NHS Reform: Will it Deliver the Desired Health Outcomes?

Dr. Keith Palmer

April 2005

Important Note: The views expressed are those of the author. They do not necessarily represent the views of any of the organisations with which the author is affiliated

EXECUTIVE SUMMARY

The government is implementing a radical reform strategy for the NHS. The key instruments of reform are PCT commissioning, payment by results (PbR), patient choice, independent sector provision, Foundation Trusts and the 'must do' access and financial targets. The introduction of these reforms is unleashing strong forces that will drive change in behaviour and in the allocation of resources. The paper asks two questions. First will the reform instruments as currently designed deliver the desired outcomes as set out in eg the NHS Improvement Plan? Second are there enhancements to the instruments that will increase the likelihood that the desired outcomes will be more fully achieved?

The strategic objectives of the NHS are:

- Shorter waiting times for consistently high quality hospital care when hospital care is appropriate
- More patient care to be provided closer to home or in the home especially for patients with chronic conditions, the elderly and other services where care in the community is appropriate
- Greater focus on prevention of illness and early diagnosis and treatment

Delivering these strategic objectives within the available resource envelope will require:

- sustained significant improvements in the productivity of hospital providers (thereby freeing up resources to expand and improve services across the NHS) and in the quality of hospital care;
- expansion of services that provide new and improved types of 'care in the community' funded in part by freeing up resources currently used to provide admitted hospital care; and
- deployment of hospital-based expertise to provide and support enhanced care in the community.

Will the Reforms Deliver the Desired Health Outcomes?

The first key issue considered is whether the changes in behaviour and resource allocation induced by the reforms will deliver the level and pattern of services required if the NHS Improvement Plan objectives are to be achieved.

The conclusions reached are:

The Good News

- So long as there is consistent downward pressure on real average tariffs over the medium term, PbR will pressure NHS providers to improve efficiency and productivity, thereby increasing the effective capacity of existing services and freeing up resources that can be spent on expanding and improving services across the NHS including enhanced care in the community
- Patient choice and GP practice commissioning will pressure NHS providers to be more responsive to patient preferences and should transmit demand signals that result over time in improvement of existing services and in expansion of the type and location of services that patients prefer
- Independent sector provision will increase capacity of certain types of services, thereby increasing choice and improving access for those services

The Concerns

- The reforms will strongly encourage NHS providers to maximise admitted and day case hospital care (sucking resources into the hospital sector) and strongly discourage collaboration across care networks and support by hospitals for improved care in the community.

- The outcome of competition between providers in response to 'market forces' cannot be expected to result in the level and pattern of services required if the NHS Improvement Plan objectives are to be realised. The demand signals generated by patient choice and GP practice commissioning will not be strong enough to induce investment by providers in new and improved types of care in the community. Nor will hospitals contribute actively to support for care in the community. Strong market management led by PCTs will be required if the desired transfer of resources from the hospital sector to care in the community is to be achieved. PCTs cannot provide that leadership as the reform instruments are currently designed.

- Currently there are no incentives to improve the quality of care. In fact there are perverse incentives which pressure providers to maximise volume of activity even at the expense of quality. There are also perverse incentives which discourage investment in clinical innovation. Yet sustained improvements in patient care will only be realised if there are strong drivers to develop and adopt new medical technologies.

- There is a significant risk that the reforms will generate a lot of 'stranded' NHS capacity. Stranded capacity is capacity that is available for patients at low marginal cost but remains unused because of a deficiency of PCT purchasing power. There is already a lot of evidence of stranded capacity in parts of the NHS today even before many of the first wave independent treatment centres are on-stream. There will be a lot more as ISPs come fully on-stream and as waiting times shorten. Stranded capacity exists only because of the particular design chosen for PbR tariffs ie fixed tariffs set equal to system average costs.

- The reforms will generate 'net winners' with large unplanned financial surpluses and 'net losers' with large unplanned financial deficits. There are considerable risks that 'net losers' will seek to restore financial balance at the expense of the quality of care of their patients; that they will defer desirable services reconfiguration and seek to minimise the transfer of activity from their hospital into the community or to another hospital in a network of care to avoid loss of revenue; and that a significant number of net losers will 'fail' with uncertain but probably adverse consequences for their patients.

Are there Enhancements to the Reform Instruments that will Improve Health Outcomes?

The paper considers whether there are enhancements to the reform instruments that would retain the predicted benefits while addressing the concerns noted above. The conclusion of the paper is that there are a number of policy enhancements that, if adopted, will markedly improve the likelihood of fully achieving the government's objectives without impeding patient choice and GP practice commissioning or entry by independent sector providers and without diminishing in any way the pressures on providers to improve productivity. The adverse consequences of the reforms identified above are not an inevitable price that has to be paid to achieve the undoubted benefits.

Two part PbR Tariffs The most important suggestion is the adoption of a two part PbR tariff. This consists of a fixed payment per unit of contracted annual capacity and an activity payment per unit of activity actually provided, with the sum of the two payments for contracted volumes equal to average system cost ie the same payment as with the one part tariff. The activity payment would make up about 50% of the total payment. The analysis shows that two part tariffs retain strong incentives to improve productivity over time while reducing the perverse marginal incentives generated by fixed one part average cost tariffs. Two part tariffs reduce the disincentives to collaborate across care networks and to support care in the community; they reduce financial instability of net losers from patient choice and ISP entry and they reduce the risk of 'stranded' capacity. They strengthen the ability of PCTs to manage hospital demand and to achieve greater control over the allocation of resources between the hospital sector and primary care. They are wholly compatible with patient choice, GP practice commissioning and entry by ISPs. They may also facilitate contracting of second wave ISPs on terms that do not include such onerous 'use or pay' provisions as were included in first wave ISP contracts.

PCT Medium Term Contracts If two part PbR tariffs are adopted then an important consequence is that there is a shift in the balance of contracting power in favour of PCTs. They will face much reduced financial risk if they enter into multi-year (say, three year) volume contracts than they would with one part tariffs. They will have much greater influence over the pattern of services provision and over the pattern of investment in new services by virtue of their ability to vary contracted volumes of hospital services at the individual service level. They can use this influence to induce hospital Trusts to contract with them to support expanded and improved care in the community in a way that is not currently possible. Hospital providers would also benefit from reduced volume and revenue risks and greater ability to plan and implement services reconfiguration. The paper shows that such contracts are wholly compatible with the operation of patient choice and GP practice commissioning and entry by ISPs.

A number of other enhancements to the reform instruments are also proposed.

Transparent, Predictable and Demanding Tariff Trajectory Currently PbR tariffs are set annually and the future trajectory of average tariffs is not predictable. Stronger signals to improve productivity and rationalise services would be transmitted if the basis for tariff setting were more transparent and predictable and if the medium term efficiency target to be built into tariffs were known. The efficiency target should be announced and maintained for at least three years ahead. This will provide important forward price signals to providers to guide current decisions about streamlining service delivery, rationalising services and investing in new services.

Linking Payments Made to Providers with Achieved Quality of Care Two options, either of which would strengthen incentives to improve the quality of care, are proposed. The first option would provide in PCT/hospital provider contracts for payment of quality bonuses and penalties set by reference to actual quality performance achieved compared to locally agreed Quality Performance Standards set out in the contract. The second option would link similar bonuses and penalties to quality performance as assessed by the Healthcare Commission against national standards. Either option would put quality improvement at the heart of the reform programme. They would reduce existing perverse incentives to maximise the quantity of services provided at the expense of quality.

Improving the Financial Regime There is no doubt that the financial regime applicable to NHS Trusts is far too inflexible for a world in which payment by results, patient choice and entry by ISPs apply. A number of proposals for improving the financial regime for NHS Trusts are set out. They are:

- Allow all NHS providers to retain and reinvest income/expenditure surpluses
- Allow deficit NHS Trusts greater flexibility to adjust to unplanned financial deficits over more than one year
- Create an improved working capital mechanism (possibly via the NHS Bank) to channel surpluses generated by 'net winners' to fund the deficits of 'net losers'
- Create more effective mechanisms to manage hospital failure to ensure that the underlying cause of failure is addressed and that the quality of care of patients at 'failing Trusts' is not compromised.
- For all NHS providers including FTs allow a portion of the dividend on Public Dividend Capital to be deferred and paid with interest at a future date.

The current mechanisms for allocating capital funding across NHS Trusts and PCTs are unlikely to either deliver the level and pattern of investment required to deliver the NHS strategic objectives or to maximise patient benefits. Currently they are biased in favour of major new hospital build schemes at the expense of mid-size capacity expansions that often have the highest patient benefits per pound of invested capital. Proposals are set out for **improving the capital allocation mechanisms**.

Proposals are also made to more **fully exploit the potential of Foundation Trusts**. In particular it is important to clarify remaining uncertainties around the role of FTs including their role in providing managed care across the acute/primary divide; their role if any in addressing NHS Trust 'failure'; and whether they are to be permitted to build and operate 'independent' diagnostic and treatment centres.

In summary, the forces unleashed by the NHS reforms contain within them strong tensions. On the one hand they will result in shorter waiting times, force providers to be more responsive to patients' preferences and strengthen pressures on providers to improve productivity. However the same reforms will discourage collaboration across care networks and support by hospitals for care in the community, will not result in the level and pattern of services required to deliver the care in the community agenda, will not encourage quality improvement and clinical innovation and risks large scale financial destabilisation of the NHS. These adverse side-effects are not inevitable. There are enhancements of the policy instruments that, if adopted, will markedly improve the likelihood of fully achieving the government's objectives without diminishing in any way the benefits of the current reforms.

Contents

1. Introduction
2. Targets and Instruments of NHS Reform
3. Key Instruments of NHS Reform
4. Framework for the Analysis
5. Will the NHS Reforms Deliver the Desired Health Outcomes?
6. Are there Enhancements to the Reform Instruments that will Improve Health Outcomes?
7. Conclusions

Appendices

1. The Impact of Payment by Results
2. The Flow of Funds in the NHS
3. Implications of Two Part PbR Tariffs

1. Introduction

'Delivering the NHS Plan', published in 2002, set out a radical reform strategy for the NHS[1]. The key instruments of reform are PCT commissioning, payment by results (PbR), patient choice, independent sector provision (ISP), Foundation Trusts and the 'must do' clinical and financial targets. The introduction of these reforms is unleashing strong forces that will drive change in behaviour of purchasers and providers and as a result change the nature and pattern of services available to patients. The key issue is whether the induced changes will deliver the desired health services and patient outcomes.

The strategic objectives for the health system were initially set out in the NHS Plan[2]. They were further developed in subsequent Department of Health (DH) publications, most recently the NHS Improvement Plan (2004) and Creating a Patient-Led NHS (2005). Whereas the NHS Plan focussed heavily on access (waiting times), the NHS Improvement Plan set out broader objectives:

- shorter waiting times for consistently high quality hospital care whenever hospital care is appropriate;
- more patient care to be provided closer to home, or in the home, especially for patients with chronic disease, the elderly and other services where care closer to home is appropriate; and
- greater focus on prevention of illness and on early diagnosis and treatment especially for high risk groups.

Delivering these strategic objectives within the available resource envelope will require:

- sustained significant improvements in both the productivity of hospital providers and in the quality of hospital care;
- expansion of facilities and services providing new and improved types of 'care in the community'[3] in part funded by freeing up resources currently used to provide admitted hospital care; and
- deployment of hospital-based expertise to provide and support enhanced care in the community, thereby giving real meaning to the concept of integrated patient care.

'Delivering the NHS Plan' is now three years old. Many of the reform instruments are now in place and their probable impact is becoming clearer. It is therefore timely to consider whether the reform programme is likely to achieve its stated goals. The paper asks - and answers - two key questions:

1. Will the reform instruments, as currently designed, drive change in directions likely to deliver the strategic objectives set out in the NHS Improvement Plan?
2. Are there enhancements to the reform instruments that will improve the chances of more fully achieving the strategic objectives within the available resource envelope?

In 2002, several months after publication of 'Delivering the NHS Plan', I wrote a paper that, while fully supportive of the strategic thrust of the reform strategy, noted that the detailed design of the

[1] 'Delivering the NHS Plan' (2002), www.dh.gov.uk
[2] 'NHS Plan' (2000), 'NHS Improvement Plan' (2004) and 'Creating a Patient-Led NHS'(2005), www.dh.gov.uk
[3] The term 'care in the community' refers throughout to services enabling patient care closer to, or in, the home. It includes, in addition to primary care centres and hospital outpatient services, such innovations as GPs with a special interest (GPSI) practices supported by hospital- based specialists and nurse-led hospital-supported outreach programmes.

reform instruments would be crucial[4]. It argued that, if the reforms were to succeed, they must provide:

- 'a national framework for allocating current and capital resources to providers that will achieve maximum patient benefits at least cost, delivering the right sorts of services in the right locations in the right amounts;

- strengthened incentives on service providers to expand the volume and improve the quality of services by paying for results and allowing good performers to retain the surpluses for reinvestment in the health economy;

- enhanced operational and financial flexibility for good performers among the NHS Hospital Trusts to enable and empower them to respond to the strengthened incentives.'

The first part of this paper is concerned with whether the reforms are likely in practice to achieve these goals.

The contents of the paper are:

- section 2 sketches out how, in the reformed NHS, the reform instruments must create incentives that drive change in behaviour and in the allocation of current and capital resources in ways likely to deliver the desired health outcomes;

- section 3 summarises the characteristics of the key instruments of NHS reform;

- section 4 sets out the analytical framework used to consider the likely outcomes of NHS reform;

- section 5 addresses, and answers, the first key question, namely - Will the reform instruments, as currently designed, deliver the desired outcomes as set out in the NHS Improvement Plan?

- section 6 addresses, and answers, the second key question, namely - Are there Enhancements to the Reform Instruments that will Improve Health Outcomes?

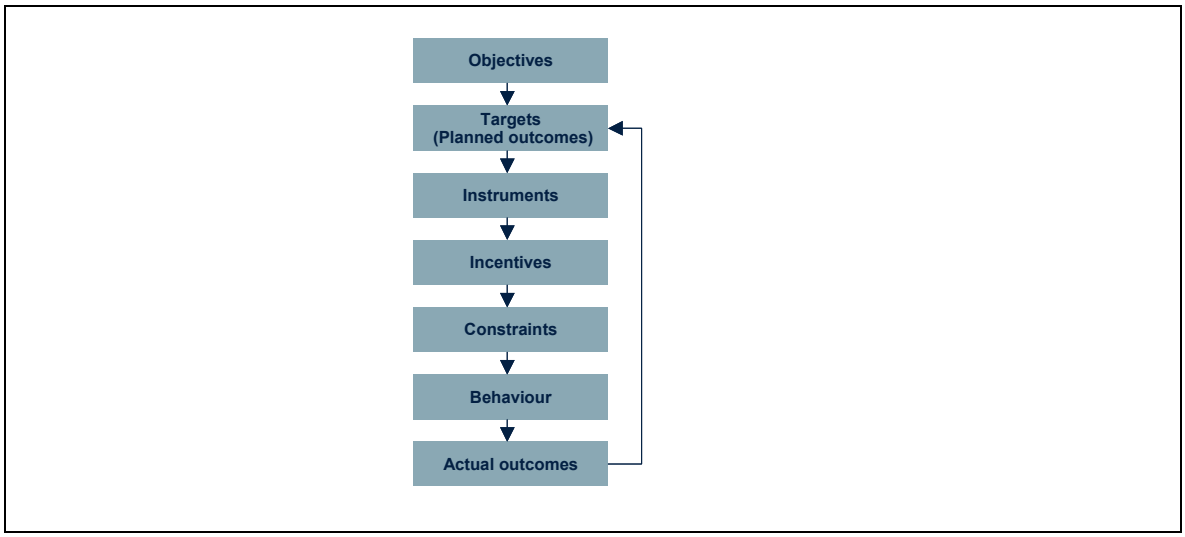
- section 7 sets out the conclusions.

[4] 'Foundation Trusts and the New NHS Architecture', Palmer, K. CEPA Discussion Paper (2002), www.cepa.co.uk

2. Targets and Instruments of NHS Reform

Figure 1 is a generic representation of what is required of any coherent reform strategy that relies for its delivery on decentralised decisions by purchasers and providers operating in a 'quasi-market'[2].

Figure 1 - Requirements of a reform strategy



Strategic objectives need to be translated into medium term targets i.e. planned outcomes. Although the ultimate target is improved patient outcomes, the proximate target has to be to deliver the level, type and location of services expected to be required by patients. In the short term the level, type and location of services available to patients is largely fixed, a legacy of past investment decisions. Patients can choose only between services that currently exist. However, in the medium term, the level, type and location of services can be changed as a result of investment by providers in new sorts of services, responding to expected future demand as reflected in PCT and GP-practice commissioning decisions and patient choice. Therefore, the future level, type and location of services from which patients can choose will be determined by capital investment decisions taken by providers today.

Implementation of the strategy is effected through the design and deployment of reform instruments. The NHS reform instruments eg payment-by-results and patient choice generate strong new incentives acting on purchasers and providers, inducing changes in behaviour and in the allocation of current and capital resources. The responses of providers to the incentives will depend on their values and objectives and on the nature of the constraints acting on them in particular in the NHS the requirement on NHS Trusts to achieve the 'must do' clinical and annual financial balance targets. The way in which providers respond to the new pressures will determine the present and future level, type and location of services available for patients to choose from.

The question for consideration is whether the pressures created by PbR, patient choice etc will result in NHS providers working actively to deliver the objectives sought by the government such as strengthening of care in the community and more patient responsive pattern of services. If the strategy 'works' then providers will deploy their resources to make available for patients the level and pattern of services that patients want and value. Most of the rest of the paper considers whether this is likely to be achieved with the current reform instruments.

[2] The term quasi-market has been used eg by the Kings Fund to refer to purchaser/provider systems which have some but not all the characteristics of a market.

3. Key Instruments of NHS Reform

The key instruments of NHS reform are:

PCT Commissioning

PCTs are responsible for purchasing primary and hospital care on behalf of their local populations. Each PCT receives a three-year funding allocation from DH. Allocations are determined by applying an adjusted capitation formula[5]. PCTs use these funds to procure or provide services in a way that is intended both to deliver national strategic objectives and meet local priorities within the constraints set by the DH including the requirement to achieve financial balance over a three year period.

PCTs enter into volume contracts with hospital providers to purchase services at PbR tariff prices. They contract for primary care and mental health services from GPs, mental health Trusts and others and in addition themselves provide certain primary care services.

In theory the PCTs' commissioning role gives them the power to shape the level, type and location of services available to patients. In practice this power is distinctly limited. PCTs are typically small, have limited commissioning expertise and experience, have very limited discretionary revenue funding and very limited access to capital funding. There is limited scope to manage demand for hospital services, yet an obligation to pay hospitals PbR prices for all activity provided. Practice-based commissioning and patient choice are likely further to dilute the ability of PCTs to shape the level, type and location of services within their area.

Payment by Results (PbR)

Payment by results is a key instrument of the reform programme. Key features of PbR are[6]:

- Fixed tariffs (prices) will be set nationally, initially for all hospital activity (procedures) and eventually for most procedures in primary care and mental health;
- The fixed tariffs will be paid by PCTs to providers for all activity provided[7]. PbR will be phased-in, replacing locally negotiated prices, commencing in 2004/5 and fully implemented by 2007/8;
- PbR tariffs are set equal to the national **average** NHS cost of providing each procedure. Average costs include depreciation of 'sunk' fixed assets and a return on the public sector capital employed (the so-called dividend on Public Dividend Capital[8]);
- The Market Forces Factor (MFF) is conceptually a tariff adjustment to take account of non-controllable regional variations in costs. In future it is likely that MFF payments will be made as annual block payments from DH to providers, rather than built into the tariffs as originally contemplated;

[5] The adjusted capitation funding formula enshrines the fundamental principle of 'social solidarity' in the NHS. In theory, everybody receives a comparable allocation - adjusted for perceived local 'need' - and PCTs use the funding to provide services free at the point of use.
[6] This is a brief summary of those aspects of PbR most relevant for this analysis. Further consideration of the short and longer term impact of PbR is set out in Appendix 1.
[7] The unit of activity is a HRG-spell, see Appendix 2 for more details.
[8] For further explanation of Public Dividend Capital see Palmer (2002)

- There are no quality differentials built into the tariffs. The same payment is made for each completed procedure, irrespective of the quality of service provided or of patient outcomes;

- PbR tariffs are set, and changed, annually and the basis for the calculation of the tariffs and the MFFs is not transparent. Nor is the basis for adjusting the tariffs from year to year for cost pressures and efficiency savings either predictable or transparent. Consequently providers face considerable uncertainty about the trend in the real average tariff for all activity provided and even more so about the trend in the average tariff that will be received by a particular service e.g. cancer, cardiac;

- The average cost of providing admitted hospital care is much higher per patient than for outpatient care or care in the community. Since tariffs are set to equal national average costs per unit of activity, average tariffs for admitted care are much higher than for outpatient care. As we see later in the paper, this has important implications for the way providers respond to the incentives created by fixed average cost tariffs.

Patient Choice

Patient choice is at the heart of the reform strategy. Choice will apply to elective hospital care from end-2005 but eventually (according to DH by 2008) it will be extended to almost all primary care, mental health and hospital services. In principle, patients will be able to choose not only between hospitals but also about where they wish to be treated eg closer to home, and about the type of care they wish to receive. However, these latter dimensions of choice will only be meaningful for patients if new types of services are developed. PbR will facilitate patient choice, with money following the patient. Since prices are fixed, competition between providers will focus on the quality and timeliness of care.

Independent Service Providers (ISPs)

The planned rapid growth of independent service providers (i.e. the private sector) is another key instrument of the NHS reforms:

- 'First wave' ISP contracts are signed and will be fully on-stream in 2005 providing significant additional elective capacity in certain services. They have 'use or pay' contract provisions which guarantee minimum annual payments irrespective of the volume of services that they provide and prices per procedure that are believed in many cases to be at a significant premium to the PbR tariff paid to NHS Trusts for the same procedure.
- A further 'wave' of treatment and diagnostic centres is planned to be contracted in 2005. Whether they will contain 'use or pay' provisions and be priced at a premium to PbR tariffs is not yet known.
- Announced plans are for 15% of all elective procedures in England to be undertaken by ISPs by 2008. In addition up to 10% of diagnostic tests are to be undertaken by ISPs. Additional involvement of ISPs in e.g. primary care is also contemplated.

Foundation Trusts (FTs)

'Delivering the NHS Plan' stated that FTs were to be a key instrument of NHS reform. They were to remain firmly within the NHS and subject to PCT commissioning, PbR and patient choice, just like NHS Trusts. However, they were to be granted greater operational and financial freedoms to enable them to lead expansion and improvement of hospital services and to develop new, more patient-responsive ways of delivering service. As it turns out, FTs have been granted rather modest operational and financial freedoms, the most significant of which are:

- FTs are allowed to retain and reinvest surpluses generated as a result of improved efficiency and higher productivity and from growing 'top line' revenues. This should have a major impact

on behaviour as FTs and their staff come to realise that improving productivity can result in immediate and lasting benefits for their patients;

- FTs have the right to borrow up to their Prudential Borrowing Limit (PBL) and they have access to the FT Financing Facility. These arrangements provide more flexible access to capital finance than is available to NHS Trusts, thereby enabling FTs to exploit more effectively opportunities to expand and improve services; and

- FTs have greater freedom to agree locally schemes to reward employees for good performance. The combination of stronger links between Trust performance and employee rewards, and the realisation that better performance at the Trust level generates greater benefits for its patients, should generate positive reinforcing drivers for further performance improvement.

Although government policy is to allow every NHS Trust the opportunity to become a FT by 2008, to date less than 30 have been created. The potential of FTs remains largely to be realised.

Must-Do Targets

Few in the NHS will disagree with the statement that the 'must do' access and financial targets have been the most important near-term drivers acting on NHS Trusts. In many cases the evidence suggests that managerial and financial resources have been committed to achieving these 'must do' targets, to the exclusion of just about everything else. Some would say that undue focus on achieving these short term targets has diverted attention and resources from achieving equally important medium term performance targets[3]. This would not be surprising since the sanctions for failure to meet them have often been draconian.

The financial regime applicable to NHS Trusts (but not FTs) has a major influence on the way they respond to the new incentives embedded in the reform instruments:

- The annual 'must do' financial targets - annual income/expenditure balance, the external financing limit, the capital cost absorption rate and the approved capital resource limit - create a strong financial straitjacket constricting all NHS Trusts[4]. All NHS Trusts are required to adjust to unplanned variations in revenues (arising because of e.g. patient choice) and costs within the year so as to restore financial balance by year-end, while also achieving the 'must do' access targets. Often this can only be done in ways that risk prejudicing the quality of patient care and/or through accounting sleight of hand eg 'brokerage'.

- The de facto situation of NHS Trusts is that they are businesses that have 100% debt financed balance sheets (because PDC is in effect debt, the fixed interest coupon must be paid each year) and operate with no free cash flow. This reality – a consequence of the financial regime within which they operate – means that they have no resilience to manage their way through financial shocks.

- NHS Trusts are wholly reliant for access to capital funding on discretionary allocations from DH channelled in part through SHAs. Most capital funding for the hospital sector is in the form of PFI schemes for large new hospital build projects. Although SHAs do allocate some discretionary capital funding to NHS Trusts in their region, this is usually small in amount, largely to finance essential operational capital expenditure (not service expansion) and is spread thinly across hospital Trusts 'equitably', rather than deployed strategically. Hospital Trusts

[3] That is not to deny that the waiting time targets have brought about important patient benefits. Later in the paper we consider whether greater progress could be made towards achieving wider NHS objectives while continuing to meet the waiting time targets.

[4] The nature and impact of these target financial ratios are discussed further in Palmer (2002).

wishing to spend, say, £20-50 million on services expansion or improvement will often find there is no way to access public (or private) capital to fund such schemes.

- Most capital funding available to PCTs is channelled through LIFT schemes. Although PCTs signing up to LIFT can decide how the available capital funding should be used, they are constrained from pursuing large schemes because the revenue expenditure consequences of the lease payments have to be absorbed within their future revenue funding allocations and they have very limited discretionary revenue funding. There is no mechanism for allocating public capital between PCTs and/or SHAs based on an assessment of what will achieve the best outcome for patients across the NHS as a whole in the future.

- Whereas current funding allocation decisions have to a considerable extent been decentralised to PCTs, decisions about capital funding in practice have not (to anything like the same extent). Hence decisions about funding of investment to change the level, type and location of services remain in reality largely under the control of DH.[11]

[11] As noted earlier, FTs in contrast have access to capital funding through the FTFF. However, even here de facto control remains with the DH since FTFF credit applications must be approved by DH and FTs cannot in practice access capital finance from the private sector.

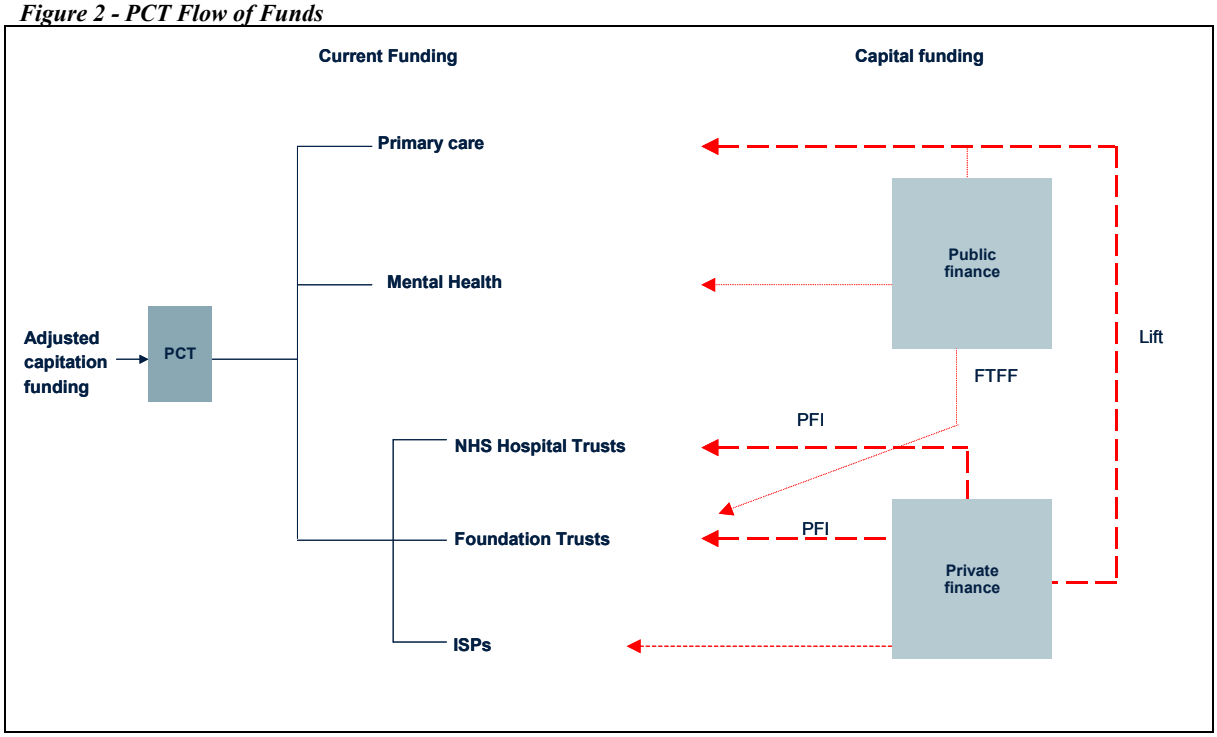
4. Framework for the Analysis

In this section, a framework for the subsequent analysis is set out. It highlights some key characteristics of the reformed NHS that are relevant when considering the likely responses of purchasers and providers to the reform instruments.

Flow of Funds in the NHS

The reformed NHS can be viewed as an economic system in which several hundred purchasers of services (PCTs) buy services from several hundred hospital providers (NHS Trusts, FTs and ISPs) and from primary care providers (e.g. GPs)[12]. **Effective demand** for services across the NHS in any period – how much, what type and in which location services are demanded – is determined by the aggregate purchasing decisions of PCTs[13]. GP- practice commissioning and patient choice do not change the monetary value of demand – which is fixed by PCT funding allocations – but they do affect the pattern of demand across providers[5]. The **supply** of services across the NHS in the same period is the aggregate of services provided to patients by all hospitals and primary care providers. The flow of funds is represented in a highly simplified schematic way in Figure 2.

Figure 2 - PCT Flow of Funds



The left side of Figure 2 shows the flow of **current** funding. PCTs spend part of their three yearly fund allocations purchasing primary care services and part purchasing hospital services. In a taxpayer funded service, where funding comes solely from fiscal transfers, **effective demand is fixed equal to the cash funding available**.

[12] This description is a simplification. It ignores mental health, the links with social services and the fact that many PCTs provide some services themselves. These simplifications do not affect the conclusions of the analysis.

[13] Effective demand should not be confused with patient need. The former reflects the real purchasing power of PCTs. Patient needs may (usually will) be much greater than effective demand ie what PCTs can afford to provide.

[5] GP-practice commissioning and patient choice change the pattern of demand eg between hospitals and may change the composition of demand eg less hospital care and more care in the community - but the monetary demand for services can never exceed available funding.

In any period if each PCT (and the NHS as a whole) is in income and expenditure balance then:

Aggregate PCT income = Aggregate PCT expenditure

Aggregate PCT Expenditure =

(Volume of hospital activity provided x Average PbR tariff for hospital activity provided)

+

(Volume of primary care activity provided x Average price for primary care activity)[14]

If NHS providers are in financial balance then:

Average revenue per unit of activity provided = Average cost of providing the services[6]

The significance of these relationships for NHS reform is that, for any given level of NHS funding:

- The total volume of activity that can be purchased by PCTs (and therefore that will be made available to patients) is fully determined by the average cost of provision of primary care and hospital care, respectively, and by their shares in total activity,
- The volume of activity provided to patients can **only** be increased **either**: by reducing the average cost of provision of primary care and/or hospital care; **or** by increasing the share in total activity of lower cost primary care and care in the community (and correspondingly reducing the share of hospital care),
- Any unplanned increase in hospital activity that is paid for by PCTs (eg higher A&E admissions and/or unplanned increases in elective hospital activity) results in a direct and immediate reduction in the volume of primary care and care in the community that can be commissioned by PCTs, and
- payments to ISPs result in a pound for pound reduction in the value of activity that PCTs can purchase from NHS primary care and hospital service providers. If the reduction in PCT demand is focused on hospital providers, it reduces demand for NHS hospital activity by an equivalent amount.

Real Purchasing Power of Budgets

The volume of activity that PCTs can purchase for patients depends on their real purchasing power. Although PCT cash budgets have increased rapidly in recent years, and will continue to increase rapidly until 2007/8, the real purchasing power of those budgets (i.e. the volume of services that can be purchased) has increased much less rapidly. This is because unit costs have also risen rapidly as a result of, inter alia, the new consultants' contract, junior doctors' pay, EWTD, Agenda for Change, the new GP contract, the national IT programme, creation of PCTs etc. These higher average costs per unit of activity reduce the volume of activity that can be purchased by PCTs, given their fixed cash budgets. Although cash budgets have increased at 9-10% per annum, the estimated increase in the real purchasing power of PCT budgets in recent years has been only about 2-3% pa and the

[14] A formal statement of these concepts and how they can be used to generate insights into the likely impact of the NHS reforms are set out in Appendix 2.

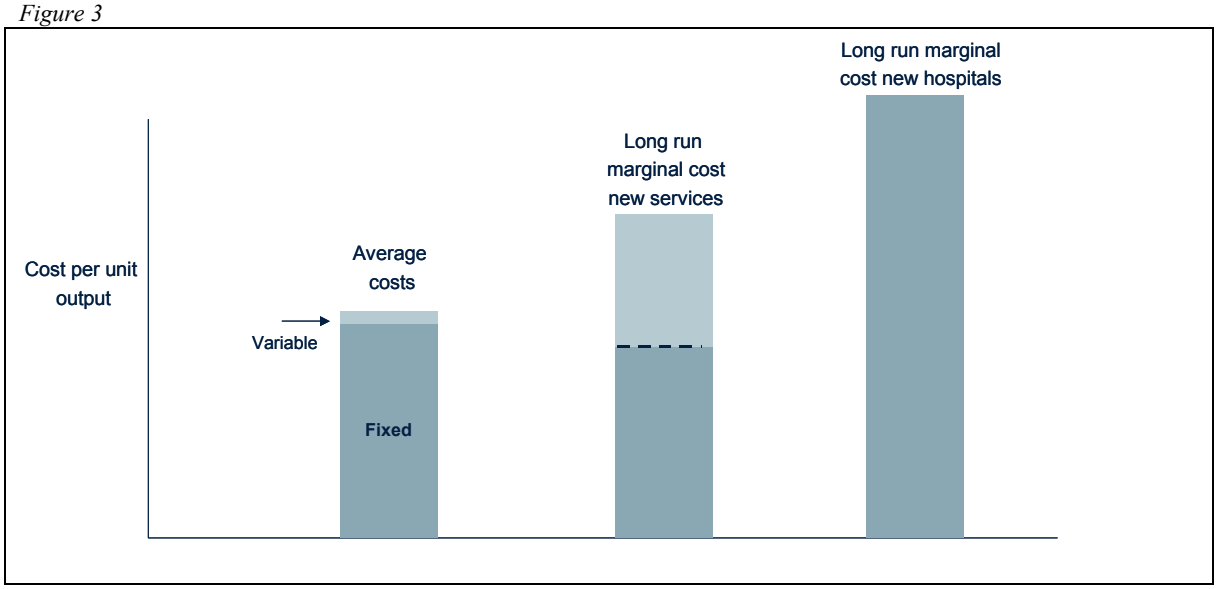
[6] Average costs include depreciation of fixed assets and the dividend on PDC.

projected increase from now to 2007/8 is estimated at about 3-4% pa. This is only slightly greater than is required to reduce waiting times in line with national access targets. There is little if any extra purchasing power available to improve the quality of services and address the challenges of the care closer to home agenda. From 2008, if the NHS budget remains a constant share of GDP ie growing in real terms at about 2.5% per annum, then the growth in real purchasing power of the NHS budget going forward is likely to be about 0-1% pa. The only way that additional or better services can then be provided will be to reduce average costs of provision of services ie improve provider productivity.

Cost Structure of Hospitals

Figure 3 describes a number of measures of cost. Average cost per unit of activity of a provider is the sum of its fixed and variable costs divided by the number of units of activity provided in the period. Average cost per unit of activity is the measure used to set PbR tariffs. Long run marginal cost (LRMC) of services expansion is a measure of the additional capital and revenue costs per unit to increase the volume of activity provided from an existing or new service. The LRMC of a new hospital is the additional capital and revenue costs per unit to build and operate a new hospital.

Figure 3



Average cost = Fixed costs (including depreciation of historic assets and dividend on PDC) plus variable costs per unit of activity provided

Long run marginal cost of new services = Additional capital and operating costs per unit of extra output from a service

Long run marginal cost of new hospital = Additional capital and operating costs per unit to build and operate a new hospital

In the short term, a very high share of a hospital's costs are fixed i.e. they remain more or less the same if the volume of activity provided increases or reduces. This is because the cost of buildings and building services, the payroll and the cost of servicing capital (ie the dividend on PDC) do not vary significantly as the volume of activity provided changes. All businesses with high fixed costs are prone to substantial swings in the surplus or deficit on their income and expenditure account if the volume of activity changes unexpectedly even a small amount.[15]

The LRMC of new hospital build is very often significantly greater than the average cost of provision of hospital services. This is because the depreciated value of 'sunk' fixed assets often does not adequately reflect the full replacement cost of those assets. When this is the case, given that PbR tariffs are set equal to average costs of service, future revenues generated by a new hospital will be insufficient to meet the costs of asset depreciation and the cost of capital.

[15] Economists refer to such businesses as having high 'operational gearing'. Small swings in 'top-line' revenues of such businesses give rise to amplified swings in 'bottom line' surpluses or deficits.

The LRMC of services expansion is usually significantly lower than the LRMC of new hospital build. This is because, within limits, the volume of activity provided by a service can be increased materially with only modest increases in certain fixed costs. Moreover, the LRMC of new services may be either lower or higher than the average cost of the service. Small increases in activity from an existing service will often be achievable at a marginal cost lower than average cost; large increases, because they require greater additional fixed costs, will often have marginal costs that are higher than the average cost of the service.[7]

Investment in Capacity Expansion

The right hand side of Figure 2 shows the flow of funds on capital account. Capital finance flows to the NHS from both the public and private sectors. Capital finance is used to expand and improve services and creates a claim over future (current) revenues of the recipients. Patient benefits will be maximised if capital finance is allocated to those individual investments where patient benefits per pound invested are highest. The financial value of patient benefits from an investment can be measured as the product of the additional volume of services provided over the life of the investment and the PbR tariff applicable to those services.[8] The higher the patient benefit per pound invested, the greater the aggregate patient benefits that can be provided within a fixed NHS capital budget. Conversely, since PbR tariffs are set to equal average costs, if the LRMC of new capacity is higher than average cost (as is the case with many new hospital schemes – see above) then the financial benefits from the investment will be lower than the costs. Consequently, providers will be reluctant to proceed with those investments. If they do proceed, the claim created over future current revenues will reduce the volume of other services that they are able to provide to their patients.

In response to changes in the pattern of demand and supply across the NHS, hospitals will sometimes need to reconfigure their services expanding in some areas, contracting in others and changing the way existing services are provided for patients. To do this requires three things:

- **time** it takes several years to plan and implement services reconfiguration if it is to be achieved without disruption to ongoing patient services
- **some degree of certainty** about the pattern of services provision that will be required over the medium term (because of the lags between conception and completion of reconfiguration projects), and
- **money** it costs money to undertake services reconfiguration and access to transitional capital and current funding is needed.

The current 'must do' annual financial balance requirements imposed on NHS Trusts do not give them either the time or the money required to reconfigure services. Nor do they have a reasonable degree of certainty about the future level and pattern of demand for services (for reasons elaborated in section 5).

The Health System as a Network of Care

The NHS is a network of care providers. Patients may be treated by GPs; they may be referred to DGHs and/or specialist hospitals and/or ISPs; DGHs may refer patients to specialist hospitals; and discharged patients often require continuing care in the community (Figure 4).

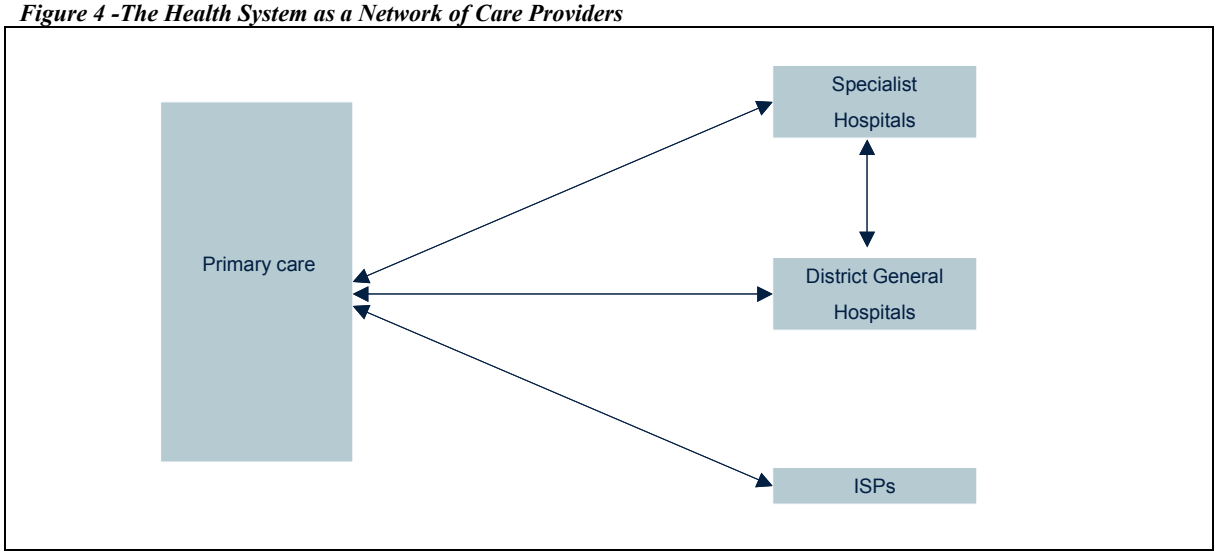
The NHS Improvement Plan highlights the importance of breaching the divide between acute and primary care and expanding provision of services closer to home. In certain services e.g. cancer and cardiac, working collaboratively across care networks involving GPs, DGHs and tertiary centres is widely regarded as best practice. In chronic disease management, care in the community provided and/or supported by hospital-based specialists is widely regarded as best practice. A key challenge

[7] In section 5 we explain why these marginal cost/average cost relationships are important.

[8] The validity of this proposition depends on the PbR tariffs adequately reflecting the value to patients of the procedures. A formal statement of the investment decision rule is set out in Appendix 2.

for the NHS is to design reform instruments that create positive incentives acting on purchasers and providers that will drive behaviour in ways most likely to achieve these objectives.

Figure 4 -The Health System as a Network of Care Providers



How Real Markets Work

The reformed NHS has been referred to as a quasi-market for healthcare. Before turning to the analysis of how the reformed NHS is likely to work, it is useful to summarise how 'real' competitive markets work. Some key relevant characteristics of 'real' markets are:

- Output prices are flexible. They both determine, and are determined by, the purchasing decisions of buyers and the response of suppliers to purchaser demand. Flexible prices are the fundamental mechanism that allocates current and capital resources efficiently in competitive markets.

- Over the long run, average output prices tend to approximate the long run marginal cost (LRMC) of new capacity. This is because if expected prices are higher than LRMC new investment increases supply, tending to push down prices, and conversely if expected prices are lower than LRMC new investment ceases, and prices tend to rise as demand increases. If the LRMC of new capacity does not equal historic average cost then there is no presumption that the average price over the medium term will equal average cost;

- In the short run, the price tends to the level where marginal revenue[9] equals short run marginal cost (SRMC). If there is surplus capacity in a sector characterised by high fixed costs and low variable costs, the price is likely to fall well below the LRMC, boosting demand and ensuring that maximum use is made of available capacity;

- Productivity improves steadily over time as a result of innovation and process optimisation and competition passes a high share of those benefits to consumers as lower prices;

- Output prices are generally higher for higher quality goods and services, thereby compensating suppliers for the higher costs often associated with providing higher quality goods and services;

- Decisions by suppliers to invest in expansion or improvement of output are based on judgments about whether the expected additional revenues (a function of the extra volumes produced and

[9] Marginal revenue is the change in revenue resulting from a small change in the volume of output.

sold and of the future prices received) exceed the additional costs (including the risk adjusted cost of capital) over the life of the proposed investment;

- In 'network' businesses and businesses with long 'supply chains', there are significant risks arising from the interdependency of actions by suppliers and customers across the network. An investment by one supplier may only be viable if a complementary investment or purchase is or is not made by another supplier or purchaser in the network. In private sector markets these interdependency risks are typically addressed either through corporate consolidation (mergers and acquisitions) or by suppliers and purchasers entering into medium term contracts prior to investing in new capacity (as the new entrant ISPs have done in the NHS); and

- In 'real' markets failure of businesses is typically handled through some or all of: takeover; change in the management team; and restructuring and rationalisation often involving significant restructuring costs, leading to restoration of a viable, often smaller, business focussed on areas of the market where it can be successful and competitive.

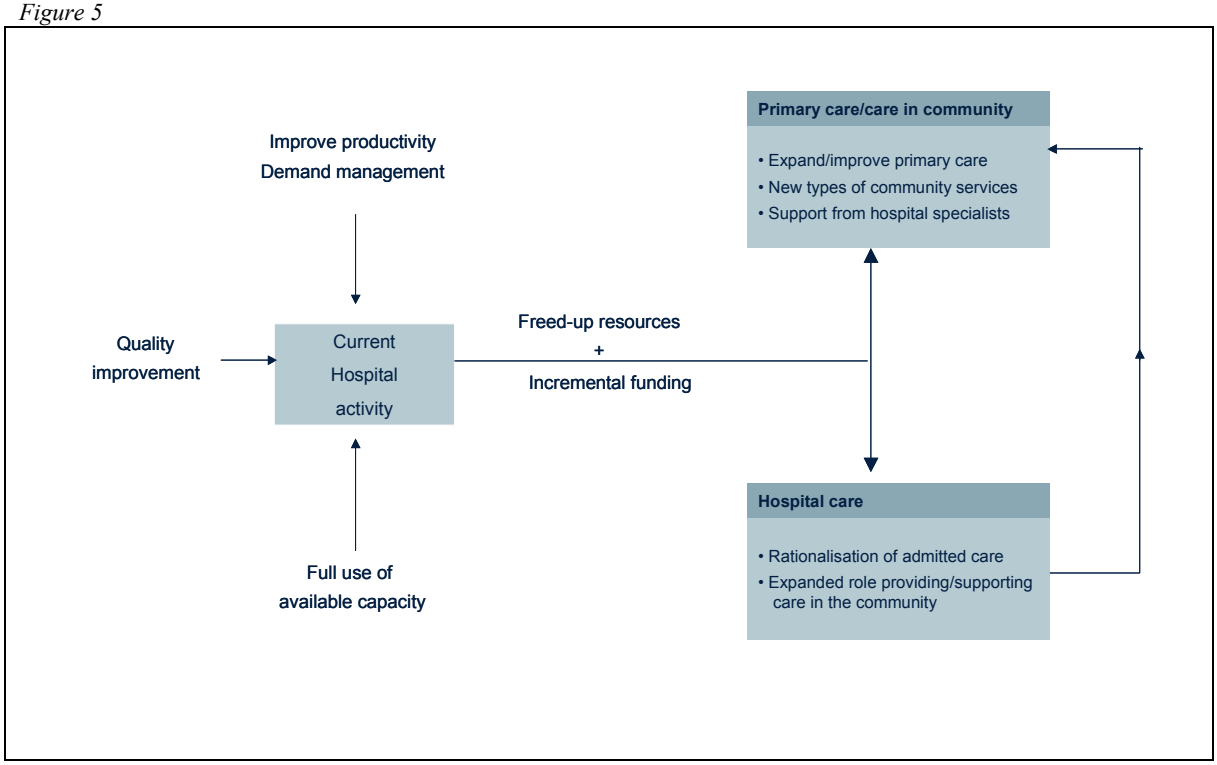
As we shall see in the next section, the reformed NHS has, in reality, few of these characteristics of a 'real' market.

5. Will the NHS Reforms Deliver the Desired Health Outcomes?

What are the reforms trying to achieve?

Figure 5 summarises how the pressures induced by reform ought to operate if the objectives set out in the NHS Improvement Plan are to be achieved. Current hospital capacity should be fully utilised, productivity should improve as should the quality of care. Effective hospital demand management should reduce the volume of high cost admitted hospital care. The result ought to be to free up resources to expand and improve hospital based outpatient care and care in the community.

Figure 5



What are the reforms likely to achieve

The analysis in this section considers whether the responses of purchasers and providers to the reforms are likely to deliver the outcomes set out above. First, it considers the responses taking as given the existing level, type and location of facilities and services. It then considers the level and pattern of investment in new facilities and services likely to be induced by the reform instruments. The responses of NHS Trusts and FTs are distinguished because they are likely to respond differently given that they are subject to different financial regimes and have different access to capital funding.

Payment by Results will impose two distinct pressures on providers. The first will ensure that each Trust seeks to keep its average costs as low or lower than the average tariff for the portfolio of services it provides. The second set of pressures result from marginal incentives inherent in fixed average cost tariffs. These are considered in turn.

Introduction of PbR tariffs will have 'introductory effects' relating to the transition from locally negotiated prices to PbR tariffs and 'sustained effects' that will persist over time. The introductory effects will give rise to 'winners' and 'losers' depending on each Trust's initial reference costs relative to the national average. The introduction of PbR will tend to suck resources into hospitals from PCTs as coding completeness and accuracy improve; and if shorter A&E waits induce more hospital attendances and admissions. There will be some risk that some Trusts will successfully game the system. These introductory effects will be dampened to some extent, in practice, by the

phased introduction of tariffs and by arrangements currently in place to limit the magnitude of gains and losses[16].

The sustained effects of PbR will be to be to generate incentives on providers to improve productivity. The extent of the pressures will depend to a significant extent on the trajectory of average real tariffs over the medium term. This will depend on the precise mechanics of the tariff adjustment process and on the value of the efficiency factor built into prospective tariffs. The efficiency factor will play a key role in driving productivity improvement. It will determine the amount of resources that are freed up for redeployment elsewhere in the NHS. If it is set 'too high' it risks causing harm to the quality of patient services and inducing widespread hospital failure. If it is set 'too low', then fewer resources will be freed up for redeployment to expand and improve care across the NHS.[17]

If a predictable and demanding reduction in real average tariffs over the medium term is set then the impact of PbR is likely to be to induce major efforts by hospital providers to reduce average length of stay, to increase day case rates and to increase utilisation of capacity. These are all effective ways to reduce unit costs which generally involve an improvement in patient care. Improving productivity will, in turn, reduce the demand for bed nights and staff required to provide inpatient services, thereby effectively increasing the annual capacity of hospital services.

Fixed average cost tariffs also generate strong marginal incentives that impact strongly on decisions by providers about the level and composition of services they choose to provide. A key observation is that, as PbR tariffs are designed, **marginal revenue for each provider equals system average cost.** Marginal revenue is the revenue gained/lost from providing one more/less unit of activity. Since (as noted in section 4), average cost is much greater for all hospitals than short term marginal cost, there will be very strong incentives on all hospital providers to maximise activity and to avoid losing activity. This is because there is a high marginal surplus (extra revenue less extra cost) for each extra unit gained and a high marginal deficit (loss of revenue less cost savings) for each unit of activity foregone.

Another key observation is that the way PbR is designed, marginal revenue is much greater for additional admitted hospital activity than it is for additional outpatient care/care in the community. This generates strong incentives on hospitals to allocate resources to maximise admitted hospital care.

These marginal incentives, together with the strong pressures on Trusts to achieve annual financial balance, will generate strong pressures on all NHS providers to:

- maximise hospital activity, particularly admitted care and day case treatment, and
- discourage working collaboratively across care networks wherever this would lead to a loss of marginal revenue (and therefore increase the Trust's deficit), and
- discourage using hospital-based clinical resources to provide and/or support care in the community wherever this would lead to a loss of marginal revenue (and particularly where it would reduce the number of admitted patients).

[16] See Appendix 1 for further discussion of the introductory and sustained effects of PbR. Here, the paper focuses primarily on the sustained effects.

[17] The Wanless report stressed the importance of sustained productivity improvements. It argued that annual productivity improvements of 2.5- 3% are necessary if maximum health benefits are to be derived from the extra funding. The 2005/6 PbR tariffs incorporated a 1.7% efficiency factor.

These incentives will be **strong** – because there is a large gap between average costs and short run marginal costs; they will be **prevalent** – because they apply to all providers; they will be **persistent** – because tariffs are fixed and there is no mechanism for 'closing the gap'; and they will be **perverse** because they will induce behaviour inconsistent with achieving the objectives of the NHS Improvement Plan.

How Trusts, and their staff, respond to these pressures is considered later, once the further pressures introduced by patient choice and ISPs have been addressed.

Patient choice and PbR Eventually patient choice will allow patients to choose which provider, type of service and the location of service they prefer. In the short term, however, choice will be constrained by the current availability of facilities and services and by supply constraints in the system. The presumption underlying the focus on patient choice is three fold. First, that patients will immediately benefit from the ability to choose more timely and/or better care. Second, that patients by choosing not to use certain facilities and services will send powerful signals to providers that are losing patients to act to make their services more attractive to patients. Third, that over time the demand signals generated by patient choice will lead to expansion of those types of facilities and services that patients prefer and contraction of those services that they do not prefer.

From a provider perspective, patient choice will change the level and pattern of demand for services in unpredictable and largely uncontrollable ways. For all providers 'volume risk' (and, therefore, given the way PbR works, also revenue risk) will be increased, manifesting itself as greater unplanned variances in revenues within-year. There will be 'net winners' - Trusts which gain more patients/revenues than they lose - and 'net losers'. Given the high operational gearing (see section 4) and the large 'gap' between marginal revenue and marginal cost, large unplanned within-year surpluses and deficits will be generated by 'net winners' and 'net losers', respectively. At the system level, however, surpluses and deficits arising from patient choice will broadly cancel out.

From a PCT purchaser perspective, the main effect of patient choice is to make it unattractive to enter into medium term guaranteed volume contracts with NHS providers. A contractual commitment to pay for a specific volume of activity could result, if patients choose to be treated elsewhere, in the PCT paying twice.

How should one expect NHS providers to respond to the combined pressures of PbR and patient choice? A useful distinction can be made between 'quality-enhancing' responses and 'quality-harming' responses. The former are planned responses that both reduce unit costs and (usually) improve the quality of patient e.g. shorter lengths of stay, higher day-case rates. Evidence from other countries where prospective payment arrangements have been introduced suggests that PbR will strengthen incentives to adopt quality enhancing, productivity improving responses over time. 'Quality harming' responses, in contrast, are short term ad hoc responses adopted in an attempt to restore end year financial balance which may increase the risks of deterioration in the quality of care and defer or prevent desirable services reconfiguration. Such responses typically include hiring freezes, deferral of clinical improvements and deferral/reduction of capital maintenance programmes.

Figure 6 sets out the likely responses of 'net winners' and 'net losers' from PbR and patient choice distinguishing potential responses by NHS Trusts and FTs.

'Net losers' among **NHS Trusts**, under intense pressure to achieve end-year financial balance, are more likely to adopt quality-harming responses. There are considerable risks that either they will restore balance at the expense of the quality of care for their patients or they will end up with large financial deficits. They are also likely to defer desirable services reconfiguration, which takes time and money to implement, and to seek to maximise admitted activity because marginal revenue is highest from that activity. They will seek to minimise the transfer of activity from the hospital into the community or to another hospital in a network of care to avoid loss of revenue. In contrast, an

FT that is a **'net loser'**, operating within a less restrictive financial regime and with access to working capital is more likely to implement planned measures to restore financial balance, reconfigure services and improve productivity over several years without recourse to ad hoc quality-harming measures. However they will be subject to the same pressures to avoid transfer of activity to the community or another hospital.

Figure 6

NHS Trusts	**Net Losers**	▪ Strong pressure to reduce unit costs quickly ▪ Pressures for "quality harming" responses ▪ Significant risk of financial "failure"
	Net Winners	▪ Use unplanned surplus within year to improve quality of care ▪ Recorded Trust unit costs ~ system average
Foundation Trusts	**Net Losers**	▪ More time/resources to adjust to unplanned deficits ▪ Planned adjustments more likely to improve productivity and quality of care ▪ Temporary deficits financed within approved PBI
	Net Winners	▪ Use unplanned surplus either within-year to improve quality of care and/or to support PBL borrowing to expand/improve services over medium term

'Net winners' among **NHS Trusts**, with no incentive to show a surplus at year-end, are likely to spend the extra revenue within-year on improving the quality of care. Paradoxically, recorded unit costs will remain close to the (MFF adjusted) national average even if they had achieved better-than-average productivity gains before spending the extra surplus. **Net winners** among **FTs**, in contrast, may choose to accrue an end-year surplus, thereby crystallising a higher Prudential Borrowing Limit that can be used to finance investment in expanded and/or improved services on a larger scale. All net winners however will be subject to the same pressures to maximise admitted activity and to discourage transfer of activity to other hospitals or into the community.

ISPs, Patient choice and PbR

'First wave' ISPs entered into 5-year use or pay contracts with DH. The contracts guarantee a minimum annual payment to these ISPs regardless of the actual use made of them by patients. Moreover, the contract prices for services are believed to be at a (significant) premium to the equivalent PbR tariffs. DH was obliged to offer use or pay provisions because in their absence new entrant ISPs would have confronted unacceptable 'demand risk' i.e. the risk that, having sunk the costs of providing the capacity, PCTs chose not to refer sufficient patients to them.

Revenues paid to these 'first wave' ISPs are, in effect, an 'off-the-top' reduction in resources available to PCTs to purchase services from NHS providers. If purchasing of primary care services remains the same then ISP contract payments reduce pound-for-pound the funding available to purchase activity from NHS hospital providers.

'Second wave' and subsequent ISP contracts may or may not contain use or pay provisions and they may or may not be priced at a premium to PbR tariffs. If they are 'use or pay' then the impact on NHS providers will be the same as with the 'first wave' contracts. If they do not have use or pay provisions and do not receive premium prices then there will be a level playing field. However, even were that to be the case, any revenues ending up paid to ISPs will reduce effective demand for NHS provided services, given that there is a fixed cash-limited budget.

If (as currently planned) by 2008 15% of all elective work is to be undertaken by ISPs (and in addition an increasing share of diagnostic services) then this must reduce by a corresponding amount the demand for these services from NHS providers, relative to what it would otherwise have been. Given that the real purchasing power of PCT budgets is expected to grow only modestly (see section 4) it is highly likely that there will be little or no growth in demand for these types of services from NHS providers in aggregate. Indeed, NHS providers located close to new ISP capacity could well see an absolute decline in demand, with the result that available capacity remains under-used and under-remunerated.

The notion of 'competition' between NHS providers and ISPs is somewhat artificial. Competition on price is ruled out. Competition of any sort is ruled out if the ISPs have 'use or pay' contracts. Even where they do not, competition based on quality is a largely artificial concept. Most NHS providers are 'demand-takers', meaning that there is little they can do to influence the level and pattern of demand for their services.

Stranded Capacity Over the next few years there will be substantial growth in both NHS capacity (to meet the ever-tightening waiting time targets) and of ISP capacity for certain services. The rate of growth of supply of services will exceed the growth in real purchasing power of PCTs. As a result there is a substantial risk of the emergence of an increasing amount of 'stranded capacity'. This is NHS capacity that is available for use by patients at marginal cost well below the tariff but unable to find a purchaser for those services. In fact, there is already clear evidence of stranded capacity in parts of the NHS even before much of the 'first wave' ISP capacity is on-stream.[19] The problem will worsen over the next few years.

How do we explain this paradox – unused but available services in a world of limitless patient need? The immediate cause is a deficiency of purchasing power. The fundamental cause is that PbR tariffs are fixed equal to system average cost. In 'real' markets, as noted earlier, prices adjust to clear the market. If there is 'excess' supply, producers produce until the price falls to the point where marginal revenue equals short run marginal cost. The fall in price boosts demand. All economic supply is sold and consumer benefits are maximised.

In the reformed NHS, because prices are fixed equal to system average cost, there is no market clearing mechanism. As noted in section 4, if the cash budget and prices are fixed, so is the quantity of activity that PCTs can purchase. As the capacity of the system to deliver services increases, and payments to ISPs reduce PCTs' purchasing power available for purchasing from NHS providers, so NHS providers find they are unable to identify a PCT able and willing to pay for the extra services. The stranded capacity just sits there unused. In effect the system becomes demand constrained despite a continuing patient need. With fixed tariffs the demand constraint can only be reduced by reducing average tariffs (thereby enabling PCTs to purchase more services with a fixed budget). However further reducing the average tariff (below average cost) risks causing financial failure at a large number of NHS Trusts.

Impact of Reforms on Investment by NHS Providers

We now consider whether NHS providers are likely to respond to the reforms by investing in new facilities and services appropriate to deliver the NHS strategic objectives. A related question is whether providers will reduce the volume of services offered in some areas eg admitted hospital care and re-allocate the resources freed-up to develop new types of non-admitted care e.g. hospital support for expanded care in the community.

There are two linked drivers of the level and pattern of new NHS investment:

[19] There have been numerous stories of NHS services with spare capacity unable to find a purchaser, see various HSJ editions 2004/5. The problem of stranded capacity is not restricted to entry by PCTs. As NHS annual capacity has increased in recent years many NHS Trusts have unused capacity unable to find a purchaser.

- Whether there is a viable business case which demonstrates that the extra revenues generated by the investment exceed the extra costs (including the cost of capital) incurred; and

- Whether finance can be sourced to fund the proposed investment.

We address two types of situation: first, where the investment will expand the **volume** of activity; and second, where the investment will improve the **quality** of the services provided, but not increase the volume.

Investment in Capacity Expansion Currently NHS providers plan service delivery within the framework of PCTs local delivery plans. The volume of activity that they provide is determined by the waiting time targets and by what the PCTs can afford to purchase. In the past demand has been reasonably predictable. However in future all NHS hospital providers will face increased volume risk because of patient choice and ISPs. Currently there is no mechanism available to them to manage these volume risks because PCTs are unable and unwilling to enter into medium term volume contracts.[20]

In addition, all NHS providers face high 'price risk' because the mechanism for setting and adjusting tariffs is opaque and unpredictable. Forming judgements about the extra revenues that can be earned by investing in additional service capacity over the medium term is very difficult. Moreover, as with volume risk, there is no mechanism to enable providers to manage the price risks.

Even if forward PbR tariffs were predictable, they will not generate appropriate investment signals because they are set equal to system average costs rather than long run marginal costs. As noted in section 4, the LRMC of new services and services expansion may be higher or lower than average cost but will rarely be the same.

The average costs of many NHS providers are similar. One effect of PbR will be to make them even more similar in future. When the industry cost curve is very 'flat' it is particularly difficult for any one provider to predict how other suppliers will respond to the same market signals. If all Trusts respond by investing, there will be over-supply and losses will be incurred by many. Risk-averse Trusts subject to a strict financial regime may seek to avoid these risks by not investing even where the patient benefits would have been high. If all Trusts respond in this way there will be underinvestment resulting in reduced availability of services for patients.

In summary, NHS providers confront increased volume and price risks, have no ability to manage these risks, are paid tariffs that do not reflect long run marginal costs and compete with other providers with similar costs. In such a world, **there can be no expectation that investment decisions made by providers in response to market forces will deliver the level and pattern of services required if the NHS Improvement Plan objectives are to be realised**.

The same perverse incentives that make it unattractive for NHS providers to allocate current resources from admitted care to support for expanded care in the community also apply to capital investment decisions. Even if the business case for a hospital owned and operated community health service is viable on a stand-alone basis, if it would reduce the flow of patients to the hospital then the indirect costs (in terms of foregone revenue) may well render the investment unaffordable for the hospital.

The DH is relying on the demand signals generated by patient choice and GP practice commissioning to induce providers to deliver the level and pattern of services required by patients. With current policies this is unlikely to work. For reasons set out earlier, NHS providers are subject to strong pressures to retain the existing pattern of care, with heavy reliance on admitted hospital

[20] Unlike first wave ISPs which did hedge volume risk by signing 5-year use or pay agreements with DH.

care, and not to support expansion of care in the community. Without a transfer of resources (both human and financial) from the acute sector to community care and without the active involvement of hospital-based clinicians supporting care in the community, patients will continue (as now) not to have much real choice about the type and location of services. Given the current pressures on the hospital sector and the very weak capacity of PCTs to control the allocation of resources, this is not likely to happen.

Even if NHS Trusts wish to invest to expand capacity, their ability to do so in practice is constrained by their limited access to capital finance. Large scale capital funding for hospital Trusts is largely restricted to 'new build' PFI schemes. These rarely add much new capacity (often replacing aged buildings with new ones of similar or reduced service capacity although of higher quality). NHS hospital Trusts seeking to invest moderate amounts to expand selected services often cannot secure funding; and therefore cannot proceed. In effect, the level and pattern of investment in services expansion by NHS Trusts is being driven not by consideration of where maximum patient benefits accrue but rather by access to capital funding. The capital allocation approach adopted by DH in effect biases investment in favour of major new build hospital schemes rather than incremental expansion of existing services whose patient benefits per pound invested will often be higher.

FTs, in contrast, have preferred access – within the limits set by their PBLs – to the FT Financing Facility which provides access to finance that can be, and is being, used to expand services capacity incrementally.

Improving the Quality of Care It is an important shortcoming of PbR that payments are the same for a completed spell regardless of the quality of the services provided or of the patient outcome. Average cost tariffs will generally under-remunerate higher-than-average cost providers of higher-than-average quality care; and over-remunerate (reward) lower-than-average cost providers that achieve lower costs by providing lower-than-average quality of care. Consequently PbR not only does not contain incentives to improve the quality of care, it contains incentives to maximise the volume of activity at the expense of quality. Hard-pressed NHS Trusts may be tempted to adopt measures that reflect these pressures – to the detriment of patients.

These perverse incentives are even stronger when NHS Trusts contemplate investments in quality improvements. Any spending to improve the quality, but not the quantity, of services will attract very little (if any) additional revenue. If the enhancement is a new (improved) procedure, it will only attract additional revenue if and when there is a new (appropriately priced) HRG. If the enhancement is use of a new (improved) medical technology, with no change in the procedure, there is no additional revenue to compensate for the extra costs. Consequently there are no drivers in the current design of PbR to invest in either quality improvement or medical technology innovation.

Summing Up

There is no doubt that the reform instruments will inject strong new pressures into the NHS. Those pressures will force purchasers and providers to behave differently. The analysis predicts that, so long as there is consistent downward pressure on real average tariffs, PbR will pressure NHS providers to improve efficiency and productivity, thereby increasing the effective capacity of existing services and freeing up resources that can be spent on expanding and improving care in the community. It will facilitate patient choice which itself should pressure providers to be more responsive to patients' preferences. Independent sector provision will increase capacity and improve choice.

However, these gains are achieved at a significant cost. The greatest concern is that the reforms will strongly discourage NHS providers from supporting expanded care in the community and the transfer of patients to other providers in a network of care, despite these being in the best interests of patients. The corollary is that hospital providers subject to strong financial balance constraints

will seek to hang onto as much admitted and day case hospital care as possible. Undue reliance on competition driven by demand signals generated by patient choice and GP practice commissioning is unlikely to deliver the level, type and location of services that are required to deliver the services that patients want. With fixed average cost tariffs, substantial amounts of 'stranded' capacity will emerge, with patients deprived of the use of those services, available at low marginal cost, because of a deficiency of PCT purchasing power. The combination of PbR, patient choice and ISPs will generate major financial deficits at a significant number of 'net loser' Trusts with unpredictable but probably adverse consequences for the quality of care at those Trusts and in some cases 'financial failure'. There are no incentives to improve the quality of care or to invest in clinical innovation. The answer to the first question posed – Will the NHS reforms deliver the desired health outcomes? - is summarised in Figure 7.

Figure 7

The Good News

- Pressure on all NHS hospitals to improve productivity thereby freeing-up resources to expand/improve services across the NHS
- Facilitate patient choice pressuring hospitals to be more responsive to patients preferences
- ISPs will increase capacity for certain services and extend patient choice for those services

The Concerns

- Reforms will discourage NHS providers from supporting care in the community and the transfer of patients to other providers in a network of care
- No incentives to improve the quality of care and risk of quality-harming responses from "net losers" from PbR/patient choice / ISPs
- Significant risk of available but unused 'stranded' NHS capacity because of deficiency of PCT purchasing power
- Financial destabilisation with major 'net winners' and 'net losers' and risk of failure of latter with adverse consequences for patients
- 'Market forces' and the capital allocation mechanisms will not generate the level and pattern of investment required to deliver the care in the community agenda

Are these perverse and unintended effects inevitable, the price that has to be paid for the undoubted benefits of reform? The simple answer is no. In section 6 we turn to the second key question – are there enhancements to the reform instruments that, if adopted, would retain the existing benefits but improve the likelihood of more fully achieving the desired health outcomes?

6. Are there Enhancements to the Reform Instruments that will Improve Health Outcomes?

Any proposed improvements must increase the likelihood that the NHS strategic objectives will be fully achieved. They must facilitate patient choice and entry by ISPs and retain the existing incentives to improve productivity. They must strengthen incentives on NHS providers to improve the quality of care, to work in partnership across care networks and to support delivery of the care in the community agenda.

Throughout most of this section, the current NHS architecture (ie PCTs, NHS Trusts, FTs, SHAs etc) is assumed to remain unchanged and national tariffs to be fixed. At the end of the section, wider reforms are briefly touched on.

Improving PbR Incentives

There are two proposed enhancements aimed at improving PbR incentives:

- A two-part PbR tariff, and
- A transparent, predictable tariff setting and adjustment mechanism.

Two Part PbR Tariffs PbR is central to the reform strategy. If well designed and implemented, it can be a powerful driver pressuring NHS providers to improve productivity, thereby releasing resources that can be used to expand and improve NHS services.

However, as noted earlier, the one part PbR tariffs have embedded within them strong perverse incentives likely to result in key NHS objectives not being achieved. The proposed two part tariff sharply reduces or eliminates many of these perverse incentives.

The two-part tariff consists of:

- a **fixed payment** per **contracted unit** of activity payable over the year if the capacity is available (whether or not used); and
- an **activity payment** paid for every unit of activity if and when provided, where
- the total of the fixed payment and the activity payment for the contracted volume equals the amount payable under the current one-part tariff (i.e. system average cost).

Figure 8 shows how the two-part tariff works. Key features of the two part tariff are:

- if actual activity provided equals the contracted volume then the total payment per unit is the same as if the one part tariff applied i.e. system average cost;
- the marginal revenue to the provider is the activity payment - hence providers that provide less than planned activity lose the activity payment but not the fixed payment and those that provide more activity than the contracted volume gain just the activity payment;
- marginal incentives are improved because the ‘gap’ between marginal revenue and marginal cost is smaller. The incentive to increase activity above contracted levels is weaker and it is less 'costly' (in terms of revenue foregone) to shift marginal resources to support care in the community;

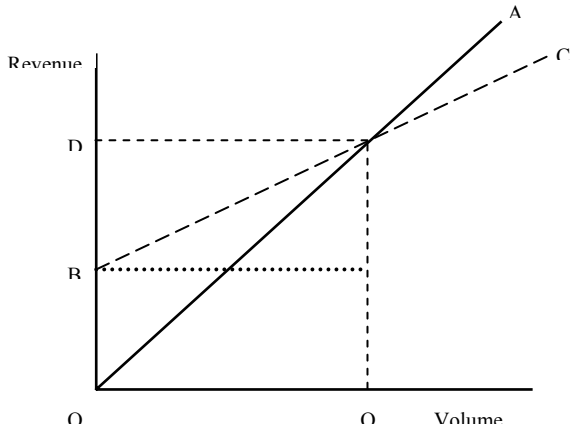
- PCTs can purchase 'above contracted' volumes of activity at a lower marginal cost (the activity payment), thereby increasing the volume of activity they can afford to purchase out of a fixed budget. This will increase real purchasing power and reduce the risk of stranded capacity;

Figure 8 - How a two-part fixed tariff works.

Fixed charge = £/unit x contracted units pa = £pa

Volume related charge = £/unit of activity provided

For contracted units pa
Total tariff = Fixed charge and volume related charge = system average cost



<u>One-part tariff</u>

Line OA = current average cost/unit PbR tariff "pay off" curve

For actual quantity of activity provided (Q) provider receives revenue = OD

Marginal revenue (gain/loss of revenue for change of 1 unit of activity given by slope of line OA)

<u>Two-part tariff</u>

PCT contracts for volume (Q) paying fixed charge per contracted unit (in this case = 50% of average cost) for Q units

PCT pays in addition variable charge per unit for activity actually provided = line BC

If actual quantity provided = Q, tariff per unit = average cost and revenue = OD

Marginal revenue is given by slope of line BC

- the financial 'shocks' are muted. Unplanned surpluses and deficits are smaller. Consequently there is reduced risk of quality-harming actions by net 'losers' and of hospital failures;

- however, there is no reduction in the pressure on Trusts to improve productivity since, for contracted volumes they receive the same revenue per unit as with the one part PbR tariff.

The relative magnitude of the fixed and activity-related payments is important. In principle, the activity payment could vary from about 10% of the total payment (being roughly the proportion of costs that are variable in the short term) to 100% (which would be identical to the one part tariff). Table 1 summarises the different impact on providers and PCTs of the two part tariff for cases where the activity payment as a proportion of the total payment varies is 10%, 50% and 100%.

Table 1

	Activity payment as % of total tariff		
	10% (= variable costs)	50%	100% (= one part tariff)
Financial instability (induced by patient choice/ISPs)	Least ……..Most		
Perverse marginal incentives ((eg too much admitted care etc)	Least ……...Most		
Extra PCT purchasing power	Most ……...Least		
Incentives to increase supply	Least ……...Most		
Stranded capacity risk	……Least ……..		

In the case where the activity payment corresponds only to the variable costs, financial instability and perverse marginal incentives are minimised, extra PCT purchasing power is greatest (because the marginal price for extra activity is smallest) but the incentive to provide extra activity is also smallest. The one part tariff is the opposite with strong perverse incentives and financial instability associated with patient mobility, the least extra PCT purchasing power and the strongest incentives to provide additional activity if a buyer can be found. When the activity payment is 50% of the total payment there is an unambiguous improvement in expected outcome compared with the one part tariff with reduced financial instability and perverse incentives, reduced stranded capacity risk and greater PCT purchasing power within a fixed budget.

In principle the two part tariff could be applied to particular services or to all services. There is a compelling case for its use in non-contestable services eg non-elective hospital care. Why should providers be paid the average cost when the number of patients increases unexpectedly when the providers' costs go up only a fraction of this amount. The one part tariff sucks resources out of primary care and 'over compensates' the provider for the marginal activity. In doing so it reduces the total amount of activity that the PCTs can afford to purchase. The two part tariff would pay the provider only the activity payment, thereby enabling the PCTs to purchase more activity in total while adequately compensating the provider for the extra activity.

However the case is equally strong for the contestable services eg elective surgery. The two part tariff does not inhibit the operation of patient choice or entry by ISPs. Nor does it subject PCTs to material 'stranded payment risk' ie the risk that they end up paying providers for services that they do not receive[10]. There is much reduced financial instability and risk of quality harming responses by net losers. However there is no reduction in the pressures to improve efficiency and productivity.

[10] Worked examples in Appendix 3 illustrate this point.

Transparent and Predictable Tariff Adjustment Mechanism The lack of transparency and predictability of the tariff setting and tariff adjustment mechanisms are significant shortcomings of PbR. If NHS providers are to plan services delivery to achieve target productivity improvements they need to know what those targets are. This requires knowledge of how real tariffs will evolve over time, which in turn requires knowledge of the mechanism that will be used to adjust tariffs year on year and the value of the efficiency factor that will be built in.

There is a strong case for introducing much greater transparency into the process and methodology used to derive prospective tariffs from lagged reference cost data. One of the benefits will be to facilitate a process of self regulation where NHS providers identify and report gaming by other NHS Trusts. There is also a strong case for announcing the medium term efficiency target that will be built into future tariffs for a minimum period of three years. This will facilitate planning and implementation of service redesign, service rationalisation and investment in services expansion.

Facilitating second wave ISPs without use or pay provisions DH has indicated that it will seek to avoid giving 'second wave' ISPs the use or pay provisions granted to first wave ISPs. With one part tariffs this will be very difficult to achieve because the volume risks are very high for new entrants. If they are required to absorb these risks it will increase the contract prices for the services. With a two-part tariff – if the fixed payment per contracted unit is about 50% of the total payment – the volume risk is reduced for ISPs as well as NHS providers. Therefore it should be easier to avoid conceding use or pay provisions in 'second wave' ISP contracts.

Strengthening PCT Commissioning

Medium Term PCT-Provider Contracts In section 5 it was argued that reliance on demand signals generated by patient choice and GP practice commissioning cannot be relied upon alone to bring about the desired allocation of resources to deliver the care in the community agenda. Currently PCTs have very little, if any, ability to manage hospital demand and therefore little control over how much hospitals get paid. Nor can they direct hospitals to reallocate resources to support the care in the community agenda.

An important consequence of adopting two-part PbR tariff is that it **shifts the balance of contracting power back to PCTs**. PCTs can now execute medium term contracts with NHS hospital providers in which contracted volumes are determined by the PCTs. By holding steady, reducing or increasing contracted volumes for particular services PCTs will, in effect, determine (or at least heavily influence) decisions by hospital providers to expand or contract services.

Figure 9 illustrates how this would work. In year 1 the PCT contracts for 100 units of activity per annum with Trust A. The contracted volume is reduced in years 2 and 3. If the Trust provides 100 units of activity in each year, it receives the fixed payment and the activity payment for a reducing number of contracted units and only the activity payment for the excess of actual activity over the contracted volume. The revenue for contracted volumes is average cost but the marginal revenue for activity in excess of contracted volumes is the activity payment. The incentive to produce more than the contracted volume is accordingly much lower than with the one part tariff.

Varying the contracted volume of hospital activity at the **service** level (rather than the aggregate activity level) becomes a powerful mechanism for PCTs to 'force' a shift in the composition of services provided to better meet changing patient needs. Moreover, unlike with one part tariffs PCTs are not materially at risk of 'stranded payment risk' if they enter into medium term contracts with two part tariffs.

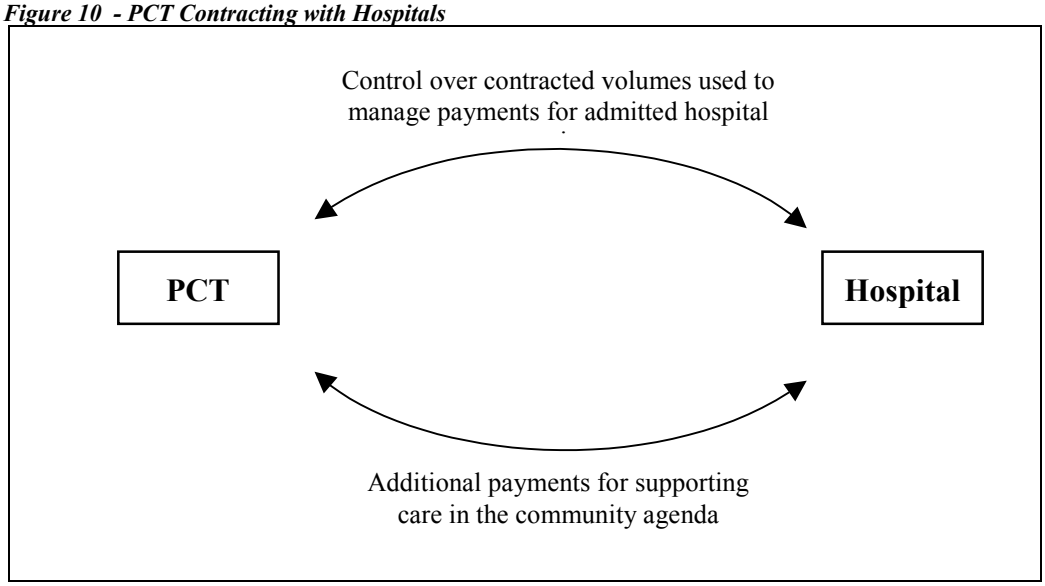
Figure 9 - Medium Term contracting with Two-Part PbR Tariffs.

Year:		**1**			**2**			**3**	
Contracted volume		100			95			90	
Actual volume provided	90	100	110	90	100	110	90	100	110
Fixed fee 5/contracted vol	500	500	500	475	475	475	450	450	450
Variable fee 5/vol provided	450	500	550	450	500	550	450	500	550
Revenue (total)	950	1000	1050	925	975	1025	900	950	1000
Average price/unit provided		10			9.75			9.5	

Assumes fixed fee = 50% of average cost

New Types of Contracts to Support Care in the Community With two part tariffs and medium term volume contracts PCTs can develop contracting relationships that drive providers to contribute more actively to support the care in the community agenda. This can be achieved by actively managing contracted volumes of admitted hospital care and simultaneously contracting with providers to pay them extra revenue in return for their providing additional support for care in the community (Figure 10). The result would be active market management by PCTs within a framework where patient choice and GP practice commissioning can operate unhindered. Contracting decisions by PCTs over time would shape the future level, type and location of services to reflect both NHS strategic objectives and actual patient choices eg stimulating new forms of care in the community supported by hospital specialists as it becomes clear that is what patients want.

Figure 10 - PCT Contracting with Hospitals



Improving Transparency of 'Market' Decision making by purchasers and providers is unduly complicated by the current lack of transparency about supply and demand of services. Improvement to the transparency of the 'market' could be achieved by requiring SHAs to report, on a consistent basis, information from purchasers and providers about actual and planned supply and demand for services. A medium term forecast of intentions by purchasers and providers and a record of what

actually happened in the past would provide important additional transparency to aid PCTs and providers when negotiating terms of medium term volume contracts and making investment decisions.[11]

Strengthening Incentives to Improve the Quality of Care

It is an important shortcoming of PbR that there is no link between tariff payments and the quality of services provided. There are no incentives to improve the quality of care. Indeed there are perverse incentives that encourage behaviour that maximises the volume of activity even at the expense of quality. As a result the whole burden of quality management falls on the health regulator, principally the Healthcare Commission (and Monitor for FTs).

Two alternative options are suggested for strengthening incentives on providers to achieve and sustain high quality of care.

Link PCT payments to achieved quality of care PCT-provider contracts, in addition to specifying contract volumes, would also set out quality performance targets. As an overlay on top of the (two-part) tariff there would be provisions for annual bonuses and discounts determined by reference to the actual quality of performance achieved. The maximum bonus/penalty would be set nationally but the quality performance targets would be set, and actual performance assessed, locally. Objective quality measures would need to be developed over time. The Healthcare Commission might have a role either assessing quality performance or providing an independent view in the event that the contracting parties were unable to agree.

Link PbR tariffs to Healthcare Commission 'ratings' An alternative approach would be for the bonus/penalty to apply in a mechanistic way by reference to the 'rating' achieved by each NHS provider in the Healthcare Commission review. This approach is less capable of calibration to address local quality concerns but involves much less local negotiation of quality standards and relies on a national process that is happening anyway. This approach has merit if there are doubts about the capacity of PCTs to effectively negotiate local quality performance indicators. If this approach were adopted it is important that the bonus/penalty provisions are set by reference to the **prospective** (i.e. future) rating rather than the most recent actual rating, if the desired quality improvement incentives are to be effective.

Either of these provisions would embed quality management as a core concern of providers and generate much stronger incentives to achieve a consistently high quality of care. This would represent an important improvement over the current situation.

The challenge of developing objective quality measures and workable mechanisms for determining penalties and bonuses clearly should not be underestimated. However, the challenges are no greater than many others that have already been grasped in the NHS reform programme.

It is true that with such arrangements in place poorer performers in financial difficulties would receive less revenue and therefore become even more mired in financial deficits. However, the appropriate response is not to reject this sort of provision but rather to calibrate quality improvement targets to ensure consistent pressures apply locally to improve quality of care over time. In addition it is important to allow greater flexibility in the financial regime to allow NHS Trusts time to achieve them. Quality provisions along these lines are more practicable if there is a two part PbR tariff because the financial destabilisation of 'net losers' is more muted.

[11] A similar approach in the gas and electricity sectors - where the National Grid is required to compile and publish a 7 year forecast of supply and demand – serves an important similar purpose adding transparency to aid decentralised supply and demand decisions.

Improving the Financial Regime for NHS Trusts

It is increasingly acknowledged that the current 'must do' financial regime applying to NHS Trusts is too inflexible for a future in which the pressures of PbR, patient choice and ISPs apply. Although DH has indicated that it recognises the need for greater flexibility no action has yet been taken. It is urgent to increase room for manoeuvre before a large number of Trusts either fail or take quality harming actions in a desperate attempt to ward off failure.

Suggested improvements to the financial regime are:

- Allow NHS Trusts the same right as FTs to retain and reinvest any surpluses generated. Only if this step is taken will there be any incentive to out-perform the 'must do' financial targets.

- Allow deficit Trusts to eliminate unplanned financial deficits over a 3-year period. This will provide more time to plan and implement unit cost reducing, quality enhancing programmes. Deficit Trusts would need to have an adjustment programme approved and the permitted aggregate cumulative deficit would be capped.

- Create a working capital mechanism (possibly via the NHS Bank) to channel surpluses generated by 'net winners' to fund deficits of 'net losers'. The working capital mechanism should charge interest for use of funds and pay interest to providers of funds. There should be aggregate limits on Trust access to working capital

- Allow a portion of the dividend on PDC to be deferred and carried forward for a maximum period eg 3 years. In effect, this would convert a portion of the fixed interest PDC into a cumulative fixed interest preference share. This enhancement should apply to NHS Trusts and FTs. In the much higher risk environment facing all NHS providers, retention of a 100% debt financed balance sheet is clearly inappropriate. The suggested additional flexibility to manage financial risk would be achieved without unduly relaxing financial discipline since the deferral would be limited in duration and carry an interest cost.

- Create effective mechanisms to manage financial failure. Although DH is reportedly working to develop failure mechanisms, at present none exist. Robust failure mechanisms are needed that ensure continuity of essential services and provide for implementation of a restructuring plan to remove the underlying cause of failure, for actors which can implement the plan and funding of the restructuring plan.

Improving the Capital Allocation Mechanisms The current mechanisms for allocating capital to NHS Trusts and PCTs do not even consider whether the proposed investments will deliver large patient benefits relative to the costs.[12] Nor do they consider whether the proposed pattern of capital investments is likely to maximise progress towards achieving the NHS strategic objectives. De facto, the only source of capital available to the hospital sector is PFI for large (often new build) schemes. There is very limited capital available for mid-size investments to expand existing services even where there is a demonstrably high patient benefit per pound invested.

There is a strong case for creating a new capital finance facility available to NHS Trusts and PCTs that would provide funding for mid-size investments to expand existing services. Assessment of bids should be made against three criteria:

- Whether they contribute to achieving the strategic objectives in the region
- Whether the patient benefits are high relative to the costs

[12] Major PFI and LIFT schemes consider whether the costs are lower than a hypothetical Public Sector Comparator but not whether the future patient benefits exceed the costs or whether the scheme is affordable under PbR.

- Whether the investment is expected to be viable in a world of payment by results.

Patient benefits should be assessed using consistent criteria which would include whether they contribute to the care in the community agenda and whether they show high patient benefits relative to the costs valuing future benefits at PbR tariff prices. Currently FTs have access to something similar – the FT Financing Facility. However the assessment criteria for FTFF bids do not include the first two of these three criteria.

Exploit more fully the potential of Foundation Trusts In 'Delivering the NHS Plan', FTs were seen as having exciting potential to lead transformation in the NHS. In the event their impact has, at least so far, been somewhat muted. There are a number of reasons for this: first, very few have yet been created; second, they are subject to the same risks and perverse incentives arising from PbR, patient choice etc as other NHS Trusts; and third, there have been, and remain, significant uncertainties about the role they will be allowed to play. Nevertheless, FTs already have been given much of the increased financial flexibility required by all NHS Trusts if they are to be able to respond effectively to the new pressures. They can be expected to play a more dynamic and constructive role if the perverse incentives embedded in the current policy instruments are removed.

The best way to increase the impact of FTs on NHS reform is to increase the rate at which FTs are created; to remove the perverse incentives acting on all NHS providers; and to clarify remaining uncertainties about the role that FTs will be allowed to play. In particular whether FTs are to be allowed to provide integrated patient care (with more active involvement in the provision of care in the community); what role they can play in addressing NHS Trust failure (are mergers to be allowed and facilitated); and whether they are to be permitted to bid to provide second wave diagnostic and treatment facilities.

Going beyond current policies

All the proposed enhancements so far remain within the current stated policies. A number of enhancements that go beyond current policies are worthy of consideration.

Structural Consolidation The fact is that PCTs are too small and too numerous. Their small size not only constrains their capacity to commission effectively, it also makes it very difficult for them to behave strategically, taking account of the wider regional picture of supply and demand. Moreover the overhead costs, while modest for each PCT, are very large in aggregate. Consolidation is currently taking place informally in some areas with joint chief executives and shared services arrangements. However there is a strong case for more pro-active consolidation of PCTs. The same is true for SHAs, at least in certain parts of the country.

Greater price flexibility The proposed move to a two part tariff will address the more immediate perverse incentives embedded in the one part tariff. Nevertheless, fixed average cost tariffs (even two-part) will not provide appropriate price signals for guiding medium term investment decisions. There is a strong case for allowing greater flexibility and 'price discovery' of prices for new capacity. There are several mechanisms that could be adopted to achieve this. For example, future tenders for new diagnostic and treatment facilities could set the PbR tariff as a cap and allow bidders to submit proposals setting out the price at which they are prepared to deliver a specified quantity and quality of services. This would ensure that the least cost supplier able to meet the quality specification would be the chosen provider; thereby maximising patient benefits per pound of taxpayers money. FTs and ISPs should be allowed to bid to supply the services. A similar approach could be adopted in the NHS with calls for NHS providers and ISPs to deliver specified new facilities and services. Bid documentation would specify the quantity and quality of services required to be provided and the price cap would be the PbR tariff but bidders would offer the price at which they are able to provide the services. This would ensure that patient benefits per pound of taxpayers money were maximised.

7. Conclusions

The government is implementing a radical reform strategy for the NHS. The key instruments of reform are PCT commissioning, payment by results (PbR), patient choice, independent sector provision, Foundation Trusts and the 'must do' access and financial targets. The introduction of these reforms is unleashing strong forces that will drive change in behaviour and in the allocation of resources. The paper asks two questions. First will the reform instruments as currently designed deliver the desired outcomes as set out in eg the NHS Improvement Plan? Second are there enhancements to the instruments that will increase the likelihood that the desired outcomes will be more fully achieved?

The strategic objectives of the NHS are:

- Shorter waiting times for consistently high quality hospital care when hospital care is appropriate

- More patient care to be provided closer to home or in the home especially for patients with chronic conditions, the elderly and other services where care in the community is appropriate

- Greater focus on prevention of illness and early diagnosis and treatment

Delivering these strategic objectives within the available resource envelope will require:

- sustained significant improvements in the productivity of hospital providers (thereby freeing up resources to expand and improve services across the NHS) and in the quality of hospital care;

- expansion of services that provide new and improved types of 'care in the community' funded in part by freeing up resources currently used to provide admitted hospital care; and

- deployment of hospital-based expertise to provide and support enhanced care in the community.

Will the Reforms Deliver the Desired Health Outcomes?

The first key issue considered is whether the changes in behaviour and resource allocation induced by the reforms will deliver the level and pattern of services required if the NHS Improvement Plan objectives are to be achieved.

The conclusions reached are:

The Good News

- So long as there is consistent downward pressure on real average tariffs over the medium term, PbR will pressure NHS providers to improve efficiency and productivity, thereby increasing the effective capacity of existing services and freeing up resources that can be spent on expanding and improving services across the NHS including enhanced care in the community

- Patient choice and GP practice commissioning will pressure NHS providers to be more responsive to patient preferences and should transmit demand signals that result over time in improvement of existing services and in expansion of the type and location of services that patients prefer

- Independent sector provision will increase capacity of certain types of services, thereby increasing choice and improving access for those services

The Concerns

- The reforms will strongly encourage NHS providers to maximise admitted and day case hospital care (sucking resources into the hospital sector) and strongly discourage collaboration across care networks and support by hospitals for improved care in the community.

- The outcome of competition between providers in response to 'market forces' cannot be expected to result in the level and pattern of services required if the NHS Improvement Plan objectives are to be realised. The demand signals generated by patient choice and GP practice commissioning will not be strong enough to induce investment by providers in new and improved types of care in the community. Nor will hospitals contribute actively to support for care in the community. Strong market management led by PCTs will be required if the desired transfer of resources from the hospital sector to care in the community is to be achieved. PCTs cannot provide that leadership as the reform instruments are currently designed.

- Currently there are no incentives to improve the quality of care. In fact there are perverse incentives which pressure providers to maximise volume of activity even at the expense of quality. There are also perverse incentives which discourage investment in clinical innovation. Yet sustained improvements in patient care will only be realised if there are strong drivers to develop and adopt new medical technologies.

- There is a significant risk that the reforms will generate a lot of 'stranded' NHS capacity. Stranded capacity is capacity that is available for patients at low marginal cost but remains unused because of a deficiency of PCT purchasing power. There is already a lot of evidence of stranded capacity in parts of the NHS today even before many of the first wave independent treatment centres are on-stream. There will be a lot more as ISPs come fully on-stream and as waiting times shorten. Stranded capacity exists only because of the particular design chosen for PbR tariffs ie fixed tariffs set equal to system average costs.

- The reforms will generate 'net winners' with large unplanned financial surpluses and 'net losers' with large unplanned financial deficits. There are considerable risks that 'net losers' will seek to restore financial balance at the expense of the quality of care of their patients; that they will defer desirable services reconfiguration and seek to minimise the transfer of activity from their hospital into the community or to another hospital in a network of care to avoid loss of revenue; and that a significant number of net losers will 'fail' with uncertain but probably adverse consequences for their patients.

Are there Enhancements to the Reform Instruments that will Improve Health Outcomes?

The paper considers whether there are enhancements to the reform instruments that would retain the predicted benefits while addressing the concerns noted above. The conclusion of the paper is that there are a number of policy enhancements that, if adopted, will markedly improve the likelihood of fully achieving the government's objectives without impeding patient choice and GP practice commissioning or entry by independent sector providers and without diminishing in any way the pressures on providers to improve productivity. The adverse consequences of the reforms identified above are not an inevitable price that has to be paid to achieve the undoubted benefits.

Two part PbR Tariffs The most important suggestion is the adoption of a two part PbR tariff. This consists of a fixed payment per unit of contracted annual capacity and an activity payment per unit of activity actually provided, with the sum of the two payments for contracted volumes equal to

average system cost ie the same payment as with the one part tariff. The activity payment would make up about 50% of the total payment. The analysis shows that two part tariffs retain strong incentives to improve productivity over time while reducing the perverse marginal incentives generated by fixed one part average cost tariffs. Two part tariffs reduce the disincentives to collaborate across care networks and to support care in the community; they reduce financial instability of net losers from patient choice and ISP entry and they reduce the risk of 'stranded' capacity. They strengthen the ability of PCTs to manage hospital demand and to achieve greater control over the allocation of resources between the hospital sector and primary care. They are wholly compatible with patient choice, GP practice commissioning and entry by ISPs. They may also facilitate contracting of second wave ISPs on terms that do not include such onerous 'use or pay' provisions as were included in first wave ISP contracts.

PCT Medium Term Contracts If two part PbR tariffs are adopted then an important consequence is that there is a shift in the balance of contracting power in favour of PCTs. They will face much reduced financial risk if they enter into multi-year (say, three year) volume contracts than they would with one part tariffs. They will have much greater influence over the pattern of services provision and over the pattern of investment in new services by virtue of their ability to vary contracted volumes of hospital services at the individual service level. They can use this influence to induce hospital Trusts to contract with them to support expanded and improved care in the community in a way that is not currently possible. Hospital providers would also benefit from reduced volume and revenue risks and greater ability to plan and implement services reconfiguration. The paper shows that such contracts are wholly compatible with the operation of patient choice and GP practice commissioning and entry by ISPs.

A number of other enhancements to the reform instruments are also proposed.

Transparent, Predictable and Demanding Tariff Trajectory Currently PbR tariffs are set annually and the future trajectory of average tariffs is not predictable. Stronger signals to improve productivity and rationalise services would be transmitted if the basis for tariff setting were more transparent and predictable and if the medium term efficiency target to be built into tariffs were known. The efficiency target should be announced and maintained for at least three years ahead. This will provide important forward price signals to providers to guide current decisions about streamlining service delivery, rationalising services and investing in new services.

Linking Payments Made to Providers with Achieved Quality of Care Two options, either of which would strengthen incentives to improve the quality of care, are proposed. The first option would provide in PCT/hospital provider contracts for payment of quality bonuses and penalties set by reference to actual quality performance achieved compared to locally agreed Quality Performance Standards set out in the contract. The second option would link similar bonuses and penalties to quality performance as assessed by the Healthcare Commission against national standards. Either option would put quality improvement at the heart of the reform programme. They would reduce existing perverse incentives to maximise the quantity of services provided at the expense of quality.

Improving the Financial Regime There is no doubt that the financial regime applicable to NHS Trusts is far too inflexible for a world in which payment by results, patient choice and entry by ISPs apply. A number of proposals for improving the financial regime for NHS Trusts are set out. They are:

- Allow all NHS providers to retain and reinvest income/expenditure surpluses
- Allow deficit NHS Trusts greater flexibility to adjust to unplanned financial deficits over more than one year
- Create an improved working capital mechanism (possibly via the NHS Bank) to channel surpluses generated by 'net winners' to fund the deficits of 'net losers'

- Create more effective mechanisms to manage hospital failure to ensure that the underlying cause of failure is addressed and that the quality of care of patients at 'failing Trusts' is not compromised.
- For all NHS providers including FTs allow a portion of the dividend on Public Dividend Capital to be deferred and paid with interest at a future date.

The current mechanisms for allocating capital funding across NHS Trusts and PCTs are unlikely to either deliver the level and pattern of investment required to deliver the NHS strategic objectives or to maximise patient benefits. Currently they are biased in favour of major new hospital build schemes at the expense of mid-size capacity expansions that often have the highest patient benefits per pound of invested capital. Proposals are set out for **improving the capital allocation mechanisms**.

Proposals are also made to more **fully exploit the potential of Foundation Trusts**. In particular it is important to clarify remaining uncertainties around the role of FTs including their role in providing managed care across the acute/primary divide; their role if any in addressing NHS Trust 'failure'; and whether they are to be permitted to build and operate 'independent' diagnostic and treatment centres.

In summary, the forces unleashed by the NHS reforms contain within them strong tensions. On the one hand they will result in shorter waiting times, force providers to be more responsive to patients' preferences and strengthen pressures on providers to improve productivity. However the same reforms will discourage collaboration across care networks and support by hospitals for care in the community, will not result in the level and pattern of services required to deliver the care in the community agenda, will not encourage quality improvement and clinical innovation and risk large scale financial destabilisation of parts of the NHS. These adverse side-effects are not inevitable. There are enhancements to the policy instruments that, if adopted, will markedly improve the likelihood of fully achieving the government's objectives without diminishing in any way the benefits of the current reforms.

Appendix 1 – The Impact of Payment by Results

The introduction of PbR will have a range of effects that can be grouped under three headings:

- Tariff design effects
- Introductory effects
- Sustained effects

Tariff Design Effects

The design of the tariffs will have an important influence on the incentives acting on providers. If the HRGs are not appropriately chosen then there will be perverse incentives acting to encourage Trusts to behave in ways that are not in the best interests of patients and/or consistent with financial viability of efficient providers. Examples include: (i) if there is not a short stay tariff then there will be perverse incentives to over-admit patients in A&E; and (ii) if there is not appropriate recognition of the higher cost of efficiently treating very complex, very long length of stay patients then there will be perverse incentives acting on DGHs to 'cream skim' the easier cases and a real risk that specialist hospitals will not be adequately funded for providing appropriate specialist care.

The mechanisms for adjusting the tariffs year to year for cost pressures, the efficiency factor and the Market Forces factor will also have an important impact on providers. Examples include: (i) if the MFF payments were included in the tariffs (as originally proposed) they would introduce perverse incentives on PCTs to steer patients to the providers with the lowest MFFs regardless of whether that was in the patients' best interests; and (ii) if the adjustment mechanisms remain opaque and the value of the efficiency factor uncertain then there will be no clear basis by which providers can plan services around a clear target for future productivity improvement.

Introductory Effects

There will be two sorts of introductory effects:

(i) The first order effect will arise from the shift by providers and PCTs from locally negotiated prices to the national tariffs. Some NHS trusts will see an increase in the average prices paid to them for the same volume and composition of activity at the expense of their PCT purchasers which will see a greater claim over their budgets for the same volume of activity provided. Other NHS Trusts will see a decrease in the average prices paid to them with a corresponding gain by their PCTs which will pay less for the same volume and composition of activity. At the system level the net effect is neutral (gains and losses in aggregate are equal) but there will be a significant shift of funding between purchasers and providers. The net effect of these first order introductory effects will depend on the details of the adjustment mechanism used to translate lagged reference costs into prospective tariffs.[13] It will also depend on the mechanisms put in place by DH to dampen the introductory effects of PbR and on whether PCT allocations are adjusted to take account of their higher costs for an unchanged volume of services provided.

(ii) Further introductory effects will be induced by the introduction of PbR. Some of them are listed below:

- **Coding completeness and accuracy** will improve immediately. All Trusts will have very strong financial incentives to ensure that coding is both complete and accurate because failure to do so will result in loss of a lot of marginal revenue. If this is not anticipated when the tariffs are set then more revenue will be paid by PCTs for the same effective volume of

[13] In general introduction of PbR should transfer revenue from above average cost NHS Trusts to below average cost NHS Trusts (where costs are as reported in the Reference Cost returns). However in practice the adjustment mechanisms particularly the MFF have the effect of offsetting this pattern of funding reallocation.

services as previously, sucking resources into hospitals leaving less for primary care[14]. If on the other hand this is allowed for when setting the tariffs then Trusts that had earlier put in place procedures to ensure complete and accurate coding will be penalised with the tariffs under- reimbursing them for the activity they have provided. Fortunately these are 'one-off' effects associated with the introduction of PbR only.

- **Gaming** There is some risk that some Trusts may be tempted to 'game' PbR, acting in a way that seeks to maximise revenue even when the actions are not in the best interests of patients. Such actions might include: (a) inappropriate early discharge and subsequent re-admission; and (b) code inflation where the procedure provided to the patient is not strictly clinically necessary but where the extra revenue is large. Although it is doubtful whether Uk clinicians would behave in this manner it is clearly a risk and is behaviour that has been observed in other countries following introduction of PbR.
- **Incidental Consequences** The successful implementation of the 4 hour stay in A&E has had the result that many more patients are presenting in A&E. It seems likely that this is because it is now more convenient for patients to go there rather than wait to see a GP. In some hospitals the increase in patients presenting in A&E has been accompanied by an increase in admissions from A&E. With PbR one consequence of this is that NHS Trusts gain the marginal revenue associated with an extra admission. Therefore an incidental consequence of PbR is further to suck resources into hospitals at the expense of primary care and care in the community.
- **Deliberate Consequences** Following introduction of PbR all NHS Trusts are faced with a payment arrangement and financial targets which create strong incentives to maximise admitted care. Where there is a clinical choice between admission, even for a short stay, or treating the patient as an outpatient there is a strong financial incentive to admit.

Sustained Effects

The sustained effects of PbR are likely to include:

- Much closer attention by providers to their cost structure and to managing the key drivers of unit costs
- Planning of services delivery to achieve sustained improvements in unit costs at least as great as the efficiency factor built into year to year adjustments of the tariffs
- This is likely to be achieved through adoption of some or all of: (a) increasing the day case rate for appropriate procedures (because the costs are lower than admitting the patient whereas the marginal revenue is the same); (b) optimising the 'patient journey' thereby shortening average length of stay by avoiding premature admission for elective care and avoiding delayed discharges beyond when clinically necessary; (c) increasing the use of clinical procedures which both benefit patients and shorten length of stay eg laparoscopic surgery; and (d) 'de-bottlenecking' clinical processes thereby increasing the utilisation rate of assets and in effect increasing the effective annual capacity of existing facilities.

Evidence from other countries where variants of PbR have been introduced strongly suggest that the sustained effects will drive productivity improvement, reduce the demand for beds and nurses for a given level of activity and permit a significant increase in the annual activity affordable within the given cash budget (because tariffs will reduce as productivity and average costs decline).

[14] Reference cost data used to set tariffs are based on incomplete coding of activity and hence are likely to underestimate the denominator in the tariff calculation.

Appendix 2 – Flow of Funds in the NHS[15]

1. The new NHS architecture creates an economic system in which buyers (PCTs) purchase Services from providers (hospitals care, primary care etc.) on behalf of patients resident in their areas.

2. The output of a hospital can be viewed as a number of services, $S_1 \ldots S_n$ e.g. cancer, cardiac etc. The output of each Service is a 'bundle' of procedures (HRGs). The volume of output of a HRG is measured in spells (patients receiving the procedure) per period (HRG spells per annum). The volume of output of a Service is the sum of all HRG spells provided by the Service over the period e.g. if service S_1 provides two procedures ($HRG_{(1)}$ and $HRG_{(2)}$) and over a year provides X spells of $HRG_{(1)}$, and Y spells of $HRG_{(2)}$ then the output of service S_1 is (X + Y) spells per annum.

   We can write :

$$Q(S_1) = \sum_{o}^{m} HRG_m \text{ spells}$$

where $Q(S_1)$ = volume of activity provided by Service S_1 over the period

m = number of HRGs provided by service S_1

   The volume of output of the hospital is the sum of the volume of activity provided by all Services :

$$Q_{h1} = \sum_{0}^{n} Q(S_n)$$

where Q_{h1} = volume of output from hospital

n = number of services provided

   The volume of output provided by all hospitals is the sum of activity across all hospital providers:

$$Q_h = \sum_{0}^{r} (Q_{hr})$$

where Q_h = volume of activity provided by all hospitals

r = number of hospital providers

[15] This analysis is a simplification. It ignores the considerable non-service revenues and costs that accrue to hospital and primary care providers. This simplification does not affect the conclusions of the analysis.

3. With PbR there is a tariff for each HRG. The revenue accruing to the provider is the product of the volume of HRG spells provided and the tariff per HRG spell. It follows from para 2 above that the revenue accruing to a service is:

$$R(S_1) = \sum_{0}^{m} (HRG_m\ spells \times T_m)$$

where $R(S_1)$ = revenue from Service S_1

T_m = Tariff applicable to each HRG spell

m = number of HRGs provided by the Service

Equally the revenue of the hospital is:

$$R_{h1} = \sum_{0}^{n} R(S_n)$$

where R_{h1} = revenue received by hospital h_1

n = number of services provided

The revenue received by hospitals in aggregate (which equals the payments made by PCTs to hospitals in aggregate) is:

$$R_h = \sum_{o}^{r} R_{hr}$$

where r = number of hospital providers

An equivalent expression is:

$$R_h = Q_h . \overline{T_h}$$

where $\overline{T_h}$ = weighted average tariff for the 'bundle' of HRG spells provided by all hospital providers

4. PbR tariffs are based on system average costs per HRG spell. If financial balance is achieved by hospital providers average revenue $\frac{R_h}{Q_h}$ equals average cost $\frac{C_h}{Q_h}$.

5. Although PbR does not currently apply to most of primary care and mental health, the same logic applies as for hospital care. Primary care providers deliver outputs all of which can be categorised as HRG like 'procedures'. The volume of these outputs can be measured in appropriate activity units (equivalent to spells) e.g. the number of patient attendances at GP surgeries etc. Implicit prices for each unit of activity are derived as funding spent per unit of activity. These implicit prices equal the average cost of providing services locally. If PbR is extended to primary care then explicit prices will be set reflecting average costs of provision of these outputs across the system as a whole.

Using similar notation as for hospitals:

$$Q_p = \sum_{o}^{r} Q_{pr}$$

where Q_p = volume of primary care activity provided to and purchased by all PCTs

Q_{pr} = volume of primary care activity provided to and purchased by PCT_r

r = number of primary care providers

Similarly, revenue received by primary care providers from PCTs is:

$$R_p = \sum_{o}^{r} Q_{pr}\ \overline{P_p}$$

where R_p = total revenue to all primary care providers

$\overline{P_p}$ = weighted average (implicit or explicit) price for the 'bundle' of activity provided by all primary care providers

6. If PCTs are in income/expenditure balance then cash inflows (adjusted capitation payments) must equal cash out flows to primary care and hospital providers :

$$B = Q_p\ \overline{P_p} + Q_h\ \overline{T_h}$$

where B = cash budget

and other notation is as above

If T_h is set by reference to system average costs and hospitals in aggregate are in income/expenditure balance then:

$$Q_h\ \overline{T_h} = C$$

where C = total fixed and variable costs of provision including historic cost depreciation and dividend on PDC

7. From these relationships it is clear that :

- total monetary demand for services is fixed and given by B
- if B, $\overline{P_p}$, $\overline{T_h}$ and $Q_p / Q_p + Q_h$ are fixed then the total volume of services $(Q_p + Q_h)$ that can be purchased is also fixed.
- the volume of activity that can be purchased by PCTs in aggregate for their patients can only be increased either by reducing $\overline{P_p}$ and $\overline{T_h}$ which - with average cost tariffs - can only be achieved if productivity improves - or by increasing the share of activity purchased from primary care and/or outpatient hospital care providers which has a lower price (because lower average unit cost) than admitted hospital

care. The conclusion is that increasing the volume of activity provided for patients within the given funding envelope requires sustained productivity improvements and/or a shift of resources in favour of non-admitted care.

- any unplanned increase in hospital activity (especially admitted care) increases the claim on PCTs budgets leaving less funding available for primary care.

- payments by PCTs to ISPs reduce pound for pound the funding available to purchase services from NHS providers. If the growth in PCTs purchasing power (i.e. the real value of their budgets) is less than the growth in supply of services by ISPs certain NHS providers may see an absolute fall in activity purchased. In any event there is unlikely to be any increase in demand across the NHS for services provided by ISPs. Any under-used available NHS services capacity will not be able to find a purchaser because the volume of demand is fixed (as long as prices are fixed).

- Reductions in waiting lists and waiting times indicate that the annual flow of services (HRG spells per annum) provided exceeds the rate at which the patients are presenting. This situation will be more pronounced in future if significant productivity improvement is achieved (because the potential annual flow of services will increase further). Even today (2005) many NHS providers have the capacity to provide more of certain services than they currently are doing. They are planning activity to meet local delivery plans (LDPs) which themselves reflect funding available to PCTs. This under-used capacity available at low marginal cost cannot find a purchaser because all PCTs purchasing power has been used up.

Investment Decisions in the Health System

Any provider contemplating investment in expansion of a service or creation of a new service must consider:

- Whether the proposed investment contributes to achieving its strategic objectives, and

- Whether the patient benefits are sufficient to warrant incurring the costs.

The latter question is answered by analysing the marginal benefits and the marginal costs. The assessment involves quantifying the following:

[Equation to follow]

where [ ] is the change in the volume of services provided and paid for in each period
P is the average price received for the extra volume of services provided
C is the change in operating costs arising as a result of undertaking the investment
I is the incremental investment cost incurred
t is the number of years over which the additional services are provided
r is the risk adjusted cost of capital of the provider

When undertaking this assessment providers must look at the total impact, not just the direct impact. The change in the volume and value of outputs is the sum of the direct effects of the investment and the indirect effects induced by it. For example, it may be that expanding outpatient services at the expense of reduced admitted elective care results in a loss of marginal revenue because the expansion in the volume of low marginal revenue work is insufficient to compensate for the loss of a small volume of very high value work.

If the incremental patient benefits exceed the incremental costs then the investment is warranted on financial grounds. This is a necessary but not sufficient condition to invest if providers are to maintain financial balance. In addition the investment must contribute to achieving the Trust's and the NHS strategic objectives.

This patient benefit/cost calculus is only valid if the prices used to estimate the patient benefits are a reasonable reflection of the true value of the services to patients and of the opportunity costs of resources.

Appendix 3 – Two Part PbR Tariffs

The specific design of the two part tariff is important. The proportion of the total payment per unit of contracted activity that is contained in the activity payment per unit of activity provided may, in principle, vary from about 10% (being roughly the proportion of costs that are variable in the short term) to 100% (which would correspond to the one part tariff).

In the case where the activity payment corresponds only to the variable costs financial instability and perverse marginal incentives are minimised, extra PCT purchasing power is greatest (because the marginal price for extra activity is smallest) but the incentive to provide extra activity is also smallest. The one part tariff is the opposite with strong perverse incentives and financial instability associated with patient mobility, least extra PCT purchasing power at the margin and the strongest incentives to provide additional activity if a buyer can be found. Stranded capacity risk is high in the 100% case; however in the 10% case although PCTs have plenty of purchasing power there is little incremental supply because the marginal revenue is very low. The least stranded capacity risk in fact is somewhere in between, say 50%, where PCTs have some additional purchasing power and providers and have greater incentive to expand supply.

Where the proportion is set, if a two part tariff is adopted, is a matter of judgment. A proportion of about 50% would appear to offer the best combination of incentives to elicit desirable responses from purchasers and providers, most likely to deliver the NHS strategic objectives.

Allocative Properties of Two Part Tariff and Medium Term Contracts

The combination of a two part tariff and medium term volume contracts between PCTs and providers has interesting allocative properties. Assume a PCT enters into a 3 year contract to purchase Q1, Q2, Q3 units of activity on a two part tariff. It pays the provider the fixed payment per contracted unit x Q1 in year 1 (spread over the year), the fixed payment x Q2 in the second year and x Q3 in the third year. It also pays in each year the activity payment x the number of units of activity actually provided.

Once the contract is signed the expected marginal revenue for Q units (= contracted volumes in each year) = average system costs (being the sum of the fixed and activity payments). The expected marginal revenue for activity provided in excess of Q is the activity payment only. Within the year the expected marginal revenue loss from providing less activity than Q is the activity payment. The combination of two part tariffs and multi-year contracts that vary the contracted volumes at the service level therefore has strong positive incentive effects. PCTs can vary contracted volumes and by doing so send strong signals to providers to expand certain types of services and contract others in light of the emerging pattern of demand consequent on patient choice and ISP expansion. If the signals are to be strong and sustained, however, the contracts must be at least three years in duration (reflecting the time required to reallocate resources to new uses) and the fixed payment per contracted unit of activity needs to be a meaningful percentage of the total price.

These contracts can be used to allocate resources between services. For example, if a PCT considers as part of the rationalisation of services across an area that certain services should be reduced in capacity at one hospital and expanded at another then it can send signals to this effect by varying contracted volumes from year to year. Likewise if it wishes to see expansion of services in the community it can induce hospitals to expand such services by reducing contracted volumes of admitted services and contracting to make additional payments for new or expanded services in the community.

How does this fit in with patient choice and practice based commissioning? The pattern of demand for services will change as patient choice and practice based commissioning take hold in ways that cannot be easily predicted in advance. As noted in the paper with one part tariffs the outcome is unlikely to be investment in the pattern of services demanded by patients and their GPs on their

behalf. With two part tariffs and medium term contracts PCTs can shape the level, type and location of services over time to provide patients with a wider choice of services eg more closer to home, than they currently have. The alternative to active service development by PCTs and SHAs is to leave investment decisions to market forces – which, as the paper indicates, is most unlikely to deliver the desired level, type and location of services.

Two Part Tariffs – All or Some Services?

Should the approach be applied to all PbR tariffs or just some categories of service? My answer is that there is a clear case for their use across all services. In A&E, for example, if the number of patients passing through A&E increases, because the 4 hour wait persuades patients to visit A&E rather than wait to see their GP, why should hospitals receive the average cost for additional patients above some base line when their costs go up rather little. This simply sucks revenue away from PCTs that could have been used to expand services in the community and help keep people out of hospital.

However the case is equally strong for the 'mobile' service categories eg elective surgery. Where patients are 'mobile' the financial instability risks are greatest with a one part tariff. Adoption of the two part tariff mitigates those risks without reducing the pressures over the year acting on providers to ensure their average costs are no greater than their average revenue (= system average cost).

Will adoption of the two part tariff land PCTs with 'stranded payment risk' ie the risk that they pay providers for services that they do not receive. The answer is no. Consider the case shown in Table 2.

Table 2

	Contracted volume	Activity Provided (Case 1)	Payment – Case 1	Activity provided - Case 2	Payment – Case 2
Hospital 1	100	100	1000	120	1100
Hospital 2	100	100	1000	100	1000
Hospital 3	100	100	1000	80	900

In case 1 the contracted volume and the actual activity provided are equal. The two part tariff pays £5/contracted unit and £5/actual activity unit provided so the revenue is 1000 (500 + 100 x 5). In case 2 although the PCT contracted for 100 units from each hospital the actual pattern of activity changed; hospital 1 provided more than the contracted volume, hospital 3 less and hospital 2 the same. In this case the distribution of revenue across providers is different but the total payments for contracted activity made by PCTs is and must be the same because the total payment for contracted activity remains the average system cost.

It follows that this approach can be adopted for 'mobile' and 'non-mobile' activity by PCTs with no risk of 'stranded payment risk' and reduced risk that higher induced activity levels in hospitals will eat into their limited discretionary cash flow.