

# **BUDGETARY DEVOLUTION IN THE NEW NHS: SOME THEORETICAL ECONOMIC PERSPECTIVES**

**Mark Dusheiko, Hugh Gravelle, Rowena Jacobs, Michael  
Kuhn, Peter Smith**

**Centre for Health Economics, University of York**

---

**Paper prepared for HESG, 3 – 5 January 2001, Oxford**

**NOT FOR WIDER CIRCULATION**

---

## **Abstract**

Primary Care Groups (PCGs) were introduced into the NHS in April 1999. Covering populations of about 100,000, they will in time become responsible for commissioning virtually all health care. Each PCG receives a unified budget based on nationally calculated expenditure targets. These new arrangements devolve considerable powers to a lower level in the NHS than before, and our paper is seeking to identify the impact of these devolved purchasing powers on NHS activity and expenditure. The purpose of the paper is to explore the issues raised from a theoretical economic perspective, as a prelude to empirical investigation. The paper examines the main components of the new NHS health care system, and discusses the themes in the economic literature that are likely to be applicable to each component. It concludes that two major economic perspectives are likely to be appropriate for modelling PCG behaviour: the public choice literature, and the principal-agent framework. The paper puts forward some tentative theoretical models based on these traditions which may be relevant to the developing role of the PCG. The intention of the paper is to provide a framework for discussion and to solicit responses as to the appropriateness of various economic models to these new institutional arrangements.

## 1. Introduction

In April 1999 the arrangements introduced in the White Paper “The New NHS: modern dependable” were put into place. In particular, Primary Care Groups (PCGs) were introduced. These cover populations of about 100,000, and in time are expected to become the principal organisations responsible for commissioning health care from providers. The new arrangements devolve considerable powers to a lower level than hitherto, and in particular make two fundamental budgetary innovations. First, financial targets are now set using a national formula for organisational units (PCGs) below the level of the Health Authority (HA). Second, PCGs are set unified budgets, based on these targets, which encompass expenditure on virtually all NHS services. Our study “Control and incentives under devolved budgeting in the new NHS” is examining impact of such developments on local NHS activity and expenditure.

The study is taking an economic perspective on the impact of PCGs on NHS activity and expenditure. It will aim to understand how the new organisations respond to budgets and how they attempt to influence their constituent practices to ensure that the PCG meets its financial constraints. The study has three components:

1. A qualitative study of how PCGs and Health Authorities are responding to the change, based on the Manchester/Kings Fund Tracker Survey.
2. The development of theoretical economic models on the responses of PCGs and general practitioners to the new arrangements.
3. Empirical tests of the models developed in stage 2, based on national and local data.

This paper presents some tentative ideas arising from stage two of the study, the development of theoretical economic models. The next section briefly outlines the salient features of the new NHS arrangements. Section 3 presents some of the economic arguments put forward for and against the sort of devolution of powers now under way, and section 4 discusses the classes of economic model that might be relevant to the analysis of such devolution. We then present outlines of the sort of models that might be relevant to PCGs, first from the public choice literature, and then from the principal/agent literature. The paper ends with some tentative conclusions.

We must emphasise that this paper is very much a report of work in progress, and does not represent our final conclusions. We therefore very much welcome comments and alternative perspectives.

## 2. Devolution in the NHS

The NHS, although nominally a unitary organization, has always devolved considerable powers and responsibilities to local geographical areas in the form of Health Authorities (HAs) or their precursors. The main managerial device for controlling this devolution has predominantly been the budget. In particular, the 1991 internal market reforms devolved purchasing powers to HAs, covering populations of about 500,000. HAs in turn devolved some of their purchasing powers to a more local level. A number of purchasing models were

developed, which lay somewhere between the practice and the HA, including general practitioner fundholders, fundholding multifunds and consortia, and local commissioning groups.

In particular, GP fundholding became much more widespread than many commentators originally envisaged. It entailed devolution of responsibility for purchase of prescribing and many elective procedures to general practices which chose to become fundholders. By 1997 over 50 percent of the population was registered with a fundholding practice, and their expenditure accounted for 15 percent of NHS expenditure on secondary care (Audit Commission, 1996). Fundholding became the model for a much more ambitious experiment in the delegation of powers to general practice, known as Total Purchasing Pilots (TPPs) (Mays *et al*, 1997). Under total purchasing, virtually all secondary care purchasing powers and finance were delegated to voluntary associations of general practices. Although TPPs were still effectively a sub-committee of the HA, they were the closest form of predecessor to the current PCG.

The principles underlying the 1997 New NHS White Paper were that “what matters is what works”. Thus aspects of the preceding arrangements (under fundholding and TPPs) that appeared to have been effective were retained, while aspects that were seen to have failed were discarded. The policy prescription of the PCG emerged from application of these principles.

The biggest changes envisaged in the reforms were in the purchasing function. Voluntary GP fundholding was abolished, replaced by the 481 Primary Care Groups. PCGs are led by local general practitioners and community nurses, and membership is compulsory for all GPs practising within the PCG’s geographically defined area. PCGs will be responsible for virtually all the NHS health care for their population and can choose to operate at one of five levels, reflecting progressively increasing levels of autonomy and responsibility:

1. Supporting in an advisory role the HA’s commissioning of health care.
2. Taking responsibility for managing the budget for local health care, although formally remaining part of the HA.
3. Becoming a free-standing body accountable to the HA for managing a budget and commissioning care.
4. As at level 3, but with the added responsibility for the provision of community health services for their population.
5. In addition, PCGs can progress to level 5 status, where they are as at level 4, but are given an additional budget for social services.

From level 3 upwards, PCGs can apply to become Primary Care Trusts (PCTs), which are empowered to deliver community health care<sup>1</sup>. In addition, PCTs can enter into Personal Medical Service (PMS) contracts with GPs giving them greater control over GP salaries, contracts and performance.

Thus at one extreme virtually all commissioning responsibility might be returned to the HA. At the other end of the spectrum a PCG might commission virtually all health care itself, making detailed reference to the HA only in respect of major service developments. In between these extremes, PCGs can exhibit great variations in the powers they adopt and the elements of health care they choose to “block back” to the HA. In the first year (1999/2000) of the new arrangements most PCGs operated at level 1 or 2, but there will be a progression towards more autonomy. In April 2000 seventeen of the PCGs became Trusts and there is strong pressure from the DH for a rapid move to Trust status.

PCGs are comprised of disparate stakeholders and their progression to Trust status may not be straightforward due to internal management difficulties. Essentially practices may have been thrown together with fundamentally different priorities and styles which will raise issues on managing a joint budget. Their challenge will be to overcome the limited management capacity and lack of information systems required for their efficient development (Bloor *et al*, 2000).

Using its usual weighted capitation approach, the NHS Executive sets budgets for virtually all health care within HAs, encompassing Hospital and Community Health Services (HCHS), prescribing, and General Medical Services cash-limited (GMS-CL). Indeed the NHS Plan envisages that even non-cash-limited GMS will be brought within the PCG budgetary regime. Using a consistent methodology, the NHS Executive also sets “indicative” expenditure targets for each of the constituent PCGs within a HA. The HA retains a good deal of freedom as to how it sets PCG budgets in the light of these targets. The HA and PCG budgets are built up service by service, and are known as unified budgets. Although a HA will be expected to adhere to its overall budget, there is no requirement to conform to the individual service elements of the budget. Thus in contrast to fundholding, which covered only limited aspects of health care expenditures, overspends on one service element will have to result in reduced expenditure on other services within the unified budget. PCGs will therefore have to develop indicative budgets for practices with mechanisms to prevent overspends.

PCG budgets are administered by a PCG board consisting of:

- 4-7 GPs
- 1-2 community nurses
- 1 social services representative
- 1 lay member appointed by the HA
- 1 HA non-executive appointee
- 1 PCG chief executive appointed by the PCG board and accountable to the chair
- 1 PCG chair accountable to the HA

---

<sup>1</sup> In the interests of brevity, we continue to refer to them as PCGs.

The PCT board membership differs in so far as it consists of the chair, 5 lay members and 5 members drawn from a PCT Executive, including the chief executive, finance director, clinical governance director, 1 GP and 1 nurse member. The PCT Executives (at levels 3 and 4) in turn consist of the chief executive, the finance director and up to 10 clinicians including GPs, nurses, a social services officer and a professional with public health or health promotion expertise. GPs may therefore have less representation on the PCT board than on the equivalent PCG board.

PCGs specify service agreements with hospitals about the quality and level of service to be provided. Accountability of PCGs is through an accounting officer to the HA and HAs will monitor and advise PCGs. At all levels the PCG will be supervised by the HA, which will remain as an instrument for strategic planning and accountability but will no longer act as a major purchaser (Hausman & Le Grand, 1999). As well as financial accountability, PCGs will be responsible for clinical governance and quality of care. PCGs will have to appoint a clinical governance lead and there will be increasing emphasis on performance measurement and associated incentives and penalties. To achieve PCT status, PCGs need to demonstrate that they are able to develop and monitor clinical standards.

The White Paper (paragraph 5.27) is not explicit about how the advisory or commissioning roles of PCGs will work, and early evidence suggests that a wide range of models will be adopted. The intention is that local circumstances should determine detailed arrangements, in the light of the National NHS Plan, the National Service Frameworks set by the NHS Executive, guidelines issued by the National Institute for Clinical Excellence, and the Health Improvement Programme developed by the local HA. There are also requirements to co-operate with local social services, and to involve the public in commissioning decisions.

PCGs will be able to keep any surpluses they make on their budgets which will be shared between the PCG itself and the individual practices, for spending on patient care. The PCG will have considerable freedom on how to spend its share of the surplus. Practices however will be more constrained. They may only spend their share of the surplus on a prescribed list of activities including the purchase of equipment for patient treatment, investment in premises and non-recurring staff costs (Hausman & Le Grand, 1999).

The devolution of powers to PCGs introduces a new incentive structure into the NHS, the precise nature of which has yet to be properly established. The choice of incentives is likely to be an important determinant of the extent to which the devolution meets policy objectives. Hausman and Le Grand (1999) summarise the key issues as follows:

If the overriding motivation for GPs was one of concern for the health of *all* the patients in the PCG, then the PCG will have an incentive to use its resources efficiently regardless of how the surplus is shared out or how power is allocated between its members. It will allocate its budget in such a way which equalises the marginal health gain from each use of its fund, whether for primary care or the purchasing of secondary care. The outcome could in principle be more efficient than fundholding since as a PCT, all the resources will be under the same budget. Under fundholding

there were separate budgets for primary care, secondary care and community services. In addition, PCGs should avoid the inequities that resulted from the fact that some GPs were fundholders and others were not.

If however GPs are motivated by the concern for the welfare of *their* patients more than others in the group, which is more likely, then there could be conflicts of interest among the GPs in the PCG and these could give rise to collective-action problems with GPs tempted to free-ride on others' activities.

The government has sought to tackle this problem by allowing each practice within the PCG to make some surplus on their budget. Practices effectively face zero tax on the first £10000 of their surplus, a 50% tax rate on the next £70000 and 100% on anything over that. Hausman & Le Grand therefore argue that practices have a strong incentive to make a small surplus, a moderate incentive to make a moderate one and no incentive to make a very large one. If practices had power over commissioning then this could be an effective incentive structure. In contrast, if each practice in a PCG had its own budget and commissioning power, then there would be no reason for the PCG to exist. Such a system is infeasible however, given the large variation in expenditure for secondary care. It is most likely therefore that PCGs will set up incentive structures and systems to restrain referral behaviour of GPs and practices to bring them in line with the objectives of the PCG.

Whether PCGs will be able to offer large enough material incentives to GPs, remains to be seen, especially for larger practices. Very successful practices may resent the 100% tax on any surplus they make over £70000 especially if it is used to offset deficits of other practices. Less energetic practices may similarly be tempted to free-ride in the knowledge that they may receive windfalls from a "good year" due to chance rather than their own efforts, especially if they know their deficits will be covered. Furthermore, if the more effective GPs' patients are penalised because of the lack of solidarity on the part of GP colleagues, then the regime may become demoralising and ineffective. Hausman & Le Grand argue that a crucial determinant of the effectiveness of the PCG will be the extent to which it is able to manage any potential disagreement and conflict between constituent practices.

### **3. The rationale for devolution**

One of the most important issues addressed by the public economics literature is the appropriate level of devolution of authority between different levels of government. In the same vein, some elements of the industrial organisation literature address devolution within large enterprises. From an economic perspective, there are numerous issues to be considered when seeking an optimal level and form of devolution, many of which are relevant to devolved purchasing in health care.

King (1984) has suggested numerous reasons for decentralised provision of public goods, including:

1. Centralised choices may lead to uniformity of provision and may therefore fail to take account of local variations in preferences (Bailey, 1999). The fiscal federalism literature suggests that for some functions, devolution of powers facilitates the matching of public sector outputs and local preferences, thus promoting allocative efficiency. Political and administrative decentralisation therefore can better reflect variations in local preferences. In the health care arena, this notion is embodied by the belief that GPs may be better able than distant bureaucrats to make purchasing decisions which reflect local preferences (Enthoven, 2000).
2. Devolution can promote the development of local competition amongst providers in a quasi-market (a public sector purchaser – provider split). Even contestability or the threat of competition can create incentives for more efficient behaviour. There was some evidence that the existence of GP fundholders stimulated higher efficiency amongst providers, as fundholders were more willing and able than larger purchasers to move contracts to alternative providers.
3. Similarly, increased devolution can enhance the potential for citizens (patients) to use purchasers which offer a preferred type and mix of public service, either through migration, or changes in GP registration. Although there was potential for this practice under fundholding, there was little evidence of the exit option being exercised.
4. Public choice theory suggests that devolution may reduce the potential for self-serving behaviour on the part of politicians, officers and bureaucrats, who through maximising their own utilities, frustrate the promotion of public interest, resulting in excessive output or low efficiency levels (Bailey, 1999). Local decision-makers might be more inclined than distant bureaucrats to provide allocatively and technically efficient services because they are more readily held accountable for local services. Similarly local accountability might enhance control of GPs over both quality and costs of the services they purchase.

Arguments against devolution include:

1. Devolution can increase the transactions and management costs of the organisation. This was indeed borne out by fundholding and other purchasing models. Transaction costs in the form of information flows, negotiating and monitoring contracts, formulating purchasing strategies and undertaking general administration and accountancy work, were not only higher in the initial set-up period but continued and possibly increased further in the longer term (Goodwin, 1998).
2. It becomes very difficult to set satisfactory budgets for small populations, as local idiosyncrasies which can safely be ignored at a higher level may become very important in determining expenditure needs (Majeed, 1999).
3. In a similar vein, the random element of health care expenditure needs becomes very important for small populations (Smith, 1999). If strict adherence to budgets is demanded, numerous dysfunctional consequences might arise when PCGs are asked to manage divergences from expected expenditure over which they have no control.
4. Purchasing inefficiencies may arise if services across jurisdictions are provided in an uncoordinated fashion. Fragmentary purchasing may result in inadequate control over large providers, and was one of the key concerns under the fundholding experiment.
5. Important inequities may arise under devolution, as different purchasers organise services in different ways with different levels of efficiency. The variations in waiting

times between fundholding and non-fundholding patients was one example of this phenomenon.

6. Devolution may give rise to important inter-jurisdictional spillovers where benefits of public services may not be limited to the geographical boundary of the local organisation providing them. This could lead to an under-supply (or indeed over-supply) of services and may require central intervention (Jha, 1998). For example, as part of its contract with a provider, a local PCG may negotiate a service enhancement (such as A&E services) from which patients from other PCGs might benefit. The other PCGs can “free-ride” on the development without bearing the expenditure consequences. Such spillover effects are inevitable given that the populations of PCGs are those people registered with its practices, not those people living in geographically defined areas.
7. There is a lack of firm evidence to substantiate claims that devolution (through competition) improves service quality (Bailey, 1999). Quality of service is crucially dependent on the ability of consumers to voice their preferences. There is for instance little systematic empirical evidence to confirm that quality of care either improved or declined under fundholding (Goodwin, 1998).

#### 4. Classes of economic models

We have sketched out the changes in the NHS associated with the introduction of PCG/Ts:

- greater use of formula finance to determine budgets
- devolution of budgets below HA level: to PCG/Ts and within PCG/Ts to practices
- unification of budgets at PCG/T level
- increased emphasis on performance measurement and incentives

We are seeking to model the behaviour of the relevant decision makers in the new NHS to inform future policy by explaining the effect of the changes in the NHS and to predict the effects of alternative policies. In this section we therefore sketch out the decision makers whose behaviour we need to model and discuss briefly the outstanding issues which arise in modelling and the areas of the existing literature we might hope to draw on. The following sections give some more detailed discussions of how two areas of the literature seem initially promising and give examples of how they might be applied.

Table 1 lists the decision makers we are concerned with, the decisions they take, the most important changes in the NHS environment which affect them, the areas of the economic literature which seem most relevant and the more obvious unresolved issues in modelling their behaviour.

**Patients.** There is a reasonably well worked set of health economic models of individuals’ decisions about their health and consumption of health care. None of the innovations associated with the new NHS seem likely to create a need for additional models. However, the increased emphasis on alternative routes into the NHS (the extension of the range of over the counter medicines and pharmacist advice medicines, walk in centres, NHS Direct and changes to out of hours arrangements) may require more work on models of choice of provider. These alternative routes may also have implications for competition for patients



between GPs and alternative gatekeepers. Interesting though these changes are we do not see them as of major relevance for the analysis of devolved budgeting.

**GPs.** Most GPs are not single handed, so that it is important to distinguish between decisions made by individual GPs in respect of individual patient care and resource use and the decisions made collectively by practices. There is a clear intention by the DH to encourage GPs and practices to switch from the standard General Medical Service contract (the “Red Book”) to Personal Medical Service contracts held by HAs, Community Health Trusts, and level four PCTs. Indeed in the case of single handed GPs the DH seems to envisage a compulsory change to PMS (Department of Health, 2000). PMS contracts permit GPs to opt for salaries so that they will have to choose the type of contract they wish to work under. The increased emphasis on monitoring of performance manifests itself at this level via clinical governance and reaccreditation.

There is a reasonably large body of health economics literature on doctor behaviour and the issues, such as the extent to which GPs are semi altruistic as regards their patients, are well known (Scott, 2000).

**Practices.** Most GPs work with other GPs in small business partnerships employing a variety of other staff and often owning their premises. Decisions must be taken at practice level about the internal profit sharing arrangements amongst the partners, the internal monitoring and control mechanisms, the numbers of partners, the numbers and types of staff employed, whether the practice should opt for a PMS contract and which PCG it should belong to. All of these affect the behaviour of the individual partners.

There is little economics literature about the decision making of practices, as opposed to individual GPs. In many instances, for example when attempting to explain referral or prescribing rates for the practice’s population as whole, treating the practice as if it were a single individual GP is a sensible modelling strategy enabling one to draw on the existing literature.

Models of individual GP behaviour do not address decisions taken by the practice as a whole, unless one believes that practices are dictatorships run by the senior partner. Alternatively one could assume that GPs sort themselves into partnerships of like minded individuals so that collective decisions can be modelled as if the practice was a single individual. The rate at which partnerships dissolve suggests that the latter assumption may not always be useful (Taylor and Leese, 1997). Models of worker cooperatives (Prychitko and Vanek, 1996) are readily adaptable for analysis of partnership size. Concepts from principal agency theory have also been used to examine the choice of internal profit sharing arrangements (Gaynor and Pauly, 1990).

**PCGs.** PCGs are the main focus of our modelling exercise: they are new organisations with a wide range of decisions with the potential to significantly impact on resource use in the NHS. In addition to commissioning secondary care, they can commission or (at level 4) provide community health services. At level 4 they will also be able to contract with practices via the PMS option. If they decide to offer salaried contracts to GPs they will in

effect be providing primary care via their employees. They are responsible for clinical governance arrangements to improve the quality of care provided by their practices. They allocate actual or notional budgets to practices and devise incentive schemes for them. They are likely, at “level 5” to have the option of holding budgets commissioning or provide social as well as health services.

It is not clear what are the appropriate models to analyse their behaviour. If we were willing to assume that for some purposes they can be modelled as if they were a single decision maker then we could adapt standard consumer models to examine the allocation of the unified budget across and within HCHS, prescribing and practice infrastructure. The choice of incentive schemes related to practice budgets or performance indicators could be tackled using the regulation and principal - agent literature. However, even in this case the question of the objectives of PCGs looms large.

A more fundamental issue is whether we need to recognise that PCGs are composed of practices rather than assume that we can use representative agent models. Level 1 to 3 PCGs are constituted so that their decision making is dominated by GPs. But although it seems reasonable to assume that they will therefore be run in the interests of (partially altruistic) GPs, it is not clear that GP preferences are sufficiently similar to justify the representative agent approach. Although practices could choose which PCG to belong to, the requirement for geographical contiguity limited their choices severely, so that we cannot appeal to assortative selection to suggest that all GPs or practices in a PCG will have similar objectives.

This suggests that at least some of the decisions taken by PCGs should be approached via the public choice and collective choice literature (see section 5). For some decisions, for example on the allocation of funds for practice infrastructure across practices, the representative individual approach seems highly inappropriate. For others, perhaps contracts with providers, there is more likely to be agreement on the merits of different types of contract.

**HAs.** Health authorities have a reduced role in the new system. They receive a global budget for all their PCGs determined by the application of the national resource allocation formulae and must decide how to allocate the budget amongst their PCGs. They do not have to do so by applying the national resource allocation formulae to their PCGs. HAs can use Accountability Agreements with their PCGs to set targets for various aspects of PCG performance. We expect HAs to have rather more influence since, especially initially with lower level PCGs, they will be taking commissioning decisions for large proportions of their PCGs' budgets. HAs will also hold PMS contracts so that they will in effect be commissioning primary care from the GPs and practices who wish to take such contracts.

The bureaucracy literature seems an obvious place to look for a template for modelling HAs. In Niskanen (1968) civil servants aim to maximise the budget, given to them by politicians to produce a set of outputs, by exploiting their superior information about the production technology. But in the NHS the total HA budget is largely determined by national formulae which cannot be manipulated by individual HAs and HAs have had a set

of performance criteria imposed on them by the centre. The more sophisticated developments of the basic theory of bureaux surveyed in Weintrobe (1997) may be more useful.

**Providers.** HAs and PCGs commission secondary and community services from independent NHS trusts which are constrained to break even after earning 6% on their capital. One approach to understanding hospital provider behaviour is that they consist of consultant firms motivated by some combination of income, patient orientated ethics, and prestige (as manifested in size or quality). An alternative is to apply one of the varieties of the managerial theory of the firm on the assumption that trusts are run by their managers and to ignore the internal principal agent mechanisms used to control the behaviour of consultant firms.

**Politicians.** For the moment we ignore the tiers of the bureaucracy above HAs and assume that decisions on NHS total funding, the resource allocation rules determining HA budgets, the Performance Assessment Framework and the other centrally determined targets imposed on HAs (as for example set out in the National Plan) are chosen by politicians. Our initial aim is to understand the behaviour of the lower level actors in the NHS within the framework created by political actors so that we do not need to concern ourselves here with models which attempt to explain the choice of the framework by politicians.

In this section we have sought to identify the key actors within the health care system, and the incentives and constraints they face. Models of other decision makers are also in need of development but these discussed here are of most immediate relevance to the devolution of budgets. In the first instance our aim is to develop simple models for each level which capture the most important features of budget devolution. Although these must first be developed separately, the hope is that they can eventually be integrated into a unified model of budget devolution.

## **5. Theoretical perspectives on PCG budget arrangements**

### **5.1 *Modelling the allocation of PCG/T unified budgets***

Although a PCG/T's budget is derived from the application of separate national formulae for HCHS, prescribing and GMS-CL it has discretion to allocate its total unified budget across these three types of expenditure. The only formal constraint is that practices cannot get less GMS-CL than the amount they received in 1998/9. Since PCG/T total budgets will grow significantly compared with 1998/9 the 'one way valve' on GMS-CL is not likely to be a binding constraint. How then should we model the allocation by the PCG/T of its unified budget across the three main budget headings?

In this section we sketch two possibilities using a simple diagram to show the relationships between alternative models:

## 5.2 PCG/Ts and consumer theory

The simplest model is to assume that the PCG/T acts as if it had well behaved preferences over the three components of the budget. The PCG/T receives a fixed total unified budget  $y$  and must decide how to allocate the budget to expenditure on prescribing ( $x_1$ ), HCHS ( $x_2$ ) and GMS-CL ( $x_3$ ). Figure 2 illustrates in  $x_1, x_2$  space. Points inside the budget frontier  $yy$  have positive  $x_3$  equal to the horizontal distance to the frontier. For a given total budget, an allocation is completely described by its  $x_1, x_2$  coordinates. The PCG/T inherits an allocation, denoted  $R$ , of its budget over the three main headings. The 'one way valve' on GMS-CL ensures that the PCG/T will not be able to choose any allocation to the right of a 45° line through  $R$  unless all GPs agree the change.

PCG boards have a majority of GPs and they have the right to a GP chair. Assume that the practices in the PCG are reasonably homogeneous so that PCG decisions are made to maximise the utility of a representative GP subject to the unified budget constraint. In Figure 2, preferences are shown by the circular indifference curves around the GP bliss point at  $x^*$ . Given the shape of preferences, the preferred allocation involves a shift of expenditure away from HCHS, into GMS-CL and prescribing.

For a testable model capable of explaining cross PCG/T differences in the allocation of the unified budget it is necessary to link observable characteristics of GPs and PCG/Ts with aspects of preferences or prices which affect allocation decisions. For example, we can hypothesise that the share of the total budget devoted to prescribing will vary with the proportion of patients for whom GPs can dispense as well as prescribe. Or the share of GMS-CL may be related to the proportion of practices that own their premises.

## 5.3 Public choice

Markets are not the only mechanism of allocating resources. Co-ordination difficulties due to information asymmetry, barriers to entry and exit or excludability and rivalry conditions imply the necessity for institutions such as clubs, co-operatives and committees to acquire and command the mechanism by which economic resources are distributed. The Public Choice and Political Economy literature has examined the economics of non-market decision making, and can enable an understanding of incentives and decision making in social institution as well as providing a framework for analysing group behaviour and collective decision making.

In examining how policy decisions are taken, three simple cases have arisen:

1. Where the set of possible policies are one dimensional
2. Multi-dimensional policies
3. Where it is possible for voters to arrange side payments

From these scenarios, three fundamental insights can be gained from the literature:

1. The median voter theorem: if the set of all policy alternatives has a simple one

- dimensional structure, then the majority principle implies that the preference of the most centrist voter should be chosen (equivalent to majority rule).
2. Multi-dimensional instability theorem: if the set of all policy alternatives does not have a simple one-dimensional structure, then the results of majority voting may still be unpredictable. This implies that any alternative can be beaten by a majority that would favour other alternatives, hence the outcome is not transitive (in the sense that  $a > b$ ,  $b > c$  does not imply  $c > a$ ).
  3. Coase theorem: if bargaining is costless and unanimous consent is required to make any policy changes, then any stable equilibrium outcomes should be an efficient policy. Hence, this theorem strengthens the notion that unanimous consent rather than a voting mechanism should be required for decision making. However, there are a number of practical difficulties with this idea. Firstly, there is an incentive for individuals to maximise payoffs (side payments) by exaggerating their costs of consenting. If there are bargaining costs such as unavoidable costs of communication, losses due to uncertainty about other peoples' true preferences and payoffs, costs of delays to getting agreements. Then unanimous agreement may be a less attractive means of decision making.

In the following two sub sections we look at bargaining and voting models of collective decision making in the PCG.

### 5.3.1. Nash bargaining

PCG/Ts may not be dominated by a single interest group. GPs are not homogeneous: single handed GPs in deprived areas and large practices in leafy suburbs are likely to have different preferences over expenditure. With two dominant groups we can model the allocation as the outcome of a co-operative bargain. Figure 3 illustrates. One interest group has allocation  $x^A$  as its preferred outcome and the other prefers  $x^B$ . The heavy line is a locus of points of tangency between their indifference curves, and  $x^A x^B$  is the contract curve: for any point off the contract curve, such as  $R$ , there is a range of the contract curve which is better for both parties. If we assume that the baseline allocation  $R$  is the default or disagreement allocation, they will agree, in this example, to an allocation on the contract curve between  $x^B$  and the intersection of  $B$ 's indifference curve through  $R$  and the contract curve.

The Nash bargaining solution depends on their preferences, the disagreement point  $R$  and their bargaining power. If the two groups are GPs, bargaining power may be related to the numbers of GPs in each group. Both groups may be subject to free rider problems if GPs who do not contribute to the cost of bargaining and lobbying cannot be excluded from benefiting from any agreement reached. Free rider problems may be easier to control for in interest groups consisting of fewer practices. Other things equal, we may find relationships between allocations and the numbers of GPs and practices in interest groups.

### 5.3.2. Voting models

If PCG/Ts have more than two interest groups matters become much more complicated. One of the simplest ways to aggregate their preferences would be to hold a vote. The median voter theorem predicts that if alternatives differ only in a single dimension then the

chosen alternative is the one preferred by the median voter. But if alternatives have two or more dimensions, the median voter result or majority rule does not apply in general. If the median voter theorem or majority rule is employed in pairwise comparisons of alternatives, neither may be a perfect mechanism (Mueller, 1989). Figure 4 illustrates. Suppose there are three groups whose bliss points, shown as  $x^A$ ,  $x^B$ ,  $x^C$ , have equal votes. Starting from the baseline allocation  $R$  a proposal to move to  $x^B$  will be accepted by  $A$  and  $B$  since  $A$ 's indifference curve through  $R$  is outside its indifference curve through  $x^B$ . In a choice between  $x^B$  and  $x^A$ ,  $A$  and  $C$  vote for  $x^A$ . In a choice between  $x^A$  and  $x^C$ ,  $C$  and  $B$  vote for  $x^C$  and in a choice between  $x^C$  and  $x^B$ ,  $B$  and  $A$  vote for  $x^B$ . Thus no allocation commands a majority against all other allocations.

Definite predictions under majority voting can be generated in a number of ways. If there is majority voting on each dimension separately then the median voter result can be applied to each dimension. However the median voter equilibrium will depend on the order in which dimensions are presented. Thus in figure 4, the PCG/T might first vote to allocate a fixed sum between GMS-CL and prescribing, holding HCHS constant at its baseline level (moving along a horizontal line through  $R$ ). With the preferences giving the indifference curves shown, the point chosen will be where  $A$ 's indifference curve is tangent to the horizontal line through  $R$ :  $B$  will prefer more prescribing and  $C$  less, so that  $A$  is the median voter. If the amount of prescribing expenditure was held constant, then the PCG/T votes on the choice between HCHS and GMS-CL. Now moving down the vertical line through the amount of prescribing chosen by  $A$ , implies  $B$  is the median voter and the equilibrium is where  $B$ 's indifference curve is tangent to the vertical line. If the PCG/T had voted first on the amount of HCHS holding prescribing constant and then on the amount of prescribing holding HCHS constant the final equilibrium would not in general coincide with the that generated by the first order of voting. Hence whoever sets the agenda can have a powerful influence on the outcome.

#### **5.4 Do the models yield different testable predictions?**

The differences in the theoretical models are of interest only if the models yield predictions which are different and are capable of being tested on the available data. Thus consider a comparison of the consumer model with the Nash bargaining model. Let  $u(x,k)$  be the utility function of GPs where  $x$  is the allocation vector and  $k^i$  is a set of preference parameters. In the representative GP model the equilibrium allocation will depend on the preference parameters, the total unified budget:  $x^e = x^*(k,y,z)$ , where  $y$  is the total budget and  $z$  the environment. But if the allocation is determined as a Nash bargain between two groups of GPs with different preference parameters  $k^1, k^2$ , then the allocation depends on the preference parameters, the baseline allocation  $R$  vector and the environment:  $x^e = x^N(k^1, k^2, R, z)$ .

Changes in  $z$  are likely to have the same qualitative effects on the allocation. The Nash model predicts that the components of the baseline allocation will affect the actual allocation whereas the consumer model predicts that only the total of the baseline will matter. But a simple extension of the consumer model to incorporate adjustment costs in changing the

allocation from the baseline means that the  $x^*$  will depend on the components of the initial allocation as well as its magnitude. The two models will not be distinguishable. In which case we should choose the consumer model as the simplest model to generate the predictions about the effects of the environment, initial allocations etc.

## 5.5 *A principal/agent perspective*

This section addresses the process of allocating budgets to individual practices within a principal-agent framework, which is a natural approach to the issue of budgeting.<sup>2</sup> As Propper (1995) notes, the health care system – of whatever design – is replete with interlocking agency relationships: that is, of relationships between principals and their agents. Figure 1 illustrates some examples within the context of the New NHS.

As Gravelle and Smith (2000) explain, delegated decision making is only unproblematic either if the principal has full information or if the objectives of principal and agent coincide. The outcomes from the agents' decisions depend both on their own actions and also on factors outside the control of the agent. For example, in providing influenza vaccinations to their elderly patients, GPs must decide on the amount of effort to put into persuading patients to be vaccinated and the choice of patients to be offered vaccination. The outcome is the health gain in the eligible population, which depends on the numbers of patients vaccinated, their personal characteristics, the type of influenza prevalent in the period, and a purely random element.

If the principal (say a PCG) knew how the actions of the GP and the uncontrolled factors affected the health gain and if it could observe both the GP's actions and the uncontrolled factors. An optimal vaccination policy could then be designed with a suitable contract to ensure that it was carried out by the GP. It could make payment contingent on the agent carrying out the action that was optimal given the uncontrolled factors. Alternatively if the principal knew the outcome (health gain) and the factors other than the GPs actions affecting health gain, it could also achieve its optimal vaccination policy by making payment to the GP contingent on achieving the health gain which was optimal given the other factors.

Unfortunately, the principal will not usually have sufficient information to implement such a policy. It will not be able to observe all the GPs actions: the PCG can monitor the number of vaccinations but it will not be able to observe the GP's efforts directly: there will be a problem of *hidden action*. The PCG is also likely to face a problem of *hidden knowledge*: it will not know enough about patients to determine whether a higher vaccination rate reflects greater effort by the GP, or a patient population who are easily contacted and persuaded to be vaccinated, or simply chance events. Furthermore, the PCG will not be able to measure the (outcome) health gain since this requires detailed information on individual patients, such as their initial health state, their valuations of alternative health states and so on.

---

<sup>2</sup> For an excellent overview on the principal-agent-problem see Sappington (1991). Chalkley and Malcomson (2000) discuss the principal-agent-problem in relation to health care systems in general and Gravelle and Smith (2000) in the context of the NHS reforms.

Thus in general a principal-agent problem exists where there is asymmetric information and non-identical objectives, and many agency relationships in health care seem to satisfy both requirements. The difference between the principal's welfare when there is full information or coincidence of objectives and when there is not is known as the *agency cost*. The principal will seek to devise managerial, information, budgetary, and payment systems to mitigate the cost of agency, taking account of the direct costs of such systems, or *transaction costs*. The principal will aim to choose a system which minimises the sum of agency and transaction costs.

Principal-agent theory is predominantly normative in asking how the principal should design incentive schemes to achieve a second-best solution under the informational constraints. The following subsection comments on some implications of this for PCG budgeting. Many of the proposed schemes involve informational requirements, which may not be satisfied in practice, or the use of instruments, which are unavailable to the PCG. It therefore seems appropriate to complement the normative approach by a positive analysis of the incentives under sub-optimal budgeting mechanisms. The second subsection addresses some of the issues involved.

#### **5.5.1. Optimal budgeting mechanisms**

It is well understood that budgeting helps to resolve the problem of 'hidden action'. Making the practice claimant of the net surplus, i.e., budget less cost, introduces an incentive to use resources efficiently. Ideally, this implies zero cost sharing. In the context of budgeting, this means that there is no reimbursement of any overspends. As we will argue below, this is, in general, unfeasible.

The problem of 'hidden knowledge' can be studied in relation to the problem of 'hidden action'. Consider the following example. The PCG wishes every practice to attain some health target at minimum overall budget. There are two types of practices, those with severe (S) and non-severe (N) case mix. The severity of case mix is a practice's private information. The cost of attaining the target and, thus, the budget to be assigned to a practice increases in severity. By exerting effort each practice can reduce the cost of attaining the target.<sup>3</sup> The utility of each practice increases in net surplus and decreases in effort. Finally, assume that practices receive an exogenous income, e.g., GP capitation according to the national contract, which is high enough to guarantee the participation of both types of practice for all allocations.

Under complete information, the PCG allocates a budget to each practice corresponding to its type. There is no cost sharing in order to elicit maximum effort. However, as case mix cannot be observed, the PCG is unable to establish the efficient allocation. If asked for their type, all practices would report S in order to obtain the larger budget.

<sup>3</sup> The problem is structurally identical to the two-type screening model considered in Laffont and Tirole (1993: section 1.3).



In order to separate the types, the PCG offers a menu of 'contracts', each specifying a budget and a degree of cost sharing. Suppose there are two such contracts: high budget and low cost sharing (HL) and low budget and high cost sharing (LH). Separation is achieved if S chooses LH and N chooses HL. The separating allocation is characterised as follows: (i) The HL contract involves no cost sharing and a budget, which leaves a positive net surplus to N. Thus, efficient effort is elicited from N but a rent has to be paid so as to stimulate the revelation of its type. (ii) The LH contract involves cost sharing and a budget, which leaves a zero net surplus to S. Hence, no rent is paid but too little effort is induced.

Thus, the restoration of symmetric information is costly and implies a loss in efficiency. Moreover, it is unlikely that budgets can be allocated so as to guarantee an equitable provision of health care. In the example, S just attains the health target, whereas N over-achieves the target if it uses some of the residual budget to provide additional care. Hence, budgeting is likely to entail a trade-off between equity and efficiency concerns.

The above example is a highly stylised model of budgeting. The following are only a few of the relevant issues in PCG budgeting to be addressed in more elaborate models.

- (1) In the example, we have assumed that staff remuneration is sufficient to guarantee participation of the practice (staff). The NHS's current initiative of raising remuneration and reduce staff shortages (NHS Plan 2000) rather suggests a binding participation constraint. This implies an upper bound for the PCG's ability to induce effort, as effort now has to be compensated for by increases in budget or in the degree of cost sharing.
- (2) Much emphasis is put on the quality of care (NHS Plan 2000). If provision of quality becomes a second task of practices besides the efficient administration of the budget, the PCG has to provide equally strong incentives for the tasks. If this is impossible because quality is difficult to observe or verify, this blunts the incentive power of budgets. In the absence of cost sharing practices would shirk on quality. The problem is alleviated if demand is reactive to differentials in quality. Improved performance measurement then does not only play a role in raising quality but also in strengthening the incentive effect of budgets.
- (3) Further instruments may be available to the PCG besides budget levels and the degree of cost sharing. These include the degree of clinical and budgetary discretion allowed to practices,<sup>4</sup> auditing, fines and rewards, and (published) performance indicators. Inefficiency arises if the number of instruments falls short of the number of tasks. Then, an increase in the number of instruments is enhancing efficiency.
- (4) Zero cost sharing is no longer efficient if risk-averse practices are exposed to uncertainty with regard to treatment cost. In this case, the PCG may use further signals of practice effort if they contribute to reducing risk. One example is the use of yardstick competition between practices. If practice cost is determined by common rather than by

---

<sup>4</sup> In a framework similar to the one presented above, Kuhn (2000) considers a PCG, which is uninformed about practice efficiency in administering budgets. He shows how the PCG can use budgets and the degree of (clinical) discretion allowed to practices to separate practice types.

idiosyncratic shocks, the PCG reduces an individual practices' risk by basing cost sharing on relative rather than on absolute performance.

- (5) The repetitive nature of budgeting gives rise to the ratchet effect. It is unlikely that the PCG is able to commit not to use information gained in the process of budgeting. Once a practice has revealed its type, it is optimal for the PCG to allocate the full information budget in all subsequent periods and extract practice rent completely. But then, the practice has no incentive to reveal its type, and, even the second-best allocation becomes unfeasible. The ratchet effect is alleviated if production is subject to random shocks.
- (6) Currently, practices receive budgets primarily for the purchase of secondary care and pharmaceuticals. This implies a three-tier hierarchy, where the practice serves as intermediate agent between the PCG and providers. A trade-off exists between the practice's superior information in the purchasing process (e.g., about providers or case-mix), on the one hand, and the superior bargaining power of the PCG, on the other. Following Caillaud et al. (1996), the optimal degree of discretion left to practices increases in their bargaining power and the concern about extracting provider rent.
- (7) Practices face multiple principals. Budgets are determined by PCG, remuneration of GPs is still substantively determined by the Department of Health, and clinical guidelines will be determined by NICE. Incoherent incentives are likely to arise. The incentive effect of budgeting relies on exposing agents to the financial consequences of their actions. The incentive is much reduced as long as GPs, as practice decision makers receive remuneration from an external source and are not accountable for overspends. Even worse, in the absence of financial accountability, the consequences of a practice's 'debt' are likely to be borne by patients. If patients are immobile between practices, the concern for equity puts much pressure on the PCG to bail out an indebted practice. In this regard, the introduction of PMS contracts is commendable for providing the PCT with an option to introduce financial incentives, likely to be more highly powered than the ones currently in place.
- (8) Practice staff may be motivated by non-economic incentives, an issue, which to limited extent has been taken up by the literature on reimbursement systems in health care (Chalkley and Malcomson 2000).

### **5.5.2. Towards a positive analysis of budgeting**

A positive analysis of budgeting focuses on the behaviour of agents under a given, not necessarily optimal, budgeting mechanism. Indeed, as argued earlier, budgeting procedures are generally sub-optimal due to lack of instruments or information. This leaves scope to agents not only to react to a given mechanism but to behave more proactively in manipulating the principal's decisions and beliefs. In the following, we consider a number of approaches, which may serve as starting points of a positive analysis.

One feature of PCG budgeting is that budgets are to some extent determined historically and then adapted over time in order to reach target levels. This could be interpreted as a game of budgeting, in which an initially uninformed principal learns the agents' types in the process.

Levaggi and Smith (1994) consider such a learning process in a game between a principal and a single agent. They show that the learning process, in which budgets are adjusted, entails efficiency losses for both players. By misrepresenting its type, the agent can distort the target allocation to its favour. Under cheating, the agent trades off an excessive loss in surplus during adjustment and the rent attained once the target is reached. Thus, the optimal amount of cheating falls in the duration of the adjustment phase and the agent's rate of time preference.

If agents are heterogeneous, some of them may actively engage in resolving the informational asymmetry by signalling their type. Reconsider the previous example, where practices differ in case mix, S (severe) and N (non-severe). If the PCG is unable to distinguish types it allocates budget and the degree of cost sharing according to some (prior) average. For S, this allocation is likely to be less favourable than the budget it would receive were the PCG informed about its high cost. Thus, S has an incentive to reveal its type by using its cost (or residual budget) as signal. N has an incentive to mimic this signal in order to claim the higher budget as well. S, therefore, has to distort effort below the efficient level. Separation is feasible only if S can profitably choose a low level of effort, which is no longer profitable for N.<sup>5</sup> Whether this is possible depends on the properties of the cost function.

The welfare implications of signalling are ambiguous in this example. On the one hand, signalling may resolve the informational asymmetry and, thus, help to attain an efficient allocation in subsequent periods. On the other hand, there is a loss of efficiency due to the downward distortion of effort. Under different circumstances practices might want to signal efficiency by distorting effort upwards. Then, there is an unambiguous welfare gain to signalling.

The PCG may be unable to adjust budget targets, for reasons relating to the absence of learning, to institutional constraints or to commitment. In such circumstances the budgeting decision could be modelled as a mechanistic bid function, which maps practice bids, or some other action, into a budget allocation that satisfies the budget constraint at PCG level. For a game, in which practices bid for budgets, one can then study the impact of the properties of the bid function on the equilibrium allocation.

One example of such a game could follow Barrow (1986, 1989) who considers local governmental competition for grants. Practices decide on the provision of two services X and Y. Each practice  $i$  receives a 'general' budget  $GB_i$ , which can be spend on X and Y, and a 'specific' budget  $SB_i$ , which is tied to the provision of service Y. The PCG commits to a  $GB_i$  for each practice and to an overall budget  $SB = \sum SB_i$  to be allocated to service Y. The following can be shown. (i) Practices over-provide service Y. When providing Y, they take into account that this raises their claim on the 'specific budget'  $SB$ . (ii) Over-

<sup>5</sup> The S practice suffers a loss in utility by distorting its effort downwards when signalling. In a separating equilibrium, this is more than offset by the ensuing gain in utility from receiving the more favourable allocation.

bidding is inefficient from the PCG and, in fact, the individual practice's point of view. In a symmetric setting, all practices receive a fixed share of the overall budget SB. Over-bidding by a practice merely offsets the over-bidding by other practices. However, as bids are linked to the actual provision of service Y, a practice distorts its allocation of real resources.. Efficiency is lost as, too many resources are spent on Y, while X may have to be rationed. (iii) Over-provision and, thus, inefficiency increase, as more funds SB become available. (iv) Over-provision decreases in the number of practices. This is because for an increasing number of practices, the individual gains from over-bidding diminish.

The model of bidding competition may be extended in a normative or positive way. From a normative perspective, one may solve for a specification of the bid function that maximises a PCG objective function. From a positive perspective, one may derive the bid function as the outcome of voting or bargaining amongst practices.

So far we have maintained the assumption that the executive, as principal, acts in the best interest of the PCG. This may, of course, not always be the case. Budgeting decisions are likely to be exposed to influence activity even in the presence of formal rules and PCG budgeting may be captured if executive members pursue private objectives.

## **6. Concluding comments**

This paper has speculated on the sorts of economic models that may be relevant to describing the institutional arrangements recently put in place in the New NHS. It has pointed to both the complexity of the health care system and the rich heritage of economic literature that may be relevant. Our task is to develop from these models hypotheses which can be tested empirically. We remain to be convinced that that this is anything but an immense challenge.

## 7. References

- Bachmann, M. & Bevan, G. (1996) Determining the size of a total purchasing site to manage the financial risks of rare costly referrals: Computer simulation model, *British Medical Journal*, 313: 1054-7.
- Bailey, S.J. (1999) *Local government economics: Principles and practice*, Macmillan Press Ltd: London.
- Barrow, M.M. (1986) Central grants to local governments. A game theoretic approach, *Environment and Planning C: Government and Policy*, 4: 155-64.
- Barrow, M.M. (1989) Local authority responses to grants: Different types of response and equilibrium, *Environment and Planning C: Government and Policy*, 7: 313-20.
- Baxter, K., Bachmann, M. & Bevan, G. (2000) Primary Care Groups: Trade-offs in managing budgets and risk, *Public Money and Management*, 20(1): 53-62.
- Bloor, K., Maynard, A. & Street, A. (2000) The cornerstone of labour's 'new NHS': Reforming primary care, in Smith, P.C. (Ed.) *Reforming markets in health care: An economic perspective*, Open University Press: Buckingham.
- Caillaud, B., Jullien, B. & Picard, P. (1996) Hierarchical organization and incentives, *European Economic Review*, 40: 687-95.
- Chalkley, M. & Malcomson, J.M. (1996) Contracts for the National Health Service, *Economic Journal*, 106: 1691-1701.
- Chalkley M. & Malcomson, J.M. (2000) Government purchasing of health services, in Culyer, A.J. & Newhouse, J.P. (Eds.) *Handbook of Health Economics*, Amsterdam: Elsevier North-Holland.
- Dixon, J. (1998) The context, in Le Grand, J., Mays, N. & Mulligan, J. (Eds.) *Learning from the NHS internal market: A review of the evidence*, King's Fund Publishing: London.
- Enthoven, A.C. (2000) Modernising the NHS: A promising start, but fundamental reform is needed, *British Medical Journal*, 320: 1329-31.
- Gilley, J. (1999) Meeting the information and budgetary requirements of primary care groups, *British Medical Journal*, 318: 168-70.
- Glazer, J. & Shmueli, A. (1995) The physician's behaviour and equity under a fundholding contract, *European Economic Review*, 39: 781-5.

- Goodwin, N. (1998) GP fundholding, in Le Grand, J., Mays, N. & Mulligan, J. (Eds.) *Learning from the NHS internal market: A review of the evidence*, King's Fund Publishing: London.
- Gaynor M. & Pauly M. (1990) Compensation and productive efficiency in partnership: Evidence from medical group practice, *Journal of Political Economy*, 98(3): 544-73.
- Gravelle, H. & Smith, P.C. (2000) The new NHS: A principal-agent viewpoint, in Smith, P.C. (Ed.), *Reforming markets in health care: An economic perspective*, Open University Press: Buckingham.
- Hao, L., Rosen, S. & Suen, W. (2000) *Conflicts and Common Interests in Committees*, mimeo.
- Hausman, D. & Le Grand, J. (1999) Incentives and health policy: Primary and secondary care in the British National Health Service, *Social Science and Medicine*, 49: 1299-1307.
- King, K.N. (1984) *Fiscal Tiers: The Economics of Multi-Level Government*, London: George Allen & Unwin.
- Kuhn, M. (2000) *Delegating practice budgets under asymmetric information*, mimeo.
- Laffont, J.J. & Tirole, J. (1993) *A Theory of Incentives in Procurement and Regulation*, London: The MIT Press.
- Le Grand, J., Mays, N. & Mulligan, J. (1998) *Learning from the NHS internal market: A review of the evidence*, King's Fund Publishing: London.
- Levaggi, R. & Smith, P.C. (1994) On the Intergovernmental Fiscal Game, *Public Finance*, 49: 72-86.
- Majeed, A. (1999) Commentary: Accurate information may be difficult to produce, *British Medical Journal*, 318: 170.
- Majeed, A. & Malcolm, L. (1999) Unified budgets for primary care groups, *British Medical Journal*, 318: 772-6.
- Martin, S., Rice, N. & Smith, P.C. (1998) Risk and the general practitioner budget holder, *Social Science and Medicine*, 47(10): 1547-54.
- Mays, N. & Mulligan, J. (1998) Total purchasing, in Le Grand, J., Mays, N. & Mulligan, J. (Eds.) *Learning from the NHS internal market: A review of the evidence*, King's Fund Publishing: London.

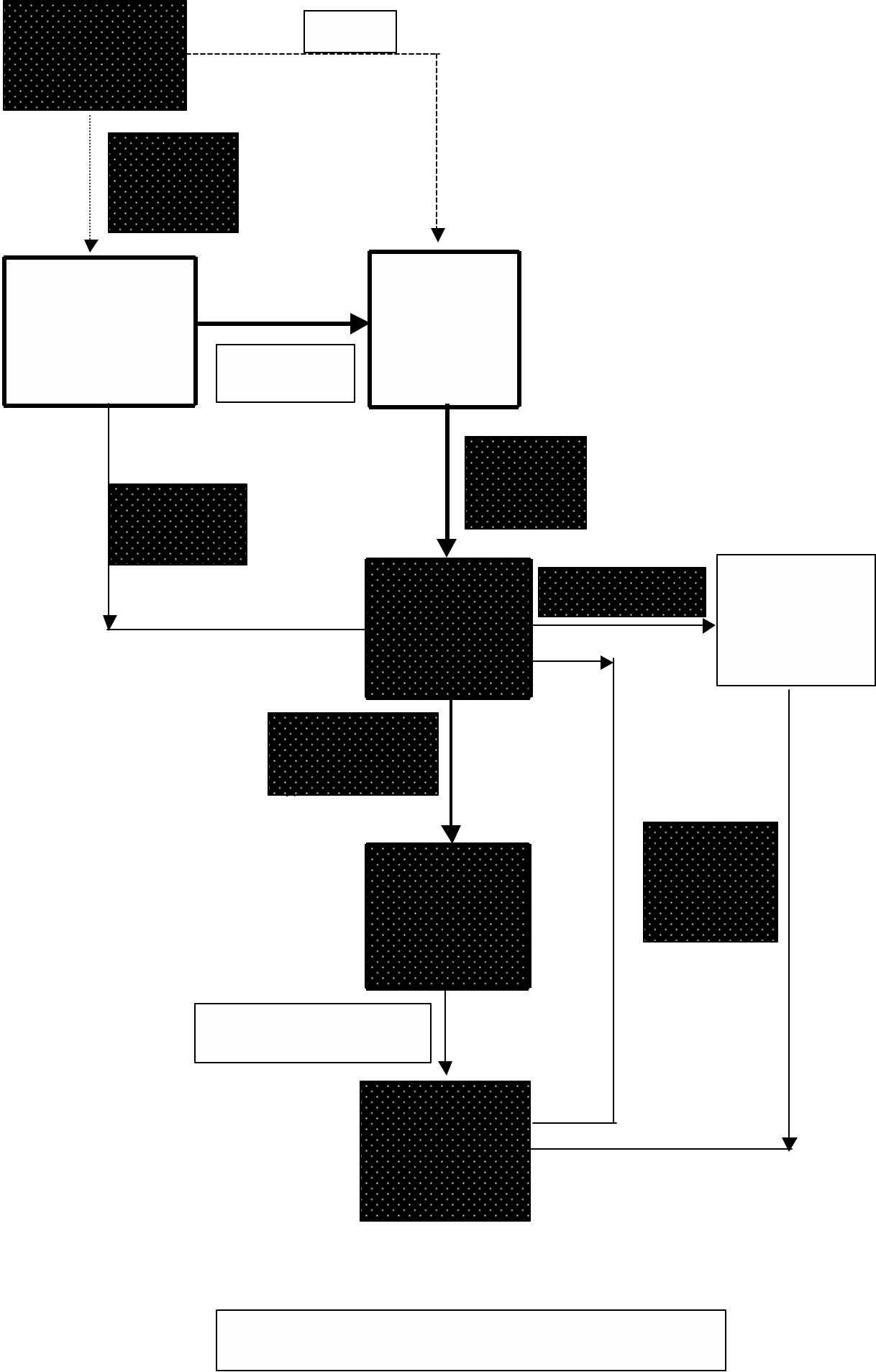
- Mueller, D.C. (1989) *Public Choice II*, Cambridge: Cambridge University Press.
- Oates, W.E. (1972) *Fiscal Federalism*, New York: Harcourt Brace and Jovanovich.
- Propper, C. (1995) Agency and incentives in the NHS internal market, *Social Science and Medicine*, 483: 1-7.
- Prychitko, D.L. & Vanek, J. (Eds.) (1996) *Producer cooperatives and labour-managed systems*, Elgar Reference Collection, International Library of Critical Writings in Economics, vol. 62: Cheltenham.
- Sappington, D.E.M. (1991) Incentives in Principal-Agent-Relationships, *Journal of Economic Perspectives*, 5: 45-66.
- Scott, A. (2000) Economics of General Practice, in Culyer, A.J. & Newhouse, J.P. (Eds.) *Handbook of Health Economics*, Amsterdam: Elsevier North-Holland.
- Smith, P.C. (1997) *Devolved purchasing in health care: A review of the issues*, The Nuffield Trust: London.
- Smith, P.C. (1999) Setting budgets for general practice in the new NHS, *British Medical Journal*, 318: 777-9.
- Taylor, D.H. & Leese, B. (1997) Partnership changes in English general practice from 1990 to 1994 NHS, *Journal of Public Health Medicine*, 19(3): 341-6.
- The NHS Plan (2000), Norwich: Crown Copyright.
- West, P.A. (1997) *Understanding the National Health Service reforms: The creation of incentives?*, Open University Press: Buckingham.

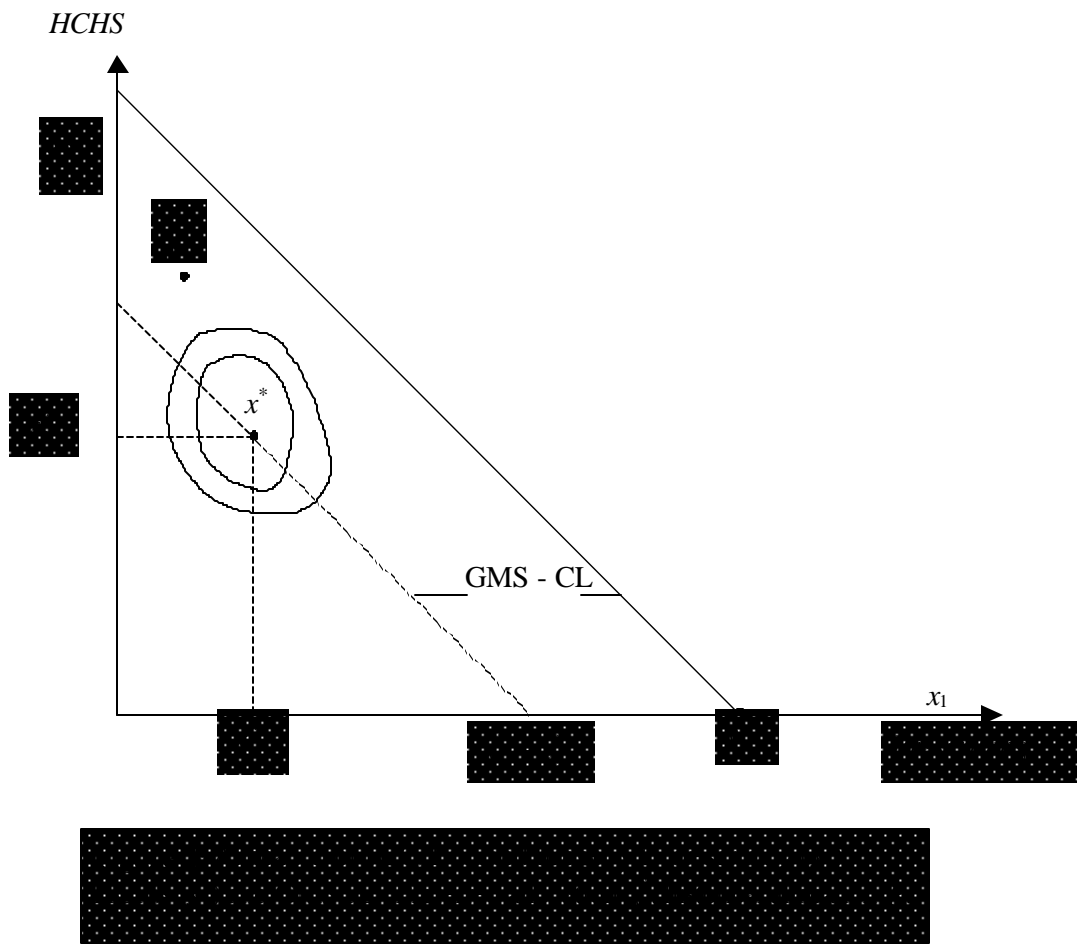
<b>Decision makers</b>	<b>Choices</b>	<b>Change in environment?</b>	<b>Relevant literature</b>	<b>Modelling issues</b>
Patients	Practice Gatekeeper Consumption of care Other health affecting activities	More gateways to care Pharmacies (OTCs), WICs, NHS Direct	Consumer theory  Demand for health	
<b>GPs</b>	Referrals Prescribing Quality of care	Re accreditation Salary contracts (PMS) Clinical governance	Imperfect agency SID P-A Labour supply	GP objectives?
<b>Practices</b>	Internal control and incentives Skill mix GMS vs PMS Choice of PCG/T	Budgets, incentives. PMS contracts. Clinical governance Nurse prescribing	P-A Worker cooperatives Bargaining Theory of firm (input mix)	Aggregation of GP preferences?
<b>PCG/Ts</b>	Commissioning: secondary, CHS, primary care (PMS). Allocation of GMS-CL. Practice budgets. Practice incentives. Clinical governance. PCG or T (I to V)	New	Consumer theory Committee decision making. Co-operative bargaining. Teams. Regulation. P-A. Transaction costs.	Whose objectives: Practices (PCGs?) Managers (PCTs?)

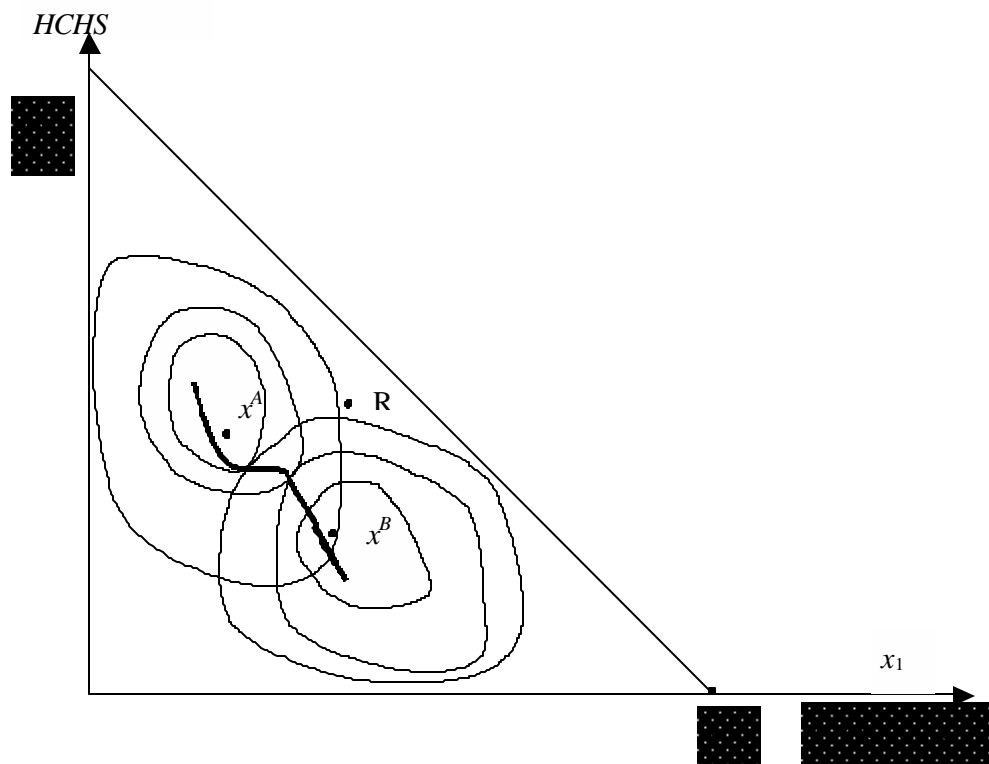


<b>HAs</b>	Budget allocation to PCG/Ts (pace of change). Accountability agreements. PMS contracts	Loss of commissioning functions to PCTs. PAF	Bureaucracy. Regulation	Whose objectives?
Providers	Internal controls and incentives. Quality and quantity of care	More commissioners with different objectives	Hospital models.	Whose objectives (managers, consultants)? How aggregated?
Politicians	Total funding. Resource allocation. Regulation framework (PAF). Pay and conditions		Public choice. Regulation.	Objectives.
Unions	Pay and conditions.		Bargaining. Trade unions	Objectives. Preference aggregation

**Table 1: The economic literature and decision makers in the NHS**







**Figure 3:** Bargaining over the allocation of PCG/T unified budget  $y$  to prescribing ( $x_1$ ), HCHS ( $x_2$ ) and GMSCL.

