ORGANISATIONAL COSTS IN THE NEW NHS

An introduction to the transaction costs and internal costs of delivering health care

Bronwyn Croxson

Office of Health Economics
12 Whitehall London SW1A 2DY

Price £7.50
ISBN 1 899040 90 0
April 1999

Acknowledgements.
I am grateful to Angela Reidy, Pauline Allen and Jon Sussex with whom I discussed parts of the paper, for the very helpful comments of four anonymous reviewers, and to the unidentified health authorities from which case-study material has been drawn. The views expressed in the paper are wholly my own.

About the author

Bronwyn Croxson spent five years teaching economics at Clare College, Cambridge. From 1996 she spent two years in a research post, working with a health authority on economic issues relating to commissioning. She is currently a research economist at the Institute of Child Health/Great Ormond Street NHS Trust, London, working on a project developing the methodology used to value research. The views expressed in this monograph are her own and not those of the ICH/GOS.

OFFICE OF HEALTH ECONOMICS

1 Terms of Reference

The Office of Health Economics (OHE) was founded in 1962. Its terms of reference are to:

- commission and undertake research on the economics of health and health care;
- collect and analyse health and health care data from the UK and other countries;
- disseminate the results of this work and stimulate discussion of them and their policy implications.

The OHE is supported by an annual grant from the Association of the British Pharmaceutical Industry and by sales of its publications, and welcomes financial support from other bodies interested in its work.

2 Independence

The research and editorial independence of the OHE is ensured by its Policy Board:

Peer Review

All OHE publications have been reviewed by members of its Editorial Board and, where appropriate, other clinical or technical experts independent of the authors. The current membership of the Editorial Board is as follows:

CONTENTS

1 INTRODUCTION

There is widespread criticism of organisational costs in the NHS. 'Reduce bureaucracy, release funds for patient care' has become a popular slogan. This so-called 'bureaucracy' may, however, be essential to delivering health care. The complexity of modern health care means that some organisational costs have to be incurred if the right health care is to go to the right people at the right time. We do not have a comprehensive estimate of organisational costs in the NHS, but their significance is indicated by the sum of £1,225m spent on managers by English NHS trusts in 1996/97 (about 5 per cent of their total income). This figure is consistent with the US Health Care Finance Administration's estimate that in 1991 $43.9bn was spent on administration in the US, almost 6 per cent of total expenditure. Understanding these costs, within a rigorous framework, is vital to the efficient allocation of health care resources because these estimates are not on their own useful. They do not tell us whether the costs were necessary to deliver health care, nor do they tell us whether they were costs incurred in ensuring that local health services met the needs of local populations. So-called 'bureaucracy' may be an end in itself, and is now commonly used as a pejorative term to suggest that it is. However, it may also comprise expenditure necessary to co-ordinate the different people who deliver increasingly complex health care, or the costs incurred in making sure that the right health care is delivered to the right people at the right time. Terms like 'bureaucracy', 'management costs', 'administration', 'red tape' and 'transaction costs' are used interchangeably by most commentators. Yet they may refer to activities making very different contributions to patient care.

The issue of organisational costs has been pushed to the fore of health care policy making in the UK, particularly by claims that such costs rose following the introduction of the NHS internal market in 1990. The Conservative government introduced a number of initiatives to reduce these costs in the internal market, and now the belief that 'bureaucracy' is too high in the NHS underlies a series of White Papers, which outline measures supposed to generate a £1bn in 'bureaucracy' during the current Labour government's term.

There is, however, little rigorous evidence about relative levels of organisational costs in the NHS, before and after 1990. It is certainly the case that we would expect the type of organisational costs incurred to change, because the internal market introduced contractual relationships into the NHS. The costs accompanying contractual relationships, transaction costs, are likely to be qualitatively different from those incurred under the pre-1990 system, internal costs, since the problems relating to uncertainty and conflict differ in the two systems. We do not, however, know that this change in type was a bad thing, since we do not have convincing evidence that the overall level of organisational costs increased in the internal market or that there were no offsetting efficiency gains elsewhere.

This monograph presents a framework that can be used to conduct the type of rigorous analysis necessary to analyse changes in the costs of organising health care. As argued above, an increase in resources devoted to organisation and management may be a good thing. Alternatively, an increase in organisational costs may not improve patient welfare. There may be waste and unnecessary expenditure in organisation and management, as in any activity. A change in organisational costs must, therefore, be subject to the same rational analysis as a change in any activity.

The monograph has been written for people working in health care, with an interest in economics but not necessarily with any formal training in the area. It begins in section 2 with a brief description of the structure of the NHS internal market. Section 3 outlines an economic framework which can be used as the basis of rigorous analysis of organisational costs, illustrated with a case study of contracting for orthopaedic services. Section 4 discusses the 1997/98 White Papers which outline reforms to the NHS in England, Wales, Scotland and Northern Ireland. This section analyses the measures outlined in these papers supposed to reduce organisational costs and concludes that many of the measures will be ineffective as well generating new organisational costs not recognised in the White Papers. Section 5 concludes with a brief discussion of the issues raised.

2 THE STRUCTURE OF THE NHS INTERNAL MARKET

In 1990 the NHS was reformed, with various structural changes made to improve efficiency and access[1]. The post-1990 structure has been criticised for increasing transaction costs and management costs, without delivering the hoped for efficiency gains. This section provides a brief introduction to the changes introduced in 1990, before addressing in subsequent sections the impact of this structure on organisational costs.

Before 1990, planning, management and provision of health care all tended to be undertaken within the same organisations. In 1990 the NHS was reorganised to separate two distinct functions: providing and purchasing health care. This separation was the basis for setting up two different types of organisations, providers and purchasers. The various provider organisations, NHS trusts, were responsible for providing acute care, mental health care and community health services. Two types of purchasing organisations were created: health authorities and GP fundholders. Each health authority was responsible for purchasing health care for citizens living in their local area, with GPs who volunteered to become fundholders being given responsibility for purchasing a specified range of services for their own patients[2].

The reforms deliberately inserted a market interface between public organisations, even calling the new structure an internal market. The provision of health care was now formally governed by contracts. Each purchaser was required to make annual contracts with providers, supposed to cover most of the health care they expected their population to need during the forthcoming year. Contracts could initially take one of three basic forms:

1. block;
2. cost and volume;
3. cost per case.

1 Working for patients (1989).
2 The role of GPs as purchasers increased over time, evident in various extensions to the GP fundholding scheme, and in policy initiatives promoting the so-called 'Primary Care Led NHS'.

Box 2.1 Types of contract

Since 1993 the NHS Executive has categorised contracts as follows (descriptions adapted from Raftery et al. 1994):

1. **simple block contracts.** Purchasers pay a fixed amount for access to a defined range of services or facilities. These contracts commonly included indicative activity levels and a maximum waiting time;

2. **sophisticated block contracts.** Purchasers pay providers a fixed amount for access to a defined range of services or facilities. In addition, contracts have indicative activity targets, and specify action if those targets are exceeded or not met. Some contracts, for example, specify that if activity varies by 5 per cent or more from the target level, a separate cost per case schedule comes into operation;

3. **cost and volume contracts.** These contracts are specified in terms of activity, with a fixed price paid for a specified volume of treatment and a price per case for activity over that volume. There is usually a ceiling placed on total activity;

4. **cost per case contracts.** The provider agrees to provide specified treatments, with each episode paid for using an agreed price schedule.

Raftery et al. (1994) found that most contracts involving a health authority were sophisticated block contracts (62 per cent). 20 per cent were simple block, 17 per cent were cost and volume and 1 per cent were cost per case.

From 1992, the NHS Executive encouraged purchasers to use more complex contractual forms than simple block contracts, and from about 1993 recognised a fourth type of contract: sophisticated block. Box 2.1 describes the characteristics of the different types of contract.

Some health care was provided without being covered by a contract. In this case it was paid for on a case-by-case basis and called an 'extra-contractual referral' (ECRs).

Since 1997 contracts between NHS organisations have been called service agreements, and the recently published White Papers suggest that they will continue to be called agreements. Agreements can be analysed using the same framework as contracts, since they will continue to be costly to write and monitor, and will continue to affect incentives.

The structure of the internal market was designed to improve efficiency in health care delivery. It was intended that purchasers should be given fixed budgets and the freedom to purchase health care from any provider. This was designed to promote competition between providers, thereby increasing efficiency in service delivery. Patients were also given full mobility, with no restrictions over changing their GP. This was designed to promote allocative efficiency in purchasing decisions, since part of GPs' incomes relates to the number of patients for whom they are responsible. Health authorities were given a responsibility to consult their local population, and were monitored on various dimensions relating to local service provision and needs. This was designed to make them directly accountable for representing local needs in purchasing decisions.

From its inception, economists argued that the structure of the internal market was likely to generate high transaction costs, not in itself a bad thing if accompanied by these hoped efficiency gains and improved access to health care[3]. The key issues are, therefore, whether the internal market delivered the hoped for gains and, if so, whether these were sufficient to offset any rise in organisational costs. Current policy initiatives are based on the premise that it did not: these initiatives will be examined in section 4 below.

3 Bartlett (1991), Roberts (1993), Robinson (1990).

3 A FRAMEWORK FOR ANALYSIS

This section is methodological, and outlines a framework that can be used to analyse the impact on organisational costs of different types of health care delivery structures. It begins by describing the types of question addressed in this framework and then discusses the nature of transaction costs and internal costs as the different types of organisational cost. Two different aspects of the role and nature of organisational costs are developed in sub-sections 3.4 and 3.5: 3.4 argues that organisational costs may be a vital prerequisite to optimal health care provision; and 3.5 that, in practice, organisational costs are generally reduced by changes in informal practice rather than in formal governance. The section concludes with a short summary of empirical work seeking to measure organisational costs.

3.1 Governance structures

Changes in the organisation of health care delivery can be analysed using a framework developed by industrial economists, called institutional economics[4]. This framework allows economists to ask: what determines whether firms buy components or make them in-house; what determines the size of firms? In health care, this type of question arises in a number of contexts, including:

- what determines whether GPs undertake minor surgery themselves or send patients to hospital?
- what are the advantages of giving primary care professionals responsibility for purchasing as well as providing health care?
- why have the functions of purchasing health care and providing it been divided between health authorities and trusts?

These questions can be asked in two different ways:

1. the so-called **positive** questions: what is the pattern of activity?

4 This framework is usually associated with Coase (1960) and Williamson (1985), accessibly introduced in Eggertson (1990). An alternative approach to the same issues has been stimulated by Hodgson (1988). There are only a few UK-based studies which explicitly address health related issues using this framework. See Bartlett (1991), Bevan (1998), Ferguson and Keen (1996).

This type of question involves description of observed phenomena, including activity, its drivers, and the incentive structure;

2. the so-called **normative** questions: what should we observe, in other words what is the optimum or most desirable pattern of activity? This type of question involves analysis of the characteristics of alternative ways of organising activity.

Institutional analysis is based on the principle that activity occurs within particular governance structures. Different governance structures comprise different sets of rules for organising economic activity. Governance structures differ according to their decision-making mechanisms, in other words, their resource allocation rules. Markets are one type of governance structure, and organisations another. In markets, prices play a greater role in decisions (i.e. in resource allocation) than within organisations, and this generally alters the nature of the incentive structures in the two types of governance structure.

It is not always possible to distinguish markets and organisations since, as will be discussed in section 3.5 below, they borrow features from each other. For the purposes of analysis it is, however, helpful to begin by setting out a distinction between the two as different types of governance structure.

The rules of stereotypical markets are that relationships are not expected to continue beyond the current exchange (i.e. interactions between individuals or organisations are said to be 'short term' and 'anonymous'), and that prices contain all relevant information. This implies that prices of high quality goods are higher than those of low quality goods and, in health care, that prices are well-defined across the whole casemix range.

The stereotypical rules of organisational governance are, by contrast, that interactions are expected to be repeated. In other words, within organisations relationships are expected to last for some time. In addition, centralised, hierarchical decision-making is expected. A number of different organisations exist within the NHS, including individual trusts, health authorities and GP partnerships.

Governance structures may be changed. An organisation may be dis-integrated, with a market interface inserted between its parts. Conversely, separate organisations, previously interacting across a market interface, may be merged into a single hierarchy. Arguably the first happened in the NHS in 1991, when the purchaser-provider split was introduced, and the second is happening now, with the policy shift towards integrated care. Integration itself can take different forms:

- **vertical integration** occurs if different stages in a process are joined within one organisation. In this case a transaction is said to have been 'internalised'. Vertical integration occurs, for example, within primary care when GP partnerships which have previously purchased minor surgery from a hospital start providing it themselves, in-house. The 1990 reform of the NHS is sometimes characterised as vertical dis-integration, with the market-style purchaser-provider split replacing the organisation which formerly governed both types of function internally;

- **horizontal integration** occurs if similar stages of a process are merged within one organisation. This occurred in primary care in the internal market when different GPs joined together as locality purchasing groups, and will occur in the future when GPs are merged into Primary Care Groups under measures announced in recent White Papers[5]. It occurs in secondary care when mergers occur between trusts.

Changing a governance structure may change costs. It is well recognised that average costs may be affected by the size of a production unit: if there are economies of scale or scope, increasing the size of a production unit may lower average costs. One way of changing an organisation's size and achieving economies of scale is to alter its governance structure by, for example, horizontal mergers. This is often presented as a rationale for mergers between NHS trusts.

5 See section 4 below.

A key insight of institutional economics is that changing governance structure affects not only production costs, but also the costs of organising activity. Market-style governance generates transaction costs, whereas internal costs are generated within organisations.

3.2 Transaction costs

To undertake a rigorous analysis it is helpful to categorise transaction costs. Those costs incurred in the process of reaching an agreement (called ex ante transaction costs) include the costs of search and of negotiation.

Search costs comprise the costs of finding another party to engage with. Purchasers must find out information about the range of potential providers, their prices, their capacity and the quality of health care they are likely to provide. In health care this is complex, since a comparison based solely on offer-prices may be misleading: price differences between providers might result from different assumptions about casemix, quality or volume of activity. In other words, prices are unlikely to contain all relevant information. Moreover, it may not be possible to 'unbundle' the price of a particular procedure from the overall range of procedures a provider undertakes and the overall 'price' offered[6].

Negotiation costs are also incurred ex ante, as purchasers and providers agree on the terms of the exchange. They have to agree the way that activity should be measured; the volume, quality and range of activity to be covered by the contract; the duration of the deal; and the price. Agreeing appropriate measures may be difficult (and therefore costly) in health care, since the output sought by purchasers is a change in patients' health status, which is often hard to measure and multi-dimensional. Box 3.1 introduces a case study illustrating the nature of transaction costs. Box 3.2 then illustrates ex ante transaction costs in the context of this case.

Ex post transaction costs arise after the agreement has been writ-

6 Dawson (1994), Leary (1997).

Box 3.1 Case study of purchasing elective orthopaedic surgery

Contracting for orthopaedic surgery provides a useful illustration of the nature of transaction costs, since it comprises a range of relatively discrete elective procedures, some of which form the basis of Patient Charter standards, and some of which can be provided in a number of alternative venues.

Orthopaedic surgery covers a range of different procedures, ranging from those that are relatively simple, inexpensive and able to be carried out as day cases in hospital or primary care (such as arthroscopy), to complex procedures often carried out only in special hospitals. Operations to replace hip joints or knee joints are common, and are commonly carried out in NHS acute trusts and in private providers.

The examples discussed in various boxes are drawn from a detailed case study carried out in 1995, involving urban English health authorities with access to a number of alternative NHS and private providers. The case study is used to illustrate the nature of transaction costs and governance structures. The example refers primarily to the costs incurred in purchasing care, although providers also incur transaction costs.

ten, while it is being executed. They include the costs of monitoring the other party to ensure that they abide by the terms of the agreement, the costs of enforcing the agreement if it is violated by the other party, and the costs of re-negotiation in the event of unforeseen circumstances. Box 3.3 illustrates these costs.

Transaction costs depend on a number of factors, each explored below: the level of uncertainty; the availability of alternatives; the nature of the relationship between the contracting parties, especially the extent of opportunism and trust; and the mechanisms available to enforce contracts. It will often be the case that these factors will generate a trade-off between ex ante and ex post transaction costs.

3.2.1 Uncertainty and incomplete information

Two factors mean that health care decision making is based on unusually limited amounts of information. First, individuals have limited ability to process the information that is available[7]. This problem is magnified in health care decision making by the complexity and anxiety accompanying many decisions. Second, there is considerable uncertainty, since rapid technological change keeps altering the environment, which is anyway characterised by hazards that are hard to predict, such as epidemics and natural disasters. Uncertain environments have high ex ante transaction costs, since it is costly to try to predict all possible events. This makes it costly to agree on contingent actions for all events, and means that some events will not be included in contracts. No NHS contract, for example, specifies who would bear the financial consequences of a cholera epidemic. The costs of writing complete contracts mean that there is often a trade-off between ex ante and expected ex post transaction costs: economising on ex ante transaction costs by leaving contracts incomplete may increase ex post transaction costs, as the parties later seek to enforce or re-negotiate an incomplete agreement.

In general, the ex post costs of monitoring and enforcing an agreement will depend on the precision with which it was written ex ante (and which, if any, unforeseen events occur). An imprecise contract may have lower ex ante transaction costs, but may lead to higher ex post costs if there is a dispute over whether the contract has actually been breached. This has been an on-going problem in the NHS internal market, with purchasers and providers arguing about whether particular activity is included in a block contract (often a purchaser's position), or whether it falls outside the contract and so should carry additional remuneration (often a provider's position). In general, block contracts may be less costly to negotiate than cost and volume or cost per case contracts, since block contracts do not require detailed ex ante assessment of casemix or level of demand, both of which are uncertain and therefore costly to predict. Block contracts may, however, be accompa-

7 Simon (1982).

Box 3.2: The ex ante transaction costs of purchasing orthopaedic surgery: search and negotiation

Search costs are incurred as purchasers locate providers able to supply the range, quantity and quality of orthopaedic procedures they wish to purchase on behalf of their local population. The purchaser has to be satisfied that providers have adequate access to intensive care and cardiology should anything go wrong. This may limit the extent to which they are willing to use private providers or specialist hospitals. They also have to ensure adequate access to rehabilitation services, particularly occupational- and physiotherapy. These services might be delivered within a provider, or the provider might have to engage in lateral 'service level agreements' with other agencies. Since most contracts between health authorities and trusts are based on activity rather than outcome (i.e. on 'number of hip replacements' or 'number of visits', rather than on 'number of people restored to full, pain-free mobility'), purchasers often spend time ensuring that providers have actually entered into appropriate service level agreements – in other words, they try to monitor inputs directly to safeguard quality.

Search costs also involve comparing and analysing prices offered by different providers. Pricing methods vary between providers:

- some offer a price per procedure, with procedures often grouped into cost bands. These are often based on so-called 'HRG' costings, now required of all providers. HRGs, or Healthcare Resource Groups, form a taxonomy of different types of health care, with different interventions and diagnoses grouped together in a way which should reflect resource usage;
- some offer a uniform average price for all orthopaedic procedures. This often forms part of a unique arrangement between a provider and particular purchaser: i.e. part of a block contract, predicated on total activity and total value.

Variations in the prices charged by different providers might reflect differences in efficiency. But, equally, they might reflect any of the following factors, which must be appraised by a health authority during its search:

- variation in the age groups usually treated. Older patients often have a longer stay in hospital and require greater input from complementary services;
- expected differences in casemix and resulting differences in the range of procedures conducted. There is wide variation in the cost of different elective orthopaedic procedures from relatively cheap day-case interventions such as arthroscopies, to the more expensive but still frequent joint replacements, to very expensive and infrequent spinal surgery. This problem is not eliminated when prices are based on OPCS codes (which were used before the introduction of HRGs) or HRG categories. One NHS trust, for example, conducts 13 different types of total hip replacement, all categorised within one OPCS code, yet ranging in price from £3,261 to £14,119 (1993 prices);
- variations in quality. In the context of orthopaedic surgery, these may alter expected costs by affecting expected revision rates. The price of artificial hips varies markedly between suppliers: it has recently been found that a type of hip supplied relatively cheaply by one company, and therefore allowing providers using it to charge lower prices, has a higher than usual risk of failure and therefore may involve higher long-term costs.

These factors mean that negotiations between providers and health authorities over prices may be extended and acrimonious, as might those over the form a contract should take. Different types of contract have different characteristics. Specialty-level block contracts, which specify total payment and level of activity, and which therefore give all orthopaedic activity a uniform average price, transfer to providers the risk of an unanticipated increase in demand for expensive procedures. If, by contrast, contracts are based on price per specific procedure undertaken, most of the risk is held by purchasers.

Box 3.3: The ex post transaction costs of purchasing orthopaedic surgery: monitoring, enforcing and renegotiating agreements

Contracts between health authorities and NHS acute trusts in the internal market were usually negotiated and written specifying cost, activity and maximum acceptable waiting times or length of waiting list. During the period covered by the contract, usually a year, health authorities monitored the volume of activity actually undertaken by the trust, its cost, and waiting times or the length of waiting lists. If any of the variables varied significantly from that specified in the contract, health authorities sought to enforce the contract and if necessary renegotiate terms. All three activities - monitoring, enforcing and renegotiating – are costly.

Monitoring involves the direct costs of obtaining information and comparing actual activity with that agreed in the contract. Obtaining information is costly. The data relating to waiting lists and activity are provided by the trusts themselves, often with a considerable time lag between when someone is treated and when a health authority is informed. Moreover, the data often omit information important to the health authority such as a patient's postcode. In the case study outlined in Box 3.1, contracts stipulated that data had to be complete, with fines to be levied if it was incomplete. These provisions were not, however, enforced. In one contract the data was such that if fines had been enforced, the trust would have had to pay £50,000. The fines were not in fact imposed, since the health authority believed that to do so would divert funds from patient care.

The ex post costs of monitoring, enforcing and renegotiating contracts are high in health care, since it is hard to predict all events in advance. The NHS Executive has identified the following as likely to cause the volume of activity to vary from that specified in the contract:

- unforeseen changes in demand;
- unplanned changes in the care delivery process within provider organisations;
- clinicians' behaviour, in other words 'whether clinicians will take note of contract targets and adjust workload accordingly';

- the number of patients treated for whom a postcode is not recorded but who actually live in another district may be higher than anticipated;
- the data or targets may have been inaccurate;
- random fluctuations in volume or casemix.

Contracts which have activity agreed in cost bands may be more costly to enforce than block contracts. In the case study referred to above, one trust insisted that activity be divided into cost bands. In this case, the health authority had to monitor to ensure that the provider did not engage in 'cost-shifting', by claiming particular patients were treated in more expensive categories than was actually the case.

Contracts can be enforced and renegotiated only if violation is demonstrable. It is often hard to prove that a contract has been violated. Checking to prevent cost shifting, for example, requires detailed assessment of casemix, which is costly to obtain.

nied by high ex post costs, if providers are not generally willing to treat an unlimited number of patients for a finite sum of money[8].

Transaction costs also increase with increases in the extent to which each party's information differs, called asymmetric information. If, for example, a provider believes that the value a buyer places on a service is higher than the buyer's actual valuation, or believe that their budget is higher than it really is, then that provider may 'hold out' for a price that is actually unattainable, thereby increasing transaction costs[9]. This type of situation is common in the NHS: in one set of contracting negotiations in 1995, for example, there was a £9m gap in the positions of a trust and a health authority[10].

8 Bartlett (1991).
9 Farrell (1987).
10 Comptroller and Auditor General (1995).

Box 3.4: Factors restraining competition in orthopaedic surgery

In large urban areas there appear to be sufficient providers to allow competitive behaviour (there are, for example, 32 potential providers of orthopaedics in what used to be the North Thames Region of the NHS), yet there is also evidence that there is little actual competition between providers. In the context of orthopaedic services, the absence of competitive behaviour even when there are alternative providers reflects all of the following factors:

- the costs of monitoring quality. Purchasers stay with providers of whom they have experience and who they believe are more likely to safeguard quality;
- preferred provider policies. Some health authorities did informal 'deals' with providers, guaranteeing long-term contracts in return for year-on-year reductions in average prices;
- political considerations. Health authorities are often effectively 'tied' to local NHS trusts for political reasons;
- local preferences. Under the arrangements of the NHS internal market, local GPs and residents may have links with particular providers. The health authority may remove their contract for orthopaedic services from a particular provider, only to find that GPs continue to refer to that provider, generating ECRs that are more costly than the original contract. The health authority is, therefore, locked into using the original provider;
- professional networks of clinicians. Orthopaedic consultants may have joint posts at two or more local providers. The health authority may find that even if patients were referred by GPs to alternative providers, consultants sometimes treated them at the first, again locking the health authority into that provider;
- transaction-specific investment. The institutional economics literature focuses on transaction-specific investment as a source of non-competitive contracting. This type of investment occurs when one

or both of the parties have to make a specific investment which has a low value in alternative transactions and which therefore ties that party to the relationship. (The investment may be in physical or human capital.) Although institutional economists emphasise this as an impediment to competition, it is hard to identify in relationships between purchasers and NHS trusts. Only two instances arose in the orthopaedic case study referred to in Box 3.1. It was evident in relationships between individuals in different organisations, who over time have developed good relationships and methods of communicating. It also had the potential to arise out of the specialised and unusual computer system of one of the trusts. This system generated output that was incompatible with the software and systems at the health authority as well as those used in other providers. Had the health authority invested in a compatible system, they might have been locked into a contract with that particular trust.

3.2.2 Competition

The transaction costs accompanying an agreement between a particular purchaser and provider are affected by whether the purchaser has viable alternative providers and vice versa, in other words, by the level of competition. Ex ante transaction costs are likely to be lower if there is competition, since if negotiators hit a bargaining impasse (in other words, get stuck) they can go elsewhere. Having viable alternatives also lowers ex post transaction costs, limiting the benefits to one party of 'holding up' the other in direct violation of the contract or of taking advantage of poorly specified terms[11]. Such behaviour is consistent with the way some NHS acute trusts demand additional payments mid-way through a financial year, threatening to suspend elective admissions unless they receive additional funding.

11 Klein (1988).

The level of competition is defined partly by the number of local providers (for purchasers) and of purchasers (for providers). It is also affected by the regulatory environment, which in the NHS restricts purchasers' mobility. Health authorities, for example, have special responsibility for local NHS trusts, and government policy now restricts the use of private providers by NHS purchasers. Evidence suggests that there is little competition between NHS trusts: that most areas have too few trusts for effective competition and that even in areas where competition might be feasible it does not occur[12]. The factors which might limit competition in the NHS are illustrated in Box 3.4, which discusses factors impeding competition in the context of orthopaedic surgery.

3.2.3 Opportunism and trust

Transaction costs depend vitally on the expectations each party has about each other's behaviour[13]. Some commentators argue that expectations about opportunistic behaviour are the fundamental determinant of the level of transaction costs[14]. Williamson argues that anticipated opportunism leads to high ex ante transaction costs, as each party tries to protect themselves from the other, and to high ex post costs as they respond to its consequences.

There is a burgeoning economic literature addressing behaviour and how it affects economic outcomes. A number of writers have analysed opportunistic and trustworthy behaviour, and argued that co-operative behaviour may lead to more efficient outcomes than competition[15]. In this literature, trust is commonly treated in one of two ways: as the antithesis of opportunism, in other words, as behaviour generated solely by ethical principles; or as rational self-interested behaviour in contexts where co-operation is more efficient than conflict[16].

12 Appleby (1993), Le Grand (1994), Propper (1996).
13 North (1981).
14 Williamson (1985).
15 cf. Croxson (1997).
16 cf. Lyons and Mehta (1997).

Box 3.5: Co-operation as efficient, self-interested behaviour

The following quotation is taken directly from a hands-on guide to negotiations, written for NHS managers. It extols the importance of short-term co-operation as a way of safeguarding the long-term interests of an individual organisation:

'Few negotiations are 'one-offs'. More usually they are part of a continuing relationship and the parties may well want to do business year after year. This is nearly always the case in NHS commissioner/provider situations. Sometimes, one party will have a particularly strong negotiating position - perhaps the provider has just lost another contract or perhaps a competing supplier cannot provide the service this year. There is an overwhelming temptation to exploit this temporary advantage and to adopt a 'take it or leave it' attitude. The inevitable consequence of yielding to this temptation is that, when the balance of power changes, the other party will seek to take advantage of its negotiating position. Negotiators have to take particular care over how much they exploit temporary differences in negotiating power. Remember that win:lose will normally be followed by lose:win.' (Faulkner, 1996, p.135).

There is also an economic literature on the prerequisites of trustworthy behaviour. One strand in the literature, based on game theory, argues that repeating an interaction, in other words having long-term relationships, allows parties to demonstrate trustworthiness. In this literature, repetition also discourages each party from acting opportunistically in the short-run if in doing so they would jeopardise a more valuable long-term relationship[17]. An example of this type of situation in the NHS is described in Box 3.5.

Another strand examines institutions directly. North argues that, in the context of manufacturing, employers may deliberately invest in measures which increase employees' loyalty, such as subsidised hous-

17 Taylor (1997).

ing, as a way of reducing the transaction costs of monitoring behaviour[18]. Others have argued that organisations foster loyalty and what might be called team spirit, which could also be interpreted as transaction cost reducing institutions[19]. Similarly, professional associations like the British Medical Association and Royal Colleges can be analysed as organisations which foster particular behaviour directly (through peer review) and indirectly (through ethical codes), effectively reducing the costs to the NHS of monitoring professional behaviour.

In conclusion, this section has argued that the transaction costs of market exchange are expected to be higher in environments characterised by uncertainty, incomplete information and lack of competition. In some contexts these transaction costs may be reduced by changing the way transactions are governed, in other words, by switching from market exchange to internal production. High transaction costs are not, however, sufficient to justify internalising production. It is also necessary to consider the impact this will have on production costs and, as will be discussed in the next section, the additional organisational costs likely to arise within organisations.

3.3 Internal costs

A GP choosing a governance structure for physiotherapy might chose market governance, in other words she might buy in the services of an external agency, or she might chose internal governance, in other words employing a physiotherapist within the organisation. Internal governance eliminates some transaction costs but generates other types of cost. The GP must negotiate wage and employment conditions; she can now monitor the physiotherapist directly; and she now has to organise and monitor support services herself.

Under some circumstances these costs will be lower than the transaction costs of market exchange, and this section begins by outlining why this might be the case. This is, however, not guaranteed and the

18 North (1981).
19 Hodgson (1993).

section concludes with a discussion of the additional costs incurred by internal governance. This is a vitally important issue since the 1997/98 White Papers announce measures purported to be transaction cost reducing, which include combining separate health care professionals into local associations[20]. Any saving in transaction costs may, however, be offset by additional organisational costs arising within these associations.

The transaction cost advantages of internalising activity rely on characteristics associated with stereotypical organisations: the existence of a central decision-making core; the greater longevity of within-organisation relationships compared with market-relationships; and the superior alignment of incentives within organisations, which reduces conflict.

Within an organisation, individuals may be less able or willing to be opportunistic, reducing the ex post transaction costs that might occur in market exchange. Some authors argue that opportunism is reduced by the existence of a central decision-maker, who resolves disputes and monitors behaviour[21]. Others argue that centralised decision-making is seldom completely effective, but that the stability and longevity of within-organisational relationships reduces transaction costs by generating loyalty and facilitating learning[22]: in other words, by facilitating co-operative behaviour.

The characteristics of organisations mean that it may not be necessary to specify in advance all details relevant to future activity (effectively economising on ex ante transaction costs). This is likely to be important in health care, characterised as it is by high levels of uncertainty. Uncertainty about demand and casemix make it hard to specify all details in advance, as required in the contracts of the NHS internal market.

Within an organisation, the existence of long-term relationships and centralised authority reduces the number of contracts needing to be made. Instead of each employee contracting separately with each other, the suppliers and the distributors, they effectively cede that responsibility to the central decision-maker in their single employment contract.

20 See section 4, below.
21 Coase (1937); Williamson (1985).
22 Hodgson (1993); Nix (1994).

Box 3.6: Monitoring versus incentives

Organisational costs, incurred when controlling opportunism, may take one of two forms:

1. they may involve direct monitoring of activity and enforcing a contract, which will be effective only if there are credible sanctions;
2. they may involve aligning incentives directly. If, for example, the goal is to increase hip replacements, this may be achieved by paying clinicians for each hip replacement they undertake.

Rational organisational design proceeds by comparing the two forms, under different governance structures, and selecting that which minimises economic costs.

Although these factors mean that some transaction costs may be reduced by internalising activity it is nonetheless costly to organise activity within an organisation. Internal costs arise because although an organisation is, by definition, one unit it comprises a number of different employees and divisions, which in turn make it costly to co-ordinate activity and deal with conflicting interests.

The activity of different individuals has to be co-ordinated, which generates costs even in the absence of any conflict of interest. Information has to flow between different individuals and divisions. These costs increase as activity becomes more complex and as organisations increase in size[23]. Co-ordination costs might be reduced if decision-making power is decentralised, or if the aims and interests of all an organisation's parts are aligned.

Internal costs also arise out of within-organisation conflicts of interest. A fundamental conflict occurs when individual employees prefer leisure time or an 'easy life' to working, and so put in less effort than would be optimal for the success of the organisation as a whole.

23 Coase (1937).

This type of conflict might be manifest within a partnership, such as a group of general medical practitioners, as well as in organisations with salaried employees. It means resources have to be devoted to monitoring behaviour or designing and implementing an appropriate incentive structure[24]. As outlined in Box 3.6, organisational costs may be incurred directly in monitoring, or indirectly in the administration of a system based on incentives.

In a market, conflict of interest can be an engine of efficiency, since self-interest and competition can generate efficient outcomes. Within an organisation, by contrast, conflict diverts resources from the organisation's primary goal, meaning resources have to be devoted to aligning the interests of the organisation's constituent parts. If the conflict is not resolved, perhaps because insufficient resources are devoted to doing so, resources may none-the-less be lost in wasteful under-performance. This is a key point, explained in the next section.

3.4 The consequences of reducing organisational costs

Just as some production costs are unnecessary, so too some organisational costs are unnecessary or wasteful. Belief that this is the case has led successive Conservative and Labour administrations to announce programmes designed to reduce 'bureaucracy to release funds for patient care'. Most of the resulting programmes have not been precisely targeted largely because what is meant by the term 'bureaucracy' has not been carefully defined. The programmes have instead usually involved general cuts in organisational costs. In 1994, for example, Virginia Bottomley as Secretary of State for Health announced that merging two regional health authorities would release £4.7m for patient care. In 1995 her successor, Stephen Dorrell, announced that the costs of employing senior managers must be reduced by 5 per cent, and in 1997 the Labour Government's Secretary of State for Health,

24 cf. Myerson (1979).

Frank Dobson, announced that the NHS would save £100m in management costs, to be reinvested in breast cancer services[25].

Allocating fewer resources to support transactions or within-organisation management will achieve the end of reducing those costs. It will not, however, necessarily free resources for achieving more fundamental objectives. In the context of the NHS, savings made by allocating fewer resources to organisational costs will not necessarily lead to more patient care. Moreover, devoting fewer resources to organising activity may mean more, not less, waste.

The activities giving rise to organisational costs are sometimes as essential to delivering patient care as conventional medical activities, such as delivering medication and carrying out surgery. Health care usually involves a significant number of different people, often from different organisations, who need to be contacted and co-ordinated.

Organisational costs may also be necessary if we are to move towards an efficient health care system. Efficiency implies that objectives have been met using minimum resources. In health care, the objective is to deliver the right health care, at the right time, to the right people. Trying to achieve this end is inevitably costly, since imperfect information makes it costly to co-ordinate activity and to manage conflicting interests, whatever the governance structure. The important question is whether these costs are worth incurring: in other words, whether the organisational costs incurred to increase efficiency are less than the benefits of that increase.

This argument can be illustrated in the context of the case study referred to above, in Box 3.1. A health authority wanted to increase the number of hip replacements and decrease the number of complex orthopaedic operations carried out by acute trusts, since they believed this would reflect local needs. They also wanted to ensure that surgeons changed the type of prosthesis used in hip replacements, given the evidence that an alternative was more cost-effective. The health authority faced two alternative courses of action:

25 Department of Health (1994), NHS Executive (1997c).

1. incurring the ex ante transaction costs of agreeing a detailed contract, with hip replacements specified and priced separately from other orthopaedic surgery, as well as the ex post transaction costs of ensuring that the trust does actually carry out the specified activity;

2. avoiding the transaction costs, either by specifying a contract based solely on orthopaedic surgery as a whole or by not enforcing a more detailed contract. Following the latter course will certainly carry lower transaction costs, but may leave the authority with no way of insisting that only cost-effective procedures are used or that surgeons carry out more hip replacements.

In selecting a course of action the health authority should compare the transaction costs of affecting trust activity with the benefits to the local population of doing so. Benefits to the local population are conceptualised in economics by the concept of social welfare. Inefficient production is said to involve a welfare loss, since existing resources had the potential to generate benefits that were not in fact realised.

Both the level of social welfare and its distribution between individuals or groups are important. A gain in welfare by one group may be achieved at the expense of another, and may be deliberately sought. If so, organisational costs may be incurred by a vulnerable group to protect themselves. In the case study referred to above, managers and clinicians believed that orthopaedic surgeons prefer undertaking complex surgery to routine hip replacements, sometimes wanting to follow their research interests, and sometimes to gain professional status. In this case the transaction costs referred to above can be analysed as organisational costs incurred by the health authority on behalf of the local population, to prevent clinicians gaining welfare at their expense.

Finally, it is important to remember that health care is delivered along a complex chain, with responsibility delegated between a number of different individuals and organisations. Taxpayers delegate to government responsibility for resource allocation, and it in turn delegates responsibility to NHS organisations. Health authorities are ulti-

mately responsible for representing citizens' interests both in the sense of consulting them and also in the sense of ensuring that providers actually deliver desired health care and health gains.

Part of this responsibility, re-emphasised in recently published White Papers, is public consultation to determine and co-ordinate public opinion. Public consultation is certainly costly. However, the welfare loss of not consulting the public is also likely to be high, and manifest as popular disaffection with the NHS. Moreover, if proper public consultation is not undertaken, special interest lobby groups may have a disproportionate influence on service provision, gaining welfare at the expense of the general public.

3.5 Firms are like markets and markets are like firms...

One way of reducing organisational costs is to alter the governance structure, as discussed in section 3.1. Empirical analysis of industrial organisation shows that firms do alter governance, either by adopting a completely new structure or by introducing new, efficiency-enhancing, arrangements within existing governance structures. This means that organisations sometimes borrow features stereotypically associated with markets, and markets sometimes borrow those associated with organisations.

This phenomenon is also observable in health care, where there is a continuum of different forms of governance, with most activity governed in arrangements borrowing aspects of both organisations and markets. Instead of only observing competition and short-term relationships in market style governance between organisations, we also observe long term co-operative relationships; and instead of only observing centralised direction, long-term relationships and direct monitoring within organisations, we also observe decentralised decisions and the use of incentives. These features may be efficiency-enhancing, insofar as they reduce organisational and / or production costs.

Organisations commonly borrow features from markets, to enhance

their efficiency and reduce organisational costs. Most organisations decentralise at least some decisions, to lower the costs both of direct monitoring and of transmitting information between the centre and periphery. Moreover, most try to motivate employees by using at least some high powered incentives, such as performance related pay which ties remuneration to productivity, and some introduce explicit competition between employees or departments to motivate individuals. These features are evident within NHS institutions. Acute trusts, for example, function as collections of specialty teams – which have even been called 'firms' since long before 1990 – and many now devolve budgets to hospital wards or to clinical directorates. Some health authorities try to motivate individuals by tying a proportion of their income to their performance, through an individually-negotiated performance-monitoring system.

Conversely, markets are like organisations in the sense that many markets – including the NHS internal market – are characterised by long-term relationships. In discussions involving health services, the practices of retailer Marks and Spencer are often used to illustrate the advantages of such relationships[26]. Marks and Spencer has long-term relationships with its suppliers, permitting it to capitalise on the advantages of external supply (it can make a credible threat to go elsewhere) without losing the transaction cost-reducing advantages of long-term relationships. This type of informal vertical integration has certainly been observed in the NHS internal market, where purchasers and providers often have long-term, co-operative relationships[27]. The desirability of long-term relationships characterised by co-ordinated planning and co-operation has similarly been noted by a number of commentators[28], and indeed they are now being actively encouraged in the NHS by the Labour government[29].

This type of relationship – co-operative but with activity governed by separate organisations – offers the possibility of lower transaction costs without completely sacrificing the incentive structure operating

26 Macara (1994).
27 Macara (1994); Redmayne (1995).
28 Allen (1995); Goddard and Mannion (1998); Maynard (1994); Nippert (1992).
29 See section 4, below.

in markets. Transaction costs are reduced, since co-operation lowers the likelihood of opportunism, repetition lowers the cost of information gathering, and the expectation that a contract will be continued may reduce the transaction-cost related risks of long-term specialised investment. Retaining a market interface, in other words continuing to organise activity in separate organisations, nonetheless allows each party to retain their mobility and therefore protects the incentive structure: if there are alternatives, each can make a credible threat to leave the relationship. It also allows each organisation to specialise, which may reduce production costs and promote innovation.

It is important to note that there are circumstances, however, when long-term co-operative relationships may not be desirable[30]. In industry, competition law exists to protect consumers from some of the undesirable effects of co-operation between firms. It used to be the case, for example, that co-operation between high street banks benefited the banks themselves but disadvantaged customers, by enabling banks to keep operating hours short and charges high.

Analogous issues arise in health care, where patients and other citizens might similarly be vulnerable to the effect of so-called co-operation between NHS organisations. Long-term co-operative relationships between, for example, acute trusts might benefit patients if it means information and resources are pooled. It might, however, disadvantage patients if it increases prices and reduces the level or quality of care.

Co-operation between purchasers and providers might also disadvantage the users of health care. It was argued above that health authorities are citizens' agents, entrusted with representing our interests when negotiating and monitoring agreements with health care providers. This agency relationship incurs costs: either health authorities must be monitored (generating transaction costs), or we as citizens

30 Even if there is no deliberate collusion, long-term relationships may not be efficiency promoting if purchasers become locked into them, losing the leverage necessary to make a credible threat to leave the relationship, and therefore becoming less able to influence provider behaviour.

might have to bear the risk of the welfare-loss of having health services that do not meet our desires[31].

3.6 Measuring organisational costs

It is obviously desirable to find some method of quantifying organisational costs and relating them to outcomes. A growing body of empirical research seeks to operationalise transaction costs in the context of manufacturing. In the context of services in general, and health care in particular, this research is however in its infancy.

In general and as will be discussed below in sections 3.6.1 - 3.6.2, two methods can be used to measure organisational costs:

1. direct counting of relevant activities, such as the amount of time spent in meetings or the number of managers employed;

2. indirect measurement of organisational costs. This type of work begins with a prediction about the level of costs likely to exist in particular environments, and then tests that prediction by observing the governance structures prevailing in different environments.

3.6.1 Measuring organisational costs by counting them directly

Various estimates have been made of organisational costs incurred by NHS organisations. Trusts and health authorities are required to publish their management costs, calculated using methods specified by the Audit Commission[32]. Trust management costs are based on the cost of people undertaking management or administrative tasks[33]. Senior management, called M1, comprises expenditure on managers with annual salaries exceeding £20,000, on senior nurses primarily undertaking management tasks and on management consultants. This was

31 Citizens' interests are safeguarded by various formal and informal arrangements. The NHS Executive undertakes direct monitoring, governed by principles outlined in a guidance on competition policy (The operation of the internal market, 1994). The risk of deleterious collusion is also mitigated by ethical standards, which are relied on to constrain self-interested behaviour in the NHS.

32 NHS Executive (1995).

33 Audit Commission (1995).

estimated at £1,275m in 1995/96, falling to £1,225m in 1997, about 5 per cent of trusts' total income[34]. Additional categories, M2 and M3, comprise all administrative and clerical staff as well as other senior nursing staff, and were designed to ensure that any change in M1 was not solely the result of a change in job title. A comment on the significance of these categories is made in Box 3.7.

Health authority management costs are calculated using a different method. They are defined as those costs remaining after excluding health authority expenditure on a specified range of items, including health care purchased from other organisations, primary care, public health and health promotion activities carried out within health authorities, and so on. The estimated level of management costs in English health authorities was £497m in 1995/96, falling to £450m in 1996/97, equivalent to about £9 per person living in England[35].

In addition to the management costs incurred by health authorities and trusts, management costs are incurred by GPs who all receive management allowances, with special allowances granted to fundholders and total purchasers. In addition, there are a number of one-off man-

Box 3.7: M1, M2, M3...

There is irony in the choice of terminology, in using M1, M2, and M3 to connote different types and levels of management cost, since it is similar to the way economists categorise money supply. In macroeconomics, M0, M1, and so on connote different ways of measuring money supply. Differences between money supply M0 and M1 are substantive, since economic theory suggests that different types of money are subject to different influences and have differing effects on the economy. By contrast, there is no clear theoretical foundation to the different measures of management cost, M1 - M3, rendering them no more than accounting entities.

34 NHS Executive (1997a), NHS Executive (1997b).
35 NHS Executive (1997a), NHS Executive (1997b).

agement-related grants, such as the extra payment of £165m allocated to support fundholding in 1995[36] and the sum of £150m allocated to develop information technology in primary care at the end of 1997[37].

A comprehensive, rigorous survey of these costs is a research project in itself, and is not attempted here. The difficulties in compiling a reliable estimate of official management costs can be illustrated by comparing the level of health authority costs presented in two government sources. In 1995/96 the NHS Executive published as total English health authority management costs the sum referred to above, £497m. In the same year, the official accounts of health authorities and regional offices of the NHS Executive included a sum for 'administration and other services' of £2,252m[38]. The two are obviously calculated from different bases, the difference between them illustrating the difficulties of deriving a single and comprehensive estimate of management costs in the NHS using official estimates.

Not only are official estimates of management costs complex, but they also probably understate actual organisational costs. There is significant political pressure on NHS organisations to minimise expenditure in this area, pressure which may lead to deliberate under-reporting of management-related costs. Moreover, even if accurately completed, the official categories do not capture all relevant costs. Two recent studies show that management allowances paid to GP purchasers understate the actual transaction costs of GP-based purchasing, since they do not capture all of the organisational costs incurred when GPs become purchasers, notably the time that primary care professionals devote to purchasing-related activities[39].

An estimate of the total cost of organising health care in the NHS does not, therefore, exist. If it did, it would include all the transaction and internal costs incurred in the delivery of health care. This type of exercise has been carried out in another context by Wallis and North (1986). They estimate that, in 1970, organisational costs across all sectors

36 Petchey (1995).
37 Milburn press release, (14 December 1997).
38 Department of Health (1998).
39 Petchey (1995), Posnett et al. (1998).

of the economy comprised 47-55 per cent of total US Gross Domestic Product, in other words that organisational costs absorbed about half of all money spent in the US economy. Their exercise is salient for the NHS, since it shows the importance of organisational costs: indeed, Wallis and North undertook the exercise to show that the efficiency gains accompanying modern, specialised economies are very costly.

As argued above, organisational costs are not meaningful in isolation, but must be accompanied by a measure of benefit. There is some work seeking to evaluate the relationship between organisational costs and efficiency in the NHS internal market. Bevan analyses changes in activity and total NHS expenditure, and concludes that the increase in expenditure since the introduction of the internal market has not been accompanied by improved productivity[40]. There are two empirical studies comparing the costs and efficiency gains of GP purchasing. The Audit Commission finds that, from its inception to the end of 1994/95, fundholding incurred additional organisational costs of £232m but generated efficiency savings of only £206m[41]. A team at the University of York evaluated total purchasing pilots and showed that, after controlling for the size of the pilot sites, those with higher organisational costs were more likely to meet their own purchasing-related objectives[42]. These studies are not conclusive, but could be the genesis of a body of work providing rigorous analysis of the relationship between the costs and benefits accompanying the type of incentive structure underlying the NHS internal market.

3.6.2 Indirect measurement of organisational costs

It is, in practice, difficult to identify transaction costs, to distinguish them from production costs, and to determine whether they have been minimised. This means that most of the empirical literature does not try to identify transaction costs directly, but rather seeks to track their footprints, by analysing the relationship between transaction costs and governance structure. Stated crudely, this literature generally posits

40 Bevan (1998).
41 Audit Commission (1996) p.7.
42 Posnett et al. (1998).

that, in competitive markets, firms facing high transaction costs will, if all else is equal, be more likely to be vertically integrated than firms facing lower transaction costs. Most empirical work has been carried out in the context of manufacturing rather than in services, although there is a growing marketing literature which may be more useful when analysing health care, since both contexts face the difficulty of distinguishing production and transaction costs[43].

There are very few health-related empirical studies in this literature. One US study shows that Health Maintenance Organisations' (HMOs) choice of governance structure for mental health care services reflects the relative organisational costs of different forms[44]. Ashton analyses the nature of contracts used to govern the provision of different types of health care in New Zealand and finds that contracts vary as predicted by institutional economics: care characterised by high levels of uncertainty and intangible outcomes is governed by block contracts; and care that can be easily specified in advance is governed by cost per case contracts[45].

This type of analysis provides useful insights into the characteristics of governance structures. It does, however, need to be accompanied by some direct measure of transaction costs, as indeed Ashton does, if it is to avoid the tautology: how do we know there are high transaction costs? Because of the nature of the observed governance structure. Why do we have this particular governance structure? Because there are high transaction costs[46].

The science of measuring the transaction and internal costs of health care delivery structures, and comparing these with production costs and outcomes, is in its infancy. This makes precise analysis of the likely impact of policy designed to reduce 'bureaucracy' difficult. The next section is, therefore, necessarily qualitative, predicting the direction of the effect of recent government policy even if not its precise magnitude.

43 Maher (1997) is an excellent example of the literature in this genre, comparing governance structures in case studies drawn from four different UK industries. Shelanski and Klein (1995) provide a useful survey of US literature.

44 Wholey et al. (1996).

45 Ashton (1998).

46 Ashton avoids the tautology by using semi-structured interviews to confirm that the choice of contract-type was in fact influenced by transaction costs.

4 ORGANISATIONAL COSTS IN THE NHS WHITE PAPERS

In 1997 and 1998 the government published four documents outlining changes to the NHS in Scotland, Northern Ireland, Wales and England[47]. This section uses the changes outlined in these White Papers as a case study, to show how the economic framework outlined above can be used to analyse the nature of organisational costs in the NHS. The section begins with a general overview of the proposed changes, before analysing their impact on organisational costs in subsequent sections.

4.1 The structures proposed in the White Papers

4.1.1 Organisations in the new structures

The main organisational components of the proposed structures are national organisations, health authorities (called health boards in Scotland, and health and social services boards in Northern Ireland), associations of primary care professionals, and NHS trusts.

National organisations are to be responsible for 'driving' performance, upholding quality standards and disseminating information within the NHS. These include existing organisations as well as some new ones. In England, for example, a new National Institute for Clinical Excellence and a new Commission for Health Improvement are being established, and in Scotland a new Scottish Health Technology Assessment centre is to advise the NHS on the cost-effectiveness of all innovations. In Wales, the nature of the national organisations is less well specified, reflecting changes that might occur following the establishment in 1999 of a new National Assembly for Wales. Among its responsibilities, the National Assembly will assume responsibility for health policy and monitoring Welsh NHS organisations, as will the new Scottish Parliament. The Northern Ireland doc-

47 Designed to care: renewing the National Health Service in Scotland (1997); Fit for the future (1998); NHS Wales: putting patients first (1998); The new NHS: modern dependable (1997). Unlike those published in the other countries, the Northern Ireland document, Fit for the Future, does not have the status of a White Paper, but is a consultation document specifying options and a consultation process. For simplicity, however, it is included in the discussions of the White Papers below.

ument states that it is likely to adopt similar institutions to those outlined in the English White Paper.

Health authorities in England and Wales and their equivalents in Scotland and Northern Ireland, are responsible for liaising with other organisations and consulting the public when making policy, and for monitoring the performance of other local NHS organisations. Health authorities' role as direct commissioners (as NHS purchasers are now called) of health care is supposed to be devolved over time to the associations of primary care professionals, although details and extent vary. In England, health authorities have been given statutory responsibility for improving the health of the local population. Northern Ireland differs fundamentally from the other countries, since health and social services are currently integrated within organisations. This integration will continue, but the consultation document suggests that the overall structure governing commissioning and providing will change.

Associations of primary care professionals are taking different forms in each country, but in all there is an emphasis on organisational integration of all professionals delivering care in the community, and on gradual devolution to them of responsibility for commissioning services on behalf of their patients. GPs retain their status as independent contractors.

- in England primary care groups (PCGs) combine existing general practices and community nursing services in areas with populations of about 100,000. A detailed set of stages is given through which PCGs may evolve, with the final stage being merger with local community trusts to form integrated primary care trusts providing primary and community care, and commissioning all types of health care;
- in Scotland, local health care co-operatives are being formed in 'natural communities' from networks of general practices. The role of nurse practitioners in primary care is emphasised in the Scottish White Paper, but it is not clear whether they will be part of the local health care co-operatives in Scotland, as is the case in England where they are part of PCGs, or part of Scottish

primary care trusts, which are discussed below;

- in Wales, local health groups similar to English PCGs are being formed, based on local authority areas;
- in Northern Ireland, various alternatives are under review.

NHS trusts continue to exist as the organisations responsible for providing health care:

- in England, mergers between acute and community trusts are discouraged. As noted above, primary care trusts will eventually be formed from mergers between PCGs and community trusts. Mergers between acute trusts will be allowed if they improve health care and reduce administrative costs;
- in Scotland, by contrast, primary care trusts will be formed by merging community hospitals, mental health services, and local health care co-operatives. Each primary care trust will comprise a number of local health care co-operatives. The number of acute NHS trusts is being reduced by merging existing trusts so that in most areas there is a one-to-one match between health boards and acute trusts;
- the number and configuration of NHS trusts in Wales is to change, to redress the current configuration, which is described in the Welsh White Paper as 'haphazard and not well placed to deliver effective health care in the most efficient manner'[48];
- the configuration in Northern Ireland may also change, although the emphasis on integrating the provision of health and social care will continue.

4.1.2 Governing relationships between organisations in the new structure

Relationships between organisations will be governed by a new set of formal mechanisms, which are discussed in turn below: national guidance is intended to form a 'virtual' link between clinicians; health

48 Putting patients first (1998). Para 6.13.

improvement plans will link local organisations; accountability agreements will link health authorities, NHS trusts, and primary care associations; Scottish primary and secondary care organisations will be linked by new Joint Investment Funds; and actual communication will be facilitated by new information systems.

Nationally circulated guidance, will be used to inform clinical practice within each country, encouraging evidence-based practice and linking clinicians across organisational boundaries. The tenor of measures set up in England and Wales is similar to that of existing institutions in Scotland, where clinical guidelines are already being formulated by national organisations and circulated through the Scottish Intercollegiate Guidelines Network. In England and Wales, the White Papers indicate that a set of national service frameworks will be developed and circulated by the National Institute for Clinical Excellence. National institutions will also promote quality assurance. In Scotland, a review of quality assurance is under way.

Health improvement programmes (HImPs) will be introduced throughout the UK. Health authorities and boards will have responsibility for ensuring that all interested parties, both health and social care organisations, are represented when these plans are developed. The programmes will provide strategic direction for services over the subsequent three to five years in England and five years in Scotland. HImPs will be implemented through various accountability agreements.

Accountability agreements are outlined in each White Paper. In England, annual accountability agreements will be agreed between health authorities and each PCG, using targets specified in the local HImP. Service agreements will be agreed once every three years between PCGs and trusts, and between health authorities and trusts for those services that health authorities continue to commission. Similar arrangements are outlined for Wales. In Scotland, annual trust implementation plans (TIPs) will be agreed between health boards and individual trusts, with an annual accountability review between the trust Chief Executive and representatives of the health board.

Joint Investment Funds will be established in each health board in Scotland, to govern relationships between primary and secondary care. The size of the fund and the purposes for which it will be used will follow from the HImP, but in general it is to be used to support services at the interface between primary and secondary care.

Linking organisations using information technology. The Scottish White Paper contains the most developed plans for using information technology 'to promote a seamless pattern of care'[49], but a detailed information strategy for England was published separately, in August 1998[50]. All countries intend developing a computer network linking NHS organisations, the use of telemedicine, and the adoption of a single patient number.

4.1.3 Principles underlying the new structures

Relationships will also be governed by a new set of principles: clinical governance; co-operation and partnership; and openness[51].

Clinical governance is a principle which will be manifest in two ways: first, in agreements between providers and commissioners, which are now to focus on evidence-based commissioning and use performance measures which promote quality and user-centred health care; and second, in greater use of guidelines to govern how clinicians provide health care. Clinical governance will be enforced by a new statutory duty for quality in NHS Trusts[52].

Co-operation and partnership are emphasised in all of the White Papers, and are presented as the antithesis of the competitive ethos supposed to underlie relationships in the internal market. Health authorities are given responsibility for ensuring that all relevant local organisations are involved in agreeing HImPs and for promoting inter-agency working in service provision, particularly between the NHS

49 Designed to care (1997), Para 19.

50 Information for health (1998).

51 I have identified the principles after analysing the White Papers, and they relate specifically to relationships. They should not be confused with the six guiding principles set out in the White Papers.

52 The new NHS (1997), p.18.

and local authorities. It is also emphasised that trusts and primary care associations must work 'in partnership' with each other and other agencies. In Northern Ireland, health and social care are already integrated in a number of cases, and organisations are enjoined to extend this to co-operation with other sectors.

Openness is required, particularly of trusts who are required to 'end the secrecy,' said to accompany competition in the internal market. Specific provisions include the requirement that they hold open board meetings and allow information about their performance, including costs, to be published. The extent to which these obligations are honestly implemented will of course depend on the extent to which competition and conflicting interests are truly abolished.

4.1.4 The cultural change implied by the new structures

The English White Paper is designed to provide a blueprint for the NHS that is consistent with the government's avowed intention to promote a 'third way' of organising society: an alternative to 'command and control' on the one hand and markets on the other. In the context of health care, they want to promote an alternative to the pre-1990 structure, which is said to have 'stifled innovation', and the post-1990 internal market, with its attendant flaws. The White Paper is designed to show how the internal market can be replaced with a system of integrated care, 'based on partnership and driven by performance'[53].

Taken as a whole, the new structure seems designed to facilitate cultural change in the NHS: in other words, to influence the informal rules governing both the relationships between organisations and the way care is currently provided. All of the White Papers refer repeatedly to the need to 'replace competition with co-operation'. Cultural change, existing as it does within individuals, cannot be achieved directly by the formal means open to government. It can, however, be influenced by changes in formal rules. By making inter-agency work-

53 The new NHS (1997), p.18.

ing mandatory, a number of measures announced in the White Papers seek to influence the quality of NHS relationships. These measures include giving English health authorities a statutory duty to promote inter-agency working; the introduction of HImPs to provide a 'shared context' within which co-operation can develop; setting up joint investment funds in Scotland and merging Scottish NHS trusts to eliminate the incentive for competition; and replacing the current annual agreements (as NHS contracts are now called) with triennial agreements as a means of providing a stable framework for developing long-term relationships. The government wishes to facilitate co-operation as an end in itself, to increase staff morale, and also as a means of reducing the organisational costs of providing health care. This is vital since, as will be discussed in the next section, the structures set out in the White Papers are on their own insufficient to produce the desired reductions in organisational costs.

4.2 A billion pound reduction in bureaucracy?

The Labour government has pledged to reduce what it calls 'bureaucracy' in the NHS by one billion pounds during its current term in office. This pledge is reiterated in the White Papers. The Scottish White Paper states that 'one of the adverse features of the internal market was the scale of the bureaucracy and the associated costs to which it gave rise', and that the measures outlined in the paper should 'lead to management savings of around £100m over the lifetime of the Parliament'[54]. The Welsh White Paper pledges that abolishing the internal market will save £10m a year by 'cutting ... bureaucracy', and that in addition NHS trusts will be reconfigured, releasing between £5m and £10m, again by reducing bureaucracy[55]. The English White Paper restates the government's pledge to reduce bureaucracy by £1bn[56]. The Northern Ireland document pledges a reduction of

54 Designed to care (1997). Paras 47, 53.
55 Putting patients first (1998). Para 2.22, 6.5.
56 The new NHS (1997). p.74.

£25m over the life of the government, stating that 'the internal market has driven up administration costs', which should instead be 'released... for better services for patients and clients'[57].

These reductions in organisational and transaction costs are supposed to be achieved by reducing the number of organisations involved in delivering health care, by abolishing contracts and ECRs[58], by extending the period of time covered by agreements between NHS organisations, and by replacing competition with co-operation. However, as will be discussed in sections 4.2.1 - 4.2.6, these savings are ephemeral, and the structures proposed in the White Papers are likely to generate additional organisational costs.

4.2.1 The impact of reducing the number of organisations

All of the NHS White Papers state that there should be fewer organisations operating in the NHS. As stated in the English White Paper:

> 'In recent years effort and resources have been diverted from improving patient services. With so many players on the field, transaction costs in the NHS inevitably spiralled'[59].

The number of English commissioning bodies is to be reduced from 3,600 to 500, to be achieved primarily by merging different GP practices into PCGs. In Scotland, as well as emphasising the role of GP co-operatives, the number of organisations will be reduced through mergers between acute trusts, leaving only one acute trust in most health boards. In all countries, reductions will also be achieved by mergers between community and primary health care providers, forming primary care trusts.

The White Papers state that reducing the number of organisations will reduce transaction costs by reducing the number of contracts needing to be made. Although the papers attribute the costs of the internal market primarily to too many commissioners, the number of

57 Fit for the future (1998). Para 3.1.
58 Extra Contractual Referrals (ECRs), defined in Section 2, above.
59 The new NHS (1997). p.14.

agreements and contracting parties is not the sole or even the greatest component of transaction costs. Reducing the number of contracting parties may actually increase inter-organisational transaction costs. As was argued in section 3.2 above, a reduction in the number of contracting parties locks commissioners and providers into specific relationships which may lead to negotiating impasse if the parties try to 'hold-up' each other. Whether this problem will arise in England depends partly on whether bilateral monopolies are created when GPs are grouped into PCGs. They will certainly arise in Scotland, between acute trusts and health boards.

Reducing the number of organisations is also supposed to reduce organisational costs by allowing economies of scale. All of the White Papers state that combining organisations will eliminate duplication of support services and allow management overheads to be shared. None, however, recognises that the impact of mergers on within-organisation costs depends on what might be called the 'managerial production function', discussed further in Box 4.1. In other words, the relative efficiency of large and small organisations cannot be predicted in advance, but depends on the characteristics of the environment and the type of health care provided. As recognised by Coase in his seminal paper, the marginal costs of organising activity within an organisation are unlikely to decrease continuously as the organisation increases in size[60]. This is borne out by a number of studies which have found no consistent relationship between the size of a trust or a primary care organisation and its management costs[61].

The costs of aligning incentives within organisations are likely to be substantial. Although reducing conflict is one of the advantages often associated with substituting internal for external governance, it is not achieved automatically simply by merging organisations. It will occur only if the interests of a new organisation's constituent parts are effectively aligned, which may require the introduction of appropriate

60 Coase (1937).
61 Audit Commission (1995), Mays et al. (1997), Posnett et al. (1998).

institutional arrangements, either those supporting centralised decision-making or those which ensure appropriate distribution of benefits.

It is possible that internal costs may be less than transaction costs, but the arrangements necessary to minimise internal costs are nowhere addressed in the White Papers. On the contrary, they imply that internalisation is itself sufficient to eliminate conflict and co-ordinate activity. The Scottish White Paper, for example, argues that removing organisational boundaries between acute trusts will generate 'collective ownership' and the development of 'mutually supportive objectives and actions'. This is supposed to facilitate rationalisation of local services, since it means that loss of services will no longer 'threaten the

Box 4.1: Economies of scale

Economies of scale exist in health care when the average cost of treating each patient decreases as the number of patients increases. Imagine, for example, a hospital where the average cost of treating each of 200 patients is lower than the average cost of treating each of 100 patients. In this case, there are economies of scale. Commentators sometimes assume that this will automatically be the case: that bigger is always cheaper. On the contrary, a recent literature review showed that economies of scale are by no means as common in health care costs as is often assumed (Ferguson, Posnett and Sheldon 1997).

In the context of management costs, economies of scale would exist if, for example, a computer system costing £100,000 could support up to 500 beds, but is currently used in a trust with only 200 beds. In this case, merger between this trust and another with under 300 beds will release economies of scale by spreading the overhead costs. Adding a third trust will not, however, realise additional savings: in fact, there may be diseconomies of scale if additional computer capacity is not purchased.

viability of a trust' as it did in the internal market[62]. Rationalisation may, however, still generate conflict within organisations if it threatens the employment of some staff groups, or comes up against the interests of professional groups.

In England it is hoped that co-operation between different organisations, sometimes called 'inter-agency partnership' in the White Papers, will generate both managerial economies of scale and collective ownership of local services. The paper states that individual organisations, including NHS trusts, must be willing to see services move to other organisations, yet the paper does not recognise any barriers to this type of mobility.

The costs of aligning incentives (i.e. ensuring common interests and co-operative behaviour) within local associations of primary care professionals are similarly overlooked in the White Papers. Local associations merely amalgamate what are currently separate GP practices and community services. No institutional arrangements are outlined which might help blur the boundaries between current organisations, suggesting that there will be internal divisions in many local associations. This will be exacerbated by the way the groups have been formed, following geographic boundaries rather than bringing together groups of 'like-minded' primary care professionals[63]. Ceding authority to a centralised decision-maker is one way that organisations commonly resolve the problem of internal boundaries and conflict of interest. The White Papers give no impetus to using centralisation in this way, as a way of resolving disputes between practices within local associations. On the contrary, they state that within local associations different practices will retain their own identity, with separate indicative budgets. This might mean that gains to the patients of one GP within a local association are losses to those of another, and will generate internal, possibly opportunistic, bargaining and the attendant costs.

62 Designed to care (1997). Para 100.
63 NHS Executive, (1998).

Opportunism between the parts of an organisation can be mitigated by aligning the interests of the parts with that of the organisation as a whole. One way of achieving this kind of alignment is to make individuals' employment dependent on their organisation's survival. In this case they are said to have a 'high powered incentive' to perform in the organisation's interest rather than in their own. However, as is emphasised several times in the White Papers, GPs will retain their status as independent contractors. They may therefore have little personal stake in the performance of the association as a whole. The people whose employment depends on their local association's success – other primary care professionals – will, however, probably have little authority.

4.2.2 Eliminating contracts

The White Papers' rhetoric consistently connects eliminating contracting with eliminating bureaucracy. 'Contracts' are to be eliminated, with relationships between organisations instead governed by HImPs and a variety of 'agreements'[64]. HImPs are to provide a blueprint for local health care and social services, and are to be used as the basis for setting targets in more detailed agreements between commissioners and providers. They will, therefore, need to include carefully specified plans if they are to be reflected in local services. Achieving the benefits associated with HImPs in the White Papers will require substantial resources, for the following reasons.

Health authorities are required to co-ordinate the process for agreeing, and then monitoring, HImPs. The authorities are required to involve a number of different organisations in agreeing local plans. They are also required to ensure that the local population is consulted. The costs of co-ordinating this process will be high, exacerbated by differences in terminology and approach of people from different professions and sectors.

64 Health Improvement Plans (HImPs) have been defined in Section 4.1, above.

The process of agreeing HImPs is supposed to demonstrate that 'we are all on the same side'. If it is indeed the case that there are no conflicts of interest, developing HImPs should be no more than an exercise in co-ordination and communication, both of which are nonetheless costly. If, however, there is a conflict of interest between different organisations, then the process of developing HImPs will involve bargaining and the attendant organisational costs. And there will be conflict, since a meaningful HImP will involve resource allocation decisions relating to alternative uses and different groups. The White Papers, therefore, enjoin co-operation and indicate that health authorities will have statutory responsibility for achieving partnership, without fully removing a fundamental source of conflict – competition for resources and the concomitant job security.

4.2.3 Long-term agreements

HImPs are presented as arrangements which will eliminate the bureaucracy associated with contracting, and long-term agreements are presented as solutions to the current bureaucracy associated with making annual contracts or agreements.

The arrangements as outlined in the White Papers suggest that there will, however, be little actual change in the period covered by agreements. In the internal market, many arrangements were effectively governed by long-term, informal arrangements. Conversely, although the White Papers frequently refer to introducing long term agreements, HImPs are in fact to be accompanied by a new set of annual agreements. Some of the new annual agreements, such as those between health authorities and PCGs in England, will require negotiation as well as monitoring; others, such as those in Scotland between health boards and NHS trusts, are primarily arrangements allowing commissioners to monitor trusts. It is not, therefore, clear how organisational costs will actually change.

4.2.4 Introducing quality-based agreements

The White Papers criticise contracts made in the internal market for focusing on the volume of activity rather than its quality. The White Papers state that contracts in the internal market were based solely on cost and activity, and that this created financial incentives which in turn distorted behaviour. The Welsh White Paper, for example, states that:

> 'Resources have been diverted from patient care into a bureaucratic process that has had little to do with the quality of care provided for patients. That process has been inefficient'[65].

This is to be rectified in the post-White Paper NHS by the use of quality-based indicators in agreements. Only in the Scottish White Paper are the costs of monitoring quality recognised, since it refers to the 'great deal of time and effort in monitoring the quality of service provision to ensure and improve standards of care'[66]. The other White Papers do not, however, recognise the impact of quality indicators on transaction costs. The ex ante transaction costs of finding acceptable trading partners and agreeing terms will be higher than if contracts were based only on activity, although this may be mitigated by model agreements and universal performance measures to be developed by the NHS Executive in England. Ashton interviewed contracting managers to try to determine the time devoted to contract negotiations, and found that the time increased as contracts became more complex:

> 'One manager of mental health services reported that a total of 1,166 hours of time had been dedicated to contract negotiations with the purchaser over the period of one year. Moreover, the time commitment was increasing, rather than decreasing, as specification of services was becoming more detailed'[67] .

The more substantial increase in transaction costs is, however, likely to occur ex post when agreements are monitored and enforced.

65 Putting patients first (1998). Para 1.9.
66 Designed to care (1997). Para 118.
67 Ashton (1998) p.363.

Agreements can be enforced only if both parties agree that a violation has in fact occurred, which requires that the agreement be made over measurable indicators. The importance of measurability is recognised in the White Papers, but not the cost of achieving it. The costs of enforcing quality-based agreements are high in health care, where an adverse outcome might result either from negligence or an 'act of god'. Although it is hard to distinguish between the two and even to measure consistently the quality of outcomes, unless poor quality can easily be attributed to negligence, a quality-based indicator is costly to enforce.

The requirement that agreements incorporate quality indicators may encourage mergers between PCGs and community trusts, in other words, the formation of primary care trusts. Measuring activity and outcomes in community care is costly given the complexity of service provision, involving as it does a number of different individuals and organisations, all providing health and social care, often in a patient's home, away from the controlled (and observable) environment available in hospitals. Moreover, the outcome is often intangible and occurs incrementally over a long period of time. In this case, the transaction costs of writing, monitoring and enforcing agreements between PCGs and community trusts may be higher than the internal costs of the merged primary care trusts, since within an organisation different types of activity do not have to be fully specified in advance.

4.2.5 Reducing invoices and paperwork by abolishing the internal market

Both ECRs and fundholding have been abolished with the abolition of the internal market. As discussed in section 2, above, ECRs referred to patients treated by a provider with whom a purchaser had no contract[68]. They generated both paperwork, in invoices and payments, as well as conflict over whether purchasers should pay for that care. ECRs are to be replaced by 'out of area treatments' (OATs) described in the

68 They also comprise people cared for at providers with whom the commissioner has a contract, but who received a type of care not covered in the contract.

White Papers as a 'simplified system' for meeting costs of patients treated away from home. This should, by definition reduce paperwork and therefore direct organisational costs, but it will not entirely eliminate these costs.

Part of the rationale for abolishing fundholding is that it is a costly system to run. The costs of fundholding are illustrated in the English White Paper by referring to the number of fundholding-related invoices handled in one health authority (60,000), one trust (40,000) and one general practice (1,000)[69]. A reduction in the number of fundholding-related invoices has been sought by the NHS Executive for some time. In 1997, for example, the NHS Executive issued instructions that providers should use multi-case invoices, that fundholders should not check every invoice sent to them, and that fundholders should organise regular monthly payments to providers and engage only in post-payment reconciliation[70]. This has already been implemented by a number of trusts.

Although fundholding undoubtedly generates organisational costs, the evidence is equivocal about whether it is, in itself, the main source of management costs in primary care. Sutherland and Cooper argue that the introduction of a new self-employment contract with GPs in 1990 has generated more administration, and Whynes et al. find that non-fundholders use more management time than fundholders[71]. Neither of these is a controlled study, but they do suggest we should be cautious before condemning fundholding per se as a source of unnecessary bureaucracy.

Moreover, ECR invoices and GP fundholding-related invoices related to episodes of health care. Eliminating ECRs and GP fundholding would not therefore eliminate invoices unless accompanied by a change in the system, as is proposed in the White Papers. Under the internal market, some form of invoicing and monitoring was a necessary part of the incentive structure, based as it was on the twin princi-

69 The new NHS (1997), p.14.
70 NHS Executive (1997c).
71 Sutherland and Cooper (1992), Whynes et al. (1995).

ples of money following patients and of budget-holding commissioners. If money is to follow patients there need to be invoices. And if money is following patients, rational budget holders need a system for monitoring how it is used and for guarding against opportunism. Although we have no evidence about the extent of opportunism in relation to ECRs and GP fundholding, fundholders give anecdotes of opportunistic behaviour by NHS trusts in the internal market. It is germane that in the US, the Government Accounting Office estimates that 10 per cent of insurance claims are fraudulent[72]. The fundamental point is that in a system with incentives such as those operating in the internal market, invoicing and monitoring may be part of a rational attempt to minimise the economic costs. In the internal market, paper followed patients and money followed paper.

The impact on economic costs of abolishing ECRs and GP fundholding will depend on the structure that replaces the internal market. As has been argued throughout this paper, we do not face a choice between on the one hand a plethora of invoices and on the other, costless delivery of the same health care. Rather, the valid choice is between two systems, with different types of economic cost. In the internal market, the resources devoted to monitoring ECRs and fundholding-related invoices may have prevented welfare loss; in the new, post-White Papers NHS the extent of welfare loss will depend on whether exhortations to co-operate are an effective way of preventing opportunism. The White Papers note the importance of vigilance against fraud, yet none recognises the costs of achieving effective vigilance.

72 Feldbaum and Hughesman (1993).

5 CONCLUDING COMMENTS

Will 'bearing down on bureaucracy' release more than £1bn for patient care over the lifetime of this Parliament, as stated in the White Papers? There is no evidence that it will. On the contrary, detailed analysis suggests that organisational costs may rise in the post-White Papers' NHS. The new structure may be more costly to run in some respects and purported cost savings are likely to be elusive.

Various features in the new structures look set to increase organisational costs. Mergers between NHS trusts and those between GPs create new, larger organisations with a concomitant increase in the problems of co-ordinating their parts and resolving internal conflict. Solving these problems generates internal costs; leaving them unsolved leads to waste. These problems, and the associated costs, will be exacerbated by failure to integrate the new organisations' constituent parts. Within the new primary care associations, individual GPs and practices retain their autonomy, increasing co-ordination costs and conflict over resource allocation. Devolving budgetary responsibility within trusts, to directorates, may similarly amplify internal costs by increased internal fragmentation.

Nor will the transaction costs of organising activity between organisations be reduced automatically by abolishing 'contracts' and introducing agreements and HImPs. The new arrangements may be as costly as those of the internal market, given the continued need to communicate and co-ordinate, and the continued likelihood of conflict over resources. Moreover, introducing quality-indicators into inter-organisational relationships may introduce additional transaction costs.

Transaction costs will be reduced only if the change in name introduced by the White Papers, from contract to agreement, is accompanied by a change in behaviour. The White Papers rely on replacing the competitive ethos of the internal market with co-operation to reduce transaction costs. It is indeed possible that transaction costs may be lower where there is co-operation rather than competition. Achieving co-operation is, however, itself costly, particularly in this context where health and social services are required to work together. The

White Papers outline a cultural change but nowhere consider the costs of achieving it and give few pointers toward prerequisite institutional arrangements.

There is, therefore, no evidence that the real organisational costs of delivering health care will be lower in the post-White Papers NHS than they were in the internal market. Indeed they look set to increase. We are, however, likely to see a reduction in nominal organisational costs. It is hard to escape the conclusion that most of the hoped-for reduction in 'bureaucratic' costs will be achieved by redefining what constitutes 'bureaucracy'. The only reliable way of reducing recorded management costs in the post-White Paper health care organisations will be to change the way costs are counted, defining fewer activities as relating to management.

Failure to reduce organisational costs is not, however, necessarily a bad thing. The presence of apparently high organisational costs, whether internal or transaction costs, does not necessarily mean that the NHS will not meet the objective of ensuring that 'every pound in the NHS is spent to maximise the care for patients'[73]. Organisational costs are inevitably incurred when health and social care is provided. It is impossible to provide health care without incurring planning costs and without co-ordinating the activities of different individuals. These costs inevitably increase with the complexity of health care.

73 The new NHS (1997). p.11.

REFERENCES

Allen, P., Contracts in the National Health Service internal market. *Modern Law Review.* 58, 321-342, 1995.

Appleby, J., Smith, P., Ranade, W., Little, V. and Robinson, R., Competition and the NHS, monitoring the market. In Tilley, I. (ed.), *Managing the internal market.* Paul Chapman, London, 1993.

Ashton, T., Contracting for health services in New Zealand, a transaction cost analysis. *Social Science and Medicine.* 46, 357-367, 1998.

Audit Commission, *A price on their heads: measuring management costs in NHS trusts.* London, Audit Commission Bulletin, 1995.

Audit Commission, *What the doctor ordered. A study of GP fundholders in England and Wales.* London, 1996.

Bartlett, W., Quasi-markets and contracts, a markets and hierarchies perspective on NHS reform. *Public Money and Management.* 11, 53-61, 1991.

Bevan, G., Has the change from hierarchy to market in the NHS been worthwhile? Paper presented to the Quasi-Markets Research Seminar, London School of Hygiene and Tropical Medicine, April 1998.

Coase, R.H., The nature of the firm. *Economica.* 4, 386-405, 1937.

Comptroller and Auditor General, *Contracting for acute health care in England.* HC 261 Session 1994/95, London, HMSO, 1995.

Croxson, B., Co-operation, competition, and the NHS internal market. University of Aberdeen Discussion Paper, Department of Public Health and Economics Discussion Paper No.02/97, 1997.

Dawson, D., Costs and prices in the internal market, markets versus the NHS Management Executive guidelines. Discussion Paper 115, University of York, Centre for Health Economics, 1994.

Department of Health, Reform of Northern and Yorkshire Regional Health Authorities will save £4.7m for patient care each year. Press Release, 1994.

Department of Health, *NHS (England) summarised accounts for the year ended 31 March 1996.* HC 127, Session 1997-98, Stationery Office, 1998.

Designed to care, renewing the National Health Service in Scotland. Cm 3811, Stationery Office, 1997.

Eggertson, T., *Economic behaviour and institutions.* Cambridge, Cambridge University Press, 1990.

Farrell, J., Information and the Coase theorem. *Journal of Economic Perspectives.* 1, 113-129, 1987.

Faulkner, M., Negotiating a contract. In Hodgson, K. and Hoile, R.W. (eds.) *Managing health service contracts.* London, WB Saunders Company Ltd, 1996.

Feldbaum, E.G. and Hughesman, M., Healthcare systems, cost containment versus quality. Financial Times Management Report, London, Financial Times Business Information, 1993.

Ferguson, B. and Keen, J., Transaction costs, externalities and information technology in health care. *Health Economics.* 5, 25-26, 1996.

Ferguson, B., Posnett, J. and Sheldon, T., *Concentration and choice in the provision of hospital services.* Report 6. University of York, NHS Centre for Reviews and Dissemination, 1997.

Fit for the future: a consultation document on the government's proposals for the future of the Health and Personal Social Services in Northern Ireland. Department of Health and Social Services, Northern Ireland, 1998.

Goddard, M. and Manion, R., From competition to co-operation, new economic relationships in the National Health Service. *Health Economics.* 7, 105-119, 1998.

Goodwin, N., Mays, N., McLeod, H., Malbon, G. and Raftery, J., Evaluation of total purchasing pilots in England and Scotland and implications for primary care groups in England. Personal interviews and analysis of routine data. *BMJ.* 317, 256-259, 1998.

Hodgson, G.M., *Economics and institutions.* Oxford, Polity Press, 1988.

Hodgson, G.M., Transaction costs and the evolution of the firm. In Pitelis, C. (ed.) *Transaction costs, markets and hierarchies.* Oxford, Blackwell, 1993.

Information for health. Department of Health, NHS Executive, 1998.

Leary, A., Provider price comparability: a purchaser view. Occasional Paper, Healthcare Financial Management Association / CIPFA, London, 1997.

Le Grand, J., Internal market rules OK. New rules take account of market behaviour. *BMJ.* 309, 1596-1597, 1994.

Lyons, B. and Mehta, J., Contracts, opportunism and trust, self-interest and social orientation. *Cambridge Journal of Economics.* 21, 239-258, 1997.

Macara, A.W., Reforming the NHS reforms. *BMJ.* 308, 848-9, 1994.

Mannion, R. and Smith, P., Trust and reputation in community care. In Anand, P. and McGuire, A. (eds.) *Implementing health care reforms. NHS internal markets.* London, MacMillan, 1995.

Maynard, A., Can competition enhance efficiency in health care? *Social Science and Medicine.* 39, 1433-1445, 1998.

Mays, N. et al., *Total purchasing, a profile of national pilot projects.* London, King's Fund, 1997.

Miller, F.H., Health insurance purchasing alliances, monopsony threat or procompetitive prescription for health sector ills? *Cornell Law Review.* 79, 1546-1572, 1994.

Myerson, R.B., Incentive compatibility and the bargaining problem. *Econometrica.* 47, 61-73, 1979.

NHS Executive, *Good practice and innovation in contracting.* Leeds, Purchasing Unit, Department of Health, 1993.

NHS Executive, *The operation of the internal market: local freedoms, national responsibilities.* London, HMSO, 1994.

NHS Executive, *Health authority costs and management costs in NHS trusts, financial year 1995-6 and planned costs for 1996-7.* Department of Health, 1997a.

NHS Executive, *Health authority costs and management costs in NHS trusts, financial year 1996/7 and planned costs for 1997/8.* Department of Health, 1997b.

NHS Executive, *Changing the internal market.* EL (97) 33, Department of Health, 1997c.

NHS Executive, *The new NHS modern and dependable: establishing primary care groups.* HSC 1998/065, London, Department of Health, 1998.

NHS Wales: putting patients first. Cm 3841, Stationery Office, 1998.

Nippert, R.P., Competition in health care delivery in the U.S. - panacea or poison? A critical view. *Health Policy.* 20, 301-308, 1992.

Nix, J., Co-operation within the entity of the firm. Is the prisoner's dilemma framework sufficient? *Journal of Economics and Sociology.* 53, 443-454, 1994.

North, D.C., *Structure and change in economic history.* London, W.W. Norton, 1981.

Petchey, R., General practitioner fundholding, weighing the evidence. *The Lancet.* 346, 1139-1142, 1995.

Posnett, J., Goodwin, N., Griffiths, J., et al., The transactions costs of total purchasing. National Evaluation of Total Purchasing Pilot Projects Working Paper, London, Kings Fund, 1998.

Propper, C., Market structure and prices, the responses of hospitals in the UK National Health Service to competition. *Journal of Public Economics.* 61, 307-335, 1996.

Redmayne, S., Reshaping the NHS. NAHAT Research Paper 16, 1995.

Roberts, J., Managing markets. *Journal of Public Health Medicine.* 15, 305-310, 1993.

Robinson, R., *Competition and health care, a comparative analysis of UK plans and US experience.* London, Kings Fund Institute Research Report, 1990.

Shelanski, H.A. and Klein, P.G., Empirical research in transaction cost economics: a review and assessment. *Journal of Law, Economics, and Organization.* 11, 335-361, 1995.

Simon, H.B., *Models of bounded rationality.* Cambridge, Mass., MIT Press, 1982.

Sutherland, V.J. and Cooper, C.L., Job stress, satisfaction, and mental health among general practitioners before and after introduction of new contract [sic]. *BMJ.* 304, 1545-1548, 1992.

Taylor, M. *The possibility of co-operation.* Cambridge, Cambridge University Press, 1987.

The new NHS: modern, dependable. Cm 3807, Stationery Office, 1997.

Wallis, J.J. and North, D.C., Measuring the transaction sector in the American economy, 1870-1970. In Engerman, S.L. and Gallman, R.E. (eds.) *Long-term factors in American economic growth.* Chicago and London: University of Chicago Press, 1986.

Wholey, D., Christianson, J. and Peterson, M., Organisation of mental health care in HMOs. *Administration and Policy in Mental Health.* 23, 307-328, 1996.

Whynes, D., Baines, D.L. and Tolley, K.H., GP fundholding and the costs of prescribing. *Journal of Public Health Medicine.* 17, 323-329, 1995.

Williamson, O.E., *Markets and hierarchies.* London, Collier MacMillan, 1990.

Williamson, O.E., *The economic institutions of capitalism.* London, Collier MacMillan, 1985.

Working for patients. Cm 555, HMSO, 1989.