 291) argues that in practice, “power devolved, far from being power retained, will be power transferred; and it will not be possible to recover that power except under pathological circumstances, such as those in Northern Ireland, after 1968”. Outside those pathological circumstances, he points out that: “Westminster found itself incapable of exercising its supremacy over the Northern Ireland Parliament”; and that it is thus even more unlikely that the Westminster Parliament could do so over the Scottish Parliament, which will speak “not for an artificially-created province”, but a country “with a history of statehood, and a national tradition embodied in concrete and ever-present institutional form” (p. 289). Furthermore, as is discussed below, the implementation of the workplace smoking ban in Scotland in 2004 illustrates that there is in practice ambiguity over what matters are reserved for decisions on a UK basis and which can be devolved.

Funding and accountability

Asymmetric devolution has resulted in two different systems for determining NHS budgets and for accountability for that expenditure. One system applies to England only: the NHS budget for England is the outcome of UK Cabinet agreements following negotiations between HM Treasury and the Department of Health for England. From 1998 to the end of the Labour Government in 2010, in principle, budgets for public services were contingent on each government...
department and its minister in England delivering performance that satisfied a set of targets agreed with HM Treasury and set out in Public Service Agreements (PSA targets) (Connolly and others, 2011). In the absence of an English Parliament, political accountability for public services is through elections to the UK Parliament, which hinge on a mix of English and UK-wide issues (such as the economy, spending on social security, and defence and foreign policy). As before devolution, the global allocations for ‘public services’ for each devolved country are largely determined by the Barnett Formula, and each devolved government then decides how much of its global allocation ought to be allocated to the NHS.

What changed with devolution is that each devolved government has direct political accountability for public services to its own parliament or assembly. But asymmetric devolution in the UK means there is no federal institution in the UK, which might negotiate arrangements for reporting performance across the countries of the UK for the purposes of accountability and learning from each other. For example, in Australia, Banks and McDonald (2010) state: “Every year Australia’s [state] governments cooperate in producing the Report on Government Services (RoGS), a comprehensive exercise in performance reporting across a wide range of services delivered by Australia’s State and Territory governments. The range of services has grown since the first Report was published in 1995 and activities included in the 2011 Report amounted to almost $150 billion, over two-thirds of total government recurrent expenditure”.

Notes

7. See https://www.gov.uk/devolution-settlement-wales.


9. In a debate on devolution in November 1977, Mr Dalyell said: “For how long will English constituencies and English Honourable members tolerate... at least 119 Honourable Members from Scotland, Wales and Northern Ireland exercising an important, and probably often decisive, effect on British politics while they themselves have no say in the same matters in Scotland, Wales and Northern Ireland.” See http://news.bbc.co.uk/1/hi/uk_politics/7702326.stm. Bogdanor (1999, p. 34) states that “this was pursued with much pertinacity by Tam Dalyell in the 1970s” (p. 228). He cites an exchange in the House of Commons (during the Committee stage of the first devolution Bill: the Scotland and Wales Bill in February 1977) when “Tam Dalyell declared that ‘The point (i.e. West Lothian question) cannot be made too often’ to which the Minister of State at the Privy Council Office, John Smith, who was in charged of the Bill replied: Yes, it can” (p. 228). This is at House of Commons Debates, 5th series, vol 925, col. 262, 1 Feb 1977.

10. The Labour Government then had a majority of only two, and there were 13 MPs from Northern Ireland, all of whom sat on the Conservative benches. The Prime Minister, Harold Wilson, argued that: “What was not envisaged, I am sure in 1920 was that those who came here, with that responsibility for representing Northern Ireland interests, should just become hacks of the English Tory Party”. House of Commons Debates, 5th series, vol 711, cols. 1560–62, 6 May 1965. Quoted by Bogdanor (1999, p. 71).

11. There were also bilateral negotiations with the Treasury for allocations to Scotland and Northern Ireland.
4 Governance of health care after devolution

The impacts on the NHS of the ‘natural experiment’ between England and the devolved countries can be split into two periods, as follows:

• 1996 to 2000, when there were marked differences in spending per head, but similar policies

• from 2000, when there were diminishing differences in spending per head, but increasingly divergent policies.

This chapter outlines how policies developed and diverged in England, Scotland, Wales and Northern Ireland following devolution.

England

For schools, after the 1997 election, the Blair Government continued the Conservative reforms in England and Wales including publication of school league tables of examination results and of school inspections by the Office for Standards in Education (Ofsted); and the quasi market (where ‘money followed the pupil’) (Chitty, 2009). But, for the NHS in England and Wales, its initial set of policies were described as offering a ‘third way’ compared with two ‘failed’ alternatives: the “divisive ‘internal market’ system of the 1990s”, and the “old centralised command and control policies of the 1970s” (the last time there had been a Labour government) (Secretary of State for Health, 1997). The government retained the organisational separation of ‘purchasers’ from ‘providers’, created for the internal market, but abandoned the rhetoric of competition so that ‘purchasers’ became ‘commissioners’: the objective of that change was that this would foster collaborative arrangements with providers. This was in effect a return to governance based on the model of ‘trust and altruism’. GP fundholding was abolished, but about 450 primary care groups (PCGs) were created within the then 90 health authorities, with the objective of securing the advantages of fundholding without its disadvantages (which included allegations of creating a ‘two-tier’ NHS). The health authorities were later abolished and PCGs were later reorganised into first, 350, and then later, 150, primary care trusts (PCTs), which in effect took over the roles of the previous 90 health authorities (Klein, 2010b). Within Wales there was the parallel creation of local health groups (LHGs) based on local authorities in place of PCGs, and these LHGs later became local health boards (see below).

A 1997 White Paper also introduced two new regulatory bodies (Klein, 2010b). First, the National Institute for Clinical Excellence (NICE; now called the National Institute for Health and Care Excellence), which issues guidelines and recommends what should be available on the NHS in a deliberative process with reference to evidence from cost-effectiveness analysis, which was how the Labour Government intended to end geographical variations in access to care (‘postcode rationing’). Second, the Commission for Health Improvement, which had a principal task of
reviewing the implementation of clinical governance in NHS providers. This was one response to a series of landmark failures of self-regulation by the medical profession (Abbasi, 1998), which contributed to the perception of a ‘NHS crisis’ of poor quality in the late 1990s (Klein, 2010b, p. 204).

A conviction grew in some quarters that one primary cause of the ‘NHS crisis’ was the long-term ‘under-funding’ of the NHS (Smee, 2008). Tony Blair, the then Prime Minister of the UK, during a television interview, made the commitment to increase spending on the NHS in the UK to the European average spend on health care as a percentage of gross domestic product (Smee, 2005). Only in England, however, did the Secretary of State for Health, Alan Millburn, make it clear that this extra funding of the NHS was to be in return for a transformation of performance. Furthermore, this extra funding was, in principle, contingent on satisfying Treasury PSA targets. The White Paper, *The NHS Plan* (Secretary of State for Health, 2000), emphasised that “investment has to be accompanied by reform” (p. 11); and announced ambitious targets for increases in capital development and staffing, reducing waiting times for access to the NHS, and improving services for patients with cancers, coronary heart disease and mental illness. Following *The NHS Plan*, capital development was financed by the Private Finance Initiative (PFI), which had been introduced by the previous Conservative Government. Under the PFI, private consortia design, finance and build projects, and run and maintain the non-clinical services over the lifetime of the agreement (typically 30 years), with the facilities being leased back to the public sector for an annual rental payment (Sussex, 2001; Pollock and others, 2002).

*The NHS Plan* was part of the Blair Government’s emphasis on ‘delivery’, coordinated by the Prime Minister’s Delivery Unit, and led by Sir Michael Barber. The aim was to tackle the problem of ‘exit’ by those with a ‘voice’ (Hirschman, 1970) by raising standards in the NHS and schools to the extent that even those who could afford to pay privately for health care or education would not need to do so to ensure a high quality of service. Greer (2008, p. 80) quoted a Labour special adviser saying in July 2006 that an unsatisfactory NHS will make the “middle classes first vote against the NHS with their feet, and then with their votes”. The policies of ‘delivery’ were based on Sir Michael’s well-developed concepts of ‘deliverology’, in which the crucial concept was the ‘delivery chain’, which requires those setting targets in central government to work out a causal chain from the target to its delivery by those who are responsible for its achievement on the ground, backed by an effective system of reporting of achievement against targets (Barber, 2007). *The NHS Plan* emphasised that, for the NHS in England, there would be a new regime of performance management to replace the current system, which, it was argued above, had been based on the model of ‘trust and altruism’, and was described as one that “penalises success and rewards failure” (Secretary of State for Health, 2000, p. 28). This was because the government; for example, bailed out hospitals with long waiting times and lists by rewarding them with extra money, and hence had inadvertently created a system of perverse incentives. This was replaced by the regime of annual ‘star ratings’ applied to NHS organisations in England, between 2001 and 2005.

In the regime of ‘star ratings’, the targets for reducing waiting times were described as ‘P45 targets’, because failure to achieve them brought threats of the sack to hospitals’ chief executives (Bevan and Hood, 2006; Bevan and Hamblin, 2009). Hence, this followed the model of ‘targets and terror’. The system of public reporting also satisfied the criterion set by Hibbard and others (2003) for
effective public reporting in that it had the potential to inflict reputational damage on hospitals performing poorly against government targets: the ranking system (from zero to three stars) made it easy to distinguish high from poor performance; this system had wide publicity; and results were published annually from 2001 to 2005. Hence, this regime also incorporated ‘naming and shaming’. One of the rewards in the ‘star rating’ system was ‘earned autonomy’ for high-performing NHS trusts, which were, from 2002, subject to satisfactory performance judged by their regulator, Monitor, eligible to become foundation trusts, and enjoy greater independence and flexibility from central controls. The way this policy won support in Westminster Parliament is discussed below as an example of the ‘West Lothian question’.

In addition to the high-powered incentives of the ‘star ratings’ regime to achieve targets, NHS hospitals were given advice on how to improve their delivery from the Modernisation Agency (2004), which was responsible for the dissemination of good practice. One of its most famous publications was 10 High Impact Changes for Service Improvement and Delivery, which would enable acute hospitals to achieve targets for waiting times by, for example, greater use of day surgery, better access to diagnostic tests, or improving the processes of admission and discharge. There is considerable evidence that the ‘star ratings’ regime backed by the Modernisation Agency led to improvements in reported performance of the English NHS for hospital waiting times (Auditor General for Wales, 2005; Alvarez-Roseté and others, 2005; Bevan and Hood, 2006; Propper and others 2008 and 2010; Bevan, 2010; Connolly and others, 2011; Bevan and Wilson, 2013) and ambulance response times to life-threatening emergencies (Auditor General for Wales, 2006; Bevan and Hamblin, 2009; Connolly and others, 2011). The accompanying empirical report (Bevan and others, 2014) adds to that evidence. Evidence for these improvements comes from looking at both England over time and in comparison with the devolved countries (in a ‘natural experiment’). However, there is also some evidence of ‘gaming’ in relation to the targets (see below).

From 2002/03, the government gradually introduced a second ‘internal market’ based on the model of ‘choice and competition’ into the NHS in England. The package of system reforms included NHS foundation trusts, as part of developing pluralism in delivery, and three other significant developments not present in the first ‘internal market’ of the 1990s (Secretary of State for Health (2002); Audit Commission and Healthcare Commission, 2008; Bevan and Skellern, 2011):

• Independent Sector Treatment Centres (ISTCs) were introduced to provide diagnostic and elective services, intended to create an independent sector market that delivers value for money.

• The choice policy aimed to enable patients, in conjunction with their GP, to decide where and how elective care is provided.

• Although both internal markets promised that ‘money would follow the patient’, only in the second did treating more cases, or a more complex case-mix, guarantee increased payments to hospitals. Payments were determined by the system of Payment by Results (PbR) so that hospitals were paid a fixed tariff (based on estimated national average costs) for different types of cases (defined by Healthcare Resource Groups) (Department of Health, 2002; Dixon, 2004). This also meant that as tariffs were fixed, hospital competition should be focused on quality, and not on both price and quality as had been allowed in the ‘internal market’ of the 1990s.
A joint Report by the Audit Commission and Healthcare Commission (2008) raised concerns about the two policies that aimed to develop the supply-side flexibility that had been lacking in the ‘internal market’ of the 1990s: there was no system of governance to ensure supply across health economies; and ISTCs were allowed generous contracts: they, exceptionally, had payments guaranteed regardless of numbers of patients treated. This meant that they were more costly than NHS providers. A review of the literature by Brereton and Vasoodaven (2010, p. 48) on the ‘internal markets’ of the 1990s and 2000s concluded that although both sets of reforms had “had unmistakable effects on the culture of the NHS”, they “have not been proven to bring about all benefits classical economic theory attributes to markets”. They asked whether the NHS was incurring the transaction costs of markets without the benefits.

Mays and Dixon (2011) summarised the findings from an independent research programme funded by the Department of Health. The strongest evidence they report of the impact of the various elements of system reform is from the evaluation of PbR by Farrar and others (2009), who compared efficiency of different types of hospitals in England with hospitals in Scotland over the period from 2003/04 to 2005/06, and estimated that efficiency had increased slightly faster in England than in Scotland. They tentatively concluded that reductions in hospital costs in England had been achieved by increases in efficiency rather than reductions in quality. Mays and Dixon (2011) highlighted the weakness of commissioning, which had been undermined by “near-continuous reorganisation while the other market mechanisms were being rolled out”; and the fact that commissioners suffered from the structural limitation that they lacked direct control over clinical decisions (such as GP referrals), which meant that they “struggled to manage spending” (p. 130). Mays and Dixon contrasted the findings from two kinds of evidence: between local studies of participants’ perceptions of change, which tended to suggest the market reforms had had relatively little impact (for example, Audit Commission and Healthcare Commission, 2008; Brereton and Vasoodaven, 2010); and national econometric studies that showed impacts that were modest in scale, but statistically significant (Gaynor and others, 2010; Bloom and others, 2010; Cooper and others, 2011) and appeared to show, controversially, that the policy of introducing competition between hospitals had improved quality of care (Mays, 2011; Stevens, 2011; Bevan and Skellern, 2011; Pollock and others, 2102; Propper, 2012).

The market reforms proposed in 2002 largely took effect from 2006, which was when another system of performance reporting against government targets was introduced to replace ‘star ratings’, namely the annual ‘Health Check’, which was published from 2005/06 to 2008/09. The Audit Commission and Healthcare Commission (2008), and Mays and Dixon (2011), both argued that the key driver of improvements in this period, given increased funding, was more likely to have been performance management rather than the pressure of provider competition. Preston (2009) argued in The Guardian that the use of central targets to put pressure on providers to improve services had paradoxically become unpopular just when there was the strongest evidence of its beneficial outcomes. A Times leader (Anonymous, 2009) commented on the 2009 Health Check by the Care Quality Commission (2009) and contrasted stories about possible ‘gaming’ with strong evidence of dramatic improvements in reported performance. The Coalition Government, elected in May 2010, sought to abandon governance by targets in favour of re-invigorating provider competition (Secretary of State for Health, 2010).
In 2011, I predicted that: “the combination of fiscal pressures on the NHS budget, organisational turmoil, and the attempt to regulate providers through self-assurance with local scrutiny and risk-based and random inspections will result in further hospital scandals. ... with hindsight, the era of target-driven performance and related regulation in the NHS in the 2000s will come to look like a golden age in comparison” (Bevan, 2011, p. 111). The regulator of quality of care in the NHS, the Care Quality Commission (CQC), was given a massive regulatory burden of health and social care. The year 2011 was an *annus horribilis* for CQC. It was criticised for weaknesses in what ought to have been its core function of inspection of quality of care by: a Panorama programme, which exposed the scandal of serious abuse of patients by staff at Winterbourne View, a private hospital registered for the purpose of providing assessment, treatment and rehabilitation for people with learning disabilities (see Department of Health, 2012); the House of Commons Health Committee (2011, p. 2) for making “the decision to divert resources from inspection and review activities towards meeting the demands placed on it by the process of registering providers”; and by the National Audit Office (2011, p. 8) for having “completed only 47 per cent of the planned number of reviews between October 2010 and March 2011”. In 2012, the CQC lost its chief executive in February (Campbell, 2012) and its chair in September (Brindle, 2012). A new regulatory unit was subsequently created in 2013 within the CQC headed by a Chief Inspector of Hospitals for the NHS using a system of inspections similar to that used by the former quality regulator, the Commission for Health Improvement (CHI), based on peer reviews (which will include patients) and organised around visits “lasting more than a week” (Campbell, 2013). The following sections outline the increasingly distinctive NHS policies developed by the governments of the devolved countries.

**Scotland**

Greer (2004, pp. 63-91) identified Scotland’s distinguishing characteristics to be a strong sense of national identity; a long tradition of high-status medical professionals closely connected to the policy process; large, scarcely populated rural areas; and relatively high levels of poor health and deprivation. Greer points out that following devolution, the SNP Opposition in the Scottish Parliament was broadly to the left of the Labour Government on key issues such as finance and long-term care (see below), and was capable of “making news” (p. 76). Steel and Cylus (2012) point out that: “The electoral system of the Scottish Parliament was in part designed to prevent any one party achieving hegemony”. This design worked for the first three elections: in 1999 and 2003 there were Labour/Liberal coalitions, and in 2007 there was a minority SNP administration. But, in the May 2011 election, Alex Salmond led the SNP to a historic victory (Bolger, 2011). That success was followed by the referendum, on 18 September 2014, on the question ‘Should Scotland be an independent country?’ (BBC News, 2013a) with 55 per cent voting ‘No’ (BBC News, 2014).

The policy of abolishing the ‘internal market’ featured in the 1997 manifesto of the Scottish Labour Party and became policy in the 1997 White Paper, *Designed to Care: Renewing the NHS in Scotland* (Scottish Office, 1997). Steel and Cylus (2012, p. 111) point out that it was published in the same week as the English White Paper, *The New NHS* (Secretary of State for Health, 1997); and although both sets of policies abolished the idea of competition and GP fundholding, the government in England retained the purchaser/provider split, but the government in Scotland
abolished it by making NHS trusts accountable to health boards. In 2000, the year of the English White Paper, *The NHS Plan* (Secretary of State for Health, 2000), which introduced the regime that became ‘star ratings’, the Scottish Government published *Our National Health: A plan for action, a plan for change* (Scottish Executive, 2000), which Greer (2004, p. 80) argues was mainly composed of lists of objectives, but also included proposals for reorganising the Scottish NHS. The NHS boards, acute hospital trusts and PCTs were brought together into 15 unified boards; and the number of ministerial appointees was reduced by one third, with the objective of leading to a greater reliance on professionals who would, it was envisaged, work together for the benefit of population health.

The White Paper, 13 *Partnership for Care: Scotland’s health White Paper* (Scottish Executive, 2003), was described by one Scottish journalist as “the mystery of the pointless White Paper”, which was published “on the same day as some bad waiting list figures and never subject to much attention” (Greer, 2004, p. 80). Greer (2004) points out, however, that it finally expunged from the NHS in Scotland all ideas of purchaser/provider split, competition and contracting: providers ceased to be independent of the purchasing side of the NHS and were organised into 15 unitary health boards. The White Paper also introduced managed clinical networks (for example based on cancers). In 2005, following the Kerr Report on the NHS in Scotland (Scottish Executive, 2005a), the Government’s response, *Delivering for Health* (Scottish Executive, 2005b), is described by Steel and Cylus (2012, p. 114) as setting out: “a programme for action designed to shift the balance of care away from episodic, acute care in hospitals, increasingly through emergency admissions, to a system that emphasizes preventive medicine, support for self-care, and greater targeting of resources on those at greatest risk through anticipatory medicine”.

Propper and others (2010, p. 320) described the regime for managing reduction of hospital waiting times in Scotland from 2000, as one in which there was no “naming and shaming”, and neither “the coupling of performance against targets and managerial sanctions that operated in England”. Farrar and others (2004, pp. 4, 20–1) observed that in Scotland, there was the perception of “perverse incentives … where ‘failing’ Boards are ‘bailed out’ with extra cash and those managing their finances well are not incentivised”. Hence the model of governance in the initial period following devolution was one of ‘trust and altruism’. Steel and Cylus (2012, p. 113) state that, following: “Unfavourable cross-border comparisons … about performance, particularly on waiting times… following a change in minister”, the government introduced a “tougher and more sophisticated approach to performance management” in the system of HEAT targets, which is organised by a delivery group within the health department, which agrees annual local delivery plans with each board, and systematically monitors and supports improvements in performance and, where necessary, intervenes (p. 114). The HEAT targets are organised in four domains (hence the acronym): 13

- Health improvement for the people of Scotland – improving life expectancy and healthy life expectancy.
- Efficiency and governance improvements – continually improving the efficiency and effectiveness of the NHS.
- Access to services – recognising patients’ needs for quicker and easier use of NHS services.
• Treatment appropriate to individuals – ensuring patients receive high-quality services that meet their needs.

This suggests that the model of governance in Scotland has moved to some extent away from that of ‘trust and altruism’ towards one of ‘targets and terror’. Information on performance by board against the HEAT targets is published,\textsuperscript{14} which also indicates a move towards ‘naming and shaming’, but this information is not organised in an explicit ranking system (as in ‘star ratings’).

In 2003, NHS Quality Improvement Scotland was established. It is described as having two key roles: “supporting, ensuring and monitoring the quality of healthcare provided by the National Health Service in Scotland ..., providing quality assurance and accreditation”; and “the evaluation and provision of advice to the National Health Service in Scotland on the clinical and cost-effectiveness of new and existing health technologies including drugs”.\textsuperscript{15} In 2009, the Healthcare Environment Inspectorate (HEI) was established to inspect all acute hospitals across NHS Scotland every three years. The Chief Inspector produces an annual report, and for 2011/12 reported that the Inspectorate had moved towards unannounced inspections, which were 80 per cent of visits in that year; a balance set to continue (Healthcare Environment Inspectorate, 2013). This report gives a sense of an organisation that is confident in its developing approach to inspecting quality in the Scottish NHS. In 2011, NHS Quality Improvement Scotland became Healthcare Improvement Scotland (HIS), which also took responsibility for Scotland’s small independent health care sector: Timmins (2013) notes that: “In 2010, less than 1 per cent of the budget in Scotland was spent in the private sector, against around 5 per cent and rising in England”. HIS describes its work programme as:

• Advice, guidance and standards – collaboration with international experts to identify evidence for improvement.

• Improvement and implementation support – to feed local improvement cycles.

• Assurance, scrutiny, measurement and reporting – to give local and public assurance and demonstrate accountability.\textsuperscript{16}

Following the 2007 election, the minority SNP Government maintained the organisational structure of the NHS, but introduced the overarching concept of ‘mutuality’. Steel and Cylus (2012) see its strategy document, \textit{Better Health, Better Care: Action plan} (Scottish Government, 2007) as confirming and extending the “direction of travel... since devolution” and confirming “the diverging paths of the NHS in Scotland and England in its rejection of solutions based on market forces or internal competition” (p. 115). They also see policies for the NHS in Scotland to be consistent with the Commission on the Future Delivery of Public Services (Christie, 2011), which was appointed by the Scottish Government in 2010 to consider how these services need to develop to respond to the fiscal pressures of austerity and called for substantial reform of public services to make them “outcome-focussed, integrated and collaborative”, “transparent, community-driven and designed around users’ needs”, and that they “should focus on prevention and early intervention” (Steel and Cylus, 2012, p. 116). The NHS in Scotland, 15 years after devolution, has developed its own distinctive form with a clear policy direction and objectives, and a stable organisational hierarchical structure (Steel and Cylus, 2012, p. 20).
Wales

Greer (2004, pp. 143-145) describes the political conflict in the Welsh Assembly as similar to that of the Scottish Parliament, with both differing from Westminster. Longley and others (2012, p. 4) state: “The post devolution governments in Wales have all been dominated by the left-leaning Labour Party, either acting alone, or in coalition with the centrist Liberals, or the nationalist party, Plaid Cymru”.

Prior to devolution, Greer (2004) points out that the Welsh Office, under the leadership of John Wyn Owen, Head of the NHS in Wales, developed a distinctive policy at the end of the 1980s as articulated in Strategic Direction for the NHS in Wales (Welsh Office, 1989). This policy aimed to develop services to achieve health gain, and be patient-centred and resource-effective. A report by the National Audit Office (1996), which evaluated this policy some seven years after its launch, concluded that: “The impact on services has been marginal so far” (p. 4). This was attributed to a lack of clarity in the status of the policy, given other later policies (including the ‘internal market’), weaknesses in the way targets had been set, weaknesses in arrangements for monitoring progress towards them (more than half of the 180 health gain targets could not be measured at the start of the initiative; p. 18), and problems in releasing savings (p. 5). But Greer (2004) pointed out that this policy had a profound effect on the thinking of those in government and the NHS in Wales, as they came to recognise that health was “an outcome to which health services contribute, rather than as a product of health services” and even in 2002 some “administrators and professionals thought the strategies were still the basis of Welsh health services” (p. 136). This legacy can also be seen to have shaped the focus of policies after devolution on health rather than health care (see below). Greer (2004, p. 13) argues that, in the absence of other constituencies (such as elites of clinical medicine, academics and think tanks), the most important political influence on the Welsh Assembly is local government, and hence he argues that the emphasis in the NHS in Wales is on localism.

The first policy for the NHS in Wales following devolution was the Welsh Health Plan (Welsh Assembly, 2001), Improving Health in Wales, “written against the background of unprecedented levels of increase in health funding” (p. 2). This set out a distinctive emphasis for health policy in Wales, a proposed reorganisation of the NHS in Wales and a new approach to performance management. Greer (2008, p. 80) highlights how policies in Wales differed from the other UK countries with its objective of improving public health in a plan that “focused on health rather than the NHS and did not confine itself to the Welsh NHS”. He argues, however, that this encountered the problem that the wider determinants “that were amenable to government action were in other departments (such as education) or in the hands of the UK government (taxes and benefits)” (p. 79), which meant that the ambitious Welsh public health agenda was reduced to campaigns giving lifestyle advice. The Welsh Health Plan of 2001 had also proposed to develop an effective system of performance management (p. 59) and included targets for national service frameworks for priority services (p. 22) and for reducing hospital waiting times: “reduce waiting times year on year until patients in Wales receive services as speedily as elsewhere” (p. 60).

The Review of Health and Social Care in Wales was the outcome of the review led by Derek Wanless (National Assembly for Wales, 2003). This followed his report for the Chancellor of the Exchequer of the UK on his review of the long-term trends affecting the health service across the UK. He was asked to follow that with
a report specifically on Wales (Wanless, 2002). The objectives of the Welsh review were to examine “how resources should be translated into reform and improved performance” given the policy direction and structural changes heralded in *Improving Health in Wales*, and the government’s commitment to a substantial increase in future resources (National Assembly for Wales, 2003, p. 9). The report (National Assembly for Wales, 2003) highlighted a series of problems:

The health of people in Wales is relatively poor. ... Large and increasing numbers are not receiving social services after assessment, and hospital waiting lists and times are unacceptably long. ... Actions to reconfigure provision, release acute capacity and raise productivity are needed alongside a rebalancing of the system to meet need earlier in the ‘care pathway’ and improvements in the way in which the parts of the system work together... Supply is involuntarily rationed by long hospital waiting lists and times and assessments without subsequent social services. Wales does not get as much out of its health spending as it should. ... Wales has significant capacity shortcomings in workforce planning, Information and Communication Technology (ICT), and estates and is unlikely to achieve the fully engaged scenario for health without (among other things) a sea-change in the quality and nature of its planning and capital and revenue investment. ... there is also widespread under-performance associated with systemic defects – we do not have the performance management and incentive systems to drive properly creation or imitation of best practice (pp. 1–2).

In April 2003, the NHS in Wales was reorganised. The five health authorities were replaced with 22 Local Health Boards (LHBs), which were geographically identical to, and included representatives from, the 22 local authorities. The objectives were to ensure local accountability and to reflect a new emphasis on joint working. This reorganisation required, over the two-year period of implementation: centralisation of some services; regionalisation to three regional offices; and encouraged the 22 LHBs to develop mergers and collaborations (Greer, 2004, p. 149). Greer criticised the reorganisation as having been “radical, untested, stressful and probably beyond the administrative capacity” of the Welsh government (p. 149).

Problems of performance management were highlighted in critical reports by the Auditor General for Wales in 2005 and 2006. In 2005 a report criticised the Welsh Assembly Government for a lack of a clear overall strategy to reduce waiting times and highlighted a series of fundamental weaknesses in the Welsh system of performance management of waiting times: targets for waiting times were included in a large set of targets (over 100 in 2003/04); they lacked clarity and consistency; they were adjusted to reflect variations in local circumstances; and breaches were allowed (Auditor General for Wales, 2005, volume 2, pp. 36–54). The report’s own analysis of expenditure on initiatives to reduce waiting times reinforced “the widespread perception that waiting time initiative funding is a reward for poor performance, with additional funding allocated to the trusts which breach targets” (p. 46). The report argued that the arrangements for performance management of waiting times “have provided neither strong incentives nor sanctions to improve waiting time performance”; and “Trust and Local Health Board chief executives consistently described their perception that the current waiting time performance management regime effectively ‘rewarded failure’ to deliver waiting time targets”. Nor was there any system of ‘naming and shaming’ (pp. 42 and 40). In 2006, the Auditor General for Wales published another heavily critical report on ambulance services in Wales, which it described as having “been let down by problems of strategy, leadership, governance, process, infrastructure, and systems, people and culture” (Auditor General for Wales, 2006, p. 8).
Until 2004, the Commission for Health Improvement (CHI) was the regulator of quality of the NHS in Wales (using the same processes as for the NHS in England). After 2004, CHI was succeeded by Healthcare Inspectorate Wales (HIW), which regulates the NHS and independent sector in Wales.\(^7\) However, the government in Wales still looks to NICE to set clinical standards (Longley and others, 2012, p. 16). An audit by the Royal College of Physicians (2006) highlighted the much worse organisation of stroke care in Wales, which meant that patients treated in Wales were more likely to die from stroke, or if they survived would have higher levels of disability than in England or Northern Ireland. Longley and others (2012, p. 65) point out that, following that critical report, stroke services have been remodelled, there has been progress in meeting targets for crucial elements in the process of care, and stroke mortality has begun to decline.

The policy response to the Welsh Wanless Report was the strategy, Designed for Life (Welsh Assembly Government, 2005), which proposed “a much more radical and focused approach to performance management” (p. 15) and use of “the model of managed clinical networks” (p. 19) – which had been proposed for Scotland in 2003 (see above). Following the formation of a coalition government in 2007, Welsh Labour and Plaid Cymru published an agenda for the government of Wales, which explicitly rejected the use of the private sector and of markets in health care, and, following consultation, went on to abolish the purchaser/provider split. This essentially means that the NHS in Wales is following the model developed in Scotland since 2004. In 2007, the government proposed another reorganisation of the NHS in Wales (Labour and Plaid Cymru, p. 2007), which was implemented in 2009 and reduced the number of LHBs from 21 to seven (Welsh Assembly Government, 2009a). Longley and others (2012, p. 58) point out, however, that there is now concern that these LHBs may be too large and remote from the populations they are meant to serve.

Following that reorganisation, the government developed the Five-Year Service, Workforce and Financial Strategic Framework for NHS Wales (NHS Wales, 2010, p. 7) in which the first of five guiding principles was: “Making LHB and Trust chief executives personally accountable for delivery with strong, transparent, specific performance management”. The government in Wales has introduced a system of performance management around two sets of core delivery targets. Tier 1 targets are described as “key priorities where immediate improvement is necessary or where performance at defined target levels must be sustained” and are “subject to very close attention by health boards, trusts and Welsh Government (through national monitoring)” (NHS Wales, 2011, p. 1). However, Longley and others (2012) point out that in the absence of governance by the model of ‘choice and competition’:

> Wales has been less clear about enunciating its alternative set of metrics and levers for change. The dominant philosophy appears to rely on a combination of exhortation to do better plus performance management of the [local health boards] by the Welsh Government. (2012, p. 66)

It seems that despite successive attempts to improve performance management of the NHS in Wales, the model of governance continues to be one of ‘trust and altruism’. Longley and others (2012, pp. 58–61) summarise a succession of policies that have been developed since 2009 that aim to improve population health (Welsh Assembly Government, 2009b); reduce health inequalities (Welsh Assembly Government, 2011a); modernise health care (Welsh Assembly Government, 2011b); and reform social services in ways that include closer
integration with the NHS (Welsh Assembly Government, 2011c). The government has also responded to a report by the Bevan Commission (2011) with a five-year plan that has set targets for five priority areas: coronary heart disease, cancer, mental health, the health of older people and the health of children (Welsh Government, 2012). Longley and others (2012) note that “developing new policies is not a substitute for making change happen” (p. 67), thus implying the lack of a clear delivery chain.

**Northern Ireland**

Greer (2004) argues that devolution in 1999 ought to have had the greatest impact on Northern Ireland, as this was where policies were likely to have diverged most from what the local population would have wanted as a consequence of “almost thirty years of ‘direct rule’ by part-time Ministers from London, and no accountability to voters” (p. 159) which had “produced an ingrained disrespect for local politicians” (p. 160). Greer argues, however, that devolution had less impact in Northern Ireland than in Scotland and Wales, with relatively little distinctive policy development. Greer (2004, pp. 159–61) attributes this to the nature of politics in Northern Ireland and the legacies of Direct Rule. Indeed devolution was suspended between 2002 and 2007. Greer describes the style of policy-making in Northern Ireland as one of ‘permissive managerialism’, which he defines as “a combination of minimal policy activity (such as quality improvement, new public health, or acute care redesign) and an emphasis on running services” (p. 159) and “Northern Irish health policy changed remarkably little with the advent of devolution and substantial money” (p. 182). He gave as an example that, whereas the three other health systems in the UK from 1997 developed policies that ended the ideas of the internal market, this did not happen in Northern Ireland, and that despite “all the Northern Irish political parties supporting the abolition of fundholding. Nothing happened until December 2001” (p. 183). The problems that Greer identified continued to mean that the process of policy change was slow and had limited impact on the structure and governance of the NHS in Northern Ireland and on tackling its excess supply of acute hospital capacity.

Appleby (2005), in his review of health and social services in Northern Ireland, criticised its system of performance management as lacking “appropriate performance structures, information and clear and effective incentives – rewards and sanctions – at individual, local and national organisational levels to encourage innovation and change” (p. 162). There was also no system of naming and shaming. Indeed, information on performance was neither made public nor part of the performance analysis/management system (p. 165). A survey of chief executives found that there was disagreement over whether the current system of performance management included sanctions or rewards (p. 166). This suggests that in the initial period of devolution the model of governance was one of ‘trust and altruism’.

O’Neill and others (2012, p. 64) argue that the desire to address what Greer has highlighted as the “democratic deficit” resulted in a “proliferation of bodies involved in the delivery and overseeing of public services in Northern Ireland”. Prior to devolution, the heads of the Department of Health and Social Services, and of the Health Service Management Executive, were both accountable to the Secretary of State. There were four Health and Social Service Boards, which were overseen by four Health and Social Service Councils; and 19 Health and Social Service Trusts. The review of public administration, initiated in 2002, and
The impacts of asymmetric devolution on health care in the four countries of the UK

implemented by Westminster (during Direct Rule), examined arrangements for the administration and delivery of all public (not just health) services in Northern Ireland and identified 150 public bodies serving a population of 1.7 million. The review’s final report was published in 2005, but Greer (2008) describes how its proposed changes were brought to a halt following the restoration of devolution, and the arrival of a new minister (p. 80). Reorganisation had to wait until legislation in 2009. The Health and Social Care (Reform) Act (NI) 2009 created one large commissioning body, the Health and Social Care Board, supported by five local commissioning groups organised geographically and five coterminous Health and Social Care Trusts to provide care (O’Neill and others, 2012, pp. 9-14). Although that structure maintains a purchaser/provider split, these bodies are intended to cooperate, and to operate consultatively and without provider competition. Timmins (2013, p. 5) points out that there is relatively little opportunity for competition between providers in Northern Ireland because of its small population, very small private sector, relative isolation from the rest of the UK and limited cross-border flows of patients with the Irish Republic. O’Neill and others (2012, pp. 64-5) too recognise that competition is likely to be ineffective in Northern Ireland, but argue that there is no clear articulation of what will bring discipline to improve provider performance and efficiency in the absence of competition. The chosen model of governance seems strikingly similar to the ‘third way’, which was tried and found wanting for the English NHS between 1997 and 2000 (see above) and was in effect governance by the model of ‘trust and altruism’.

Greer (2004) implies that the arrangements for governing the NHS in Northern Ireland are designed to shelve controversial problems. O’Neill and others (2012, p. 64) state that following the 2009 reorganisation, it is not clear: “Exactly where and by whom decisions are made”, hence there is a lack of an obvious delivery chain for policy implementation. One serious problem facing those running the NHS in Northern Ireland is the excess capacity of acute hospital care. O’Neill and others (2012, pp. 65–6) highlight the stark finding of the review of health and social care (Department of Health, Social Services and Public Safety, 2011), which “pointed out that in England a population of Northern Ireland’s size would be serviced by four acute hospitals rather than the eleven that currently exist”. They also point out that: the recommendations of that review echo those made by another review in 2001 (Department of Health, Social Services and Public Safety, 2001); Northern Ireland has approximately one quarter more acute beds per 100 population than England (Appleby, 2011a), uses them less intensively than England (Connolly and others, 2011; O’Neill and others, 2012, p. 74); and the formula used to guide the allocation of resources to the NHS in Northern Ireland is designed to perpetuate inefficient small hospitals, whereas the formula in England seeks only to fund unavoidable variations in costs (mainly from differences in labour markets; see Bevan, 2009).

In England, Wales and Scotland, there has been a considerable focus on how to regulate the quality of health care. This appears to be much less of an issue in Northern Ireland. The regulation of the quality of public and private health and social care in Northern Ireland is the responsibility of the Regulation and Quality Improvement Authority (RQIA). The RQIA registers and inspects a wide range of health and social care services based on minimum care standards; and aims to assure the quality of services provided by Health and Social Care (HSC) Board, HSC Trusts and agencies.
Notes


17. www.hiw.org.uk/.

5
Comparing governance of health care in the four countries

This chapter outlines differences between the four countries in: models of governance; choice by patients; the approach to the common goal of integration; funding in the period of austerity; and recent reports of financial pressures and concerns over quality. As was explained earlier, prior to devolution, the policies for the NHS in England essentially applied throughout the UK. Table 5.1 shows how far the policies of the devolved governments currently differ from the English policy based with its NHS organised for governance by ‘choice and competition’.

Table 5.1: Current differences in policy and organisational characteristics between the four countries of the UK

<table>
<thead>
<tr>
<th></th>
<th>England</th>
<th>Scotland</th>
<th>Wales</th>
<th>Northern Ireland</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population (million)</td>
<td>50</td>
<td>5</td>
<td>3</td>
<td>1.7</td>
</tr>
<tr>
<td>Organisational characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Purchaser/provider split</td>
<td>Yes</td>
<td>Abolished 2004</td>
<td>Abolished 2009</td>
<td>Yes</td>
</tr>
<tr>
<td>• Money follows the patient</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>• Competition between public and private providers</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>• Integration of health and social services</td>
<td>No</td>
<td>Various initiatives</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Models of governance

The initial period of devolution provided a ‘natural experiment’ because, in each country, there were: similar substantial increases in NHS funding; and targets for waiting times for hospitals (Auditor General for Wales, 2005a, p. 16) and for response times to life-threatening emergency calls by ambulances (Bevan and Hamblin, 2009); but only in England did the government abandon the model of ‘trust and altruism’. Neither the NHS in Scotland nor Wales was subject to the Treasury’s PSA targets (Connolly and others, 2011). Timmins (2013, p. 13) observes that, although the Department of Health, Social Services and Public Safety of Northern Ireland performed comparatively poorly against its PSA targets when these were reviewed in 2011, “it is not clear that any penalties were incurred, or indeed that much action followed, as a result of that poor performance”. In addition, the NHS in each devolved country was not subject to scrutiny by the Prime Minister’s Delivery Unit, or the threats of middle-class ‘exit’ to independent health care providers (and schools) that is present in large parts of England. Given these differences in performance management in the initial period of devolution, not surprisingly, studies that examined performance across the four countries over that period, in terms of hospital waiting times and ambulance response times,
found that the NHS in England performed best. However, after the end of the regime of ‘star ratings’ in 2006, the differences in models of governance become more complex. In England, from 2006 to 2010, the annual Health Check replaced the regime of ‘star ratings’ with elements of ‘naming and shaming’ and ‘targets and terror’, but the government emphasised the re-introduction of the model of ‘choice and competition’, in the form of a revised ‘internal market’; and from 2010, the Coalition Government stopped publication of the annual Health Check and hence appeared to be relying solely on the model of ‘choice and competition’. The shifts in policy in 1997, 2000 and from 2002, on the English NHS, have been compounded by the insatiable appetite of successive Secretaries of State for structural re-organisations. Timmins (2013, p. 6) observes that this has reached the point at which “‘organisation, re-organisation and re-disorganisation’ might almost be dubbed the English NHS ‘disease’”. All three devolved governments have abandoned governance by ‘choice and competition’, but the accounts of the different countries suggest that only in Scotland has the government developed an effective alternative in a system of strong performance management around HEAT targets (Steel and Cylus, 2012, p. 113; Longley and others, 2012, p. 66; O’Neill and others, 2012, pp. 64–5).

Patient choice

As England alone has opted for the model of ‘choice and competition’, this raises the question of how choice varies across the four countries. Table 5.2 summarises a study across the four countries undertaken in 2008 and 2009 by Peckham and others (2012) on the development and implementation of policies related to patient choice. This showed that patient choice, as an aspect of the responsiveness of the system to the needs of individual patients, was apparent in the operation of all four countries. However, the only elements of choice that potentially applied in all four countries were over the time and date of appointment, and site within NHS trust; and, except in Northern Ireland, choice of specialist, both at the discretion of the provider organisation. In practice, in each of the four countries, the study found that there was a lack of clarity about the options available and in the way choices were offered to patients, with limited discussion of choices between referrers and patients, and tension between offering choice and managing waiting lists. Only in England did patients have a guaranteed choice of provider organisation with a system of funding in which ‘money followed the patient’ (known misleadingly as ‘Payment by Results’ or PbR) (Department of Health, 2002; Dixon, 2004).

<table>
<thead>
<tr>
<th>Choices</th>
<th>England</th>
<th>Scotland</th>
<th>Wales</th>
<th>Northern Ireland</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provider</td>
<td>Yes</td>
<td>No (exception basis only)</td>
<td>No (exception basis only)</td>
<td>No (exception basis only)</td>
</tr>
<tr>
<td>Consultant</td>
<td>At provider’s discretion</td>
<td>At provider’s discretion</td>
<td>At provider’s discretion</td>
<td>No</td>
</tr>
<tr>
<td>Time/date</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Site</td>
<td>At provider’s discretion</td>
<td>At provider’s discretion</td>
<td>At provider’s discretion</td>
<td>At provider’s discretion</td>
</tr>
</tbody>
</table>

Source: Peckham and others, 2012.
Integration

A justification used by governments in Scotland and Wales for abolishing the purchaser/provider split has been to enable the development of integrated care. However, this objective, particularly the integration of health and social services, is common to all four governments, and is argued as being necessary for better organised care for an ageing population. Curry and Ham (2010) describe three levels at which integration may occur for providers, either together or with commissioners: *macro*, at the level of populations, for example, as in the Kaiser Permanente health maintenance organisation (described as Kaiser) in the US; *meso* at the level of a particular care group or population with the same disease or condition, for example, as in the managed clinical networks in Scotland; and *micro*, at the level of individual service users and their carers. Goodwin and others (2012, p. 2) point out that integration at the macro level “appears to be neither necessary nor sufficient to deliver the benefits of integrated care”. Feachem and others (2002) compared the NHS in the UK (for about 60 million) with Kaiser (for about six million Californians) and reported that the two systems used similar levels of resources per head, but that Kaiser performed substantially better than the NHS with, for example, faster access to both primary and secondary care doctors. In a follow-up study, Ham and others (2003) found that the use of hospital beds in the NHS for 11 leading causes of admission was three and a half times that of Kaiser’s standardised rate, because of Kaiser’s combination of low admission rates and relatively short stays. They concluded that the NHS in England could learn in general from Kaiser’s integrated approach, which links funding with provision, inpatient with outpatient care, and prevention, diagnosis and treatment (in particular, for the effective management of chronic diseases).

Studies of integration suggest reductions in use of hospital beds over time in England: in the long-standing pilot of integrated health and social care in Torbay (Thistlethwaite, 2011); and on a country-wide basis in Scotland and in Wales (Ham and others, 2013). But none of these developments has been subjected to a well-designed comparative evaluation. A review by Bardsley and others (2013) of controlled studies of pilots of recent attempts to develop integrated care in England found that for none was there good evidence of a reduction in rates of emergency hospital admissions. In contrast, they did find good evidence of the impact of the Marie Curie nursing service at the end of life, “a well-established and widely used model of care service delivered in a standard way” (p. 12). Furthermore, Ham and others (2013) concluded that: “Despite having the longest history of administratively integrated care, Northern Ireland has been slowest to exploit the potential benefits” (p. 78). Heenan (2013, p. 5) explains this in terms of the distinctive problems of governing Northern Ireland, which “already had sufficient political controversy without adding reform of health and social care to the mix”. Indeed, as explained above, the reason for moving social services out of local government in 1973 was not to achieve integration with health services, but because they were part of systematic discrimination by local governments (Greer, 2004, p. 165). Although the organisational and policy environment in Scotland and Wales ought to enable integration, Ham and others (2013, p. 79) point out that in each of these countries, “it has been difficult so far to shift resources within local health boards from hospitals to the community, let alone from health to social care”.
The pressures of austerity

The Organisation for Economic Co-operation and Development (OECD, 2013) reported that “Health spending has fallen in the United Kingdom in 2010 and 2011 for the first time since the 1970s... spending in real terms per head fell by 1.1 per cent in 2011, following a 2.5 per cent decline in 2010. It had increased by an average of 5.3 per cent per year over the previous decade” (p. 1). Appleby’s analysis of governments’ policies for expenditure on health services from 2010/11 identified Wales as an outlier with planned reductions in NHS spending of nearly 11 per cent by 2013/14 (Appleby, 2011b). Timmins (2013) questioned the ability of the government in Wales to achieve its planned reduction in spending, citing the evidence from the House of Commons Health Select Committee that “no country has managed to keep spending on healthcare flat for four years let alone cut it” (p. 7). It is hard to see how such reductions could be achieved without rationalisation of acute hospital services, and Longley and others (2012) state that the NHS in Wales is embarking on proposals to do that, but these are, as expected, proving controversial and facing vocal opposition. In October 2013, the Finance Minister Jane Hutt announced an increase of nearly 3.6 per cent in cash terms (and 1.7 per cent in real terms) to “help the NHS in Wales to avoid a scandal such as the one in Stafford Hospital” (BBC News, 2013b). Although there is no systematic basis for comparing how each country is managing its NHS in the period of austerity, there are reports that indicate pressures and concerns over meeting targets and quality of care.

Each country has experienced problems with accident and emergency (A&E) services. In England, the House of Commons Health Committee (2013) reported on “the failure of emergency departments to meet national waiting time targets in the early months of 2013”, which was attributed to “a broader failure resulting from fragmented provision of emergency and urgent care and a structure that is confusing to patients”. In Wales, the Health and Social Services Committee (2013) reported in December 2013 that: “Waiting times at hospital emergency departments have generally increased over recent years, with some patients, particularly older people, spending longer than 12 hours in these departments” (p. 5). News reports suggested improvements in the second half of the winter in early 2014 in England (Triggle, 2014) and Scotland (Puttick, 2014), but serious problems were reported at hospitals in Wales (BBC News, 2014b) and Northern Ireland (BBC News, 2014c).

There have been public concerns over quality of care and inspection regimes in England, Scotland and Wales. In England, following the publication of the report of the (second) Francis Inquiry into the scandal at Mid Staffordshire NHS Foundation Trust, in February 2013 (Francis, 2013), Sir Bruce Keogh, the Medical Director of the NHS, was asked by the Secretary of State for Health and the Prime Minister “to conduct a review into the quality of care and treatment provided by hospital trusts with persistently high mortality rates” (Keogh, 2013, p. 3). He selected 14 trusts with high rates following inspections and 11 of these were placed in special measures. The new Chief Inspector of Hospitals, Sir Mike Richards, recommended two other NHS trusts be put in special measures in November and December 2013 (Care Quality Commission, 2013a; 2013b). In Scotland, following higher than predicted Hospital Standardised Mortality Ratios in NHS Lanarkshire, the Cabinet Secretary for Health and Wellbeing commissioned Healthcare Improvement Scotland to undertake a rapid review of the safety and quality of care for acute adult patients in their hospitals. The report from that review...
made 21 recommendations, including on better management of data, medical and nurse staffing levels, and handling complaints (Healthcare Improvement Scotland, 2013). In Wales, a review of 2013 by the BBC Wales health correspondent highlighted a series of concerns over poor care in various hospitals in Wales, which included cases of *Clostridium difficile*, high death rates on cardiac waiting lists, and avoidable deaths of eight patients with liver disease (Clarke, 2013). In a review by the Welsh Health and Social Services Committee of the HIW, the independent inspectorate and regulator of all health care in Wales, the chief executive of HIW was unable to reassure the committee that there was no likelihood of a scandal in Wales like that of Mid Staffordshire in England (BBC News, 2013c).

---

**Notes**

19. PbR creates financial incentives for hospitals to reduce costs and increase the numbers of cases they treat, with concerns over incentives to skimp on quality or discharge patients too early. Farrar and others (2007; 2009) compared efficiency of different types of hospitals in England with hospitals in Scotland from 2003/04 to 2005/06. In Scotland, there was no tariff system for funding hospitals (other than for cross-boundary flows), and hence only weak financial incentives to reduce costs and treat more numbers. They estimated that efficiency had increased slightly faster in England than in Scotland. They tentatively concluded that reductions in hospital costs in England had been achieved by increases in efficiency rather than reductions in quality.

20. This study standardised for age and sex, but did not take account of differences in morbidity (Talbot-Smith and others, 2004).

21. In Northern Ireland, there is little that is reported publicly over concerns over quality of care. The Inquiry into Hyponatraemia-related Deaths after the deaths of five children in Northern Ireland hospitals between 1995 and 2003 is expected to report in 2014 (Inquiry into Hyponatraemia-related Deaths, 2014).
6

Policy outcomes of asymmetric devolution

This report ends by highlighting the outcomes of asymmetric devolution as illustrated by the differences in what has happened in each NHS over nearly 15 years of devolution.

The absence of an English Parliament

The absence of an English Parliament gives rise to the ‘West Lothian question’ outlined earlier. This has implications for the NHS. For example, Greer and Trench (2008, pp. 21–2) describe how the Blair Government secured a majority in July 2000 for the vote on NHS foundation trusts, which was opposed by the majority of English MPs, by relying on the party loyalty of Labour MPs from Scotland and Wales, who would not face any hostile constituency reaction as their (Labour-led) governments were against that policy. They also point out that “the Scottish National Party, which normally abstains on England-only policies, chose to vote against foundation hospitals, with the declared justification that it might have Barnett Formula implications”.

Hazel (2000, p. 36) argued that a pragmatic solution to the ‘West Lothian question’ would be for Scottish and Welsh MPs to stay away from matters and debates on public services for England, and that: “Some may follow the self-denying ordinance of Tam Dalyell and decline to vote on such issues”. The UK Government’s Coalition Agreement of 2010 made a commitment to “establish a commission to consider ‘the West Lothian question’” (HM Government, 2010, p. 27). This Commission reported in March 2013, and recommended that “decisions at the United Kingdom level with a separate and distinct effect for England (or for England-and-Wales) should normally be taken only with the consent of a majority of MPs for constituencies in England (or England-and-Wales)” (McKay, 2013, pp. 8–9). Hence the McKay Commission has recommended, in effect, to enforce Tam Dalyell’s self-denying ordinance. Wright (2013) reported that this principle is “part of sweeping constitutional reforms being drawn up by ministers”. If this were implemented, then it would raise the issue of MPs at Westminster having different voting rights.

What is reserved for the UK and what is devolved?

Cairney (2007, p. 77) points out that: “An unanticipated consequence of devolution is the often fluid boundary” between matters that are reserved for the UK and those that are devolved. He argues that this was powerfully illustrated by the way that the official decision to legislate to ban smoking in workplaces was made in Scotland at the end of 2004, when this went against the government position in England, and it was made clear in the concordat between the Health and Safety Executive of the UK and the Scottish Executive that such legislation would impinge on reserved territory. Cairney describes the enacted legislation as
marking “one of the most significant policy divergences between Scotland and England since devolution”. This was “one of the Scottish Executive’s ‘flagship’ policies’ and seen by the First Minister, Jack McConnell, as the most important achievement in his term of office” (p. 73). Cairney also argues that the smoking ban shows the capacity of Members of the Scottish Parliament (MSPs) who are not members of the government to shape legislation: in January 2004, McConnell had “voiced public opposition to comprehensive legislation” (Cairney, 2007 p. 73). Before devolution, the UK government could impose the disastrous poll tax on Scotland; after devolution, the success of the smoking ban in Scotland influenced similar legislation in England, and indeed: “By 2006 all four countries in the UK had legislated to ban smoking in public places” (Cairney, 2007 p. 73).

**Charges**

One way of illustrating the way asymmetric devolution has allowed differences in charges for each NHS is by considering two fictional families, the Stewarts and the Smiths. Each family has a child away at university and an elderly mother who moved to be close to them. William and Catherine Stewart are Scots, who now live in London, and have to pay for their prescriptions; pay for the social care of Catherine’s elderly mother, Agnes; and their son Jack has taken out a loan for his tuition fees for his four-year undergraduate degree at Edinburgh University of about £36,000. In contrast, Richard and Elizabeth Smith, who are English, and now live in Edinburgh, have free prescriptions; free personal social care for Richard’s elderly mother, Mary; and their son Martin is enjoying free tuition fees for his undergraduate degree at Edinburgh University. It seems that in Canada, for example, differences in charging for health care would be proscribed by the Canada Health Act. Table 6.1 shows the differences in charges for health and personal social services between the four countries.

<table>
<thead>
<tr>
<th>Charges and entitlements</th>
<th>England</th>
<th>Scotland</th>
<th>Wales</th>
<th>Northern Ireland</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free personal social care for the over 65s</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Free prescriptions</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

The first country to introduce free prescriptions was Wales. The Labour Manifesto for the 2002 elections to the Welsh Assembly had proposed the abolition of prescription charges as a symbol of the differences in policies between Wales and England, and part of “what the First Minister, Rhodri Morgan, had referred to before the election as the ‘clear red water’ between old and new Labour” (Longley, 2004, p. 10). This policy was announced in 2004. Prescription charges were also abolished in Northern Ireland, in 2010, and in Scotland, in 2011. Timmins (2013, p. 8) points out: “it was clear in both cases that overall health spending was coming under severe pressure” and “In Northern Ireland, abolition went ahead despite advice from the civil service not to proceed”. In England, about 90 per cent are exempt from these charges, but for those who do have to pay, the charge is nearly £8. Timmins (2013) raises the question of whether at a time of austerity, prescription charges might be re-introduced (p. 18). The governments of Wales...
and Scotland have also abolished charges for parking cars at hospitals (unless these are required under contracts) (Timmins, 2013, p. 8).

The implementation of free personal and nursing care for people aged 65 years and over in Scotland further illustrates how fundamental policy differences can be implemented in the absence of arrangements for federal governance among the UK countries and with funding by the Barnett Formula (Greer, 2004, pp. 87-90). In 1997, before devolution, the New Labour Government asked Professor Sir Stuart Sutherland to chair a Royal Commission on Long-Term Care for older people. Although the Sutherland Report (Sutherland, 1999) recommended free universal provision of personal care for older people, the UK government’s response was to introduce means testing for this care; this was also the policy in Scotland when the report was published (Scott, 2000). In January 2001, however, Ritchie and Dinwoodie (2001) describe how a majority of MSPs “united to assert the supremacy of the parliament over the executive” and voted for free universal provision in “unprecedented scenes of passion”. Hence, this decision, like the ban on smoking in workplaces, is another illustration of the power of MSPs. This policy was estimated to have an annual cost of about £110 million for Scotland (Scott, 2000) and is said to have been funded as a ‘Barnett consequential’ from the massive increases in spending on the NHS in England that followed Tony Blair’s commitment to increase NHS spending to EU levels (Greer, 2004, p. 89). Greer (2004, p. 88) reports one Scottish Labour politician being told ‘angrily’ by English colleagues that the higher funding of public services was “because of the deprivation you have in Scotland, not so you can do things we can’t do”. But, as Timmins (2013) points out, the costs of this ‘free’ care have more than doubled (in cash terms) in seven years from £219 million to £450 million, and there are now waiting lists for access to it (Scottish Government, 2012).

**Funding**

Commissions have been appointed and reported on the original arrangements for devolution in Scotland and Wales, which have criticised, and recommended changes to, the limited revenue-raising power of the Scottish Parliament and the absence of any such power of the Welsh Assembly. The report from the Commission on Scottish Devolution (Calman, 2009, p. 8) recommended changes in which devolved tax revenue would be substituted for some of the block grant from the UK Parliament to allow the Scottish Parliament to decide on taxes for devolved current spending, which would allow over one third of that spending to be funded by taxes decided and raised in Scotland. The Scotland Act 2012 devolved a number of new taxation powers to the Scottish Parliament, including the provision to set a new Scottish rate of income tax from 2016 (BBC News, 2012). The UK Government’s Coalition Agreement made a commitment to establish a Commission similar to the Calman Commission to review arrangements for devolution in Wales (HM Government, 2010, p. 28). The first report from the Commission on Devolution in Wales (Silk, 2012, p. 4) recommended “that the Welsh Government should share responsibility for income tax at all rates with the UK Government”. Wright (2013) reported that “as part of further constitutional changes, ministers are also likely to agree to devolve many more powers to the Welsh Assembly, including the ability to vary income tax rates, alter stamp duty and borrow money backed by Treasury guarantees”. The UK government has confirmed it will consult businesses on one of the recommendations of the Silk Commission, namely that the Welsh government should get control of stamp duty.
land tax (BBC News, 2013d). It was mentioned earlier that working out finance in devolved countries is a ‘wicked problem’, as experienced in financing Ireland during Home Rule in the nineteenth century, because “the areas on which most money needed to be spent had the lowest tax base” (McLean, 2000, p. 81). Indeed, in health care, this problem of the mismatch between need and capacity to raise the revenue to meet that need was vividly captured by Tudor Hart’s ‘inverse care law, which he argued meant that: “The availability of good medical care tends to vary inversely with the need for it in the population served” (Tudor Hart, 1971). The purpose of formula funding, which is based on the relative needs of populations, is precisely to remedy this ‘inverse care law’ (Bevan, 2009).

Insofar as there is a formal basis for funding public services, this is still through the Barnett Formula, which the Holtham Commission (Holtham, 2009) pointed out had been described by Lord Barnett himself as no more than “a temporary expedient not expected to last ‘a year, or even twenty minutes’” and that:

Politically it was not found possible to secure agreement on a formula that allocated resources on the basis of needs, although the Treasury had devoted time and effort to developing such a formula, the results of which were published in 1979. The outcome was simply to take expenditure per head as it was in Wales, Scotland and Northern Ireland as a baseline. (Holtham, 2009, p. 14)

The lack of weighting for need was one reason why the House of Lords Select Committee on the Barnett Formula (2009) concluded that it should “no longer be used to determine annual increases in the block grant for the United Kingdom’s devolved administrations” (p. 7) and that “the resulting per capita allocations are arbitrary and unfair” (p. 8) between countries. The Select Committee recommended that (p. 8):

Public spending per head of population should be allocated across the United Kingdom on the basis of relative need, so that those parts of the United Kingdom, which have a greater need receive more public funds to help them pay for the additional levels of public services they require as a result.

The Select Committee also recommended that (p. 8): “the Commonwealth Grants Commission (CGC) in Australia offers a useful institutional model of an independent body that has responsibility for making recommendations about the allocation of finance” and that: “An independent body, similar to the CGC, should be established in the United Kingdom... to recommend the allocation of public monies based on population and through a new needs based formula”. The first report from the Holtham Commission (Holtham, 2009) recognised that the Barnett Formula “must ultimately be superseded by a needs-based formula” (p. 30), but recognised that such a change:

... will need to be accompanied by an adjustment mechanism since the formula may imply substantial changes to block grants and it would be both disruptive and politically difficult to introduce those rapidly. ... [they] would need to be seen to be equitable to all parties. ... [and] therefore be jointly agreed by Ministers from both the UK Government and all the devolved administrations concerned. (p. 30)

The report from the Holtham Commission recognised that “any significant changes to the status quo will require a process of consultation, which will take time and will inevitably induce political difficulties” (p. 30). The government’s response to the Select Committee’s report (HM Treasury, 2009) rejected the proposal to develop a needs-based weighting for the Barnett Formula, so the Barnett Formula still fails to take account of relative needs. Paradoxically, for over
30 years, each country has developed and implemented complex formulas to ensure that within its borders, resources for the NHS have been allocated with the objective of securing equal opportunity of access for equal need (Bevan, 2009). Although measuring relative need is complex and contentious, there is obvious scope to introduce measures that would make the Barnett Formula fairer.

**Accountability**

The House of Lords Select Committee on the Barnett Formula (2009, p. 30) highlighted the problem of inadequate comparable data published by the Treasury:

> Despite its importance, the Treasury only publish limited data about devolved public spending, and the published official data appear in a number of places – in the *Statement of Funding Policy*, the *Public Expenditure Statistical Estimates*, and the annual reports of the Scotland and Wales Offices. Older published data do not distinguish clearly which level of government is responsible – United Kingdom or devolved – for particular spending in the breakdowns published in the *Public Expenditure Statistical Estimates*. There is no time series showing how expenditure has changed as a result of spending decisions made in previous years or spending reviews. It is difficult to establish comparable levels of spending in England for devolved functions as they are different in each part of the United Kingdom.

The Select Committee called for greater transparency and recommended that (p. 30):

> ... the Treasury publish their statistics of the workings of the Barnett Formula, or its successor, in a single, coherent and consistent publication. This annual publication should contain all material data on devolved finance, showing the allocations of grant to the devolved administrations, changes from previous years and explanations for any changes made. We recommend that the statistics be monitored by the UK Statistics Authority.

The government’s response to that report (HM Treasury, 2009) did accept that “the allocations of grant to the devolved administrations, changes from previous years and explanations for any changes made, as provided to the Committee, should be included in PESA (Public Expenditure Statistical Analysis) supplementary material” (p. 6).

I have, however, been unable to find any such supplementary material. Furthermore, what would also be required are not just data for the devolved governments, but also comparable data on expenditure in England.

More recently, the National Audit Office (NAO, 2012, p. 10), in its comparison of the NHS in the four countries, concluded that the current state of data is inadequate both for the retrospective purpose of accountability at the level of each country and for comparing value for money across the four countries:

> we found limited availability and consistency of data across the four nations, restricting the extent to which meaningful comparisons can be made between the health services of the UK. For this reason, and without a single overarching measure of performance, we cannot draw conclusions about which health service is achieving the best value for money.

The NAO also highlighted the potential for learning from the ‘natural experiment’ following devolution:
The shared history and similarities between the four health services mean they offer a natural starting point to better understand the factors that affect value for money and the impact of divergent health policies and systems on performance. We consider there would be value in the four health departments carrying out further comparative work to evaluate the variation in, and understand the drivers of, value for money.

It made recommendations that the health departments should (NAO, 2012 p. 10):

- confirm that there is a desire at a national level to compare performance with a view to learning lessons and identifying good practice
- agree the specific indicators that would provide the most insight
- establish what data would be required to make comparisons and identify how to collect and collate these data proportionately and cost-effectively
- use the comparisons as a starting point to draw out key factors that drive performance and value for money.

Timmins (2013), in his overview of policy in the four countries, makes the same points as the NAO, but in much stronger language, which seems to be driven by his frustration at trying to make comparisons using the separate reports on each country from the European Observatory on Health Systems and Policies (Boyle, 2011; Steel and Cylus, 2012; Longley and others, 2012; O’Neill and others, 2012). He also emphasises the opportunity provided by the UK’s ‘natural experiment’: “From the point of view of anyone interested in policy – politician, civil servant, policy adviser, academic, member of the public – this should be a unique opportunity to compare, contrast and learn. It is an almost perfect test bed” (p. v). The central argument of his paper is that there is far too little comparative work taking place and that: “Something needs to be done to change this” (p. v). Timmins points out that it is difficult to do comparative work using the data that are collected routinely (though not impossible), and that when this was done by Connolly and others (2011), the response to their report “… appeared to be a greater willingness to pick holes in the data, or seek reasons, even excuses, for less good performance rather than confront the fact that there might be a real message here, despite the problems” (p. 1). His concluding section pulls together what can be gleaned from comparative data across the four countries, and leads to what he sees as “the most striking conclusion”, which is that the problems of doing so mean that “there is a huge opportunity going to waste” (p. 22). His concluding statements echoed those of the NAO. In the meantime he argued that “academics, their funders and others should do whatever is possible with what is available. And they should do so without fear of the answers. It is a public duty”.

The empirical report that accompanies this report, The Four Health Systems of the United Kingdom: How do they compare? (Bevan and others, 2014), which has been funded by the Nuffield Trust and the Health Foundation, is our response to his call to arms.
Notes

22. These arrangements become bizarre for students who are from one country and choose to study in another: under European law, students from any EU country are, like Scottish students, entitled to ‘free’ tuition at Scottish universities with three exceptions: students from England, Wales and Northern Ireland are required to pay annual fees, which at Edinburgh were £9,000. See BBC News (2013e).

23. The Scottish Parliament had the power (under Part IV of the Scotland Act) to vary the basic rate of income tax applying in Scotland by up to plus or minus three pence in the pound (Calman, 2009).
The impacts of asymmetric devolution on health care in the four countries of the UK

References


www.timesonline.co.uk/tol/comment/leading_article/article6875253.ece.


Barber M (2007) *Instruction to Deliver: Tony Blair, the public services and the challenge of achieving targets*. Politico.


news.bbc.co.uk/1/hi/uk_politics/3984387.stm.

www.bbc.co.uk/news/uk-scotland-17903145.


www.bbc.co.uk/news/uk-scotland-29270441.


The impacts of asymmetric devolution on health care in the four countries of the UK


Care Quality Commission (2013b) ‘Chief Inspector of Hospitals recommends that Barking, Havering and Redbridge University Hospitals NHS Trust is placed into special measures following in-depth inspection’, 18 December. www.cqc.org.uk/content/chief-inspector-hospitals-recommends-barking-havering-and-redbridge-university-hospitals-nhs.


The impacts of asymmetric devolution on health care in the four countries of the UK


The impacts of asymmetric devolution on health care in the four countries of the UK


Mays N and Dixon J (1996) *Purchaser Plurality in UK Healthcare: Is a consensus emerging and is it the right one?* The King’s Fund.


Timmins N (2013) *The Four UK Health Systems: Learning from each other.* The King’s Fund.


The impacts of asymmetric devolution on health care in the four countries of the UK


Wright O (2013) ‘English revolution in House of Commons: plan to give England’s MPs right to veto on issues not affecting Scotland, Wales or Northern Ireland’, The Independent, 10 July.

About the author

Gwyn Bevan is Professor of Policy Analysis in the Department of Management at the London School of Economics and Political Science. His current research includes: comparisons of performance of health care and schools across the countries of the UK; developing a new approach to priority setting: STAR (socio-technical allocation of resources), which enables stakeholders to understand how best to reallocate resources to benefit patients; and reducing unwarranted variations in health care, with the Wennberg International Collaborative. Gwyn is a member of the Department of Health’s two advisory groups on resource allocation.

Acknowledgements

I am grateful to Benedict Rumbold for letting me see his Summary of Key Developments in Health Care in the Devolved Countries, which covers the period from devolution to 2011. I am also grateful to anonymous referees and Rudolf Klein for general comments, and Vernon Bogdanor and Nicholas Mays for their close reading and detailed comments on earlier drafts. The usual disclaimer applies: any remaining errors are my own.
The impacts of asymmetric devolution on health care in the four countries of the UK

The Nuffield Trust

The Nuffield Trust is an authoritative and independent source of evidence-based research and policy analysis for improving health care in the UK. Our vision is to help provide the objective research and analysis that boosts the quality of health policy and practice, and ultimately improves the health and health care of people in the UK.

The Health Foundation

The Health Foundation is an independent charity working to improve the quality of healthcare in the UK. We are here to support people working in healthcare practice and policy to make lasting improvements to health services. We carry out research and in-depth policy analysis, fund improvement programmes in the NHS, support and develop leaders and share evidence to encourage wider change.

Download this publication:

www.nuffieldtrust.org.uk/compare-uk-health
www.health.org.uk/compareUKhealth

Follow us at:

Twitter.com/nuffieldtrust
Twitter.com/healthfdn