

EXPLAINING NHS DEFICITS, 2003/04 - 2005/06

Corporate Analytical Team Department of Health

DH INFORMATION READER BOX

Policy	Estates
HR / Workforce	Performance
Management	IM & T
Planning	Finance
Clinical	Partnership Working

Document Purpose	For Information						
ROCR Ref:	Gateway Ref: 7883						
Title	Explaining NHS Deficits, 2003/04 - 2005/06						
Author	Chief Economist, Corporate Analytical Team, DH						
Publication Date	20 Feb 2007						
Target Audience	PCT CEs, NHS Trust CEs, SHA CEs, Care Trust CEs, Foundation Trust CEs, PCT PEC Chairs, NHS Trust Board Chairs, Directors of Finance						
Circulation List							
Description	This report is a response to the Director of Finance's commission of the DH's Chief Economist to consider factors that may have caused the emergence of deficits in 04/05 and contributed to their apparently uneven geographical distribution (as stated in the Director of Finance's June 2006 pulication 'NHS Financial Perfomance, 2005/06'						
Cross Ref	Report was commissioned in 'NHS Financial Performance, 2005/06, published in June 2006 Gateway no. 6684						
Superseded Docs							
Action Required	N/A						
Timing	N/A						
Contact Details	Professor Barry McCormick Corporate Analytical Team Room 558C, Skipton House 80 London Road London, SE1 6LH 020 7972 5220						
For Recipient's Use							

CONTENTS

SECTION 1: EXECUTIVE SUMMARY OF THE ISSUES AND FINDINGS	Page 3
SECTION 2: INTRODUCTION	8
SECTION 3: FACTS ABOUT NHS DEFICITS	11
SECTION 4: WORKFORCE GROWTH AND THE HEALTH DIVIDEND: CONSEQUENCES FOR "HEADROOOM" AND DEFICITS	39
SECTION 5: HAS THE REVISED RESOURCE ALLOCATION FORMULA CAUSED THE EMERGENCE OF DEFICITS	53
SECTION 6: COST PRESSURES	63
SECTION 7: THE IMPACT OF SERVICE DELIVERY TARGETS: ELECTIVE AND A&E WAITING, CANCER AND CHD	67
SECTION 8: ACCOUNTING PRACTICES	77
SECTION 9: FINANCIAL MANAGEMENT AND CONTROL	83
SECTION 10: DEVELOPING AN EMPIRICAL MODEL OF THE IN- YEAR DEFICIT	87
SECTION 11: CONCLUDING COMMENTS AND POLICY IMPLICATIONS	100

Annexes A to F

EXPLAINING NHS DEFICITS, 2003/04-2005/06

Executive Summary

This report draws together various evidence, and explores six explanations, concerning NHS deficits. A compelling account of the emergence of deficits will be consistent with not only their occurrence, but also the timing and any new structural patterns associated with the deficits. For reasons discussed in the text much of the analysis is based on PCT Health Economies rather than individual organisations. In a twenty-week commission we can do no more than summarise the available evidence, and review how well the proposed explanations account for the most prominent facts - elevating some explanations and deflating some others - the incompleteness of the available data and the temporal proximity of the events disallow a definitive answer to all the questions raised.

The hypotheses considered are: that increases in national wage rates created deficits by providing no local "headroom" for capacity growth to meet clinical objectives; that changes in the allocation model may have re-directed resources to the "needy areas"; that emphasis upon workforce targets, and concern to avoid unspent resources being transferred to other organisations, may have encouraged excess or inappropriate employment growth 2001-4; that changes in accounting regulations which prohibited capital to revenue virement from 2004/5 onwards were recognised too slowly by the NHS in some areas; that service delivery targets had uneven financial consequences across geographic areas; and that management weaknesses may have added to and reinforced the economic drivers of deficits.

A full explanation for NHS deficits must account for their sudden emergence in 2004/5, their new tendency to be moderately more likely in the south and east and low age-needs areas, and the increased heterogeneity of financial outcomes, in which about 90 PCT Health economies move into larger deficits as about 50 acquire larger surpluses.

The view that the *additional* NHS resources – which we shall call the Health Dividend – were absorbed by generous national wage contracts, leaving little for capacity growth, is not consistent with the various evidence, including in particular: the substantial increase in capacity, reflected in about 128,000 additional NHS staff that were hired, 2000-4, *above long term trend employment growth*, the tendency for unit labour cost growth to be similar in the four year period up to 2004/5 as it was 1996-2000, and the inability to explain the new geographic patterns and increased dispersion of financial outcomes. This hypothesis is therefore not convincing.

The second hypothesis points out that a new allocation model increased target resources relatively more in less prosperous areas, but it is also the case that this effect was offset, most notably by the greater distances between allocations and targets. Perhaps for this reason, when the new allocation model was introduced in 2003/4, deficit problems did not ensue. Moreover, the deficits in 2004/5 are uncorrelated with the *change* in PCT resources 2003/4-2004/5. While the changes provided by the allocation model post-2004/5 may make the restoration of financial balance more difficult, the second purported explanation does not provide a sharp enough change to

budgets to make a significant contribution to explaining the emergence of deficits in 2004/5.

The most important explanation of the emergence of deficits is probably the slow adjustment to the accounting change, which disallowed virement from capital to revenue accounts from 2004/5 onwards. This change may best explain the sudden emergence in 2004/5 of deficits primarily in the East and South and low age-needs areas, when many other areas enjoyed increasing surplus. These virements had amounted to about £380 millions in 2003/4, and had this scale of virement been allowed in 2004/5, the aggregate deficit would not have been uncovered.

Three other factors may have contributed to the emergence of deficits. First, by as early as 2004, there had been a tendency in certain areas for employment to increase to levels well above those consistent with national targets, and perhaps more significantly, with a mix of additional staff that may have been unsuited to the pattern of clinical demand pressures that were to emerge in 2004/6. This in turn may have been characteristic of an expenditure/reserves strategy in the period to 2004 that failed to anticipate the flexibility needed to meet the various developments in the 2004/6 period. A further contributing factor may have been the greater distance to certain secondary sector clinical targets in Southern and Eastern England – particularly the "four hour A&E wait" - reflecting a weaker secondary sector performance in these areas prior to 2004 that was only revealed in the drive to reach a national standard of various elements of secondary sector care. The consequence of this was for the distance travelled by organisations in the South and East to be on average greater, with concomitant cost pressures. The preceding three economic drivers were to place management in certain areas under particular pressure, which in some cases was met satisfactorily, and in others management weakness was magnified into substantial deficits. In this way the weak management found by the Audit commission at organisations with substantial deficits, are consistent with the emergent geographic patterns of deficits. It is highly unlikely that managerial weakness by itself explains either the timing or geographic features of the emergence of deficits.

More Detailed Summary of the Key Findings and Lessons

Facts:

- The NHS reported an aggregate deficit in 2004/05 following four years in which it ran a small surplus. The in-year deficit persisted at a similar level in 05/06.
- When aggregate deficits emerged in 2004/05 they tended to be concentrated in a cone-shaped area above a line from Bristol to Southampton and below a line from Bristol to the Wash, whereas in 2003/04 deficit areas are more equally distributed across regions.
- Deficits are a problem for a minority of NHS organisations 36% were in accumulated deficit in 2005/06. When organisations' balances are combined to create health economies defined on the basis of patient flows, a larger share is shown to be in accumulated deficit 48% in 2005/06.
- A growing heterogeneity of financial outcomes has been observed. About 90 health economies have an increasing deficit in 2004/05 but also about 50 health economies have an increasing surplus.

- Amongst PCT Health Economies with deficits over 5% of PCT allocations, the share of the Health Economy deficit with the PCT is 75% whereas in all other deficit PCT Health Economies the share is significantly less.
- Rural areas are more likely to be in deficit.

Causes of Deficits:

- Health dividend: How was it spent? Increased NHS revenue growth 2001/2-2004/5 - by 2.7% per annum above trend - provided about £13 billion of additional resource. After allowing for national wage settlements and pharmaceutical contracts, a substantial share of the 'Health Dividend' remained as 'headroom'. The argument that high wage rate growth left no headroom to provide increased capacity is much weaker than commonly believed, with employment by September 2000/2004 rising almost 130,000 full time equivalents above trend, and unit labour cost growth rising similarly during 2000-5 as for 1996-2000. (The actual growth in NHS staff 2000-4 including GP and practice staff is 178,973 ftes. Probably no other organisation in OECD countries added as many staff in this period.) The conversion of about 70% of the Dividend into volume employment increases is inconsistent with the view that the Dividend was spent on wage inflation. About 14% of the HD is estimated to have been spent on above trend wage growth, with the residual 16% spent on volume growth of non-labour inputs. On available evidence, the NHS largely avoided demand-pull cost inflation - one of the major macro-system threats posed by the Dividend and succeeded in translating the Health Dividend into a substantial increase in capacity.
- Resource allocation while a significant relationship between areas with low age/needs and deficits arises in 2004/05, this can neither be attributed to a change in the estimation of needs (since this occurred in 03/04 and was largely 'offset' by changes to "distance from target" and the MFF) nor by movement to allocation targets (since pace of change from 03/04 to 05/06 was historically slow and there is no relationship between growth in allocations and deficits in 04/05). Rather, the primary reasons behind the emergence of deficits in low age/needs areas are likely to be explained by their increased expenditure growth rather than relative reductions in income growth, or by loss of revenue from the non-availability of capital revenue transfers. (The persistence of deficits beyond 2004/5 to the present may however be aggravated by movement to allocation targets).
- Cost pressures while revised labour contracts affecting NHS staff have cost more to implement than originally expected, pharmaceutical prices have been lower than expected. This means that unanticipated growth in unit costs is small in comparison to the growth in the volume of staff employed.
- Performance targets all areas of the country did not start equal with respect to their distance from achieving key performance targets. In 2003 areas of low age/needs tended to be further from secondary care targets, particularly the A&E 4-hour wait, and these areas made significantly larger improvements in service delivery, 2003/4-2004/5. Areas that in 2002/3 had larger distances to travel to secondary sector targets tend on average to have larger deficits.
- Changes in accounting practices in particular, the removal of between £200-£300 million of local virement flexibility have contributed to the emergence of

aggregate deficits in 2004/05, and together with performance targets, may well help explain the uneven geographical distribution.

• Organisations with large deficits are frequently found to have had weak management capacity. However, geographic and age-needs patterns of deficits are not easily explained by management alone. Management skills may interact with economic drivers, so that in areas experiencing adverse economic circumstances good management may moderate the consequences, but weak management may exacerbate the underlying causes and result in large deficits.

Lessons

Targets:

- Be aware of the geographic unevenness of starting positions when setting uniform national targets and the variable financial consequences implied by such unevenness.
- Once a broad decision to target an area of service has been made, conduct detailed analysis of the marginal costs and benefits of varying the timing and application of targets. For example do the benefits of moving from 98% to 100% of A&E attendances seen with 4 hours outweigh the costs? What are the marginal costs and benefits of achieving a certain target within 3 months as opposed to 6 months?
- Avoid using input targets, as they prejudice the optimal mix of inputs at a local level and send a signal that the achievement of input targets is an end in itself rather than the means to increasing output and improving outcomes.

Productivity:

- Local decisions about the hiring of labour need to a) be integrated with an organisation-wide approach to assessing where a given labour input is likely to deliver the most output and b) to consider the associated financial implications of taking on additional staff.
- Enthusiasm for making productivity improvements is diminished in an environment of rapid growth in resources. Therefore, mechanisms for driving through productivity improvements should specifically be strengthened at the time extra resources are made available.
- Institutions need time horizons which allow/facilitate the accumulation of balances to provide buffers to absorb shocks and which allow spending to be postponed when high value-added spending is unavailable.

Resource allocation:

- The estimation of local needs (demand) based on observed utilisation should investigate controlling for geographical variation in the supply and quality of <u>all</u> types of service, not just for admitted patients¹.
- Since deficits are found to be more prevalent in rural areas, further investigation of the costs and organisational aspects of servicing rural populations is recommended.

¹ The Advisory Committee on Resource Allocation (ACRA) is conducting a review of the funding formula (including further research on the MFF and needs parts of the formula) that is due to report in Autumn 2007

- The relationship between the revenue costs of younger capital and deficits needs to be investigated further.
- The impact of weaknesses in the resource allocation process may in the past have been softened by flexible accounting practices such as brokerage and capital to revenue transfers. Such weaknesses may in future be exposed under a stricter and more transparent financial regime, raising the significance of the resource allocation process. Further analysis of how the SHA may offset a weakness in the allocation process, without undermining incentives for subsidiary organisations to make efficient decisions appears appropriate.

Employment:

• At a time of financial consolidation, innovative employment arrangements may be required to ensure that graduating clinical staff are offered employment in the NHS. Relying upon decentralised behaviour by individual organisations may fail to capture wider 'external' benefits from offering NHS employment to those completing training, with too few staff retained for the long run strength of the NHS.

SECTION 2: INTRODUCTION

2.1. The prospect of uninterrupted exceptional growth in health sector expenditure 2000/01-2007/8 was scarcely expected to produce substantial national level deficits in the middle of this period. This report has two objectives: to draw together critical facts concerning these NHS deficits, and to attempt to explain those facts.

2.2. In an issue as wide ranging as that addressed, and with a premium on completion within 20 weeks of commissioning, the methodological intention is to offer evidence to bring to the fore, and give quantitative shape, to some issues that have gone unnoticed, to deflate a few issues that have perhaps received undue attention, and to give an understanding of how the various elements might cohere to explain the facts we uncover. However, it should not be supposed that the analysis of these various issues is any sense definitive and complete.

2.3. 'Explanations' can be constructed on several levels, and we shall be specific about the sense in which we shall provide an answer. We shall also need to be thoughtful about how we define deficits, as they relate to both time, and to the accounts of "linked" organisations. To address this latter issue we construct accounts for the PCT health economy, which entails appropriately folding back the accounts of the provider organisations.

2.4. Section 3 begins by providing a description of the evolution over time of NHS deficits, distinguishing between the accumulated deficit and the "in-year" deficit that arises because expenditure exceeds income within a particular accounting year. Unlike the accumulated balance, which reflects circumstances in many previous years, one may aim to explain the in-year balance using a more limited range of explanatory factors. By combining these factors with the accumulated balance at the year beginning we can hope to account for the accumulated balance (IB) for PCTs, NHS Trusts, and "health economies". We explore the latter given the arbitrary way in which deficits may reside within the health system in either PCTs or NHS Trusts and enable us to gain a more reliable indicator of the deficit/surplus obtaining in any given geographic area.

2.5. While we shall discuss the picture of deficits at length in Section 3, it is worth noting at the outset the change in revenue growth that begins in 2001/2. The expansion of resources reflected an average increase in the growth rate of allocations to PCTs from 6.7% (1999/00 and 2000/01) to a mean of 9.4% from 2001/02 to 2004/05, an increase of 2.7% in the rate of growth.

	0	
		%
1999-00	Health authorities	6.60
2000-01	Health authorities	6.78
2001-02	Health authorities	8.91
2002-03	Health authorities	9.88
2003-04	Primary Care Trusts	9.24
2004-05	Primary Care Trusts	9.50
2005-06	Primary Care Trusts	9.32
2006-07	Primary Care Trusts	9.20
2007-08	Primary Care Trusts	9.40

Table 2.1- Nominal growth in HA/PCT allocations

2.6. Various explanations of the deficit paradox have been offered and we shall attempt to consider most of them through the prism of a particular structure. In the absence of surprises in either expenditure or income, or the introduction of institutional or accounting methodology change, and without costly adjustment of prevailing levels of employment and capital, it might be expected that deficits would only arise in the event of failures of managerial control. These classes of explanation – surprises, institutional change, costly adjustment (particularly downwards) of employment structures and levels, and failures of financial control - are the four organising concepts that underpin the analysis in Sections 4-10. We shall not, however, engage in lines of analysis which put deficits down to expenditure which turned out, in the estimate of the protagonist, to be "wasteful". Had this expenditure not been undertaken, some other expenditure, hopefully less wasteful would have replaced it, and with identical consequences for the budget. Hence it does not seem helpful to open that conceptual distinction.

2.7. Following a description of various facets of NHS deficits up to 2005/6 in Section 3, we present in Section 4 a picture of inputs the health dividend was used to purchase. This raises one possible "explanation" for the deficits, linked to the implication of workforce targets, and the consequences of inputs which are not costlessly adjusted in the event of shocks. Section 5 gives an analysis of the impact of changes in the Allocation Model which were introduced in 2003/4. Whilst a single section accommodates a discussion of the issues arising from income considerations, Sections 6 and 7 address hypotheses concerning the pressure on expenditure provided by the role of national employment contracts, pharmaceutical cost over-runs and high level targets. In Section 8 we discuss several accounting innovations that impacted on this period, and in Section 9 analyse how weaknesses in financial management and turbulence in the PCT/SHA organisational structures may have influenced the financial outcomes. In Section 10 we bring the arguments together into an econometric model of in-year deficits for 2004/5 to test how far the arguments identified still remain in place and an indication (no more) of their comparative empirical importance, in a multi variate context.

2.8. Section 11 provides some tentative conclusions and policy indications.

2.9. Figure 2.1 presents a graphical representation of the conceptual approaches that have been adopted in this paper to explain the emergence of and geographical

distribution in NHS deficits.

.



Figure 2.1: Conceptual model of possible drivers of NHS deficits

SECTION 3: FACTS ABOUT NHS DEFICITS

The aggregate national picture

3.1. The most commonly used measure to describe the financial position of the NHS at any given point in time, is the accumulated² balance (AB). In simple terms, a financial balance is income minus expenditure, and an accumulated balance is the sum of all balances over time. The top panel of Table 3.1 shows that by the end of 2004/05 the NHS as a whole had overspent its accumulated budget by just over £250 million.

3.2. Since the accumulated balance at any time includes historical as well as current balances it does not reveal anything about the immediate past. For example, when viewed in isolation, it is not clear in which year the £250 million of accumulated deficit in 2004/05 was incurred. In understanding the reasons why the NHS might systematically under or overspend its budget, it is necessary to observe changes in accumulated balances over time. When such changes are viewed on an annual basis they are referred to as in-year balances (IB). We can think of this in-year balance and its determinants as driving the change in the AB, and this is how the identity between accumulated and in-year balances can be interpreted:

$$AB_t = IB_t + AB_{t-1}$$

Table 3.1: Accumulated and in-year balances across the NHS

£ 000s	NHS as a whole	PCTs	NHS Trusts and FTs	SHAs
2001/02	71,259	14,694	-40,229	96,794
2002/03	42,273	92,784	-146,903	96,392
2003/04	72,530	3,867	-137,607	206,270
2004/05	-255,772	-272,392	-356,116	372,736
2005/06	-590,688*	-492,059	-624,871*	526,242

Accumulated Balance (AB)

In-year Balance (IB)

£ 000s	NHS as a whole	PCTs	NHS Trusts and FTs	SHAs
2002/03	-28,986	78,090	-106,674	-402
2003/04	30,257	-88,917	9,296	109,878
2004/05	-328,302	-276,259	-218,509	166,466
2005/06	-334,916*	-219,667	-268,755*	153,506

*Figures for Whipps Cross NHS Trust are provisional

3.3. The figures in Table 3.1 show that the NHS moved into accumulated deficit in 2004/05 following three years in which it reported a small accumulated surplus. When viewed on an in-year basis it is clear that a step change occurred in 2004/05. The size

² Or 'cumulative'

of the in-year deficit was roughly ten times the absolute level of the surplus in the previous two years. The in-year deficit was similar in size in 2005/06.

3.4. At an operational level, the way in which accumulated balances are accounted in individual NHS bodies is by carrying forward any surplus/deficit from the previous year. However, when the surplus/deficit is carried-forward, the accounts from the previous year (on which the carry-forward is based) are still provisional. This is rarely an issue since the audited position is similar to that presented in the provisional accounts.

3.5. However, in 2004/05 the aggregate difference between audited and provisional accounts was of the order of £112 million. This meant that significantly less of the aggregate deficit in 2004/05 was carried forward into 2005/06. To avoid having to restate the 2005/06 position the NHS makes the adjustment for the carry-forward error in the 2006/07 accounts. This is in line with standard accounting practice, but makes calculating an in-year balance for 2005/06 difficult because some of the deficit that should have been carried forward from 2004/05 appears two years later in 2006/07.

3.6. Therefore, for the purposes of calculating an accurate measure of in-year balances in 2005/06 the audited figures for that year have been adjusted to include the full 2004/05 carry forward, as shown in Table 3.2. The revised figures show that 2005/06 was worse than 2004/05 in terms of the aggregate in-year deficit. These figures will be used for the remainder of this paper.

£ 000s	NHS as a whole	PCTs	NHS Trusts and FTs	SHAs
2005/06 accumulated balance	-701,597	-563,103	-656,266	517,772
2005/06 in-year balance	-445,825	-290,711	-300,150	145,036

Table 3.2: Revised figures for 2005/06

Table 3.3: Gross and net surpluses and deficits

Accumulated balance	Gross surplus £m	Gross (deficit) £m	Net surplus / (deficit) £m
2003/04	484	(411)	73
2004/05	508	(764)	(256)
2005/06	833	(1,535)	(702)
In-year balance			
2004/05	347	(675)	(328)
2005/06	602	(1,048)	(446)

3.7. While the NHS has only recently moved into significant aggregate overspend, deficits at the local level have always existed. Table 3.3 disaggregates the net balance into a gross surplus and gross deficit. The gross surplus (deficit) is the sum of all NHS organisations' surpluses (deficits), and the net surplus (deficit) is the sum of the gross surplus and the gross deficit. While the gross deficit has grown significantly over the

recent past, so too has the gross surplus implying that financial outcomes have become more diverse over time. This increased financial diversity is an issue we attempt to explain in Sections 4-10. The following parts of Section 3 focus on the prevalence, scale, and distribution of deficits among PCTs, NHS Trusts, SHAs and PCT health economies.

SHA, PCT, and Trust deficits in cross-section

3.8. For the period 2003/04 to 2005/06 there were roughly 600 NHS organisations, including SHAs, PCTs, NHS Trusts, and Foundation Trusts. Table 3.4 below shows the number of organisations in accumulated deficit and surplus (and those in balance) for the three year period of 2003/4-2005/6. The number of organisations in deficit steadily increased between 2003/04 to 2005/06, although the total number of organisations in surplus was far greater than the number of organisations in deficit for all three years.

3.9. Table 3.4 also shows the numbers in deficit and surplus by type of organisation. It can be seen that SHAs were almost invariably in surplus. Most PCTs were in surplus in 2003/04, but experienced increasing financial difficulties during the three year period, from 41 PCTs in deficit in 2003/04, to 125 PCTs in deficit in 2005/06. The number of Trusts in deficit increased by a small proportion across the time period, while Foundation Trusts also deteriorated financially in 2004/05. The increase of organisations in deficit is therefore concentrated at the PCT level.

	2003/04	2004/05	2005/06	2003/04-04/05 change	2004/05-05/06 change
All organisations in deficit	105	169	216	61%	28%
All organisations at breakeven	36	25	11	-31%	-56%
All organisations in surplus	457	404	371	-12%	-8%
SHAs in deficit	0	1	1	N/A	0%
SHAs at breakeven	0	0	0	N/A	N/A
SHAs in surplus	28	27	27	-4%	0%
PCTs in deficit	41	90	125	120%	39%
PCTs at breakeven	10	6	1	-40%	-83%
PCTs in surplus	252	207	177	-18%	-14%
NHS Trusts in deficits	60	64	77	7%	20%
NHS Trusts at breakeven	23	18	10	-22%	-44%
NHS Trusts in surplus	152	153	148	1%	-3%
FTs in deficit	4	14	13	250%	-7%
FTs at breakeven	3	1	0	-67%	-100%
FTs in surplus	25	17	19	-32%	12%

Table 3.4: Number of organisations in accumulated deficit and surplus 2003/04 to 2005/06, by type of organisation

Note: An organisation is defined as being at breakeven if its deficit/surplus is within £1,000 of being in balance

3.10. Table 3.5 shows that the deterioration in the financial position 2004/5 was associated with an increase in the number of organisations (to 33) responsible for 50% of the gross deficit; this number increased further to 39 in 2005/6. In 2003/04, 70% of

the gross deficit is in 5% of organisations, whereas in 2004/05 and 2005/06, it spreads to 10% and 12% of organisations respectively. The percentage of organisations with any kind of deficit doubles, from 18% in 2003/04 to 36% 2005/06.

Cumulative % of gross deficit 2003/04	Number of orgs.	% of all NHS orgs.	Cumulative % of gross deficit 2004/05	Number of orgs.	% of all NHS orgs.	Cumulative % of gross deficit 2005/06	Number of orgs.	% of all NHS orgs.
50	18	3	50	33	6	50	39	7
60	24	4	60	44	7	60	52	9
70	32	5	70	58	10	70	70	12
80	42	7	80	75	13	80	92	15
90	59	10	90	99	17	90	121	20
100	105	18	100	169	28	100	216	36

Table 3.5: Concentration of deficits (accumulated balance) across all NHS organisations

3.11. This picture is confirmed below. Table 3.6 shows the distribution of the varying sizes of accumulated deficits and surpluses across all organisations from 2003/04 to 2005/06. In 2003/04, over two-thirds of organisations (434) had a surplus of up to \pounds 5m million, and there were only two organisations with deficits over £10 million. By 2005/06, just over half the organisations (339) had a surplus of up to \pounds 5 million, and there were 50 organisations with deficits over £10 million. This represents a sharp shift from a concentration of organisations close to breaking even, to a large proportion in deficit, and many with deficits of considerable size.

Table 3.6: Distribution of size of deficits across all organisations (accumulated balance)

	2003/04	2004/05	2005/06
Deficit over £20m	0	2	13
Deficit £15-£20m	1	4	9
Deficit £10m-£15m	1	10	28
Deficit £5m-£10m	14	41	53
Deficit up to £5m	89	112	113
Breakeven	36	25	11
Surplus up to £5m	434	378	339
Surplus £5m-£10m	16	8	10
Surplus over £10m	7	18	22

See note to Table 3.4

3.12. Chart 3.1 and Table 3.7 below show the distribution of PCT accumulated balances for 2003/04 and 2005/06. The number of PCTs with a balance within a range of £4 million is signified by the height of the bars, which read off against the left hand side of the chart. In 2003/04, around 250 of the 303 PCTs have a surplus of less than £4 million. By 2005/06, there is a large shift from PCTs in surplus to many more with deficits ranging from less than £4 million to over £16 million. There is also a general widening variation in the balances. This suggests that achieving balance for PCTs became increasingly difficult over time.





 Table 3.7: Distribution of size of deficits across PCTs (accumulated balances)

	2003/04	2004/05	2005/06
Deficit over £16m	0	1	5
Deficit £12m-£16m	0	3	6
Deficit £8m-£12m	1	4	15
Deficit £4m-£8m	7	27	37
Deficit up to £4m	33	55	62
Breakeven	10	6	1
Surplus up to £4m	250	206	174
Surplus £4m-£8m	1	0	3
Surplus over £8m	1	1	0

3.13. Chart 3.2 and Table 3.8 show the same as Chart 3.1 and Table 3.7, but for Trusts rather than PCTs. A striking feature of Table 3.8 is that the number of Trusts in surplus scarcely declines with the occurrence of deficits elsewhere in the health economy -152 in 2003/4 and 148 in 2005/6. The challenges posed by the various sector developments, 2003-6, were financially accommodated by one group of Trusts, just as a second, smaller number of Trusts at breakeven or in deficit in 2003/4 were increasingly in financial difficulty in 2005/6. The data here do not illustrate whether the Trusts in surplus were the same throughout the period, but the considerable "persistence" in surplus and deficits is discussed in 3.41-3.45. By 2005/6, there were 63 Trusts with a deficit in excess of £4 million, as compared with only 8 in 2003/4.





Table 3.8: Distribution of size of deficits across Trusts (accumulated balances)

	2003/04	2004/05	2005/06
Deficit over £16m	1	4	13
Deficit £12m-£16m	1	1	14
Deficit £8m-£12m	1	14	10
Deficit £4m-£8m	13	14	14
Deficit up to £4m	44	31	26
Breakeven	23	18	10
Surplus up to £4m	151	151	144
Surplus £4m-£8m	1	0	4
Surplus over £8m	0	2	0



Maps 1-2: PCT accumulated balances as a percentage of allocation for 2003/04 and 2005/06 (with support)

3.14. The two maps above show the geographical distribution of PCT accumulated balances as a percentage of their allocation in 2003/04 and 2005/06. The areas in light and dark grey have small and larger surpluses, and light and darker red have small and larger deficits. In 2003/04, most of the PCTs in the country are in surplus or small deficit. Only Cornwall and the East of England (and two PCTs in London) have large deficits, relative to their allocation. In 2005/06, the situation has changed, and there are many PCTs with larger deficits, which are fairly evenly spread across the country. However, for reasons outlined below, focusing on PCT balances may be misleading. The following section investigates financial balances at the health economy level.

Comparative insights from considering health economies

3.15. The preceding sub-sections presented details of the prevalence, scale and distribution of surpluses and deficits across individual NHS organisations. This is important to understand the challenges that managers and other staff in these organisations face. However, in trying to understand the variation in financial outcomes at a local NHS level, there are a number of reasons why analysing the accounts of individual NHS organisations may lead to misleading results.

3.16. Firstly, the NHS' structure is based on a purchaser-provider split. Purchasers (PCTs) commission the bulk of services for the population they serve from NHS providers such as hospitals and ambulance trusts. Over the period studied, most decisions about the volume and price of services to be commissioned have been the outcome of a series of bilateral negotiations between purchasers and providers. This

has led to a situation in which, depending upon the relative negotiating strengths and bargaining positions of the respective parties, surpluses and deficits have resided at either a purchaser or provider level.

3.17. Since payment terms between NHS bodies can be viewed as internal transfer prices (that are the subject of a series of bilateral negotiations), it is preferable to remove the effect that variable negotiations have on producing variable financial outcomes by combining the balances of purchasers and providers together to form a local health economy. This has been done by using the Purchaser Provider Matrix³ to attribute the balances of NHS Trusts, Foundation Trusts and Ambulance Trusts to PCTs in proportion to current commissioning patterns⁴.

3.18. It is important to make clear that we create health economies to arrive at a measure of the financial health of local areas that is defined on the basis of observed patient flows rather than merely on the basis of geography. This is not to suppose that health economies in any sense 'exist' as tangible entities. They are an artificial construct designed to enable us to understand the pattern and underlying causes of NHS deficits. Individual NHS organisations are entirely responsible for their own financial performance and the estimation of health economy balances should in no way be interpreted as suggesting otherwise.

3.19. Until 2005/06, SHAs were able to move resources around the organisations in their jurisdiction in an effort to ensure broad financial balance at an organisational level. This had the effect of masking the true financial position of individual organisations. Since it is known who receives planned financial support it is possible to subtract this from organisations in receipt. It is not known who provides support, but it is reasonable to assume that only organisations in surplus would be in a position to provide such funding and that they tend to do so on the basis of their size, i.e. larger organisations in surplus provide more than smaller organisations in surplus. Throughout the remainder of this paper, financial balance figures are presented either with SHA support or adjusted to remove support.

3.20. Removing the distorting influence of planned financial support is the most appropriate way of analysing the underlying financial situation. However, it is important to note that accounting for such support payments involves a degree of estimation and, at the level of individual organisations, presents financial balances that may be at variance to audited accounts.

3.21. Finally, there are reasons why organisations in an SHA area may choose to 'bank' their surplus with their SHA. For example, it may improve an organisation's negotiating position if it is shown to have a small surplus or no surplus at all. This is one reason why SHAs tend to report surpluses at the end of the financial year (the others being underspent central budgets and SHA running cost efficiencies). These SHA surpluses have been redistributed to organisations, in the same way as for SHA planned support. Since the net surplus being held by SHAs in 2005/06 was £518

³ A matrix of providers and commissioners showing the proportion of secondary care expenditure commissioned by each PCT from each NHS trust

⁴ In other words, the financial balance of a PCT health economy is equal to the income a PCT receives through its allocation minus its spending on secondary care (commissioned services) minus its own primary care costs.

million (see table 3.2), failure to redistribute these balances to local health economies would seriously overstate the prevalence of NHS deficits.

3.22. For these reasons, we must also qualify any account of how far aggregate deficits are the consequence of specific management failure in a small number of organisations.

The distribution of health economy finances.

3.23. Table 3.9 below shows the number and proportion of PCT health economies with in-year deficits in 2004/05 and 2005/06. Looking across PCT health economies, 19% of health economies account for 70% of the deficit in 2004/05 and 22% of health economies in 2005/06. Therefore, the majority of gross deficits sit with roughly a fifth of health economies. However, in 2004/5, 56% and in 2005/6, 59% of health economies have an in-year deficit of any size.

3.24. Since a PCT is likely to commission services from several different Trusts, a health economy is comprised of a number of organisations. The deficits of individual organisations are often larger sums than the surpluses, so by including even a portion of a Trust heavily in deficit the PCT health economy will emerge in deficit itself. This means a larger proportion of health economies are in deficit than are individual organisations. There are also a greater number of PCTs than NHS trusts, so the deficits of the trusts are spread across a larger number of units.

Cumulative % of gross deficit 2004/05	Number of PCT HEs	% of PCT HEs	Cumulative % of gross deficit 2005/06	Number of PCT HEs	% of PCT HEs
50	31	10	50	37	12
60	42	14	60	50	17
70	56	19	70	66	22
80	73	24	80	86	28
90	98	32	90	112	37
100	171	56	100	178	59

Table 3.9: Concentration of deficits (in-year balance) across PCT health economies

3.25. Chart 3.3 below compares in-year balances for 2004/05 and 2005/06. This shows a shift to further variation of in-year balances between the health economies. In 2004/05, there is a concentration of health economies with a balance within £5 million of breaking even. In 2005/06, these are distributed more evenly across the range, particularly between the surplus balances.

Chart 3.3: PCT health economy in-year balances 2004/05 and 2005/06



Movement between deficit and surplus of health economies

3.26. In addition to understanding the concentration and distribution of health economy financial balances, it is important to know how organisations have tended to react to financial imbalance at the health economy level. Chart 3.4 below shows the extent to which health economies have been able to adjust their in-year balance in 2004/05 in response to their historic balance in 2003/04.

3.27. Health economies denoted by a dot to the left/right of the vertical line through the origin (0,0) were in accumulated deficit/surplus in 2003/04. The dots above the horizontal line through the origin had an in-year surplus in 2004/5.

3.28. Health economies appearing on the backward sloping diagonal line have fully adjusted by 2004/05 to a financial imbalance in 2003/04, with those in the top left box moving from a accumulated deficit to balance, as a result of an in-year surplus equal in value to the deficit. Those in the bottom right quadrant move from surplus to balance, given an in-year surplus in 2004/5. Points on the backward sloping line are therefore equivalent to an accumulated balance in 2004/05 of zero. Organisations in the top right box are moving further into surplus, while those in the bottom left are moving further into deficit. The line of best fit is slightly shallower than the backward sloping line, indicating that in general health economies show a tendency to run an in-year surplus if they entered the year with an accumulated deficit. This effect is confirmed (and robust) in multivariate models discussed in Section 10.

Chart 3.4: 2004/05 in-year balance against 2003/04 accumulated balance for health economies



3.29. The following chart is the same as that shown above but plots the 2004/05 accumulated balance against the 2005/06 in-year balance. This shows a different pattern from the previous year. The health economies are more widely spread, indicating a general loosening of financial control. The trendline is almost perpendicular to the backwards sloping line, with a much greater number of health economies going further into deficit or surplus. The aggregate accumulated balance in 2003/04 was very different to 2004/05 (\pounds 73 million surplus and \pounds 256 million deficit respectively). In 2005/06, health economies were more likely to turn already large deficits into even bigger ones. Thus, there seems to have been an element of inertia with regard to the in-year balance in 2005/06: it became much more difficult to bring a deficit or surplus into balance than in 2005/06.



Chart 3.5: 2004/05 accumulated balance against 2005/06 in-year balance for health economies

Make-up of health economy finances

3.30. Chart 3.6 below disaggregates the balances at health economy level into the PCT balance on the one hand and the balances of all other organisations on the other. It explores whether in deficit HEs, both PCTs and provider organisations are in deficit, and conversely for surplus organisations. Health economies in the top right and bottom left quadrant can be thought of as containing organisations that are in some sense 'sharing the burden', since PCTs and "other organisations" either both have a "surplus" or both a deficit. In other words, the PCT's balances and those of other organisations have the same direction of balance. The random allocation of blue diamonds into the four quadrants indicates that there is no general tendency for PCTs and their providers to share the same pattern of surplus or deficit. (However, we discuss below how this may arise in certain SHAs.) This remains true after adjusting for planned support. Unsurprisingly, the impact of allocating planned financial support is to reduce the number of health economies with divergent financial outcomes.



In-year surplus/deficit for rest of organisations as % of allocations

Table 3.10: PCT balances compared to the balances of all other organisations in the health economy

	With support	without support
Both surplus	81	69
Both deficit	87	84
PCT surplus, the rest deficit	49	57
PCT deficit, the rest surplus	86	93
Same direction	168	153
Opposite direction	135	150

3.31. Since planned support is a means of evening out variations in financial outcomes, areas of the country with opposing financial outcomes (once planned support payments have been made) may have problems in terms of financial control.

3.32. This is explored further in the following chart (Chart 3.7). This shows the absolute percentage point difference in the balances between PCTs and other organisations in a health economy at SHA level, as a measure of the variation in financial outcomes. SHAs are ranked in order of the absolute difference in this measure in 2005/06. The variation across SHAs is striking, ranging from an average difference between PCT's and other organisations' balances of almost 11% to about 1%.



Chart 3.7: Variation in financial outcomes by SHA, 2003/04 and 2005/06 (without

3.33. If the chart is produced again but having accounted for one of the instruments SHAs can use to control the finances in their area - i.e. planned financial support - the pattern in the following chart emerges.



Chart 3.8: Variation in financial outcomes by SHA, 2003/04 and 2005/06 (with

3.34. Unsurprisingly, the level of variation in financial outcomes after planned support reduces significantly. What is particularly interesting is that some SHAs choose to use planned support as a means of financial control, while others did not. If the use (or not) of planned support (as a means of controlling the finances of an SHA area), is used as a proxy for the overall level of financial control in an area, the variation in financial outcomes shown above may be associated with deficits in certain parts of the country. This is investigated further in section 10.



Map 3: Measure of financial control in 2003/04

3.35. The measure of financial control described above is represented in map 3, to show the geographical spread of financial control in health economies across England. The areas in green are with better financial control, through to red with worsening financial control. There are areas with relatively poor financial control in the South and East, and also the South West and the West Midlands



3.36. Chart 3.9 shows the make-up of health economy finances for the most indebted health economies (i.e. those with in-year balances in 2004/05 of less than -5%). The part of the health economy's in-year deficit accounted for by the PCT is shown in blue and the part made up of all other organisations is in green.

3.37. The first point of note is that in all but two of the 28 most indebted health economies, both the PCT and the rest of organisations are in deficit. This suggests that heavily indebted health economies experience system-wide financial failure.

3.38. However, despite this tendency for the PCT and all other organisations to be in deficit, Chart 3.9 shows that the share of the health economy deficit accounted for by the PCT's deficit is generally larger than the contribution of all other organisations. Notable exceptions are Crawley, East Surrey and West Gloucestershire.

3.39. In fact, the average PCT share of the health economy deficit for the 28 most indebted health economies (in terms of their in-year balance in 2004/05) is 75%. The equivalent figure for the health economies with deficits less than 5% is -103%, which suggests that on average the PCT is running a surplus roughly equivalent to double the health economy deficit.

3.40. Interestingly, there is a tendency for these health economies to be located in the South or East of England, a theme to which we will return later in the paper.

Chart 3.10: Accumulated balance at health economy level versus the accumulated balance that

Sensitivity of health economy balances to the most indebted organisations



% accumulated balance that would result from bringing 15 most indebted organisation into balance

3.41. The chart above is another way of presenting the prevalence of NHS deficits. It plots the health economy accumulated balances in 2005/06 against the balances that would arise from bringing the 15 most indebted organisations into financial balance. If this would eliminate health economy accumulated deficits the points on the graph would lie on the vertical axis; if it has only marginal impact the majority of points would lie on the 45° line. The chart suggests that a relatively small number of health

economies would be significantly affected by reducing the 15 largest debts: only eight health economies would be brought from accumulated deficit into accumulated surplus and another 17 have material reductions in their deficits. This suggests that the most indebted organisations do not contract with a particularly wide range of other NHS bodies, since their balances are shown not to be widely distributed across health economies.

Descriptive statistics on characteristics of health economies

3.42. This section explores the nature of health economies, and relationships with various indicators one might *a priori* expect to find correlated to deficits. Some of these are presented in the tables below as a preliminary examination of underlying correlations and potential causes of the deteriorating financial position. The health economies have been split into three ranges, with a lower quartile, an upper quartile, and the two inter-quartiles as the middle range, based on a series of different indicators.

3.43. Table 3.11 shows three averages of accumulated balances of the health economies, split into the three ranges according to the PCTs ranking on the needs index, for 2003/04 to 2005/06. Generally, the less needy health economies appear to have been in greater deficit. The differences between the three ranges became more exaggerated across the period, as all the averages are significantly different to the mean in 2004/05 and 2005/06. The needs index therefore seems to have been significantly negatively correlated with deficits in 2004/05 and 2005/06, an issue we return to below.

Needs index	2003/04 £k	2004/05 £k	2005/06 £k
Lower quartile	135	-2,899*	-6,856*
Middle range	-242	-1,174	-2,783
Upper quartile	1,301*	1,867*	3,154*

Table 3.11: Needs index

* Indicates quartile mean is statistically significantly different from middle range mean at the 5% level

3.44. Table 3.12 shows the same ranges and means table, but split according to the PCT ranking in the Market Forces Factor (MFF). In 2003/04, deficits are higher in areas with low MFF, although differences across ranges are not statistically significantly different. In 2004/05 and 2005/06 deficits are higher in both high and low MFF areas, relative to the middle range.

Table 3.12: Market Forces Factor

MFF	2003/04 £k	2004/05 £k	2005/06 £k
Lower quartile	-476	-1,453	-3,531*
Middle range	531	6	-258
Upper quartile	375	-1,296*	-5,187*

3.45. The rurality index, in table 3.13, is only significant for the most rural organisations (in the upper quartile) in 2004/05 and 2005/06. Deficits are significantly higher on average in the most rural health economies.

Table 3.13: Rurality index

Rurality	2003/04 £k	2004/05 £k	2005/06 £k
Lower quartile	724	-55	-1,713
Middle range	312	-153	-945
Upper quartile	-390	-3,007*	-5,640*

3.46. Table 3.14 explores whether deficits emerged in areas experiencing a low allocation growth. In 2003/04, areas in the highest quartile of allocation growth have balances that on average are tending to be in larger surplus than either the middle range or the areas with lowest allocation growth. This is not so in 2004/05 or 2005/06, where the differences between the quartiles are not statistically significant different.

Table 3.14: Allocation growth 2003/04 to 2005/06

Allocation growth	2003/04 £k	2004/05 £k	2005/06 £k
Lower quartile	491	953	-4,022
Middle range	-240	-998	-2,086
Upper quartile	940*	-430	-1,064

3.47. The PCT HEs with the smallest population size tend to have the larger surpluses/smaller deficits although this is only significant for 2004/05 and 2005/06.

	-	• -	D	1	. •	•
Table	e 3.	5:	Pot	วบปลา	tion.	size.
				- min		OIL C

Population	2003/04 £k	2004/05 £k	2005/06 £k
Lower quartile	801	661*	-388*
Middle range	-13	-1,510	-2,966
Upper quartile	179	-1,026	-2,952

Persistence





3.48. To illustrate the persistence of balances over time, chart 3.11 above shows the accumulated balances at health economy level in 2003/04 plotted against the accumulated balances in 2004/05. Dots in the top right quadrant represent PCT HEs (138) in surplus in both years, and those in the bottom left quadrant, are PCT HEs in deficit in both years (92). The bottom right quadrant contains dots reflecting (52) PCT HEs that were in accumulated surplus in 2003/04 but in accumulated deficit in 2004/05. The dots in the top left quadrant describe the 21 PCT HEs that turned about an 03/04 deficit into an 2004/05 surplus. The 45° line corresponds to the same balance in both years – in other words an in-year balance of zero. Points to the left of the 45° line run a positive in-year balance in 04/05. Points to the right of the line run a negative in-year balance.

3.49. Interestingly, most points are shown to be in the bottom left and top right quadrants (230 health economies out of a total of 303), demonstrating a high degree of persistence (in terms of both surpluses and deficits) over time. Of those 138 with repeated surplus (top right), 88 managed to achieve an in-year surplus in 2004/05 adding to their accumulated surplus. Unfortunately, amongst those 92 in repeated deficit, 69 ran a 2004/05 in-year deficit that increased the accumulated deficit. However, a small number of health economies have moved from large accumulated deficits into accumulated surplus, suggesting that turnaround is feasible for some health economies within a short time period. To see this, note that amongst the PCT HEs with accumulated deficits in 2003/04 of greater than 5% (dots to the left of a vertical line through -5%) a large proportion arise above the 45° line and thus are running an in-year surplus in 2004/05. Many of the largest accumulated deficits in 2003/04.



3.50. For completeness the 2004/05 accumulated balances are plotted against the 2005/06 balances. Compared to the prior two year period, there is more divergence in the balances across 2004/05 and 2005/06. Again, there are a significant number of health economies with worsening accumulated balances (178 out of 303 health economies). This begs the questions as to whether the worsening deficit health

economies in 2004/05 are the same as in 2005/06. This is investigated by plotting the 2004/05 in-year balance against the 2005/06 in-year balance, to give an indication of the persistence in balances. Chart 3.13 implies a degree of persistence in in-year balances with most points in the bottom left and top right quadrants.



Maps 4-7: Health economy accumulated balances 2003/04 and 2004/05, and in-year balance 2004/05 (without support)





3.51. The four maps above show health economy accumulated balances as a percentage of allocations in 2003/04 and 2004/05, and also the 2004/05 and 2005/06 in-year balance. Map 4 shows that most health economies have either a small surplus or deficit (indicated by light grey or light pink). The health economies with larger deficits are randomly spread across the country. Map 5 shows the accumulated balance in 2004/05, with more health economies in larger deficit emerging in the South and the East of England particularly.

3.52. The 2004/05 accumulated balance is partly a function of historic balances, so in order to see what change occurred in 2004/05, it is necessary to observe Map 6, the in-year balance for 2004/05. This shows that in 2004/05 the majority of health economies experiencing a worsening financial position were concentrated in a cone-shaped area, below a line from the River Severn to the Wash, and above a line from the Severn to the Test. In 2005/06 (Map 7) however, the pattern changes. There is still a concentration of health economies in the cone shaped area, but there are also other areas emerging with large in-year deficits, such as in the South West, and areas in the North of England.

Timelines

3.53. As well as responding to the needs of their local population, NHS organisations (and the health economies of which they form part) are also called upon to meet the requirements placed on them at the centre, from the Department of Health. The timing of these central demands and their variable geographical impact may be important in understanding the timing and geographical distribution of NHS deficits.

3.54. In the following section we present, in the form of a series of timelines, some of the key targets and implementation milestones imposed by DH on the NHS since the year 2000. We have grouped important dates and targets into four separate categories as follows:

- a timeline for the major labour and pharmaceutical contracts and an important piece of legislation affecting NHS labour;
- a timeline for input targets;
- a timeline for output targets;
- a timeline for process targets.

3.55. The relationship between the NHS and Department of Health since the late 1990's is characterised by the setting of national targets which have been used to judge the performance of local NHS organisations, in particular the managers responsible for running them.

3.56. For each timeline, we give an indication of the point at which the service was informed of a target or a commitment to implement a certain policy or contract. This is easier for some targets than it is for others. We also indicate phasing of implementation and target or contract end points.

3.57. We present the four timelines to give the reader a feel for the central pressures being imposed on the NHS over the period of interest. While the dates and milestones presented have been set at a national level, their impact on NHS health economies may be variable if certain parts of the country are further away from meeting the targets at the point at which the targets are set, or they are simply more exposed to costs associated with their achievement. We will return to this issue in more detail in section 7.



Timeline 1 – Contracts and WTD

3.58. The intention to re-negotiate contracts affecting three major NHS staff groups was communicated to the service in the July 2000 NHS Plan. With respect to the majority of NHS employees there is specific reference to, 'a new pay system – one that rewards staff for what they do, for their own particular skills and abilities, rather than simply being based on their job title'. In relation to doctors, there is mention of a 'move to new quality-based contracts for GPs' and a 'new contract for consultants'. Both the Department of Health and the NHS will have been forming expectations about the outcome of these contract re-negotiations from the point at which they were first mentioned. But costs of implementation are shown to fall predominantly over the period from April 2003 to September 2005.

3.59. The ruling that applied the European Working Time Directive to NHS staff was made in October 1998. However, special arrangements with regard to junior doctors were applied over a longer time horizon. A new Pharmaceutical Price Regulation Scheme started in January 2005, with significant reductions in the price of branded drugs.



3.60. The second chart focuses on workforce targets set by the NHS Plan in June 2000, and additional targets made in the document 'Delivering the NHS Plan' in April 2002. The Department of Health therefore set crucial input targets that were not directly linked to improvements in outputs, to be achieved by a time that coincided with the arrival of large deficits. In fact, employment for clinical staff increased well in excess of that outlined in the NHS Plan. Between 2000 and 2004, the NHS exceeded the NHS Plan target for nurse employment by approximately 28,500.

3.61. For Allied Health Professionals (AHP) the Plan promised 6,500 more staff, but between 2000 and 2004, the NHS recruited an additional 19,000 AHP staff. The employment of GPs exceeded the NHS Plan level by 460, and that for consultants was less than the NHS Plan level by 1,702. Overall, it is clear that the emphasis in over expansion was amongst nurses, AHPs, and hospital doctors. Further discussion of workforce growth is found in Section 4.


<u>Timeline 3 – Output targets</u>

3.62. A series of important waiting time targets were announced in the NHS Plan, relating to inpatients, outpatients and A&E. Of these, the earliest target achievement date was set in terms of the A&E waiting time target of 4 hours, to be achieved by the end of 2004. This target was revised down from 100% of attendances to be seen within 4 hours to 98%. Although central data collection does not span the entire period, the chart below shows the national movement towards the target's achievement. The service can be seen to have advanced more quickly to target during 2004/05 than 2003/04.



3.63. Further major milestones in terms of meeting generic waiting time targets were also set to be achieved in 2004/05 and 2005/06. Disease-specific targets were also imposed on the service in two of the Department of Health priority areas: coronary heart disease and cancer (and these are described in more detail in Annex D). Many of these were set to be achieved during 2003/04 and 2004/05.



Timeline 4 – Process targets

3.64. The fourth set of timelines relate to broader issues of process which are integral parts of NHS system reform. The announcement of the new Payment by Results (PbR) system of financing providers, with the aim to improve efficiency and support patient choice, came as part of the 'Delivering the NHS Plan' in March 2002. Significant implementation of PbR began in April 2005 with elective care being commissioned at national tariff (although a small number of Foundation Trusts were early implementers of PbR in April 2004). Choice of provider for elective care was implemented in January 2006, after being announced as part of the NHS Plan in 2000, along with electronic booking, which had a target date of full implementation by 2005.

3.65. Other dates of significant events in the period concerned relate to structure and financing of the NHS. The creation of Primary Care Trusts in early 2001, which took on the responsibilities of Health Authorities in April 2002, along with the allocation for the first time of a three year resource settlement to the new PCTs in April 2003. April 2001 marked the introduction of Resource Accounting and Budgeting across Government, although as discussed further in Section 8 its impact on DH and the NHS was more gradual. This meant that capital to revenue transfers were no longer possible at a local level from April 2004, and at a national level from April 2006.

SECTION 4: WORKFORCE GROWTH AND THE HEALTH DIVIDEND: CONSEQUENCES FOR "HEADROOOM" AND DEFICITS

4.1. The deficits that erupted in 2004/5, and which are described in Section 3, occurred after four years of exceptional revenue growth. To explain this paradoxical outcome we begin by studying the size of the Health Dividend (HD) and the type of expenditure it facilitated. In particular we ask how far the dividend was absorbed by generous national wage and pharmaceutical contracts, and how far by the creation of extra capacity.

4.2. We estimate below that the HD, up to April 2005, was worth about £13.0 billion⁵. The evidence that the HD encouraged both employers to add to their employee capacity, and also enabled bargainers to secure wage growth that otherwise would not have occurred, is assessed drawing upon the relevant data. Before analysing (in Sections 5 to 9) the external drivers to the local health economy that might have prompted deficits, we consider how far the context in which the HD was spent – which included workforce targets for clinical and related staff groups – may have separately influenced the outcome. We consider how far employment – both levels and mix - despite being a "derived demand" that is explained in the long run by external drivers, may nevertheless in the short to medium term, be an influence on near term expenditure decisions, and thus deficits, due to the costliness of reducing employment levels. In summary, in this section we review the scale of the HD, the expenditure that followed, and how the nature of these expenditures may have impacted on the deficits in 2004/5.

The Health Dividend

4.3. Annex A describes the methodology adopted to calculate the health dividend, which is estimated to be ± 13.0 billion. To place this in context of other large projects, it is about 50% greater than the cost of modernising the West Coast Rail line.

We may contrast the value of the dividend with the impact of the dividend on 4.4. the accumulated expenditure on employment growth. Annex A gives details of a simple framework for projecting the change in wage-bill costs arising from the health dividend. First, we note historic employment growth, 1995-2000, was 1.187% per annum. We take the impact of HD on employment growth to be the difference between actual growth and 1.187%. Next we might ask whether there exists evidence that cost inflation has been greater during the years of high growth in allocations, and in particular whether wage settlements have tended to be greater. We offer only a first-pass at this question in the following way. We note the debate concerning the impact of the Dividend on unit cost (of labour) growth. Interestingly, actual unit labour cost grew at a very similar rate, 1996/7-2000/1, of 5.59% as it did 2000/1-2004/5, 5.85%.⁶ Since underlying RPI growth and unemployment rates are similar over the two periods, the underlying wage rate pressures were probably not dissimilar. This is less straightforward to confirm than might be thought. While the HCHS pay and price index for the secondary sector is well developed, the only index for the primary sector, FHS, is non-continuous in 2005 because of the changed basis of GP pay. The relevant data for the HCHS and FHS are given in the Tables below.

⁵ In nominal terms

⁶ HCHS data, which covers the "secondary sector", and includes 90% of NHS staff.

Year	Pay %	Prices %	Pay & Prices %
1997-98	2.5	0.4	1.7
1998-99	4.9	2.5	4.0
1999-00	6.9	1.2	4.5
2000-01	7.2	-0.3	4.2
2001-02	8.3	0.1	5.1
2002-03	5.0	1.0	3.5
2003-04	7.3	1.5	5.2
2004-05*	4.5	1.0	3.2

HCHS Pay and Prices Inflation

FHS inflation

Year	GMS/PMS	GDS/PDS	PhS	GOS	FHS Total
	%	%	%	%	%
1997-98	5.1	0.3	2.8	2.8	3.0
1998-99	2.3	4.6	2.9	2.9	3.0
1999-00	10.4	1.0	2.3	2.3	4.1
2000-01	3.7	4.0	1.9	1.9	2.7
2001-02	1.0	2.8	2.5	2.5	2.1
2002-03	5.2	4.0	3.2	3.2	3.9
2003-04	9.7	1.8	2.6	2.6	4.6

GMS/PMS – General Medical Services/Personal Medical Services GDS/PDS – General Dental Services/Personal Dental Services

PhS – Pharmaceutical Services

GOS – General Ophthalmic Services

Source: DH evidence to Health Select Committee, 2005

Data for 2004/05 not available.

4.5. The HCHS pay and price index has an average growth of 3.6% 1997-2001 and 4.25% 2001-5. The FHS index has an average of 3.2% for 1997/8-2001/2 and 3.5% for 2001/2-2003/4. The HCHS pay index averages 5.38% 1997-2001, and 6.27%. While the primary sector suggests identical pay inflation in the pre and post HD periods up to 2003/4, mindful of the cost of the GP contract in 2004/5, we estimate the most reasonable assumption appear to be to assume that wage rate growth was about 0.8% per annum greater as a result of the HD – largely reflecting relatively strong wage growth in the secondary sector in the later period, and the predominance of the secondary sector. We shall assume the net effect to be +0.8% per annum, and monitor the robustness of our arguments to a range from 0.5-1.0% of the HD impact on unit cost growth.⁷ (We shall discuss separately the increased use of temporary staff and the concomitant wage rates in Section 6 and Annex C).

4.6. We estimate that during the period 2000/1-2004/5 the NHS recruited 128,000 fte (excluding practice staff) additional staff above what would on a historic growth basis have been hired. Only 840,000 fte staff would otherwise have been employed in 2005, whereas the actual figure was 968,000 ftes, excluding GP and practice staff. The accumulated additional cost of this employment was about £9.2 billion. Assuming 0.8% per annum impact of the HD on unit labour costs 2001-5, the accumulated additional cost of the estimated greater unit cost growth during this period is £1.8 billion. These two factors acting together give a total incremental wage bill increase of £11.0 billion.

4.7. Therefore, to a broad order of magnitude we estimate that by 2004/5, a HD of ± 13.0 billion was largely absorbed in an increased volume of staff growth (71%) and greater wage cost (14%). Since there is no evidence that the HD prompted higher prices of other inputs, the residual 15% can be assumed to be volume growth of other inputs. The various other NHS developments 2001-5 that were funded from non-capital revenues can be thought of as being absorbed within the 15% remaining and the baseline historic funding that would have occurred had the dividend not been paid. These conclusions are not materially altered by allowing the unit cost growth 2001-5 associated with the HD to be 0.5 or 1.0%. (See Annex A).

⁷ The NHS Staff Earnings Survey are conducted every 2 years – 2004 is the latest available - and give
insight into basic salary and earnings growth. Whilst higher % earnings growth 2000-2004, relative to
1998-2000 is a feature for junior doctors and consultants, it is not valid for nurses.

		Qualified nurses	Grade D nurses	Consultants	Junior Doctors/ Doctors in training and their equivalents	Ancillary
sɓu	Growth 98-00	7.38%	7.92%	4.69%	5.91%	9.00%
Total Earni	Growth 00-04	3.96%	3.60%	7.13%	8.16%	4.84%
	Growth 98-00	6.51%	6.95%	4.52%	4.47%	4.34%
Basic Salary	Growth 00-04	4.01%	3.15%	3.50%	3.50%	6.10%

Total Earnings and Basic Salary data supplied by the Information Centre

The Earnings Survey is based on a sample of approximately 50% NHS trusts in August. Figures are from SPS payroll data from all trusts in the survey for the month of August.

Figures calculated for 2004 have included national, national clone and local payscales.

Earnings figures are based on monthly payments in August, multiplied by 12 to give annual equivalent amounts. The August 2004 figures are the relevant benchmark for the financial year 2004/5.

The sample used in the survey is not necessarily geographically or structurally representative of all trusts in the English NHS. All figures for 2000 and 2002 have been calculated using national or national clone payscales alone.

Doctors in training and their equivalents refers to the registrar group, senior house officer, house officer and other staff grades at these grades that do not hold an educationally approved training post.

Were deficits a consequence of lack of local "Headroom"?

4.8. One of the deficit debates which has arisen concerns how far the DH has provided "headroom" for expenditure to create volume growth and how far increases in national wage and pharmaceutical prices have absorbed the HD, and thus given the NHS little scope to meet various cost shocks in 2004/5. The evidence above suggests that the headroom 2001/2-2004/5 was substantial, and that it may be compared to the increase in employment – 128,000 – that is estimated to have occurred as a result of the HD. This approaches 850 staff per PCT health economy. In rejecting the "DH left no headroom" explanation of deficits in its least qualified form, we must also view the issues from the context of the one year financial planning horizons within which many local decisions were taken, and hence the practical limits that existed to turn apparent headroom into a buffer against future financial shocks.

4.9. It can be argued that management possessed, in the HD, resources that were sufficiently large to offset 2004/5 shocks to local expenditure, but the rules and organisational framework within which NHS Trusts and PCTs made decisions were not conducive to long term financial planning. In an important sense the debate concerning "headroom" is a minor element of a larger issue: the (in)adequacy of the organisational framework for optimizing the use of health sector resources.

4.10. Apart from failing to accumulate reserves, given the emphasis on employment and activity growth, conventional employment is a less reversible expenditure than certain other inputs, so that scaling back current employment spend to meet financial shocks is costly and carries potentially adverse effects on morale and productivity.

Employment: A "driver" or just a large component of purchases?

4.11. There are two major reasons why employment decisions may have proven a separate "driver" for deficits beyond the intermediate role that arises as a "derived demand" following changes to external factors. First, the workforce <u>targets</u> have themselves carried considerable weight in "building capacity" in the growing service, both in terms of concrete effect of numbers specified and for the "atmosphere" engendered, so that we might think of workforce numbers as an end in themselves, potentially driven by targets to recruit beyond the level that optimizing organisations would ordinarily choose. Secondly, given that firing workers is costly and morale-reducing, an organisation that happens to over-employ or chose a mix of skills that are unsuited to demands that arise later in time, will allow the level or mix of current employment to influence on-going optimisation decisions. In particular, given firing costs, and an inappropriate skill-mix it may; economise on under-provided factor inputs, retain current workers, and possibly incur a deficit.

Workforce Targets

4.12. Table 4.1 gives the various workforce targets and the outcomes. They include targets that appeared in the NHS Plan (2000), the 2001 Election Manifesto, and Delivering the NHS Plan (2002). All targets have been met with the exception of two from Delivering the NHS Plan, which concern on-going 2008 targets: to deliver 15,000 more consultants and GPs, and to deliver 30,000 scientists and therapists, 2001-08.

Table 4.1 –	- NHS Plan	and Manifesto	Commitments
-------------	------------	---------------	--------------------

Staff in Post (headcount)	1999	2000	2001	2004	2005	Increase 1999- 2005	Increase 2000-2005	Increase 2001-2005
Commitment								
NHS Plan target 20,000 more nurses and midwives over a 1999 baseline by 2004	329,637			397,515 (achieved in September 2001)	404,161	74,524		
Manifesto commitment, 20,000 more nurses and midwives over the 2000 baseline by 2005		335,952			404,161 (target achieved in September 2002)		68,209	
Delivering the NHS Plan forecast 35,000 more nurses and midwives over a 2001 baseline by 2008			350,381		404,161 (achieved in September 2004)			53,780
NHS Plan target 2,000 more GPs (excluding registrars and retainers) over a 1999 baseline by 2004	28,467			31,523 (achieved in December 2003)	32,738	4,271		
NHS Plan target 7,500 consultants over a 1999 baseline by 2004	23,321			30,650 (met in December 2004)	31,993	8,672		
Manifesto commitment, 10,000 more doctors (consultants and GPs) over 2000 baseline by 2005		52,994		62,173	64,731 (achieved in March 2005)		11,737	
Delivering the NHS Plan forecast 15,000 more consultants and GPs over a 2001 baseline by 2008			54,584		64,731			10,147
NHS Plan target 6,500 extra therapists and other health professionals over a 1999 baseline by 2004	47,920			58,959 (achieved in September 2003)	61,082	13,162		
Delivering the NHS Plan forecast 30,000 therapists and scientists over a 2001 baseline by 2008			110,241		134,534			24,293

4.13. The workforce targets relating to each of the 6 main staff groups are discussed in more detail below:

• Nurses

The NHS Plan target set a staffing level for 2004 of 349,637 nurses. This was achieved early and the actual number in post in September 2004 was 397,515, increasing to 404,161 in September 2005.

The Manifesto commitment sets out the target staffing level for 2005 of 355,952 nurses. This was achieved early and the actual number in post in September 2005 was 404,161

The Delivering the NHS Plan forecast of 385,381 by 2008 was achieved early. There are nearly 20,000 more nurses than was originally forecast.

• GPs

The NHS Plan target set a staffing level for 2004 of 30,467 GPs. The target was achieved early, and in September 2004 the number of GPs in post was 31,523. This increased to 32,738 GPs in September 2005.

• Consultants

The NHS Plan target set a staffing level for 2004 of 30,821 consultants. In September 2004 there were 30,650 consultants in post, this increased to 31,993 in September 2005.

• Consultants and GPs

The Manifesto commitment set out the target staffing level for 2005 of 62,994 consultants and GPs. In September 2005 there were 64,731 consultants and GPs in post.

The Delivering the NHS Plan forecast of 69,584 consultants and GPs by 2008 has not yet been met. In September 2005, there were 64,731 consultants and GPs.

• Allied Health Professionals

The NHS Plan target set a target staffing level for 2004 of 54,420 AHPs. This was achieved early and in September 2004 there were 58,959 AHPs in post, this increased to 61,082 in September 2005.

• Scientific, Therapeutic and Technical Staff

The Delivering the NHS Plan forecast of 140,241 more therapists and scientists by 2008 has not yet been met. In September 2005 there were 134,534 qualified scientific, therapeutic and technical staff.

4.14. All plan targets concern headcount which as a measure of pressure targets placed on recruitment could be criticised for disregarding changing preferences towards part-time work. It is helpful therefore to look at 2002 Spending Review supply projections which is based on Full Time Equivalent (FTE) data, for various groups of clinical staff.

4.15. Table 4.2 gives actual employment levels and projected supply figures for recent years. These tables provide major "surprises" that were not anticipated in 2002. Whereas the employment levels of GPs and consultants are broadly in line with expectations, those for other hospital doctors, nurses, and the broad AHPs category are well in excess of the numbers that it was believed would be available and also, in the case of nurses, well in excess of what was thought needed to achieve the targets laid out in the Departmental objectives. In the case of nurses, the 2005 target was achieved early in 2003. Whereas on the basis of supply projections the Department expected to hire **22,500** extra nurse FTEs between 2001 and 2005, the NHS actually hired **44,200** extra FTE nurses- almost exactly double that projected. Similarly it has hired about **10,000** more hospital doctors – almost entirely below consultant level – than was anticipated, partly due to lower wastage rates and partly to overseas recruitment.

	Full Time Equivalents (FTEs)									
	All Hospi	tal Doctors	(excludi and r	GPs ing retainers egistrars)	Nurses AHP scientific, therapeutic staff)			therapeutic 8 staff)	(qualified & technical	
Year	Actual	Projected supply *	Actual	Projected supply *	Actual	Projected supply *	Projected demand	Actual	Projected supply *	Projected demand
1997	57,100		26,360		256,090			81,600		
1998	58,750		26,460		257,600			84,560		
1999	60,340		26,560		261,340			86,840		
2000	62,090		26,560		266,990			89,630		
2001	64,060		26,630		277,330			93,090		
2002	68,260	66,850	26,830	27,540	291,290	281,710		98,400	95,330	
2003	72,260	68,710	27,620	28,580	304,890	287,150	289,190	102,910	99,210	99,350
2004	78,460	69,880	28,310	29,510	315,440	294,010	300,510	108,590	103,790	104,210
2005	82,570	71,190	29,250	30,720	321,540	299,870	309,730	113,210	108,330	109,220
2006		72,710		31,880		318,870			113,310	
2007		74,590		33,150		329,060			118,510	
% increase 00-03	16.4%	10.7%	4.0%	7.6%	14.2%	7.6%	8.3%	14.8%	10.7%	10.8%
% ncrease 00-05	33.0%	14.7%	10.1%	15.7%	20.4%	12.3%	16.0%	26.3%	20.9%	21.9%

 Table 4.2: Actual and projected FTEs (supply and demand) for various groups of clinical staff

Figures have been rounded to the nearest 10.

Actual Figures are taken from the General and Personal Medical Services; Medical and Dental Workforce Census;

Non-medical Workforce Census: The Information Centre for Health and Social Care

* Projected supply and demand figures are taken from 2002 Spending Review projections.

4.16. Taking stock, we have found that all workforce targets thus far expiring have been met, and that measurement in FTEs suggests that NHS employment

demand pressure has exceeded supply projections, albeit not obviously the trajectory towards the two 2008 targets for GPs/Consultants, and scientific/ therapeutic staff. Given the emphasis placed on staff recruitment targets to build capacity it is difficult not to recognise the possibility that targets may have contributed to a frame of mind in which recruitment becomes an objective in itself that in some HEs might have outpaced affordability. In this context the integration of the recruitment process with that of deployment to the area of work of greatest value, might have been jeopardised. These considerations may also help to explain the disappointingly flat trend in NHS productivity over recent years.

4.17. To help make the cost implications concrete, in Table 4.3 we calculate estimates of the additional cost of employment above projected supply. The figures should not be attributed exact precision but are to give ballpark estimates of the orders of magnitude. Accumulated over four years from 2001/02, the cost of nurse employment above that projected is estimated at approximately £2 billion, the cost of hospital doctor employment at about the same amount, the additional AHPs at about £0.5 billion and a small saving on GPs, who do not reach projected supply, of approaching £300 million. The net accumulated cost to 2005 is about £4 billion. Patently, given that the accumulated deficit to 2004/05 is only £330 million almost all of the employment above projected supply was affordable within the headroom created, but a small amount, in a small proportion of HEs, was not within budget.

Table 4.3: The difference between projected supply figures and actual FTEs, and the associated extra in-year and accumulated costs in £millions.

	2002			2003			2004			2005		
	Difference between projected and actual	Estimated Extra Cost from Projection in year **	Estimated Accumulated Extra Costs from Projection from 2002 onwards	Difference between projected and actual	Estimated Extra Cost (£s) from Projection in year **	Estimated Accumulated Extra Costs from Projection from 2002 onwards	Difference between projected and actual	Estimated Extra Cost (£s) from Projection in year **	Estimated Accumulated Extra Costs from Projection from 2002 onwards	Difference between projected and actual	Estimate d Extra Cost (£s) from Projectio n in year **	Estimated Accumulate d Extra Costs from Projection from 2002 onwards
	FTEs	£m	illions	FTEs	£n	nillions	FTEs	£n	nillions		£millions	
Nurses (including practice nurses)	9,580	248.3	248.3	17,740	482.8	731.1	21,430	626.0	1,357.1	21,670	657.9	2,015.0
All Hospital Doctors AHP (qualified scientific, therapeutic & technical staff)	1,410 3,070	102.5 90.4	102.5 90.4	3,550 3,700	271.1 114.4	373.6 204.9	8,580 4,800	703.2 159.3	1,076.8 364.2	11,380 4,880	969.4 168.4	2,046.2 532.6
GPs (excluding retainers and registrars)	(710)	(43.1)	(43.1)	(960)	(61.3)	(104.5)	(1,200)	(84.1)	(188.5)	(1,470)	(103.0)	(291.5)
Total	13,350	398.2	398.2	24,030	807.0	1,205.2	33,610	1,404.5	2,609.6	36,460	1,693	4,302.3

Workforce figures and projections have been rounded to the nearest 10.

Actual Figures are taken from the General and Personal Medical Services; Medical and Dental Workforce Census; Non-medical Workforce Census: The Information Centre for Health and Social Care

* Projected figures are taken from 2002 Spending Review

** Estimated extra cost from projection (£s) has been calculated using an estimate of NHS staff earnings for the various professional groups from various sources, adjustments have been made to take Into account changes in pay over time and where relevant increases have been made for employers NI and pension contributions (15%). 4.18. Central workforce targets, and the pressures they present, have been a major concern throughout the period 2000-4, although these have been de-emphasised in the past two years with a move towards credible local plans.

4.19. For PCT HEs attempting to achieve financial balance labour force flexibility would become an important concern. Down scaling and restructuring a workforce is a costly exercise. In this situation, employment levels may become an influence beyond the present period's expenditure and input supply, but also determine levels of employment in subsequent periods. In other words, over short time periods, labour is not a variable cost but effectively a fixed factor of production. We shall explore this issue in Section 10, within a multi-variate context.

4.20. In Section 7 the mix of staff added 2000-5 by clinical area is reviewed to establish how far the staff mix added before 2004/5 fitted the areas of high service delivery priority in 2004/6. Many of the key delivery targets involve the secondary sector. The "drill down" to the nursing profession which is central to the NHS suggests a fairly even pattern of appointments across specialities (Table 7.2). In 2003/4 it might be thought that the growth of nurses in acute, elderly and general (2.1%) is rather modest, relative to the overall growth (3.5%), given the various secondary sector targets. In this section we now review the relationships between staff growth in the years before deficits erupt in 2004/5 and (i) a measure of deficits and (ii) the age/needs index.

4.21. For all staff, including those employed by the PCT, employment growth, 2002-2004 shows a significant negative relationship with deficits (2004/05 in-year balance as proportion of allocation, as seen in chart 4.1 below and corresponding table in Annex B). This relationship can be dissected further. For PCT HEs with all staff growth above 15% (and not less than 0.01 of proportion of the running balance), there are only five PCT HEs in surplus while 20 HEs were in deficit. The average 2004/05 balances were £3.7m and -£6.3m for each group respectively. There were 14 HEs with an in-year deficit as a proportion of allocation above 0.04, but 41 with a proportion less than -0.04. The average workforce growth for these two groups are 7.9% and 13.2% respectively. It would appear that workforce growth for deficit PCT HEs has been markedly greater.



All staff percentage growth 2002-04 at health economy level

4.22. In table 3.11 of Section 3, it was observed that a lower level of age/needs was associated with higher deficits. This motivates us to investigate a possible relationship between staff growth and the age/needs index. Chart 4.2 is a plot of all staff growth at the health economy level between 2002 and 2004 against the age/needs index. There is clearly no observable relationship.



Fixed workforce costs

4.23. Also in Section 3, one of the events taking place in the time period preceding the emergence of the deficits was the creation of PCTs. Each of the PCTs is smaller than the Health Authorities (HAs) they replaced, and may require a fixed level of non-medical staff for the effective running of PCT operations, irrespective of the level of

activity which the PCT commissions. This would imply there are increased fixed costs of establishing and running PCTs, which would adversely impact PCTs that received smaller absolute budgets. The establishment of PCTs may have had a significant impact on the in-year balance for 2004/05 at the PCT health economy level. If PCTs with relatively small budgets are located in the less needy South and East, then the existence of fixed costs associated with establishing and managing a PCT's operations may impact disproportionately in parts of the country with smaller budgets.

4.24. To test whether the employment of non-medical staff rises in proportion with PCT size, the log of PCT non-managerial staff in 2003 was regressed against the log of PCT allocation in 2003. A slope co-efficient close to one would imply a proportionate relationship – i.e. that non-medical costs rise in proportion to the size of organisation. The resulting relationship was highly significant and suggested that 96% of a 1% increase in non-medical staff is explained by increasing allocations to PCTs (see chart 4.3 and the related table in Annex B). We cannot reject the hypothesis that non-medical staff numbers rise in strict proportion with PCT allocations. Thus we do not conclude that fixed costs associated with establishing and running PCTs have contributed to the emergence of deficits in certain parts of the country.



Regional Workforce Growth

4.25. In Table 4.4 we disaggregate the pattern of staff recruitment across England into four regional categories: London, South East, East of England and the rest of the country. Across England as a whole (row 1 of the table) staffing increased by 23.7% across the NHS between 2000 and 2005. In Eastern England this grew by 28.3%, in London by 25.7%, in the South East by 21.6%, and in the rest of the country by 23%. Overall the three areas in the "cone" have an average employment growth rate about 1% greater than the rest of the country. In particular, we note that Eastern England contains many deficit organisations as shown in Section 3.

					Full Time	Equivalent				
		ļ	Γ					_ I	% increase	increas
		ļ	2000	2001	2002	2003	2004	2005	00-03	00-0
ENGLAN	D Total Staff		892,620	931,048	978,376	1,027,284	1,071,462	1,104,008	15.1%	23.7%
		Professionally Qualified Clinical Staff	460,972	477,169	501,659	525,196	549,836	566,420	13.9%	22.9%
		· · · ·	1					1		
		Hospital Doctors	62,094	64,055	68,260	72,260	78,462	82,568	16.4%	33.0%
		GPs (excluding retainers and registrars)	26,557	26,628	26,833	27,624	28,308	29,248	4.0%	10.1%
		Nurses	266,987	277,334	291,285	304,892	315,440	321,537	14.2%	20.4%
		Qualified scientific therapeutic and technical statt	89,632	93,085	98,397	102,912	108,585	113,214	14.8%	26.3%
		Other qualified clinical statt	15,701	16,066	16,885	17,508	19,041	19,852	11.5%	26.4%
LONDON	Total Staff		140,537	145,578	152,908	164,114	172,001	176,609	16.8%	25.7%
		Professionally Qualified Clinical Staff	79,516	82,820	86,797	92,657	97,262	100,561	16.5%	26.5%
		User-ital Destars	10.000	10 110	11001	15.010	10 500	17.100	10.00/	07.10
		Hospital Doctors	12,696	13,412	14,084	15,019	16,539	17,406	18.3%	37.1%
		GPS (excluding retainers and registrars)	3,990	3,977	4,015	4,162	4,255	4,424	4.3%	10.9%
		Nurses	44,/2/	46,870	49,344	52,825	54,406	55,/21	18.1%	24.6%
		Qualified scientific therapeutic and technical star	16,001	16,419	17,016	18,145	19,311	20,167	13.4%	26.0%
		Other quained cinical stan	2,101	2,141	2,335	2,000	2,700	2,043	19.270	30.370
SOUTH	Tabel Obe#		101.150	100.101	100 500	140.001	110.010	454 077	14.00/	01.00
EASI	Total Stan	Professionally Qualified Clinical Staff	124,458	130,181	136,539	142,981	146,648	151,377	14.9%	21.6%
		Professionary quanned Clinical Stan	63,206	65,929	69,066	/2,/52	/5,045	/6,403	15.1%	21.07
		Hospital Doctors	8,434	8,474	9,304	10.036	10,860	11,399	19.0%	35.2%
		GPs (excluding retainers and registrars)	4,263	4,270	4,310	4,340	4,444	4,570	1.8%	7.2%
		Nurses	36,238	38,248	40,048	42,121	42,446	42,860	16.2%	18.3%
		Qualified scientific therapeutic and technical staff	11,711	12,278	12,838	13,375	14,204	14,577	14.2%	24.5%
		Other qualified clinical staff	2,560	2,659	2,568	2,880	3,092	3,076	12.5%	20.2%
			L							_
EAST OF										
	D Total Staff		79,498	84,805	88,800	92,916	99,891	102,023	16.9%	28.3%
		Professionally Qualified Clinical Staff	41,349	43,636	45,865	47,577	50,705	52,028	15.1%	25.8%
		User-itel Destage	5 071	5 007	0 107	0.450	0.000	7 100	20.29/	20.00
I		Hospital Doctors	5,371	5,627	6,127	6,459	6,696	/,180	20.3%	33.07
i -		GPS (excluding retainers and registrars)	2,030	2,817	2,803	2,952	3,033	3,202	4.3%	13.27
		Nurses	23,/00	25,214	26,4/9	27,302	29,400	29,302	10.5%	23.17
i -		Qualified scientific therapeutic and technical star	/,460	7,906	8,407	8,916	9,444	10,113	19.5%	35.6%
		Other qualified clinical start	1,931	2,072	2,048	1,869	1,982	2,145	-3.2%	11.1%
OTHER	Total Staff		548,127	570,483	600,129	627,273	652,921	673,999	14.4%	23.0%
		Professionally Qualified Clinical Staff	276,900	284,783	299,928	312,210	326,825	337,348	12.8%	21.8%
i.		Hospital Doctors	35 593	36 542	38 745	40 746	44,366	46 578	14.5%	30.9%
		GPs (excluding retainers and registrars)	15.473	15.564	15,705	16,170	16.510	17.052	4.5%	10.29
		Nurses	162,264	167,002	175,414	182,564	189,104	193,573	12.5%	19.39
1		Qualified scientific therapeutic and technical staff	54 460	56 482	60 135	62 476	65 627	68,358	14.7%	25.5%
1		Other qualified clinical staff	9 109	9 193	9 930	10 254	11 217	11 788	12.6%	29.49
1		Other quanted on total stat.	0,100	0,100	0,000	10,201		11,730	12.070	20
Notes:			L							
Total staf	if includes all pr	rofessionally qualified staff, those that support clinical staf	ff and NHS	infrastrutı	ure and su	pport.				
Professio	onal qualifed s'	taff includes all hospital doctots, GPs, Nurses, QST&TS	and other o	ualified cli	inical staff					

Table 4.4: Regional workforce data, FTEs by professional groups, 2000-2005.

Other qualified stafff includes qualified ambulance staff and other qualified staff not included in above groups such as GP registrars. Source: Information Centre for Health and Social Care - General and Personal Medical Services; Medical and Dental Workforce Census; Non-medical workforce Census

Source: Information Centre for Health and Social Care - General and Personal Medical Services; Medical and Dental Workforce Census; Non-medical Workforce Censu

4.26. The considerable ability of the NHS to surpass all expired workforce targets suggests that in any naïve sense workforce targets are not constraining the service, but rather the service is, of its own accord exceeding the targets. The most notable example is the rise in nurse numbers. We should not expect, therefore, any simple relationship between the targets and deficits. SHAs have set both targets <u>and</u> corresponding workforce outcomes: other factors have influenced workforce outcomes and deficits rather than targets. This is a complex relationship that we continue to explore but as of this point the evidence we have is in Section 10. We note here that planned staff group growth varied considerably by SHA – we give a chart (Chart 4.4) of consultant targets and outcomes below, and note the strong correlation between planned and actual growth of consultants in comparison to nurses.⁸

⁸ We note the absence of correlation of planned consultant growth with the age/needs index, which contrasts with the modestly greater growth of consultant numbers in low age/needs areas. This deserves more exploration than we have been able to provide.



EXPLAINING THE FACTS

In the following sections we review the evidence concerning five lines of argument which have been deployed to explain the various facts concerning deficits. We begin by discussing the possibility that the allocations provided by the revised resource allocation formula have contributed to the emergence of deficits in 2004/5. We consider this first since this argument appertains to the income side of NHS accounts and the remaining arguments that we discuss largely appertain to influences driving the aggregate level and geographic pattern of NHS expenditures.

SECTION 5: HAS THE REVISED RESOURCE ALLOCATION FORMULA CAUSED THE EMERGENCE OF DEFICITS?

5.1. A new resource model – developed for ACRA at the University of Glasgow – was used to produce the 03/04-07/08 allocations. Subsequently, the 03/04-05/06 allocations were announced in December 2002. This model gave an increased emphasis on "health needs" of the population, whilst continuing the interpretation of need based on the costs of providing NHS services. Some observers have ventured the view⁹ that this model has contributed to deficits, and they have subsequently proposed reforms to the allocation model based on the occurrence of deficits. How might a revision of the resource model, which reallocates funds, but does not reduce the overall provision of funds, explain deficits? Combined with the relatively anodyne assumption of sluggish adjustment of expenditure to income, it can directly explain a new pattern of deficit, but to explain increased <u>aggregate</u> deficits would require asymmetric expenditure responses from gaining and losing PCTs. Namely, that adjustments are more sluggish in the losing PCTs.

5.2. The Chief Economist of DH presented evidence to the HSC (October 19th), which did not support the view that the sudden emergence of deficits in 2004/5, with emphasis in the South/East and a slight propensity to be more prominent in low "age/needs" PCTs, was a result of changes introduced by the new resource allocation formula. This section will discuss the evidence in more detail than was possible at the HSC. We therefore explore below how far the allocations of the resource model can be linked to the timing, geographic or age-needs picture of deficits. This is not the only question concerning the resource model that might be addressed, and before proceeding it is helpful to be clear about what we do not address here. We focus on whether the new Resource Model is a significant explanation of the emergence of deficits, during 2004/5 and 2005/6, and do not consider whether the geography of resources available prior to 2004/5 were "appropriate": this is beyond the remit of the report. Similarly, the consequences of funding patterns beyond April 2006 for the evolution of deficits are not explored.

5.3. In the analysis that follows, details of statistical relationships presented in the various charts are presented in Annex B.

⁹ Sixth report, NHS Deficits, written evidence to the HSC 3rd July 2006.

Timing

5.4. As can be seen from the Tables 3.1 and 3.3 in Section 3, the major change to the NHS deficit pattern arises in the financial year 2004/05, with a substantial turnaround from a £30m in-year surplus in 2003/04 to a £328m deficit in 2004/05. The *timing* of the emergence of aggregate deficits in 04/05 does not fit easily with changes to the resource model which were introduced <u>one year prior</u> in 03/04, with seemingly little consequence for deficits. Thus, the arguments regarding sluggish adjustment to change given in 5.1 appear not to have had a significant influence on deficits. This innocuous introduction of the new resource model is not accidental. Any adjustment to funding, if done too quickly, may result in financial difficulties in the short run. This is because in the short term PCTs have few instruments with which to reduce hospital activity and expenditure. As a consequence the DH aims to ameliorate the effects of changes to allocations.

5.5. There were at least two elements of DH policy, 03/04-05/06, which explains why the increased recognition of "health needs" and the corresponding movement of PCT target allocations may not have significantly impacted on the budget and hence the deficits of PCTs/NHS trusts in the period 2003/4-2005/6. First, the model provided new target allocations, but the DH did not move the resources actually allocated for 2003/4-2005/6 swiftly towards the targeted allocation levels – allowing instead, increased "distance from target". Chart 5.1 below shows the range in distance from target (DFT) allocations since 1998/99. The step increase in the range of DFTs in 2003/04 is largely explained by this being the first year that allocations were made to PCTs, which were (at that time) one third of the size of their predecessor Health Authorities. Despite this step change, it is clear that there was a relatively slow pace of change to target allocations over the period 03/04 to 05/06.



Chart 5.1: Range in % DFT from 1998/99 to 2007/08

5.6. Second, changes were made to the Market Forces Factor – the element of the formula which ensures that unavoidable cost differentials between geographic areas are centrally funded - which tended to offset the allocations which would otherwise have been redirected as a consequence of the resource model. An indication of this

"offsetting" effect is given in chart 5.2 below. The vertical axis gives the change in the index of target allocation (before MFF and minor adjustments) and on the horizontal axis the change in the MFF factor associated with the same PCT. There is a broad tendency for the MFF weight to be reduced (increased) for those PCTs receiving an increase (decrease) in target allocation.

5.7. What is more, chart 5.2 might also provide insight into the relationship between deficits and rurality described in section 3.45. In terms of the 59 health economies that experience both a reduction in age/needs and MFF (located in the bottom left quadrant) an average of 45% of the population are defined as living in a rural area, as opposed to an equivalent figure of just 26% for all other health economies. This issue deserves more investigation than can be undertaken here.



Chart 5.2: Change in MFF versus change in age/needs index 02/03 to 03/04

Note: the DFT figures presented above are the average DFT for PCTs in each of the four quadrants

Geography

5.8. In Table 5.1, we give the growth in allocations by SHA (column 1) and the percentage deterioration in their financial circumstances up to April 2006. The similar growth allocations across the SHA areas of the country, and the sharply different rates of financial deterioration – measured as given in the footnote to Table 5.1 – suggest little relationship between allocations and financial balance. SHAs such as Essex enjoyed a comparatively high rate of resource growth but experienced a prominent rate of financial deterioration. In contrast, West Yorkshire received one of the smallest rates of resource growth and achieved an improvement in financial balance. Norfolk, Suffolk and Cambridgeshire received an average rate of growth of resource but also the second largest deterioration of financial balance. In summary, this pattern of growth in resource allocation scarcely appears to explain the emerging regional patterns of in-year deficits that obtained in 04/05, but below we take the argument further by exploring at PCT health economy level both the link between deficits and allocation growth, and also the putative prominence of deficits in low age-needs PCT areas.

5.9. We have observed that deficits emerge suddenly in 2004/5. To explore the sensitivity of deficits to changes in resource allocation in the immediately preceding year, we document the relationship between PCT health economy deficits in 2004/5 and the growth of resources to the PCT area 2003/4 to 2004/5. This is summarised in Chart 5.3. All PCTs received at least 8.8%. Whilst an outlier in the bottom left of the chart has some influence on the statistical outcomes, there appears no significant relationship up to the 10% level (see statistical output in Annex B) between 2004/5 deficits and the growth in allocation.

accumulated balance 2005/04 to 2005/00 as	/ of anotations by	
SHA	Growth in allocation 2003/04- 2005/06	Change in financial balance 2003/04- 2005/06
Bedfordshire and Hertfordshire	31.3%	-7.6%
Norfolk, Suffolk and Cambridgeshire	30.9%	-5.8%
North West London	30.7%	-4.0%
South West London	29.9%	-3.9%
Surrey and Sussex	30.0%	-3.3%
South East London	31.2%	-3.1%
Leicestershire, Northamptonshire and Rutland	30.1%	-3.0%
Thames Valley	30.8%	-2.8%
North and East Yorkshire and Northern Lincolnshire	28.8%	-2.7%
West Midlands South	30.1%	-2.7%
Shropshire and Staffordshire	30.3%	-2.7%
Essex	32.9%	-2.2%
Avon, Gloucestershire and Wiltshire	30.5%	-1.6%
North Central London	31.4%	-1.6%
North East London	35.8%	-0.9%
Hampshire and Isle Of Wight	30.3%	-0.6%
Cheshire & Merseyside	33.2%	-0.6%
Kent and Medway	30.9%	-0.4%
South West Peninsula	29.8%	-0.3%
County Durham and Tees Valley	30.8%	-0.3%
Dorset and Somerset	29.5%	-0.2%
Trent	30.4%	0.3%
Northumberland, Tyne and Wear	29.7%	0.4%
Birmingham and The Black Country	31.6%	0.7%
West Yorkshire	29.5%	0.8%
Cumbria and Lancashire	29.9%	0.9%
Greater Manchester Strategic	31.1%	1.0%
South Yorkshire	30.2%	2.2%

Table 5.1 Growth in PCT allocations 2003/04 to 2005/06 and change in accumulated balance 2003/04 to 2005/06 as % of allocations by SHA area

Note: Change in financial balance measured by percentage change from 2003/04 accumulated balance as a proportion of 2003/04 allocation to 2005/06 accumulated balance as a proportion of 2005/06 allocation. SHA financial balances are the sum of constituent Health Economy balances – i.e. they are defined on the basis of patient flows rather than geographical location.



Chart 5.3: In-year balance as % of allocation against % growth in allocation 2004/05

5.10. We return to consider funding growth and PCT deficits in a multivariate analysis in section 10, but next explore the link between funding and the age-needs index.

Deficits and the funding of age-needs

5.11. The deficit problem has been alleged to arise from under-funding of areas adjudged to be "less needy". We explore the link between the age-needs index of a PCT, which determines the target PCT allocation per head (before allowing for unavoidable differences in regional input costs and certain minor adjustments), and the local accumulated deficit during 2003/4 (Chart 5.4) and then 2004/5 (Chart 5.6). First, we illustrate how in 2003/4 - after the new resource allocations have been introduced – there is little evidence that health economy balances were associated with the age/needs index. That is, no evidence of any tendency for the revenues of PCT health economies in less needy areas to be any more likely to be in deficit.



Chart 5.4: Health economy accumulated balance in 03/04 as % of allocation (with support) against

Age/needs index

5.12. There is further evidence that in 2003/04 there was no link between deficits and age/needs: we have compared SHA-administered planned financial support to PCT health economies in 2003/04 with the age/needs index. Planned financial support payments are made by SHAs to ensure financial balance across all organisations in their jurisdiction. Such support is sometimes referred to as 'brokerage'. In the chart below, those organisations above the x-axis make a net contribution to planned support, whereas those below the x-axis are in net receipt of support. A relationship between deficits and age/needs in 2003/04 may be hidden if less needy areas were being supported by more needy areas.

5.13. Chart 5.5 suggests that in 2003/04 there was no general tendency for local health service managers to move resources away from the needier areas suggesting the lack of a relationship between deficits and age/needs in 2003/04 [i.e. it is unlikely that a substantive relationship was hidden by accounting practices].

Note: the age and needs index is calculated by dividing each PCT's target unified weighted population (before MFF) by their actual population



5.14. Whilst in 2003/4 there is no correlation between accumulated balances and the age/needs index, by 04/05 the health economy deficits <u>are moderately</u> correlated with the age/needs index (see Chart 5.6). The PCT health economies with lowest needs/age index are more likely to be in deficit. In the rest of the paper we shall try to account for this, and understand whether this change after 2003/4 is linked in an interesting way to the emergence of aggregate deficits. First, we begin by considering the resource model.



Chart 5.6: Health economy accumulated balance in 04/05 as % of allocation (with support) against 03/04 to 05/06 age/needs index

5.15. To pursue how far this association between deficits and the age/need index in 2004/05 (Chart 5.6) is due to the resource allocation model we address directly the key evidence: to what extent has allocation growth been greater in high age/need index areas? The following chart plots the increase in actual allocations from 2002/03 (the year before new model) through to 2005/06 against the age/needs index. A group of 9 PCTs of relatively high age/needs are found to obtain a growth in allocations 2002/03 to 2005/06 in excess of 39%, but for the large body (294) of PCTs that remain, there is not a clear link between the % growth in allocations and the age/needs index.¹⁰

5.16. It is therefore hard to maintain the case that the revised resource model had substantially reallocated resources towards the "high-age" needs areas by 2005/6, when the deficits were well established.





5.17. The preceding charts suggest that the tendency for deficits arising in 2004/05 to be more prominent in both the South/East and the low age/needs index areas appears not to arise from shocks to the income side of PCT budgets propagated by the resource model, and instead these effects must arise from increases in the expenditure side of the budget for the South/East and less needy areas, or from 2004/5 changes to accounting practice that impact in regionally uneven ways. This inference is consistent with either a positive shock to total expenditure in 2004/05 or amendments to accounting practices causing the <u>aggregate</u> deficit in that year.

Summary

5.18. Some observers have argued that deficits have been prompted by the revised resource model, and are more commonplace in less needy areas. It has been suggested

¹⁰ These 9 PCTs are: Ashfield, Barking and Dagenham, Central Liverpool, Easington, Heart of Birmingham Teaching, Knowsley, North Liverpool, Tendring, Tower Hamlets

that this reflects the consequence of the resource model moving resources to needy areas. Proposals to reform the resource model have been advanced to the Health Select Committee (HSC). This section explores whether the allocations from the revised resource model have in fact played a significant role in explaining (i) the emergence of deficits, 2004/5, and (ii) any tendency for deficits in 2004/5 to be more or less common in "needy" areas as measured by the age/needs index.

The work in this Section finds no significant relationship between the 5.19. emergence of deficits in 2004/5 and the actual allocations to PCTs induced by the revised formula. In reaching this conclusion the paper discusses the putative empirical link, mentioned by observers, between the age/needs index - which determines the target allocations before MFF adjustment - and deficits. It finds no significant relationship between deficits and the age/needs index in 2003/04. However, in 2004/05 – the year that deficits first emerged as an aggregate problem – deficits appear to become slightly more common in less needy areas. The paper then explores various evidence which together fail to support the view that this tendency is a result of the resource model bringing about relative reductions in these areas' incomes. (The link from target allocations - set in part by the age/needs index - and actual allocations is broken by policy which increased the distance of allocations from target following the 2003/04 revised formula). Instead the deficits in these areas are most likely explained by an unexpected rise in the expenditures of less needy areas. Possible reasons why expenditures may particularly have risen in less needy areas, 2004/5, will be explored in Sections 4-10.

The Funding of Marginal Services by Region

5.20. The preceding Sections have argued that the emergence of deficits is not readily explained by relative reductions in funding. There is, however, an argument which suggests that whilst the overall pattern of deficits is unlikely to have been resource model driven, certain labour shortage areas may have been under-funded. This factor is unlikely to explain the sudden emergence of deficits but may contribute to the relative financial under-performance of the South East and London since the late 1990s when the role of temporary staffing became more significant.

5.21. Essentially, the relatively high use of agency staff and higher levels of bed occupancy in the South and East suggest that the marginal cost of additional activity may be significantly higher in these parts of the country. While variation in average costs of labour are reflected by the Market Forces Factor (MFF), if the costs of increasing activity at the margin are significantly higher than average cost in the South and East, then these parts of the country will be relatively underfunded.

5.22. This issue is discussed in more detail in Annex C.

SECTION 6: COST PRESSURES

6.1. Section 4 discusses actual unit cost growth and suggests that substantial "headroom" was provided to the local service. This may nevertheless be the case, but upward revisions to labour and pharmaceutical contracts may have been sufficiently large to create financial imbalance. This section sets out an estimate of the cost pressures arising from revised labour and pharmaceutical contracts. It aims to show the anticipated costs of the contracts and provides estimates of how far actual costs have diverged from these expectations.

6.2. Our measure of 'anticipated' costs arises from assumptions made in the SR2002 settlement between DH and HM Treasury, as these at least reflect the agreed sums of money that were ear-marked for specific purposes.

6.3. In terms of the labour contracts, it is possible to disaggregate additional costs into those associated with unit cost growth (i.e. the anticipated and unanticipated costs of paying the existing workforce on the basis of the new contracts) and the anticipated growth in staff volumes. It is not possible to disaggregate the anticipated costs of pharmaceuticals into a unit cost and volume effect.

6.4. We will show that anticipated unit cost and volume growth and unanticipated unit cost growth in relation to the roughly £9.6 billion health dividend in 2003/04 and 2004/05 leave a significant degree of "headroom" to meet residual unanticipated volume growth and other 'discretionary' expenditures.

Labour contracts

6.5. Managers in the NHS and those responsible for making recruitment decisions have always been called upon to formulate expectations about future pay inflation affecting the staff they employ. This is because most pay negotiations are conducted at a national level by professional bodies and/or trade unions on behalf of their members. Perhaps the degree of uncertainty around pay growth was higher during a period in which contracts affecting almost all staff groups were being totally renegotiated rather than simply uprated. But one thing is clear: the NHS will have made expectations of the outcome of the contract negotiations and made planning decisions accordingly.

6.6. The three main contract re-negotiations affecting NHS employees or contractors over this period were the GMS contract (for providers of General Medical Services), the consultant contract (for NHS consultant appointments) and Agenda for Change (which applies to all directly employed NHS staff, except very senior managers and those covered by the Doctors' and Dentists' Pay Review Body).

6.7. It is difficult to know exactly when managers will have been forming their expectations about pay inflation but we know that assumptions about the cost of the new contracts were made by the Department in its discussions with the Treasury during the 2002 spending review. It seems reasonable to suppose that these central assumptions will have been incorporated into local planning decisions. These are what we might describe as the anticipated costs of the new contracts.

6.8. The distinction between anticipated and unanticipated costs can also be extended to include the indirect costs of implementing the new contracts.

6.9. The new General Medical Services contract came into effect from 1^{st} April 2004. The new contract makes payments to practices rather than to individual GPs. There are also additional incentives on practices to provide a local range of 'enhanced' services and payments linked to the achievement of levels of quality in particular disease areas (known as the Quality and Outcomes Framework, or QOF).

6.10. While there is a significant relationship between QOF payments as a share of PCT allocations and the age/needs index (i.e. higher QOF payments in low age/needs areas) once account is taken of the funds allocated specifically to meet QOF payments no such relationship exists. QOF payments are on average higher than QOF funding but this does not vary in relation to the level of age/needs.

6.11. Unanticipated costs of the new GMS contract (i.e. those not met by ear-marked funding) have been estimated to be $\pounds 155$ million in 2004/05.

6.12. Agenda for Change implementation began on 1^{st} December 2004 with almost all staff assimilated by 30th September 2005. Pay terms and conditions were backdated to 1^{st} October 2004. A provision for any staff for whom a new level of pay had not been agreed by the year ending 31st March 2005 was included in the 2004/05 accounts.

6.13. Monitoring of the costs of Agenda for Change in 28 sample sites in 2005 suggested that in the first twelve months from October 2004 to September 2005 direct earnings costs exceeded those originally estimated by 0.5 per cent of the Agenda for Change pay bill, or around £120 million a year in cash terms. In the same period, this data suggested that the indirect costs of replacing additional hours and leave arising from Agenda for Change exceeded those originally estimated by at least £100 million a year. Assuming an even phasing of these additional costs over the 12 month period suggests that £110 million of unanticipated additional costs impacted in 2004/05. However, these indirect costs are based on trust estimates rather than actual payroll records, and are susceptible to management action.

6.14. A new contract for NHS consultants was negotiated in 2003. Consultants giving a formal commitment to the new contract by 31^{st} October 2003 had their pay backdated to 1^{st} April 2003. Those making a formal commitment between 1 November 2003 and 31^{st} March 2004 had their pay backdated by three months. Additional funding of £130 million in 2003/04 rising to £220 million by 2005/06 was allocated to PCTs to meet the costs associated with the new contract.

6.15. By late 2004, NHS trusts reported the cost of the contract had exceeded expectations by as much as £150 million, mainly due to higher levels of programmed activities. However, a DH survey (conducted in October 2004) suggested that, while the levels of programmed activities were higher than expected, the excess cost was £90 million rather than the £150 million being reported by hospitals. Again, assuming even phasing of costs over the 18 months since the effective contract start date suggests unanticipated costs of £120 million in 2003/04 and 2004/05¹¹.

¹¹ Calculated by converting the £90 million 18 month cost over-run to a 2 year figure, pro-rata.

6.16. The anticipated cost of two of the new contracts were estimated as part of the 2002 Spending Review. They are shown in the rows marked as bold in the table below and take account of general pressures in terms of pay drift, trend settlement growth and workforce growth.

	ana in	2004/5	000516
	2003/4	2004/5	2005/6
	over	over	over
£ million	2002/3	2003/4	2004/5
SR 2002 Submission to HMT			
HCHS_General Settlement	880	900	920
HCHS_Pay Drift	210	220	230
Addition to Settlement and			
Drift due to increase in			
Workforce		60	125
Agenda for Change	0	490	460
Consultant Contract	130	50	40
TOTAL:	1,220	1,720	1,775

Table 6.1 SR2002 workforce spending assumptions

Source: DH finance

6.17. These figures imply an anticipated unit cost (i.e. excluding costs associated with volume growth and wage drift) for Agenda for Change and the Consultant Contract of £670 million up to 2004/05, which when added to the anticipated saving of £300 million on the new GMS contract equals a cost of £370 million up to 2004/05.

In explaining the emergence of significant aggregate deficits in 2004/05, it is important to note that costs associated with two of the three main labour contract negotiations did not impact on the service until 2004/05, and in the case of Agenda for Change not until mid-way through 2004/05.

Pharmaceutical prices

6.18. Table 6.2 below compares SR2002 assumptions on prescribing and drugs to actual and latest estimates.

£ million	03/04 over 02/03	04/05 over 03/04	05/06 over 04/05
FHS Total Pressure SR2002	762	870	1,011
FHS Total Pressure Actual	618	410	-50
difference between actual and expected	144	460	1061
HCHS Total Pressure SR2002	239	260	277
HCHS Total Pressure Actual	298	284	Not Avail.
difference between actual and expected	-59	-24	

Table 6.2 SR2002 anticipated pharmaceutical spend versus actual spend

Source: DH finance

6.19. The steep reduction in FHS prescribing expenditure started in February 2005 (last two months of 04/05) when a new five-year PPRS agreement was negotiated which included a 7% price reduction in branded prescription medicine. This will produce savings to the NHS over the next five years. The impact of these contributed to a reduction in expenditure growth of over 3% in 2004/05 compared with the

2003/04 outturn. This impact continued throughout 05/06.

6.20. The fact that HCHS expenditure on drugs has been higher than expected while FHS drugs spend has been significantly lower than expected, implies that the cost pressures over the period may have impacted differentially on the primary and secondary care sectors, with secondary care bearing more of the cost overruns and fewer of the cost savings.

6.21. Taking FHS and HCHS drugs spend together, expenditure has been shown to be \pounds 521 million lower than anticipated up to 2004/05.

6.22. When added to the £385 million worth of unanticipated costs associated with the new labour contracts, the NHS as a whole actually faced about £140 million of unanticipated savings up to 2004/05.

Overall

contracts up to 2004/05 (on a 2002/03 baseline)									
	£ million								
	Anticipated	Unanticipated							
GMS contract ¹²	-300	155							
Agenda for Change	490	110							
Consultant contract	310	120							
<i>Volume and wage drift¹³</i>	3300	-							
FHS pharmaceuticals	2394	-604							
HCHS pharmaceuticals	748	83							
Balance	7242	-136							

<u>Table 6.3</u> – Cumulative cost pressures associated with labour and pharmaceutical contracts up to 2004/05 (on a 2002/03 baseline)

6.23. Section 4 presents estimates of the Health Dividend from 2001/02 up to 2004/05. To be consistent with figures presented in Table 6.3 above we estimate the cumulative Health Dividend over a shorter time period – i.e. for the first two years of the 2003/04 to 2005/06 Spending Review period – as being roughly £9.6 billion. Subtracting the roughly £7.2 billion of unit cost and volume growth anticipated from the main labour and pharmaceutical contracts over the period and adding a roughly £100 million unanticipated unit cost saving leaves somewhere in the region of £2.5 billion of unanticipated staff volume growth (plus expenditure on labour not covered by the three contract re-negotiations described above – e.g. PMS – and other non-labour costs). This is the share of the Health Dividend that can be thought of as 'discretionary'.

6.24. The scale of unanticipated cost pressures is estimated to be less than £200 million, and the 'discretionary' element of spending to be £2.5 billion. This suggests that the scale of surprises was modest and adds weight to the argument that the size of the health dividend allowed for considerable headroom at local level, after planned staff growth and higher than expected unit costs are accounted for.

 $^{^{12}}$ Based on the GMS component of the Gross Investment Guarantee. Figures for 2002/03 are £3.4 billion, for 2003/04 are £3.2 billion and for 2004/05 are £3.1 billion. This gives a cumulative reduction on the 02/03 baseline of £300 million

¹³ Excluding Primary Medical Care (GMS and PMS)

SECTION 7: THE IMPACT OF SERVICE DELIVERY TARGETS: ELECTIVE AND A&E WAITING, CANCER AND CHD

7.1. We discussed in Section 3.46 how many delivery targets were imposed on the NHS during the period 2003/4-2004/5. In this section we shall explore in more detail the financial consequences of the timing and local impact effect of four of the most important targets: those relating to inpatient and outpatient waiting; A&E 4 hours wait; three cancer targets and three CHD indicators. We want to investigate the evidence that pursuit of these targets contributed to financial deficits.

7.2. To provide a macro-perspective we begin by examining whether the high profile targets in cancer and CHD were a sufficiently important factor for the growth of resources 2003/4-2004/5 to exceed that in other areas. In Table 7.1 we find that growth in CHD is marginally below the national average and that for cancer, a little above. However, there is not a major redirection of resources when these programmes are taken together.

£	2003/04	2004/05	growth	% growth
1 Infectious Diseases	878,516,000	989,607,000	111,091,000	13%
2 Cancers & Tumours	3,173,245,000	3,692,449,000	519,204,000	16%
3 Blood Disorders	689,874,000	832,397,000	142,523,000	21%
4 Endocrine, Nutritional and Metabolic Problems	1,413,095,000	1,558,041,000	144,946,000	10%
5 Mental Health Problems	6,513,111,000	7,146,630,000	633,519,000	10%
6 Learning Disability Problems	1,851,427,000	2,067,667,000	216,240,000	12%
7 Neurological System Problems	1,457,358,000	1,717,669,000	260,311,000	18%
8 Eye/Vision Problems	1,202,990,000	1,354,633,000	151,643,000	13%
9 Hearing Problems	280,223,000	308,562,000	28,339,000	10%
10 Circulation Problems (CHD)	5,378,049,000	5,999,406,000	621,357,000	12%
11 Respiratory System Problems	2,666,838,000	3,075,930,000	409,092,000	15%
12 Dental Problems	526,025,000	652,731,000	126,706,000	24%
13 Gastro Intestinal System Problems	3,104,259,000	3,589,070,000	484,811,000	16%
14 Skin Problems	1,024,407,000	1,218,692,000	194,285,000	19%
15 Musculo Skeletal System Problems (excludes trauma)	2,996,258,000	3,502,650,000	506,392,000	17%
16 Trauma and Injuries (includes burns)	3,042,807,000	3,534,237,000	491,430,000	16%
17 Genito Urinary System Disorders (except fertility)	2,703,244,000	3,054,052,000	350,808,000	13%
18 Maternity and Reproductive Health	2,555,569,000	2,689,007,000	133,438,000	5%
19 Neonate Conditions	573,405,000	682,350,000	108,945,000	19%
20 Poisoning	472,756,000	603,749,000	130,993,000	28%
21 Healthy Individuals	992,658,000	1,069,009,000	76,351,000	8%
22 Social Care Needs	1,214,054,000	1,234,239,000	20,185,000	2%
23 Other Areas of Spend/Conditions:	6,691,023,000	7,607,803,000	916,780,000	14%
Total	51,401,191,000	58,180,580,000	6,779,389,000	13%

Table 7.1 Changes in national programme spending

7.3. We shall return to consider programme expenditure data on CHD and cancer in Section 10.

7.4. To provide a second macro-perspective we explore the changing distribution of nurses – a major source of employment growth – between clinical areas. This is provided using data given in Table 7.2 and asking how far nurse employment grew differently in particular disciplines.

full time equivalent	2000	2001	Growth 2000- 01	2002	Growth 2001- 02	2003	Growth 2002- 03	2004	Growth 2003-04	2005	Growth 2004-05	Total growth 2000- 05	Average annual growth 2000-05	Actual growth 2002- 05	Average annual growth 2002-05
All qualified nurses (including practice nurses) Qualified nursing, midwifery & health visiting staff	266,987	277,334	3.9%	291,285	5.0%	304,892	4.7%	315,440	3.5%	321,537	1.9%	20.4%	3.8%	10.4%	3.3%
All areas of work	256,276	266,171	3.9%	279,287	4.9%	291,925	4.5%	301,877	3.4%	307,744	1.9%	20.1%	3.7%	10.2%	3.3%
Acute, elderly & general	138,120	144,597	4.7%	154,057	6.5%	162,671	5.6%	166,098	2.1%	168,759	1.6%	22.2%	4.1%	9.5%	3.1%
Paediatric	13,644	13,997	2.6%	14,546	3.9%	14,825	1.9%	15,258	2.9%	15,412	1.0%	13.0%	2.5%	6.0%	1.9%
Maternity	22,776	22,684	-0.4%	23,043	1.6%	23,758	3.1%	24,463	3.0%	24,750	1.2%	8.7%	1.7%	7.4%	2.4%
Psychiatry	35,804	36,973	3.3%	38,176	3.3%	39,383	3.2%	41,585	5.6%	42,529	2.3%	18.8%	3.5%	11.4%	3.7%
Learning disabilities	8,398	8,440	0.5%	8,323	-1.4%	7,824	-6.0%	7,526	-3.8%	7,367	-2.1%	- 12.3%	-2.6%	- 11.5%	-4.0%
Community services	36,871	38,221	3.7%	39,302	2.8%	41,850	6.5%	44,989	7.5%	46,917	4.3%	27.2%	4.9%	19.4%	6.1%
Education staff	662	760	14.8%	819	7.8%	968	18.3%	1,140	17.7%	1,119	-1.9%	69.0%	11.1%	36.6%	11.0%
Practice nurses	10 711	11 163	1 20%	11 008	7.5%	12.967	81%	13 563	16%	13 703	17%	28.80%	5 2 0%	15.0%	18%
I TUCHUC HUISCS	10,711	11,105	 2 /0	11,770	1.5 /0	12,707	0.1 /0	15,505	 .0 /0	15,775	1.7 /0	20.0 /0	5.4 /0	13.0 /0	T.0 /0

Table 7.2 Qualified nurses (FTE) by area 2000 to 2006

7.5. Unfortunately more disaggregated data are not available. The evidence in Table 7.2 suggests that nurses have been increasingly employed in various areas since 2000, including for example, about 2000 more nurses in maternity care. Although acute care is one of the fastest growing specialties it is not so exceptional that it suggests additional expenditure is self-evidently concentrated into that targeted area.

7.6. It is not apparent that targets have had noticeable effects on the distribution of resources between specialties, however many specialties have had targets and therefore targets may impact on over-spending without significantly altering the pattern of spending. Furthermore, the deficit PCTs may have been particularly vulnerable to aggregate increases of spending in certain areas.

7.7. We now turn to evidence at PCT health economy level. We begin by exploring the link between the distance from a particular target, which may indicate the implicit pressure on a PCT health economy to undertake expenditure on improvements, and geography as captured by the age/needs index. We do this to attempt to understand the tendency, originating in 2004/5, for deficits to be linked to this index. We consider a range of targets and find that, in general, the less needy areas were considerably <u>further</u> from target in the year(s) before 2004/5. To illustrate this we examine data for in-patient waiting time; outpatient booking; Rapid Access Chest Pain Clinics (RACPC); CHD waits; Accident and Emergency 4 hour wait target. Detailed statistical evidence of the relationship between the age/needs index and opening distance from meeting targets can be found in Annex B.

7.8. Let us first look at the result for <u>patients waiting over 6 months for an inpatient</u> <u>admission</u>. As we wish to take into account size considerations, we plot the number of inpatients waiting over 6 months as a percentage of the total inpatient waiting list. We use the data as at March 2003, as this is when the trajectories came in effect. The 6 month target was introduced as part of the NHS Plan, 2000.



7.9. We find a significant correlation between PCTs with low age/needs and PCTs with far to travel to target at greater than a 1% level of significance. Those 0.1 units lower on the age needs index tend to have 5% more patients waiting over six months.

7.10. Turning now to Outpatient Booking we find a similar, statistically significant outcome. The target on outpatient booking is that all first consultant led outpatients appointments should be booked for the convenience of the patient by December 2005.



7.11. The PCTs with further to travel on the RACPC target also tended to be low age-needs areas.



7.12. The percentage of CHD patients who in March'03 were waiting more than 13 weeks was also greater in the low needs areas. (The target is discussed in more detail in Annex D.)



7.13. Turning now to the A&E data we find a highly significant relationship between age/needs and distance from target with a greater percentage of patients waiting in low age/needs areas.



7.14. These data show a strong relationship between the age/needs index and the scale of improvement required to reach the A&E target (distance from targets). Those PCTs with a 0.1 lower age/needs index (10% of mean index level) we estimate on average had approximately a 12 percentage point greater distance from target.

7.15. We look now at the primary care access (PCAS) targets. These targets are that patients should have guaranteed access to a primary care professional within 24 hours and to a primary care doctor within 48 hours by December 2004.

7.16. Let us look first at the target for primary care professionals.



7.17. The analysis on this target shows a significant relationship between high age needs areas and areas having further to travel to the target on access to a primary care professional.

7.18. We now examine the target for access to a primary care doctor within two working days.


7.19. Again, we find a significant relationship between higher age needs areas and areas with a large distance to travel to the target.

7.20. We have found a number of indicators that show a significant relationship with age needs. Of the 13 targets examined 5 show a negative correlation, 2 show a positive correlation and 6 show no statistically significant relationship.

7.21. Interestingly, the two targets that show a positive correlation also are the two targets in the primary care sector. While in the secondary care sector 5 of 11 targets explored produced negative correlations, and the other 6 no correlation.

7.22. This appears to show a split between primary and secondary care in relationship to age needs. In secondary care low age needs areas appear to have had a greater distance to travel to targets. In the primary care sector, areas with low age needs had less distance to travel to the PCAS.

7.23. Within the secondary sector, the group of targets that were adopted gave a quite asymmetric geographical impact, placing much greater financial pressure on the "less needy" PCT health economies to reach national standards in a short period of time.

7.24. Addressing each of these targets reflects only a small fraction of health care expenditure, even though cumulatively their impact may be substantial, so that demonstrating that "distance from target impacted on deficits" presents a difficulty. We are presently exploring the link of "distance to target" with changes in local programme expenditure data for the relevant specialties.

7.25. We have also explored whether there is a direct link between distance to target prior to 2004/5 for each of these targets, and the 2004/5 in-year deficit. The only target giving a clear link is that for A&E, although for thrombolysis there is the suggestion of a relationship. In the following chart we give the relationship between health economy deficits and the % point rise in patients seen at A&E within 4 hours.



7.26. The health economies which achieved the greatest % point rise in patients seen at A&E within 4 hours were also those with the greatest in-year deficits in 2004/5. A one percentage point improvement is correlated with a 0.16% increase in deficit as a percentage of the allocation. For the mean PCT allocation this amounts to about £250,000.

7.27. It would be a mistake to suppose that this demonstrates beyond doubt that improvements to A&E performance, with their focus on expenditure in 2003-5, "caused" deficits, but rather should be viewed as one of a number of factors that prompted considerable expenditure at relatively short notice to raise standards, and that the less needy areas may have been those that incurred the greatest financial pressure of this kind.

Maps 7.1-7.2: A&E attendances seen within 4 hours 2003/04 and A&E attendances seen within 4 hours change 2003/04-2004/05



Maps 7.1-7.2 demonstrate the geographical pattern of A&E performance for the 4 hour target, presented in charts 7.5 and 7.8 across health economies. Map 7.1 shows the percentage of A&E attendances seen within 4 hours as a percentage of all A&E attendances in 2003/04. This is in effect the starting position in terms of distance from target for the year preceding the emergence of deficits. It shows that most of the areas with the greatest distance from target, shown in red, are found in the cone-shaped area in the South and East, first discussed in Section 3. Map 7.2 shows the change in the distance from the 4 hour target. The health economies with the greatest change, shown in green, are found in the same cone shaped area. This is consistent with the high correlation between change in distance from A&E 4 hour target and the health economy in-year balance 2004/05.

7.28. The evidence for thrombolysis is available at SHA level and gives the following results.



7.29. In this case there is a statistically small, but not significant negative correlation. The role of target variables is reconsidered in the multivariate analysis in Section 10.

7.30. It was shown above that low age/needs areas had greater distances from secondary sector targets. Chart 7.10 shows that relatively low age-needs health economies tend to have made greater improvements towards the four hour A&E target in 2004/5. The line of best fit indicates that a 0.1 reduction in the age-needs index is associated with a roughly 5 percentage point increase in the numbers of patients seen within four hours between 2003/4 and 2004/5. About 30 PCTs have travelled more than an eight percentage point increase in A&E performance 2003/4-2004/5, but none of these have age-needs indices over 1.2, and over 20 have indices under 1.



7.31. The empirical model presented in section 10, indicates that a 1 percentage point improvement in the numbers of patients seen within four hours in A&E implied a decline in the in-year balance of around £250,000. This section discusses evidence

showing that moving from a typical health economy in the top quartile of age-needs to one in the bottom quartile is associated with a greater improvement in A&E 4 hour wait performance of about 2.2% (5.9%-3.7%), which may have worsened the PCT financial balance in the low age-needs areas by about half a million pounds. However, the modelling in Section 10 is subject to qualification.

SECTION 8: ACCOUNTING PRACTICES

Background to the NHS financial regime

8.1. The NHS financial regime is the system of funding, accounting and control applied to NHS organisations. The main elements of the regime are:

- 1. funding of purchasers/commissioners (PCTs) on a capitation basis weighted for age, additional needs and geographical variation in factor inputs;
- 2. funding of providers through income earned from purchasers/commissioners;
- 3. accounting on an accruals (i.e. resource) rather than cash basis.

8.2. The history of the financial regime can be traced back to the introduction of the internal market in the early 1990s and can be split into three distinct phases. The final phase – i.e. the one the NHS is currently operating in – relates to the third element of the regime listed above, the introduction of Resource Accounting and Budgeting (RAB). RAB was introduced across government in 2001/02, applying resource controls as well as cash controls on Departmental budgets.

8.3. To ease the implementation of RAB, certain flexibilities that could previously be used to mask underlying financial problems were slowly removed. For example, resource limits (as well as cash limits) were applied, the ability to transfer from capital to revenue budgets was gradually removed, unplanned financial support payments were abolished and, finally, planned support payments were no longer permitted.

8.4. In this section we discuss two of these flexibilities in more detail.

Capital to revenue transfers

8.5. Prior to the introduction of RAB, the Department of Health (and by association the NHS) had relied heavily on capital to revenue transfers to fund overspending on the revenue account. This can be seen more clearly in the table below.

Financial Year	Capital to Revenue £m
1997-98	123
1998-99	290
1999-00	367
2000-01	437
2001-02	320
2002-03	350
2003-04	388
2004-05	250
2005-06	200

Table 8.1: DH Capital to Revenue transfers 97/98 through 05/06

Source: Departmental Reports (1998-2006)

8.6. Up to 2001/02, virement between capital and revenue accounts was unlimited. However, following the introduction of RAB, (and the setting of the 'Golden Rule' by HM Treasury) limits on the use of capital to revenue transfers were set. The Department was able to negotiate a period over which transfers between capital and revenue were permitted up to a certain limit. These limits were reduced over time such that 2005/06 was the last year in which resources were permitted to be moved from capital to revenue budgets.

8.7. However, in 2004/05 and 2005/06 capital to revenue transfers were used specifically for the purpose of joint PCT/LA funding of certain parts of the Connecting for Health IM&T programme. Therefore, to all intents and purposes, 2003/04 was the final year in which transfers from the capital budget were used specifically to support the revenue account at a local level.

8.8. It is interesting to note that the £388 million of transferred capital resources in 2003/04 (£200 million of which was allocated to SHAs) is similar in scale to the £328 million in-year deficit that is revealed in 2004/05. This suggests that the removal of local virement flexibility in 2004/05 may in part explain the <u>timing</u> of aggregate deficits in 2004/05.

8.9. Up to and including 2003/04, the Department allocated some of its virement limit down to the NHS. This was usually done in proportion to the share of the capital budget held by each area. Local NHS areas were free to pass this virement flexibility down to commissioner organisations in their patch. The table below gives an example of the virement limits devolved to the NHS in 2002/03.

Table 1: 2002-03 Virement Limits	2002-03 Virement Limit	Less Capital Charges on Retained Estate	Remaining Virement Available
	£000	£000	£000
Northern & Yorkshire	13,005	0	13,005
Trent	10,020	0	10,020
Eastern	9,975	0	9,975
London	16,035	93	15,942
South East	16,536	4,241	12,295
South West	9,345	0	9,345
West Midlands	11,220	1,789	9,431
North West	13,864	0	13,864
TOTAL	100,000	6,123	93,877

Table 8.2: Allocation of £100 million local virement limit to NHS Regional Officesin 2002/03

Source: DH finance

8.10. Unfortunately, it is not possible to say how ROs or SHAs passed their virement limits down to the commissioner organisations in their patch. This means that we cannot definitively implicate the removal of capital to revenue transfers in both the timing <u>and</u> geography of deficits from 2004/05. However, if we look at the size of strategic capital allocations to SHAs in 2003/04 ranked by their share of the

total capital budget (as a proxy for the share of the £388 million worth of transferred resources allocated to the NHS in 2003/04) there is a slight tendency for the parts of the country in the top half of the list to have greater negative running balances, 2003/04 to 2005/06.

Org Code	Strategic HA Name	2003-04 Strategic Total	% of total	running balance 03/04 to 05/06 as % of allocation
		£000		
Q27	Birmingham and the Black Country	45,762	6.69%	0.73%
Q04	North West London	39,819	5.82%	-4.09%
Q15	Cheshire and Merseyside	37,805	5.52%	-0.51%
Q19	Surrey and Sussex	35,021	5.12%	-3.30%
Q14	Greater Manchester	34,333	5.02%	1.08%
Q06	North East London	33,376	4.88%	-0.76%
Q24	Trent	31,885	4.66%	0.36%
Q07	South East London	31,067	4.54%	-3.07%
Q12	West Yorkshire	29,119	4.26%	0.84%
Q01	Norfolk, Suffolk & Cambridgeshire	27,344	4.00%	-5.91%
Q05	North Central London	26,377	3.85%	-1.43%
Q08	South West London	23,819	3.48%	-3.84%
Q20	Avon, Gloucestershire & Wiltshire	23,693	3.46%	-1.67%
Q17	Hampshire and the Isle of Wight	21,777	3.18%	-0.70%
Q18	Kent and Medway	21,182	3.10%	-0.42%
Q03	Essex	21,028	3.07%	-2.12%
Q16	Thames Valley	20,612	3.01%	-2.65%
Q09	Northumberland and Tyne & Wear	19,840	2.90%	0.54%
Q13	Cumbria and Lancashire	19,522	2.85%	0.90%
Q25	Leicestershire, Northamptonshire & Rutland	19,319	2.82%	-2.93%
Q11	North and East Yorkshire & Northern Lincolnshire	17,860	2.61%	-2.76%
Q28	Coventry, Warwickshire, Herefordshire & Worcestershire	17,521	2.56%	-2.67%
Q21	South West Peninsula	16,552	2.42%	-0.50%
Q23	South Yorkshire	16,442	2.40%	2.33%
Q26	Shropshire & Staffordshire	14,196	2.07%	-2.71%
Q10	County Durham and Tees Valley	13,703	2.00%	0.23%
Q02	Bedfordshire & Hertfordshire	13,548	1.98%	-7.62%
Q22	Dorset and Somerset	11,777	1.72%	-0.10%
	Strategic Health Authority Totals	684,300		

Table 8.3: Strategic Capital allocations 2003/04

Source: DH finance

Financial support

8.11. Since NHS organisations have a duty to breakeven, SHAs (or their predecessor organisations) have a history of acting as 'honest broker' in moving resources around the organisations in their patch to give an impression of overall financial balance. This has the effect of masking the true underlying financial position of individual NHS organisations. It also creates an atmosphere in which organisations are not entirely responsible for balancing their books. Organisations that run a surplus are likely to have it removed, and those with a deficit see it disappear (as if by magic) at the end of the financial year.

8.12. To remove some of the distorting effects that the provision of financial support created, in 2003/04 the Department took away SHAs' powers to give organisations unplanned financial support. This meant that any support given to organisations had to be planned at the beginning of the financial year as part of the organisation's overall recovery plan. In addition, 2005/06 was the final year in which planned support payments were permitted, increasing NHS financial transparency still further.

The removal of planned support could not realistically have taken place sooner 8.13. than 2006/07 because this is the first year in which a significant share of activity has been commissioned on the basis of a national tariff. Up to this point, it will always have been possible to make *de facto* support payments by manipulating local prices.

Since a system of planned support operated up to and including 2005/06 and 8.14. because the net impact on the aggregate NHS deficit of moving resources from one part of the system to another is zero, it cannot be implicated in the emergence of aggregate deficits.

8.15. However, planned support may have created a changing geographical distribution in deficits if the pattern of planned support payments have changed over time. Since we identify a relationship between age/needs and deficits in section 5, we show in the following graphs (charts 8.1 through 8.3) the net contribution of planned support for each health economy against their level of age/needs, for the years 2003/04, 2004/05 and 2005/06.

8.16. If planned support was contributing to the relationship between age/needs and deficits referred to above we would expect there to be an increase in the net contribution of the less needy areas over the period concerned.



Chart 8.1: Net contribution to planned financial support 2003/04 against needs index by health

Age/Needs index 03/04 to 05/06



Chart 8.2: Net contribution to planned financial support 2004/05 against needs index by health economy

Age/Needs index 03/04 to 05/06





8.17. There is a very slight tendency for the needy areas to contribute more in planned support from 2003/04 to 2005/06 but this is <u>not</u> statistically significant.

8.18. The evidence suggests that the operation of a system of SHA administered planned support (or its removal from 2006/07) neither caused the arrival, nor significantly influenced the geographical distribution, of deficits in 2004/05.

8.19. In summary:

• the removal of local virement flexibility in 2004/05, the imperfect and gradual recognition of this during 2003 by the NHS, and the substantial scale of usage in 2003/4, are consistent with the arrival of significant aggregate deficits in that year;

- informal discussions with various managers and the underlying land/asset values suggest that capital to revenue transfers may have been used more in parts of the country south and east that subsequently went into significant deficit in 2004/05, although in the analysis above this relationship is not a strong one nor has it been tested at an organisation level (due to unavailability of data);
- the system of planned support neither caused the arrival of nor significantly influenced the geographical distribution in deficits in 2004/05.

SECTION 9: FINANCIAL MANAGEMENT AND CONTROL

Introduction

9.1. Preceding sections, have considered the contribution of external factors such as cost pressures, service performance targets or resourcing differences to generating deficits. This section focuses on another possible factor, briefly raised in Section 3, financial management and control. Specifically it considers how difficulties in internal financial management and control in responding to cost, service or resourcing pressures, exacerbated by rapid change and organisational turbulence, might adversely affect local financial balances.

9.2. It considers how this factor could contribute to the creation of deficits, how it could contribute to large deficits being concentrated in a small minority of NHS organisations – described in Section 3 - and how it could have contributed to deficits increasing in recent years. Arguments in this section may also be regarded as helping to understand why the deficit consequences of the expenditure shock variables are to be explored: good management would reduce the sensitivity of deficits to such shocks. It brings in evidence from various sources including Departmental simulation modelling of financial control, the Audit Commission's local evaluations, and the NHS Chief Executive's Panel. (A fuller discussion of the subject of this section is at Annex E.)

Financial Management

9.3. The National Audit Office has made its views on this quite clear. In its most recent report ¹⁴ on NHS finance, NAO said " the scale of variation in financial performance implies that some NHS bodies have financial management and governance arrangements which mean that, when faced with [external] financial pressures, they have coped better than others" (para 3.31).

9.4. The Audit Commission's report on local evaluations of NHS organisations ¹⁵ provides some illuminating information about the relationship between problems with financial *standing* (for which running a deficit is a key element) and quality of financial *management*; it found that financial management inadequacies are a good (almost a *sufficient*) indicator of financial standing problems (nine-tenths of NHS organisations with financial management inadequacies have financial status problems), and that financial standing problems are quite frequently (but not *necessarily*) associated with financial management inadequacies (two-fifths of NHS organisations with financial status problems have financial management inadequacies)¹⁶

 ¹⁴ Financial Management in the NHS, National Audit Office and Audit Commission, June 2006
 ¹⁵ Auditors' Local Evaluation, Summary Results, Audit Commission, October 2006

¹⁶ These measures of *financial management* and *financial standing* were not in use before 2005/6. There is some work related to earlier periods - the Audit Commission has found that while NHS organisations rated as weak in terms of their *financial stability* in 2003 (i.e. the year before aggregate deficits emerged) were found to be more likely to be in deficit in 2005/06 (at least in the case of PCTs), no such relationship existed in the case of organisations rated weak in terms of their *management capacity* - but the stability and management measures used there appear less closely related to the focus of this work.

	Financial Standing		
Financial Management	Organisations scoring adequate or better	Organisations scoring "inadequate"	
Organisations scoring adequate or better	62%	20%	
Organisations scoring "inadequate"	2%	16%	

Auditors' ratings of performance of NHS Organisations on financial management and financial standing for 2005/6

Source: Audit Commission Base: 538 NHS organisations

9.5. A key problem in financial management is that it often has to operate with imperfect information and imprecise levers. Departmental simulation modelling work ¹⁷ has demonstrated how large local annual imbalances could, indeed under plausible assumptions, *inevitably would*, be created *in a small proportion* of Trusts through difficulties in internal financial management generated *purely* from uncertain information about, and imprecise control ¹⁸ of, monthly financial flows even when funding was sufficient to meet requirements. This is consistent with observation - most Trusts achieve or approach financial balance, while a small proportion have large imbalances.

9.6. While the above provides a clear indication of the important role that problems in internal financial management and control are likely to play in generating local financial imbalances, it does not explain how the impact of this factor may have become more important and therefore contributed to increasing imbalances. It might be expected that financial management would be more difficult in an environment of rapid change, partly because of the effects this would have in disrupting information flows and control levers and partly because of its effects on managers' response to perceived imbalances, and that, therefore, the greater the rate of change and the level of organisational turbulence in the NHS the harder it would be for Trusts to maintain financial balance. This hypothesis is explored below.

Rapid change and organisational turbulence in the NHS

9.7. There has certainly in recent years been rapid change in NHS *finance* – from 2002 the annual real rate of financial growth increased from a historic average of around 3% to an unprecedented 7%, a rate of increase planned to run until 2008 and which will result in a NHS budget double what it was at the beginning of the decade.

¹⁷ G Royston, *System change and financial imbalance: the dynamics of deficits*, DH Corporate Analytical Team, May 2006

¹⁸ There is an extensive literature on the difficulties in achieving precise control of complex dynamic systems, see e.g. P Senge, *The Fifth Discipline*, Century NY, 1992.

This financial growth has allowed a considerable increase in the NHS *workforce* (and *pay*) and in *capital investment*– the NHS Plan announced 100 new hospital schemes by 2010 – nearly one every month for ten years.

9.8. Alongside this growth in finance, workforce and new building there has been major and rapid *organisational* change in the NHS, especially since around 2002, with for instance the creation of Foundation Trusts, the replacement of 100 health authorities by first 28 and then 10 Strategic Health Authorities, and the number of Primary Care Trusts reduced from about 300 to half that number.

Change, turbulence and financial imbalance in the NHS

9.9. It is not difficult to see how rapid change could precipitate breakdowns in financial control through creating a vicious circle - information flows becoming degraded or disrupted, key staff moving under the pressure of managing a rising number of initiatives in difficult conditions, so financial knowledge and expertise becoming even less available. A recent survey of members of the Healthcare Financial Management Association found that 15% had had to re-apply for an existing finance post in the last 18 months and that 75% of respondents felt that there were NHS shortages in finance skill-sets.

9.10. Rapid change and organisational turbulence could also affect behaviour in regard to deficits. Trusts (other than Foundation Trusts) are not permitted to retain surpluses. (Indeed the Health Select Committee has in the past publicly chastised officials for NHS underspending, on the grounds that to carry a surplus was to deprive the community of services that could otherwise have been provided.) In these circumstances a greater local tendency would be expected towards spending surpluses than towards recovering deficits¹⁹, especially when particular emphasis has been given to non-financial objectives such as reducing waiting times in A&E departments. Such a bias is likely to be increased in conditions of rapid financial growth, which tend to engender financial optimism, and organisational turbulence, for instance managers whose jobs are disappearing may not see a risk of overspend as their greatest concern.

9.11. There is little routine statistical information that helps us investigate the impact of change and turbulence on financial imbalances. There is however illuminating evidence from auditors' reports. NHS auditors issue *public interest reports* on Trusts for which they have concerns about their financial position. The recent Audit Commission report ²⁰ on financial failure in the NHS reviewed the public interest reports issued by auditors for 25 NHS Trusts in 2005/6. These Trusts had a combined deficit of £173m in 2004/5, approximately 70% of the total net overspend for that year.

9.12. The Commission found that *internal* factors contributed crucially to financial failure and noted lack of board cohesion and inability to challenge, frequently compounded by a high turnover of board directors "*a consistent theme among NHS*

¹⁹ These effects could be reduced by the management action of Strategic Health Authorities acting as banker for surpluses; but SHAs, like Trusts, will face information uncertainty and control precision difficulties over operating the banking function

²⁰ "Learning the lessons from financial failure in the NHS", Audit Commission, July 2006

organisations facing financial failure appears to be a lack of stable, permanent appointments at chief executive and finance director level". Another factor was weaknesses in the information available to the organisation particularly in financial monitoring. One Trust Medical Director said "We'd lost our feedback loop- we knew we had lost it but there was so much going on". One PCT had such poor monitoring data that it was unaware that nearly £1m had been paid to the local hospital for extra work.

9.13. A further source of evidence comes from NHS Chief Executives. The NHS Chief Executives' panel was asked for views on drivers of deficits. A typical response was "A significant proportion of deficits is due to a combination of organisational turbulence together with turbulence in the financial regime. In particular where organisations have had a significant level of turnover in key Executive positions it is difficult for such organisations to implement a coherent financial strategy to address the financial problems."

Conclusions

9.14. There is good evidence that difficulties in internal financial management and control have contributed to there being an increasing number of Trusts with sizeable deficits, and some evidence that these difficulties have been aggravated by rapid change and organisational turbulence in the NHS. However, given that these factors do not coherently explain the timing and new geographic patterns of deficits that are this Report's major concern, the role of management financial control is seen as an interactive one, amplifying deficits where management control is week and ameliorating/eliminating financial imbalance when controls are strong.

9.15. The *proportion* of total financial imbalance due to this effect remains uncertain. Firstly, as noted at the start of this paper, not all financial imbalance is likely to be a result of internal problems in financial management. For example PCTs with deficits tend to be in the more affluent and rural areas ²¹, which suggests other factors are also at work. Secondly, not all uncertainty and imprecision in financial information and control will arise from change and turbulence; some may arise from say, long standing inadequacies in information systems. Determining the contribution of each likely factor would require a major research investigation, and even then might not be fully resolvable. However, Departmental modelling work showed that that uncertainty and imprecision in monthly financial flows for Trusts could by itself generate annual imbalances comparable in number and size to those seen recently in the NHS, and the Audit Commission found that over two-fifths of NHS organisations with financial standing problems also had inadequate performance on financial management. It is likely that the contribution to deficits from difficulties in financial management and control is appreciable.

²¹ P Badrinath, RA Currell and P M Bradley, Characteristics of Primary Care Trusts in financial deficit or surplus- a comparative study in the English NHS, BMC Health Serv Res, 2006, 6: 64

SECTION 10: DEVELOPING AN EMPIRICAL MODEL OF THE IN-YEAR DEFICIT

Some principles of empirical models

10.1. This section of the paper brings together various arguments and evidence discussed in the previous sections to explore in a single framework the consequence of a group of factors for in-year Health Economy (HE) deficits. There are several purposes of the model:

- First, to explore the separate importance of each concept under discussion.
- Secondly, to give an indication of the quantitative importance of each effect. An important output of the model is a measure of the average impact on the in-year deficit of a unit change in a specific explanatory variable. As will be shown below, the average impact on HE in-year deficit of such a change can be expressed in financial (£) terms.
- Thirdly, to help identify how far there exist separate geographic effects unexplained by the explanatory variables we have discussed.

10.2. It is important to understand the limitations of the statistical models that are developed. Some variables are not directly observable or recordable; some data will be subject to measurement error, and; some will not be available to us for the period of interest. Furthermore, confirming that strong relationships exist as presupposed does not prove that one factor causes the other. Such a result from the statistical empirical model may suggest that a causal link exists but it cannot confirm that it does. For these reasons, the results in this section are used only as one piece of evidence concerning a view regarding deficits, and not as a decisive factor.

The Model

10.3. Building a model of the in-year balance rather than the accumulated balance offers one key advantage. The accumulated balance reflects the cumulative effect of health organisations' decision-making over time, the impact of historic demand-side shocks and the organisations' responses to them. A model, which sought to explain the differences in accumulated balances across health economies, would have to incorporate variables that captured those decisions, impacts and effects across time. The dataset required to support such a model would be vast. The in-year balance, on the other hand, measures health economies' response over the year to the position in which they find themselves at the start of the period as well as shocks which occur inperiod. The historical context is summarised by the accumulated balance, which is used as an explanatory variable alongside variables summarising characteristics of the health economy and others measuring shocks. So while the data requirements of a model of the in-year balance are considerably less, the model still provides us with insights into the impact of shocks on the health economies' finances.

10.4. The explanation of HE features, such as deficits, post a problem for the model – builder in that there is no single "decision-taker" with a coherent objective function, constrained in various ways, that can be readily summoned as the organising concept. Instead we imagine the Chief Executives of the PCT and provider organisations together making decisions at the end of one financial year about spending in various areas in the following year, subject to uncertainty about the cost implications of

various labour contracts, delivery and workforce targets, patient-referrals for elective care, A&E admissions and certain other factors. While certain choices are available to them – staffing being a critical expenditure – there are also certain influential factors that are known and "given" to the organisation in the short-run. These include: the accumulated deficit from the previous year, the degree of competition between the suppliers, the age of the building stock, the geographic context (which we think of as both SHA and degree of rurality), the increase in allocation from the previous year, the extent to which the PCT is receiving an allocation close to the target allocation, and recent past staff growth relative to allocation growth – to reflect how far, as a result of workforce targets or local hiring propensity in the recent past, workforce growth may have run ahead of the level or mix which is optimal to meet current demand. Given that a failure in either level or mix of labour skills is costly to adjust, organisations may incur deficits as a "least worst" alternative. In general, these factors will determine whether a PCT HE has sufficient "headroom" to absorb financial shocks without deficits in the following period.

10.5. As well as "given" variables to the decision takers there are also "shocks" in the form of patient demands at A&E and elective care, as well as the uncertain cost implications of new labour contracts and achieving improved standards of patient care in certain specific areas. We think of our decision takers incurring a certain level of such invoices at short notice and accepting deficits as the least worst outcome given the various pressures at hand

10.6. The variables used in the implementation of the models have been categorised into a small number of groupings:

- *State of the world* or 'given' variables: variables reflecting the state of the world at the end of 2003/4,
- *Shock* variables: which try to reflect in-period demand-side shocks that impact during 2004/5 upon the Health Economy and the costs of which may not have been fully anticipated,
- *Expenditure* variables: it can be shown that including both levels and growth in expenditure in a model will capture the impact of unanticipated expenditure upon the Health Economy's financial position. These variables measure the unanticipated financial consequences of demand-side shocks and therefore serve as alternatives to the *shock* variables.

10.7. The key explanatory variables used in the model and their presupposed relationship to the in-year balance are described in turn below (for a fuller description see **Annex F**):

• Accumulated balance: The historic balance as at the end of the previous financial year is thought to be negatively related to the in-year balance because of the requirement that NHS organisations breakeven over a financial year (or a small number of financial years). The onus will be on organisations with surpluses to spend down the following year and those in deficit to make cuts in order to get back towards financial balance.

- *Concentration index*: Health economies which buy services from a relatively small range of providers ought to be better able to coordinate and organise their finances. We include a measure of the extent to which PCTs commission from a range of providers in order to capture any such effect.
- Average age of health economy building stock: Investment in new buildings may impose a financial pressure on budgets, which could lead to deficits if maintenance costs for older buildings are relatively lower, for example. We include a measure of the age of the health economy's (or more specifically the providers' serving the health economy) building stock to test this hypothesis.
- *Rurality*: An adjustment is made in the NHS resource allocation formulae for the additional costs associated with delivering services in rural areas. However, it applies solely to Ambulance Trusts, but it has been argued that applying this adjustment to ambulance services alone is not sufficient. We would expect a measure of the degree to which a health economy serves rural populations or areas to be associated with deficit if this argument is valid.
- *Growth in allocations*: relatively large growth in allocation over the previous financial year should, in principle, make it easier for organisations to stay within budget or record surpluses. However, the counter-argument is that it is the growth in allocation relative to growth in demand for services that dictates the pressures on health economy finances, not growth in allocations per se.
- *Financial control*: The degree of overall control that health economies have over their finances will have a major impact on the in-year balance although it is hard to measure directly. A variable measuring the degree to which health economies move resources between organisation to control deficits is included as a proxy measure. The measure is constructed using data for 2003/4 to avoid any problems of circularity of argument.
- *Impact of formula changes*: Changes to the resource allocation formula will impact on the distribution of resources between Health Economies. Although some of the changes are dampened, it remains likely that some Health Economies will effectively face reductions in funding. Any changes to the formula at the start of the funding cycle may impact on the financial position in later years if the pattern of service provision is slow to respond to changes in funding. If this were the case we would expect large swings in the distribution of resources away from particular areas to be associated with deficits in those areas and vice versa. To test this we include a variable which measures the difference between what the Health Economy would have got under the 2002/3 distribution of resources (grossed up according to growth in total resources 2002/3-2004/5) and what they actually got in 2004/5.
- Level of expressed demand for healthcare: In each Health Economy there will be both an underlying level of demand for healthcare services and an expressed or met level of demand. Those Health Economies that meet a relatively larger fraction of the underlying demand (controlling for differences in population etc) would be expected to incur relatively more costs. A measure based on the total number of GP written referrals plus emergency admissions

divided by the age and needs adjusted population was included in the model to test this idea.

- Staff growth 2002-04 relative to allocations growth 2002/3-2005/6: Health economies which take-on staff at a faster rate than their growth in resources will be reducing the proportion of their income available for other expenditures. In addition, it is relatively difficult to reduce or change the skill-mix of the workforce quickly because of the nature of their contracts. As a result, those health economies which take on staff at a relatively faster rate in the earlier part of the funding cycle will restrict their ability to respond to unanticipated demand pressures within budget. A measure based on the percentage terms growth in staff in the earlier part of the funding cycle (2002-04) relative to the percentage terms growth in funding over the cycle (2002/3 to 2005/6) was included in the model to account for this effect.
- *PFI value*: New PFI schemes may impose an additional cost pressure on providers. The share of priority PFI costs that fall to the health economy is then likely to be negatively related to the in-year balance.
- *Reference Cost Index (RCI)*: The Reference cost index is a measure of the cost of providing services within a given health economy. Relatively high scores on the index imply that the health economy costs are higher than in other areas, in part attributable to inefficiencies in the provision of services. A high reference cost index would imply additional pressures on the health economy budget, particularly if actual demand for services is greater than planned for.
- *Growth in GP written referrals*: PCTs have imperfect controls on demand as they cannot oblige GP practices to curtail their referrals. Relatively rapid growth in GP written referrals in a given health economy may indicate unanticipated demands that will put pressure on budgets and may therefore be associated with in-year deficits.
- *Percentage of patients seen within 4 hours*: The target that 98% of patients presenting at A&E should be seen within 4 hours is one of the most well-known NHS targets. It is possible that good performance against this target in the previous year may have reflected investments that could not be reversed quickly in the face of demand shocks and would therefore be associated with a tendency towards deficit.
- *Improvement in percentage of patients seen within 4 hours*: As for the previous variable, an improvement in compliance with the central target between 2003/4 and 2004/5 was assumed to remove headroom to deal with demand side shocks and so be associated with a tendency toward deficits.
- *Growth in emergency admissions*: Patients presenting at A&E and requiring treatment cannot be turned away and so any unanticipated increase in emergency admissions will imply additional costs for providers beyond what was planned for.

- *QOF payments*: Payments under the Quality and Outcomes Framework (QOF) are thought to have been larger than anticipated and so may have constituted a demand side shock associated with a tendency toward deficit.
- *Outpatient waiting times target breaches*: Health Economies, which make relatively larger improvements in performance against waiting times targets, could be more likely to record deficits because of the cost of the improvements. To test this notion, a variable measuring the improvement (reduction) in the number of breaches of the 17-week waiting time target for Outpatients over the course of 2004/5 was included.
- *Inpatient waiting times*: A variable measuring the improvement in the percentage of patients on the in-patient waiting list waiting less than six months was included to test whether those Health Economies making the largest improvements were doing so by spending above budget.
- *Cancer's share of total spending*: Improving cancer care is a priority for the Department of Health including improvements in access to care, with particular emphasis in 2004/5. A variable measuring the percentage growth in Programme Budget spending on cancer relative to growth in total PB spending was tested in the model. Substantial changes in the proportion of spend on cancer could reflect pressures to spend that were not fully anticipated leading to deficits.
- *Regional dummy variables*: To test whether different parts of the country tended to have different in-year balances even when all of the other variables of interest had been controlled for, regional dummy variables were included. A regional dummy variable takes the value 1 when the health economy is in the region to which it relates and a value of 0 otherwise.

10.8. Two classes of statistical models have been developed. The first class of models seek to explain differences in the absolute in-year balance across health economies in terms of a series of explanatory variables (or *drivers*) where all financial variables are also expressed in absolute terms (*levels* models). A variant of this model aims to explain in-year deficits as a proportion of local allocation and also constructs relevant expenditure variables on a proportionate basis (*proportions* models). Both variants use *shock* variables rather than expenditure data to capture the impact of demand-side shocks.

10.9. Typically, we do not know which variant should be used and depend on being able to make a suitable case for preferring one over the other. If we believe that the impact of an explanatory variable in absolute terms varies with scale (big changes have bigger impacts) then we would employ a proportions model. If on the other hand we believed that the impact of an explanatory variable is independent of scale then we would employ a levels model. In the work reported here we have tested both levels and proportions models and found that in terms of the variables that feature in the best-performing of each type, there are no material differences. Constraints on the available data also mean that the type of model employed will never be a pure levels or pure proportions model. With this in mind we have opted to present only the results

from the levels model. A full set of results from the proportions model is included at Annex F.

10.10. The second class of models use expenditure data in place of the shock variables as an alternative (*expenditure* models). In econometric terminology we think of the first set of equations as "reduced form" models, and the second as structured equations. Using different classes of models allows us to explore whether the specific data we have used to represent the states of the world and demand side shocks and the nature of the model employed have an impact on our results. Which model we prefer will depend on the use to which we intend to put it. The levels model has the advantage over the expenditure model in that it allows us to examine the impact of specific, defined shocks on the in-year balance. However, we do not have complete data on all shocks to allow us to examine all of the impacts in detail. Using expenditure data gets round this problem as it reflects health economies' responses to all experienced shocks. The major disadvantage is that the expenditure data cannot be disaggregated beyond programme budget categories. Given the focus of this work, the *levels* model is likely to be preferred over the *expenditure* model.

10.11. In this section we present the results of the best *levels* model (**Table 10.1** presents the impacts of different key drivers on the in-year balance). Results from the best-performing proportions and expenditure models are presented in **Annex F**. Results from other key models in each class are also presented at **Annex F**. Some variables have not been included in these tables because they were found to have no influence on the in-year deficit once the effect of other explanatory factors had been controlled for. Other factors were omitted because the model showed them to have an influence on in-year deficits that was counter-intuitive. Such effects indicated that the detected association was likely to be down to a quirk of the data rather than a real underlying relationship. A fuller discussion of the results is available at **Annex F**.

Key Findings

Accumulated balance

10.12. The accumulated balance was shown to be an important driver of deficits. In both the levels and the expenditure model, we observe that an extra £1 of surplus brought forward from the previous year implies an extra 40-50p worth of in-year deficit. This is consistent with the argument that the requirement for organisations to be in balance on a yearly basis leads to organisations spending down any surplus the following year and cutting back to make good any deficit. Given a mean 2003/4 balance of +£239,000 and a mean in-year balance of -£1.08m we can deduce that the requirement to get back into balance in 2004/5 only accounted for around 10% of the 04/05 in-year deficit on average.

Log concentration index

10.13. The concentration index was intended to measure the degree of coordination between providers and purchaser in the Health economy with the presupposition being that higher concentration led to better financial controls. Ultimately, the data

show that there is no relationship with the 04/05 in-year deficit once the influence of other key drivers has been controlled for.

Average age of building stock

10.14. The average age of the health economy's building stock is positively related to the 04/05 in-year balance. In the levels and expenditure models, an extra year on the average age of the building stock implies around an extra £100k surplus, all other things being equal. The mean average building age across health economies was 37 years with a standard deviation of 7 years. A change from -1 to +1 standard deviation from the mean implies an impact of around £1.4m extra on the in-year balance, underlining the significant contribution of this explanatory variable.

Rurality

10.15. The data reveal strong evidence that health economies serving a more rural population had worse in-year balances than those with relatively more urban populations, all other things being equal. In the levels model we see that raising the number of people that lived in areas classed as rural by 10% implies a decline in the in-year balance of about £300k.

Growth in allocations 2003/4 to 2004/5

10.16. There is no evidence of a relationship between growth in allocation and inyear deficit once the influence of other key drivers have been controlled for.

Financial control

10.17. The financial control measure was found to be strongly, negatively related to the in-year deficit. As this variable was intended to measure the degree of financial control of the health economy, as a whole, the result is unsurprising. The levels model implies that the mean score on this indicator (0.9) would be associated with a deficit of around £550k-£600k or, to put it in context, around 50% of the mean 2004/5 in-year deficit.

Staff Growth

10.18. Staff growth 2002-4 relative to growth in allocation 2002/3 to 2005/6 is an important contributor to deficits. Although there is a degree of imprecision around the estimated impact on deficits, we estimate that a mean relative growth rate of 0.34 is associated with \pounds 500-600k additional in-year deficit. This would account for around half of the in-year deficit for an average health economy.

Impact of formula changes

10.19. There is some evidence of a relationship between the variable measuring the impact of the formula change and the in-year deficit, but the relationship is not statistically significant at even the 10% level. The larger the difference between what the Health Economy would have got had the 2002/3 distribution of resources applied

and what they actually got, the smaller the in-year balance. An extra £1m difference (loss) implies a decline in the in-year balance of around £100,000, on average.

Level of expressed demand for healthcare

10.20. The level of expressed demand for healthcare as measured by the fraction of the age-needs adjusted population accessing healthcare services (GP written referrals plus emergency admissions) was found to be negatively related to the in-year balance. An extra percent of the age and needs adjusted population accessing services implies a reduction in the in-year balance of some £100,000 all other things being equal.

PFI value

10.21. The health economy's share of priority PFI costs was not found to be related to in-year balances once the influence of other key drivers was controlled for. This may be because the PFI schemes were carefully planned and budgeted for or because any effect on deficits of PFI schemes are captured through the building age variable.

Reference cost index (RCI)

10.22. The reference cost index was not found to be related to in-year balances once the influence of other key drivers was controlled for.

Growth in GP written referrals

10.23. There is limited evidence of a negative relationship between the growth in GP written referrals and the in-year balance (relatively higher growth implies lower in-year balance), particularly so when regional dummy variables (see below) were excluded from the models.

Percentage of A&E patients seen within 4 hours

10.24. There is some evidence of a negative relationship with in-year balance as supposed but the evidence is weak.

Improvement in percentage of A&E patients seen within 4 hours, 2003/4 to 2004/5

10.25. There is strong evidence of the presupposed negative relationship in the data. In the best performing *levels* model we see that an additional percentage point improvement in compliance with the target implies a decline in the in-year balance of around £250k on average.

Growth in emergency admissions

10.26. As expected, the data support the hypothesis that growth in emergency admissions is negatively related to in-year balances. However, the implied impact of growth in emergency admissions upon the in-year balances appears relatively small. The levels model indicates that a growth in admissions of 10% would imply a decline in the in-year balance of around £50k.

QOF payments

10.27. There is some evidence of a relationship in the presupposed (negative) direction although it is not strong. QOF payments do feature in the best-performing levels model, which suggests that an extra £1,000 of QOF payments by the health economy implies a decrease in the in-year balance of around £700 on average, all other things being equal.

Outpatient waiting times target breaches

10.28. An improvement of one percentage point in the percentage of patients breaching the outpatients waiting times target is associated with around £350,000 of additional in-year deficit, all other things being equal.

Inpatient waiting times

10.29. The analysis shows little evidence of a relationship between improvements in the percentage of inpatients waiting less than six months and the in-year deficit. What evidence there is suggests a small negative relationship, if any.

Cancer's share of total spending

10.30. Growth in the share of total health economy spending directed at Cancer is associated with a larger in-year deficit. There is a degree of imprecision around the estimated impact of higher relative growth of cancer programmes relative to everything else, but we have estimated that growth of cancer of 1% higher than growth in total programme budget spend is associated with an decline in the in-year balance of around £150,000.

Regional dummy variables

10.31. Regardless of the specific model employed, the data seem to tell the same story. Health Economies based in the East of England²², the West Midlands²³, London or the South Central²⁴ SHA regions tended to have relatively worse in-year balances than other areas once all other factors are taken into account. Health Economies in the East of England tend to have in-year balances of £3m lower, and those in West Midlands, London or the South Central have in-year balances around £2m lower, than the rest of the country. The empirical model works by looking at how the key drivers relate to the in-year deficits, on average. **Table 10.2**, below, shows us that for the key shock variables the East of England, West Midlands, London and South central health economies have larger values than the rest of the country and it may be the case that their impact is not reflected appropriately by the model based on the average effect of these variables.

Discussion

²² Norfolk, Suffolk, Cambridgeshire, Essex, Bedfordshire and Hertfordshire

²³ Birmingham and the Black Country, Shropshire, Staffordshire, Herefordshire, Worcestershire, Warwickshire and Coventry

²⁴ Thames Valley, Hampshire and the Isle of Wight.

10.32. In this section of the paper we have brought together a range of arguments and evidence about how in-year deficits are formed to derive a single model which explains differences in in-year balances between health economies in terms of differences in a set of key explanatory variables or *drivers*.

Variable	Relationship with	Statistical Significance ²⁵	Operational/Economic Significa	
	(£'000s)	Significance	Change in Variable	Resulting (average) change
1. State of the world variables			4	9
Accumulated balance 2003/4 (£'000s)	Negative	***	Extra £1,000 of surplus in 2003/4	Decrease in-year balance in 2004/5 by around £450
Average age of building stock (years)	Positive	***	Extra 1 year on average age	Increase in-year balance in 2004/5 by around £100,000
Rurality (% of population living in rural areas)	Negative	***	Extra 10% of population living in rural areas	Decrease in-year balance in 2004/5 by around £300,000
Financial control measure 2003/4 (% point difference between PCT balance and balance of rest of HE)	Negative	**	Extra 1% difference between PCT and HE balance	Decrease in-year balance in 2004/5 by around £600,000
Impact of formula changes (loss in £'000s)	Negative	Insignificant at 10% level [†]	Extra £1,000 difference between allocation under old and new formulae	Decrease in-year balance in 2004/5 by around £100
Level of expressed demand (% of age-needs adjusted population accessing healthcare)	Negative	**	Extra 1% of adjusted population accessing healthcare services	Decrease in-year balance in 2004/5 by around £100,000
Staff growth (2002-04) relative to growth in allocation (2002/3 – 05/6) (%)	Negative	*	Staff growth increases by 1% between 2002 and 2004 but growth in allocations remains constant	Decrease in-year balance in 2004/5 by around £1.6m
2. Shock variables				
Change in % of patients seen within 4 hours in A&E between 2003/4 and 2004/5	Negative	**	Improvement of 1% in the % patients seen within 4 hours	Decrease in-year balance by around £250,000
Percentage terms growth in emergency admissions	Negative	**	10% growth in emergency admissions	Decrease in-year balance by around £50,000
QOF payments (£'000s)	Negative	*	Extra £1,000 payments made to practices	Decrease in-year balance by around £700
Change in % of patients breaching the 17 week waiting times target between 2003/4 and 2004/5	Negative	***	Improvement of 1% in the % patients breaching target	Decrease in-year balance by around £350,000
Change in % of patients on in- patient waiting list waiting less than six months		Insignificant at 10%	level, and negligible impa	ct [‡]
% Growth in Cancer PB spend relative to % growth in total programme budget spending	Negative	Insignificant at 10% level [†]	Increase in growth of cancer spend of 1% (holding growth of	Decrease in-year balance by around £150,000

Table 10.1: Model of 2004/5 PCT Health Economy in-year balance (Levels)

 $^{^{25}}$ * indicates significance at the 10% level, ** at the 5% level and *** at the 1% level

			total spend constant)	
3. Regional dummy variables				
East of England	Negative	***	HE PCT falls within boundary of East of England SHA	in-year balance in 2004/5 around £3.0m lower than if HE was in region 1,2,3,4,8 or 10
West Midlands, London or South Central	Negative	***	HE PCT falls within boundary of West Midlands, London or South Central SHAs	in-year balance in 2004/5 around £2.0m lower than if HE was in region 1,2,3,4,8 or 10

†:Whilst this variable is statistically significant at the 10% level, it has an associated t-statistic of greater than 1 implying some operational/economic significance.

‡: The coefficient on this variable has a t-statistic of less than one, so we treat the coefficient as being insignificantly different from zero implying that there is no economic significance of this variable. Therefore, no impacts are presented in the table.

Table 10.2: Mean values of levels model key drivers

Variable	"East of England" and "West Midlands, London or South Central"	Rest of Country
Accumulated balance 2003/4 (£'000s)	214.8	257.1
Average age of building stock (years)	39.0	36.1
Rurality (% of population living in rural areas)	25.3	33.1
Financial control measure 2003/4	1.1	0.8
Impact of formula changes (loss in £'000s) [†]	-172.5	124.5
Level of expressed demand (% of age-needs adjusted	48.8	47.3
population accessing healthcare)		
Staff growth (2002-4) relative to growth in allocation	0.4	0.3
(2002/3 – 2005/6) (%)		
Change in patients seen within 4 hours in A&E between	5.2	3.9
2003/4 and 2004/5		
Percentage terms growth in emergency admissions	13.0	10.5
QOF payments (£'000s)	2350.6	2158.3
Change in number of patients breaching OP waiting	0.1	0.4
target (%)		
Change in cancer's share of total spend $(\%)^{\dagger}$	1.5	1.5

†:although the coefficient on the variable was insignificant at the 10% level it had an associated t-statistic greater than 1 which is taken as evidence of some economic or operational significance.

10.33. The empirical model confirms that demand-side shocks will have impacted upon the observed in-year deficits. In **Table 10.3**, below, we present the mean values (and standard deviations) of our variables of interest for the health economies in the bottom quartile of the age-needs index and for those in the top quartile of the age-needs index. The difference in the mean values has been calculated and then using the estimated impacts on in-year balances, shown above, we can estimate what the financial impact would be of moving a health economy from the top to the bottom quartile.

Table 10.3: Mean values of key drivers for lowest and highest quartile age-needs health economies and the associated impact on the in-year balance

Variables	Lowest quartile age-needs health economies Mean (sd)	Highest quartile age-needs health economies Mean (sd)	Impact on in-year balance (moving high to low a-n) (£'000s)
Age needs	0.83	1.17	

	(0.04)	(0.07)	
In-year deficit	-3,616.62	796.90	-4,413.52
	(6,478.20)	(3,292.66)	
Accumulated balance 2003/4 (£'000s)	320.02	817.76	+217.71
	(5,204.63)	(4,461.89)	
Average age of building stock (years)	39.28	37.81	+165.77
	(6.37)	(6.57)	
Rurality (% of population living in rural areas)	39.23	17.74	-687.22
	(31.86)	(28.15)	
Financial control measure 2003/4	1.19	0.65	-337.16
	(1.60)	(0.74)	
Impact of formula changes (loss in $\pounds'000s)^{\dagger}$	949.11	-2,368.72	-250.16
	(2,175.00)	(5,747.10)	
Level of expressed demand (% of age-needs adjusted	50.92	44.68	-585.93
population accessing healthcare)	(5.23)	(4.25)	
Staff growth (2002-4) relative to growth in allocation	0.36	0.30	-88.86
(2002/3 – 2005/6) (%)	(0.24)	(0.15)	
Change in patients seen within 4 hours in A&E between	5.88	3.69	-513.83
2003/4 and 2004/5	(2.49)	(1.84)	
Percentage terms growth in emergency admissions	18.13	4.64	-90.22
	(75.78)	(6.95)	
QOF payments (£'000s)	2,269.31	2,208.23	-42.04
	(828.43)	(856.40)	
Change in the number of outpatient waiting times	0.23	0.43	+72.69
breaches (%)	(0.65)	(2.55)	
Growth in cancer spend relative to growth in total spend	2.00	1.54	-59.14
$(\%)^{\dagger}$	(3.29)	(4.07)	
Net impact			-2,198.39
N	75	75	

†:although the coefficient on the variable was insignificant at the 10% level it had an associated t-statistic greater than 1 which is taken as evidence of some economic or operational significance.

10.34. The Table shows us that moving a health economy from the top quartile to the bottom quartile would be associated with a decline in the in-year balance of around \pounds 4.4m. Roughly \pounds 650k of this deterioration in the financial position (about 15%) is attributable to differences in the mean scores on the shock variables. Low needs areas tended to make larger improvements in the A&E target (5.9 percentage point improvement versus a 3.7 improvement in high needs areas) and had much higher growth in emergency admissions.

10.35. A sizeable proportion of the difference in in-year balances between low and high needs areas can be explained by key state of the world variables. Differences in mean values for variables measuring rurality, financial control, impact of formula changes, level of expressed demand and staff growth between high and low needs areas account for around £2m of the difference in in-year balances (roughly 45% of the difference in mean in-year balances).

10.36. Differences in the key drivers discussed above do not explain all of the differences in in-year balances between Health Economies. There are a number of reasons why an empirical model such as the one presented will not explain all of this difference including measurement error on the key variables, over-simplification of the model or the omission of unobservable drivers such as efficiency. The model reported here systematically under-predicts extreme in-year balances (large surplus or large deficit) and this suggests that we are missing a key driver (or drivers) of deficits such as managerial ability. It is not unrealistic to assume that poor (good) management amplifies (diminishes) the effect of unexpected demand side shocks. An adverse shock such as large growth of emergency admissions may tend, on average, to increase deficits in all Health Economies as indicated by our model. However, the

emergency admission growth may impact much more heavily on the poorly managed Health Economies than on the well-run ones. In addition, **Section 8** raised the issue of virement as a potential explanatory factor yet we do not have access to suitable data to test empirically its impact on the observed in-year balances.

10.37. Some results of the model are surprising and warrant further work in the future. PFI value does not come through as being an important factor possibly because of its relationship with the average age of building stock variable but also perhaps because of difficulties in mapping the variable to health economies. The strength of the relationship between building age and the in-year balance is also surprising and further work should explore the reasons behind this.

Summary of findings

10.38. In this section of the paper, we reported on efforts to construct an empirical model, which explains variation in the in-year balance in terms of differences in key explanatory factors between Health Economies.

10.39. There are some key factors for which data are unavailable meaning that they could not be included in the model. In particular, PCT level data on the relative managerial ability and on capital to revenue transfers are missing. As a result, the model is vulnerable to the criticism that it is mis-specified and does not represent a complete account of the formation of in-year deficits. Therefore, it should be seen as a modest contribution to the evidence, rather than as providing a complete story.

10.40. Developing an empirical model allows us to test some of the hypotheses about the causes of deficits raised in preceding sections. We employed statistical tests to assess the economic or operational significance of key factors in causing in-year deficits. The empirical model also allowed us to quantify the impact of these key factors on the Health Economy in-year balance in order to provide a better understanding of their relative importance.

10.41. The empirical model results indicate that there are a number of inherited or *state of the world* variables that explain a proportion of the variation in in-year balances. Those Health Economies that serve a more rural population tend to have larger in-year deficits, which can be attributed to higher costs of providing services in these locations or perhaps reflect an inability to capture scale economies. Higher levels of expressed demand relative to underlying need are also associated with a tendency toward in-year deficit. Those areas that saw high growth in staff numbers relative to the growth in their resource allocations across the funding cycle also tended to have relatively worse in-year financial positions.

10.42. The results of the model indicated that as hypothesised, demand-side *shocks* do have a negative impact upon balances with larger shocks resulting in a relatively poorer in-year financial position. We observe that key *shock* variables include improvement in performance against the 4 hour A&E target, growth in emergency admissions and QoF payments - although the level of statistical significance for the latter is low.

SECTION 11: CONCLUDING COMMENTS AND POLICY IMPLICATIONS

Concluding comments

11.1. This report attempts to provide an explanation why, following a health dividend which we estimate to be worth about £13 billion from 2001/2 to 2004/5, the NHS incurred an in-year deficit of £328 million in 2004/5, with substantial financial pressure in following years. Evidence concerning six broad lines of explanation are presented in the preceding sections. This section attempts to draw together the arguments and evidence into a narrative explanation of deficits, before discussing a few of the key policy lessons.

11.2. It is important to bear in mind that until 2004/5 the NHS had been in approximate balance for four years with no particular region in deficit, not withstanding local clusters. Moreover, as the financial turmoil of 2004/5 became manifest, many health economies remained in surplus during this period and indeed over 50 managed to increase their surplus. Thus both increased financial heterogeneity and a tendency for deficits to be suddenly more commonplace in the South and East are features that must be consistent with a cogent explanation. The financial outcome in 2004/5 was sufficiently different in aggregate balance, and with new geographic features, for managerial failures alone to be unlikely to provide a satisfactory account of the evolution of deficits.

11.3. The argument that national wage and pharmaceutical contracts absorbed the Health Dividend, leaving little local scope for capacity growth without deficits, is inconsistent in particular with the massive capacity growth that actually occurred. This amounted to an extra 179,000 fte staff 2000-4 and 128,000 above trend growth, and was proportionately similar for clinical and non-clinical staff. It was perhaps the largest of any organisation in OECD countries, with an increase not dissimilar in size to the workforce of Bristol. Indeed the conversion of about 71% of the Dividend to 2004/5 into labour capacity and a further 15% into non- labour inputs, with only about 14% into higher wages, would appear a major success given the threat that such a large increase in sector-specific expenditure brings to cost inflationary pressures. Paragraph 9 and 10 below look more closely the consequence of the disproportionate emphasis on employment growth 2000-3 which may have left too little flexibility to meet delivery objectives in 2004/5 and thereby contributed to deficits.

11.4. Why then did deficits erupt in 2004/5? First, it should be noted that the allocation of income for PCTs in 2004/5 was known in December 2002 so that at an aggregate level deficits did not arise from an unanticipated loss of income, accounting conventions and capital to revenue apart, but must have been due to either unanticipated or "optimal" high expenditure given costs of rapid adjustment of spending, to which we return below. Sections 5 and 10 discuss how the evidence does not support the view that the new geographic or age/needs pattern of in-year deficits in 2004/5 is due to the inception of, or sluggish adjustment to, the allocations provided by a new resource allocation model in 2003/4. The explanation of the aggregate eruption, and geographic pattern, of the NHS deficit in 2004/5 appears to lie either with accounting changes or on the expenditure side of the budget for these low-needs areas, although a very small (under 10%) of the new deficit pattern may

have arisen from the new allocations.

11.5. The absence of data relating to two key factors in the emergence of deficits – a) the local use of capital to revenue transfers in 2003/04 and b) objective measures of the quality of local management – makes it difficult to quantify the impact that the remaining factors have played in causing deficits. The model of deficits that is described in section 10 cannot predict health economies with the most extreme financial outcomes (large deficits or large surpluses), possibly implying (as has been noted above) that important information relating to the use of capital to revenue transfers is missing and/or that unobserved management quality amplifies the impact of the factors for which reliable evidence exists.

11.6. HM Treasury introduced a new system of government accounting in 2001/2, and while the DH was allowed dispensation to delay its full introduction from 2003/4 until 2004/5, DH has sought to eliminate various opaque accounting practices. The issue that appears to have been of particular consequence for the aggregate level, as opposed to location, of deficits is the amendment of arrangements which permit virement from capital to current accounts – a practice which had been commonplace and which amounted to over £300 million in 2003/4, the financial year of their last availability at a local level. Whilst the NHS managers and clinicians gradually became aware of this development, there may have been a period which included 2004/5, during which medical and administrative leaders adjusted expectations concerning the financial "conjuring" that local financial managers might previously have undertaken on their behalf. (One consultant vividly described how before 2004/5 claims from local managers of impending deficits had been discounted by clinicians in anticipation of financial wizardry, whereas in 2004/5 clinicians had recognised too late that financial managers had largely lost their unnatural powers.)

11.7. The accounting practice change effectively withdrew between £200-300 million that had in 2003/4 been available for current expenditure. We have not had access to data concerning the scale of virement at local level, and so it has not been possible to contrast whether the PCT health economies now in deficit are those which previously have relied on virement from the capital account. We have given some modest evidence of this in Section 8, which suggests these transfers were probably more significant in 2003/4 in those regions that run into deficit in 2004/5. In these areas land values have increased more rapidly generating more valuable local assets. The NHS became aware of this loss of accounting freedom during 2003 and in some PCTs this would amount to as much as 1% of their projected expenditure. It must remain a major concern that the deficits partially reflect slow adjustment to tighter and more transparent accounting arrangements, and this issue is considered as probably the most important contributor to 2004/5 deficits.

11.8. The evidence in Sections 6, 7 and 10 also points towards a "proximate explanation" that NHS organisations were in part obliged, in part decided to choose, to undertake additional expenditure in 2004/5 to achieve a range of targets designed to improve patient welfare, and that the NHS was unable to address these objectives by "doing things differently" or reallocating resources provided by the HD from other specialties, but instead increased expenditures on inputs- largely staff, to meet these objectives. The difficulty here is not so much targets per se but the unequal distance to

target in different areas that was not recognised.

11.9. We have used a simple model of deficits to estimate the implied costs of one of the targets – that of reducing the percentage of patients waiting more than four hours in A&E. While amounting to a significant sum, it is not particularly large relative to the £13 billion health dividend. However, about £8 billion of the dividend was already expended by the beginning of 2004/5 and the large number of extra staff employed 2000/1-2003/4 to raise the quality of patient services, continued to be employed. For many organisations the trajectory of 2000/2004 hiring fitted perfectly with the exceptional additional costs that were required in 2004/5 to meet targets and wage bill pressures. For others, there was insufficient additional resource retained to meet the exceptional 2004/5 expenditure pressures without moving into deficit. Our evidence suggests that 2003/4-2004/5, the mean PCT achieved a 4.4 percentage point improvement in the proportion of A&E patients seen within four hours and that this may have impacted on expenditure as much as £1.1 million per PCT. Around this mean, however, there is evidence of considerable variation. This accounts for a tentative additional national estimated expenditure of about £330 million. However, we have described several other exceptional factors – albeit mostly less costly – and to some extent this additional expenditure would have been accommodated within the growing health dividend.

11.10. We have estimated the Dividend-related employment growth at 128,000 FTE staff, which we have noted provides simple testimony refuting the suggestion that the incremental resources have merely provided higher wage rates. However, a slightly less rapid growth of employment prior to 2004/5 in certain areas, would have allowed the accumulation of balances to meet future contingencies. It would also have given the PCTs a more flexible set of forward looking options to meet changing patterns of staff demand, given the costliness of adjusting both employment levels and skill mix. Sections 4 and10 give modest and tentative evidence that PCTs which allocated a larger share of their 2003-06 allocation to wage bill growth 2002-4, experienced greater financial difficulties in 2004/5. Given the large share of the wage bill in total expenditure, an expenditure-driven deficit is likely to imply over-employment, but the circumstances which promote an especially rapid growth of employment may in so-doing, also promote a less flexible health economy in the following years.

11.11. From the NHS perspective rapid workforce growth would be regarded as critical to the National Plan and Manifesto promises, albeit perhaps less emphasised since 2004 than other objectives. Employment growth would also be regarded as crucial to meeting the targets in 2004/5 although following strong employment growth 2000-04, it can be argued that greater emphasis in 2004/5 should have been placed on "doing things differently" and the productivity advances that are also part of the NHS Plan. Commentators have questioned how well joined-up the financial and human resource decisions have been in NHS organisations, which may have undermined both productivity gains and the maintenance of financial balance. A critical influence on decision taking appears to have been the concern that resources might be lost if not spent within the financial year; and to spend a significant sum quickly normally resulted in recruitment. It appears hard to over-emphasise the difficulty created by one year time horizons for rational budget planning and possibly their importance in explaining the recent emphasis on recruitment, rather than the design of new approaches to treating patients. The advent of the FT financial regime

should help address this issue.

11.12. In section 9 we see that poor financial management is almost an inevitably found to be present in the event of a large deficit, but that poor management is not necessary for large deficits since some poorly managed organisations do not have large deficits. The divergence of financial outcomes in 2004/05 (and the continued divergence in 2005/06 – larger deficits at the same time as larger surpluses) is only partially explained by the economic modelling which may suggest that managerial capacity interacts with and magnifies the impact of the (unevenly distributed) factors discussed in sections 4 through 8. As a measure of the importance of management failure in a few large deficit organisations, we discuss in Section 3 how only a small proportion of health economies are restored to financial health if the 15 most indebted organisations are returned to financial balance, *and that a good number are left with large deficits*.

11.13. It would be remiss to overlook the difficult budget management challenges that the programme of reforms and revenue increases posed 2001/2-2004/5. On the one hand the £13 billion or so increase in resource to 2004/5 greatly expended the available opportunities, but on the other the combination of major policy pressures and workforce targets, set in a context of organisational turbulence as commissioning powers were devolved from HAs to PCTs, and SHAs restructured, provided a context in which the probability of financial control withstanding the various pressures and achieving balance, was diminished. This was discussed in Section 9. In some regions, and perhaps Eastern England is the strongest example, financial control appears to have been a particular problem which should be viewed as additional to (and perhaps magnifying) the pressures discussed above. Certainly the NHS Chief Executives and the Audit Commission regard the rapid growth of resources and organisational change as having been accompanied by a reduction in the availability of reliable financial information and in the ability (and possibly also the motivation) to exert tight financial control. (One manager advised that the extra resources made saying "no" to doctors more difficult, since it was no longer possible to argue that additional resources were unavailable).

11.14. The discussion in Section 5 noted the tendency for deficits to become more commonplace in low-age needs areas in 2004/5 – which had not been the case in 2003/4. It has also been noted that deficits became more common in the South and East. How can this geographic correlation be related to the equally sudden emergence of aggregate deficits in 2004/5? The answer we provide is the following: national targets gave quite uneven "distances to travel" in different parts of the country. In the low age-needs areas - frequently in the South and East - there was a greater "distance to travel" for 5 of the 11 secondary sector targets - including inpatient and outpatient waiting and the A&E four hour wait, and no correlation for the other secondary sector targets. Emergency admissions growth was higher in low age-needs areas. Moreover, given that primary sector QOF payments were also higher in low age-needs areas, these incremental costs were largely higher in low age-needs areas. As a group, these expenditures may explain a significant proportion of the higher deficits in low ageneeds areas. It is important to be clear that this relationship between deficits and ageneeds is not driven by age-needs but largely by the weaker 2002 performance of the secondary sector, prior to national targets, in the low age-needs areas. The second

answer we offer – and perhaps was more important - is that the consequences of the change in accounting practices and in particular capital to revenue transfers were also more likely to have been more important in the South/East and low age-needs areas. These impacted in 2004/5. A third influence on the advent of regional patterns concerns the implications of London and parts of the South and East, in order to raise staffing, to rely on recruiting labour at a cost above average cost – to an extent less prevalent in other regions. Finally, noting the large share of PCT as opposed to NHS Trust deficits that arise in the highest deficit health economies (>5% of allocation) all of which are located in the Southern half of England. This may indicate that PCT management in the low age-needs areas was less strong than that elsewhere and when exceptional financial pressures were applied in 2004/5, were less able to avoid financial difficulties.

11.15. It was not perhaps fully anticipated how far national targets would impose unequal burdens in different regions of the country in such a way as to generate cumulative additional spend in certain areas that may have been a significant contributing factor towards the greater preponderance of deficits in these areas. This raises the question of how far unequal national targets should have received differential funding. The answer to which appears to be "no", if the funding model is already reflecting patient "needs", and the costs of meeting those needs, in a symmetric way to that in other regions. However, this raises the question of whether targets should more appropriately be conditioned upon the local circumstance, and acceptable costs of adjustment. It also raises the question concerning how far needs based on usage can be accurately modelled, if in different parts of the country, the secondary sector is offering quite different standards of service and thus attractiveness to patients in areas such as A&E. It appears, for example, that the rise in A&E attendance and admission in certain areas may in part be explained by improved A&E performance in those areas.

11.16. The evidence also points to high deficits in the more rural areas, holding constant various other factors. This deserves careful further examination.

11.17. The correlation between increased building age and greater tendency to surplus in the local economy raises challenging questions about the incentives such a financial regime implies for investment in new buildings. As with rurality, it deserves more careful analysis than is possible here.

The key lessons to be learned from the NHS' paradoxical experience of deficits at a time a rapid growth in resources are summarised in the Executive Summary section.