dpp

DEVELOPMENT POLICY AND PRACTICE

THE OPEN UNIVERSITY

dpp

DEVELOPMENT POLICY AND PRACTICE

THE OPEN UNIVERSITY

Managing Public Sector Reform:

The Case of Health Care

by
Maureen Mackintosh

DPP Working Paper No 37

Development Policy and Practice Research Group

Faculty of Technology

The Open University

April 1997

Managing Public Sector Reform:

The Case of Health Care

by
Maureen Mackintosh

DPP Working Paper No 37

Abstract

This paper considers the problem of regulation in liberalising health care systems in a development context. Drawing on very preliminary evidence from Kerala and Tanzania, the paper argues that regulation needs to be understood broadly in health care systems, to encompass both formal and informal regulatory mechanisms. This implies that research intended to support the development of effective regulation needs to address culture as well as incentives, and to examine ways of generating mutually reinforcing forms of ethical behaviour among the whole range of participants in a health care system. The paper was written as a lecture, and retains a rather informal style and a limited bibliography in this only slightly amended version.

The conventional wisdom

There is now a conventional wisdom about a 'leading edge' set of reforms of the management of public services. The catch phrases are the 'new public management' and 'managed competition'. These concepts are being actively promoted as the right way to manage public money, and the key elements of this 'competitive state' model can be rapidly summarised:

- centralised cash limited programme budgeting at the centre, tight fiscal constraints and no off-budget expenditures (especially not on social security);
- tight targeting of social spending to primary provision and 'safety net'-style activities;
- decentralisation of management of the *use* of funds to (sometimes competing) semi-autonomous agencies, and cost centres within public services;
- contractual and semi-contractual relations between central funders ('purchasers' in some fields) and those spending the funds;
- 'market testing' of in-house public provision against external suppliers;
- use of public funds to 'lever in' private and charitable funds at agency and project level;
- floating off of financially viable activities to the private sector.

The advantages of this 'competitive state' model are generally stated to be: spending control, cost reduction, innovation at 'local' level to produce higher quality and more responsive services, and higher levels of service provision (Mackintosh 1995). While

the limited applicability of so-called 'quasi-market' mechanisms to small and cash-restricted public services in Africa used to seem evident (I remember a Tanzanian participant at a seminar in the Hague a few years ago making just that point), there are several pressures which generate a version of the 'new public management' in aid-dependent countries. Pressure from multilateral institutions to control or reduce public expenditure totals, through the use of cash limits on centrally defined and controlled expenditure programmes, has formed a key element of structural adjustment programmes. Pressure has also come from bilateral and multilateral donors, and from international NGOs, to employ aid funding within an organisational framework of semi-autonomous or autonomous projects, rather than providing support for central government programme budgets, as part of a shift away from the support of governmental spending programmes. Hence, internal government resources contributing to the aid-supported projects may need to be passed over to such project budgets.

This context can be rationalised and reinforced - indeed justified - with reference to the 'new public management' ideas, which in turn appear to offer mechanisms for handling the mix of centralisation and decentralisation, such as contractual frameworks for decentralisation, and project accounting frameworks for user charging systems. The framework then takes on a status of prescription - a management model propagated by consultants - hence my description of it as a 'conventional wisdom'. Once this happens, it takes on an aura of coherence. It is also sometimes suggested that the model improves equity, a claim rarely attempted by the model's advocates in industrialised countries.

This paper makes the following arguments about this public management model. First, it seeks to cast some doubt on its technical coherence, drawing on research on British public sector experience in recent years - and as a result perhaps to undermine some of its 'leading edge' claims. Second, it considers two economic models of some of the conflicts set up by the new public management, taken from the economics of regulation and from institutional economics. And third, it uses the institutional model to argue that the key issue which emerges from the 'new public management' reforms for regulating mixed systems of social provision is path dependency: countries make their own fate according to the kind of independent sector which is created. Countries get the social markets they deserve. Throughout, the paper uses the health sector as the main illustration of the more general argument about regulation of social sectors, because its reform is well documented in industrialised country contexts, because its reorganisation is actively under way in many developing countries, and because it poses perhaps the greatest regulatory challenges of all.

A coherent model?

From the perspective of one of the two countries that started it all (Britain, the other being New Zealand) the new public management model looks neither as coherent nor as widely accepted and consensual as the model's international proponents often imply. This paper concentrates on technical issues, but I want to start with just two more political points. The political 'clout' of the model in some countries has arisen in part from some confusion around notions of 'empowerment': the 'reinventing government' polemic, (that is, the idea that government is there mainly to create a framework for others to do things), mixes the notion of decentralisation as a necessary condition of initiative (which it is), with the notion that this is always best done through market-like mechanisms (which is a much less solid generalisation). The worst case of this type of elison has been found in concepts such as 'decentralisation to the market' (Bennett 1990).

The other point is that a key idea in the new public management is the separation of policy and management. This has immense attractions to some politicians - because it protects them from some of the political consequences of hard financial decisions, and infuriates others - because they cannot influence detail of decision making. I suggest later on that the model overstates the extent to which it is in fact possible or proper to create a policy/management split, and that this is particularly true in mixed systems requiring regulation.

The key technical issues in the new public management model centre around accounting systems, incentives and economic behaviour in response to incentives. The key move which shifts a system of public financial management from an 'old' integrated to a 'new' model is financial decentralisation. And the key problem can be quickly summarised: how active are the decentralised units expected to be? Which implies: how are they expected to manage their funds?

The reforms instituting programme budgeting in central government have often - as for example in Mozambique (Wuyts 1995) - implied very large shifts from old line-item budgeting systems to financial programming. The accounting systems at central level however remain *financial* accounting system. Their aims are to monitor spending effectively, track commitments, prevent overspending and close the accounts at year end. They are control systems aimed at budget balance. Once we go down to project level however, the purpose of accounting can become very different under the 'new' system.

Compare for example managing a budget within an integrated system, and managing funds under an internal contractual system. Under a budgetary system, including

'devolved' budgets but within integrated spending programmes, the aim of budget management locally is much the same as it is nationally: to ensure that the money is spent properly, not over-spent, and properly accounted for at year end. There is no reason why there cannot be quite a lot of devolution of rights to 'vire' funds between headings, but nevertheless, the purpose of the local budgetary system is to control spending and to feed into the higher level accounting processes.

Now contemplate the role of financial manager once the system moves over from budgets to contracts. Two things typically happen. One is that your budget acquires an 'income side'. That is, you become responsible for finding your own budget. This may mean that you have to bid for funds from a variety of 'purchasers' internally; it may mean that you are formula-funded (so much per child in your school); it may mean that you have to collect charges yourself from clients; it may mean that you have to look for charitable or other contributions to your budget. So your responsibilities as a manager change greatly even if it is not clear - as it usually is not - just what happens if you fail to meet your income targets.

The second thing that happens is that you have to start thinking about your costs in a very different way. That of course is the 'selling point' of the new public management reforms: they make managers worry about costs. You may be required to set your prices in a way which requires you to define 'unit costs' - of which more later. You may need to have more detail on costs in order to work out how much income you need and various hypotheses about how much work you may get: in other words you may need to work out break-even points and assess risk. You may have to demonstrate that you are not making a profit - despite being asked to assume more risk. The moment you have to do all this you are into management accounting rather than financial accounting. You are using accounting information to run - in some sense and even if you dislike the language - a business.

There are several problems with this move. One is that the central and local accounting categories may no longer coincide. You need a considerably more sophisticated financial information system if you are to assemble data from devolved businesses - with considerable powers to move cash around - into categories needed for financial accounting for the public activity or department as a whole. And during the year, it makes control and reporting harder. This is one thing that public managers mean when they say, as they often do, that the new public management is a very 'demanding' way to run a public sector. Trying to do this kind of complex financial management - which is not functioning particularly well in much of the reformed British system so far - as well as introducing tighter national financial accounting, *simultaneously*, may be asking too much. The reaction in Britain has

been in part not even to try to provide the kind of broken down national expenditure data which was previously available in, for example, local government.

The other main problem, furthermore, is worse. The moment people are having to compete for the income side of their activities, they have new incentives, specifically to manage financial information in a strategic manner, for competitive advantage. So there is an incentive to shift overhead costs, for example, to areas which are perceived as less competitive - easier to fund, perhaps because a local organisation has a monopoly on an expensive service - and to cover only variable costs in an area where you are having to fight off competitors (Ellwood 1996). This, of course, is how private profit seeking firms use accounting information.

This pattern of strategic accounting practice is not easily compatible with an integrated public sector financial accounting system. (It can also be seen as problematic for other reasons to which I return.) There are two possible reactions from the centre. One is to impose detailed cost accounting rules at local level. This typically results at local level in the two-set-of-books strategy: one set for public/ central consumption; one set for the information of the managers to decide how to manage the system. It also results in local organisations being decreasingly willing to provide financial information to any level of detail at all.

The second policy reaction is to 'stop paddling about in the books' of local organisations at all[1], but just to fund them as you might fund, say, non-profit organisations, on the basis of their bids. This is a logically plausible strategy: it just runs against all the instincts of politicians and top civil servants, to ask them to keep their hands off what is going on. And it means the end of integrated public financial accounts, and an end to being able to answer questions like, 'how many people does the public sector employ?' without endless qualification. What tends to happen in practice is an edgy compromise, which means lower level unit managers never really know where they are.

All decentralised public finance systems are however ambiguous, and all of them subject managers and staff to risk. Decentralisation has undoubted benefits, but risk needs to be managed at a high level, and public finance can never be guaranteed year on year. What makes the new-style systems particularly complex is the pattern of competition within them. Competition potentially brings benefits, in innovation and perhaps cost reduction. It also has several major costs: transactions processing costs, including the loss of unpaid exchange; loss of efficient externalities, especially the

[1]This phrase comes from an interview with a social services purchasing manager in British local government.

free flow of information and informal training; loss of economies from team working and scale, including efficient risk spreading; and perverse incentives for cost-raising and self-seeking behaviour.

I consider some of these costs further below. Their measurement is still little developed, and what has been done is largely on transactions processing costs including costs of contract management. It seems quite likely however that the key issues for the long term are perverse incentives and the loss of risk spreading. If this is so, then, as the next section argues, it is particularly important to understand how those problems in the public financial management system interact with the other main financial pressure on the system: how to manage the interface at local level between tax-funded income and other income, from charges, charitable funding, commercial sources, trading. And how to monitor and control, in turn, the feeding back of those influences onto the national public accounts, through commitments to repay loans, contractual commitments to match outside with public funding, and so on.

The new public management, then, sets up new tensions within public financial management, by its introduction of more 'business-like' and 'market-like' forms. It is not an integrated model, but one with some strong pressures towards fragmentation. I now turn to analysing some of those tensions.

Two models of internal public economic and financial relations

I want to contrast two economic modelling approaches to thinking about internal economic relations in the 'new' public sector. They both generate interesting results, and they overlap in some ways, but they are methodologically distinct and, probably most important, they have different implications for policy. Both are very different from 'old' neo-classical perfect-information economics.

Incentive contracting

The first approach forms the basis of the new economics of regulation, and is rooted in the principal/agent framework. The assumptions are straightforward - and problematic. They model the internal economic relations of government as a central 'principal' facing decentralised 'agents': policy departments vs. devolved agencies. Originally developed and extensively applied for the purpose of analysing the problems of regulation of private firms (as well as employment relationships) it is now being applied to government itself (Tirole 1994, Barrow 1996). The model characterises the principal as capable of defining the required outputs - but unable to achieve them by simple fiat. Instead, the principal has to contract with agents who

are self-interested, and may well not share the principal's aims at all. Furthermore, the agents typically know more than the principal - for example about their costs or how much effort they are putting in (hidden information). Worst of all, the agents may also be the only ones who know what output they are supplying: for example how good the quality of a service is (hidden action).

In this model, the government principal then has to try to specify contracts - called incentive contracts - to align the motivations of the agents as closely as possible to their own. One of the major contributions of this literature has been to clarify the distinction between high-powered and low-powered incentives in contracts. Doing violence to a complicated literature, we can say that approximately, high powered incentives shift risk to the agent, while low-powered incentives leave risk with the principal. Performance-related payments are high-powered; so are fixed price contracts. Conversely, salaries and cost-plus contracts are low-powered. Key conclusions of the literature are that, where it is hard to identify individual contributions to output, where only some objectives can be monitored, or there is no way to measure output effectively, incentives should be low-powered; furthermore, low-powered incentives are efficient where the principal is much more capable than the agent of bearing risk.

In this framework it is therefore interesting that most applications of the 'new public management' framework have moved towards higher-powered incentives. Performance pay, payment by results and fixed price year-on-year contracts are all quite common features of the new harder nosed public sector. So are contracts which, while looking low-powered, are not: for example, contracts priced at average cost, but with an uncertain case mix. (It is hard to criticise these incentives since the critics tend to be labelled politically as the soggy and sentimental party, interested in wasting public money.)

The theory strongly suggests however that the high-powered incentives embodied in, for example, competitive tendering at fixed prices, are most suitable for easily monitored and measured activities. And indeed, these have been the areas where the market-like reforms in Britain have been the least controversial: for example, contracts for grass cutting and collecting rubbish from households. Conversely they suggest that the moment you arrive at less easily specified activities - such as medical care - there is likely to be more controversy, as is indeed again the case.

The snag is that, once the outputs become hard to observe and uncertain, and hidden action (or 'moral hazard') becomes widespread, and especially once you add to this the importance of team work and commitment, then high-powered incentives can

become perverse and ethically corrupting. There are many examples from health and education. If you pay for 'output' in education - and if teachers both choose and assess students - then grades will inflate and students needing a lot of teaching will be selected out. If you pay for cheapness in health care you eventually get cheapness - at the expense of quality. If you demand achievement of measurable targets based on data which can be manipulated at agency level, eventually the agency staff will manipulate them.

Worst of all, the more people respond in this way, the less they respect the principal who is paying them. That is part of what I mean by 'ethically corrupting'. And the more playing those games works, the more people's commitment to a service for its own sake can be eroded. Note that the model is built strongly on the notion of self-interested agents, so it acknowledges that if low-powered incentives are the only suitable ones, then some other motivations - such as 'career incentives' - must be put in place. But it has less to say in practice about non-financial incentives.

There is one other interesting aspect of these models. They acknowledge that things become substantially more complicated when information is, not just asymmetric, but frankly incomplete - when no one really knows what may happen. In that case, contracts are necessarily incomplete, and the parties face the problem of who gets to decide what to do when the unforeseen circumstances arrive. Most professional working relationships are like this. Incentive theorists call this the distribution of 'control rights' between principal and agent, and have much less in the way of agreed theory about this. This is another factor which limits the application of incentive contracts in professional areas of government: there are not that many problems of unforeseen circumstances in refuse collection short of a major urban crisis; but hospitals get lots because they constantly have to use professional judgement, and public hospitals in particular generally have rather low levels of influence over the pattern of demand they face.

Institutional economics and professional norms

A great deal of the incentive contracting analysis is conducted in terms of game theory. I am not going to pursue that line of analysis here. Instead I want to use game theory to consider a distinct approach to analysing internal relations of government, derived from the institutional economics of co-operative behaviour. This approach also emerges from the industrial economics literature, but with particular emphasis on analysing corporate culture and intra-firm relationships. There are overlaps with the previous approach, but some key assumptions are different.

This modelling approach asks, not, how does one party get the other to do something, but, how can co-operation between two or more parties be induced? What mixture of rules, norms and incentives are needed? The reason I think this is promising for analysing internal economic relations in government is that what I called above the 'allocation of control rights' is an enormous problem in the public sphere. And once the problem is met, the relationships among 'contracting' parties become key to its resolution: people need constantly to collaborate to resolve unforeseen problems. Furthermore, if transactions processing costs are a problem, reducing the number of things which operate as 'transactions' through collaboration may be an answer. Finally, and I think this by far the most important point, this model endogenises norms and culture: it treats them as real forces which influence behaviour, but which are also subject to change.

Games are useful tools to think with in this field, because they can be broken down into three aspects, all of which we are interested in: outcomes or 'payoffs'; rules of the game; and motivations and expectations. (That's not the only way to classify the elements of game theory of course.) The game presented here is called an 'assurance game' (Sen 1973, Collard 1978). An assurance game is defined as a game - say between two parties - where one person will behave collaboratively if, and only if, they believe the other person will collaborate too. So the outcome depends on expectations of someone else's behaviour.

For example:

A \ B	Cooperate	Behave with self-interest
Cooperate	20 / 20	16 / 0
Behave with self-interest	0 / 16	12 / 12

This is a straightforward assurance game for two players. If A thinks B will behave collaboratively, then so will A. If A thinks B will act instrumentally, then so will A. The motivations are symmetrical, and the game has two equilibria, the top left and bottom right quadrants. Where the players end up depends on their view of each

other. If A believes that B will collaborate more than 75% of the time, for example, then A will collaborate too: the expected value for A of collaborating is higher above that probability than A's expected payoff from behaving instrumentally. The same argument works for B. The specific numbers in the matrix are unimportant, what matters are the inequalities which provide the motivation and the required probabilities. I have made the payoffs symmetrical for the two parties but that is not fundamental either - indeed in many situations asymmetric payoffs would be more plausible.

The attraction of this game theory framework is that it doesn't assume that people will behave instrumentally. It asks how people think others are motivated, what influences that and what the implications are? You can tell a quite plausible story here. It's a situation where the highest output is available if people behave collaboratively. Everyone wins. The next best collectively is for everyone to behave instrumentally - the system works, but a lot is lost, say in bargaining and litigation costs, or private payments. Worst for each party is if they try to collaborate while the other behaves instrumentally. Then the system really stops working, but what is left goes mainly to the self-interested party. That pure self-interest payoff is not large enough, though, to make this choice preferable, even for that party, to other options. (If it was, this would be a prisoners' dilemma.)

Faced with this situation, what determines expectations and behaviour? One quite plausible answer is professional norms: once you've got to the most favourable situation, the norms are self reinforcing since the alternatives are not better. But you've got to get there: otherwise you will stick at the lower level self-interested equilibrium. Professional ethics may get you to the upper left quadrant: if you believe that the probability of B behaving co-operatively is high enough, you will co-operate, and the system will settle at the higher equilibrium.

If your expectations of the others' behaviour are lowered however, you may jump first into self-interested behaviour in order to avoid being 'suckered'. So the system will 'flip'. Similarly, the higher level equilibrium would be destabilised if the individual payoff to behaving instrumentally, while the other co-operated, went up even slightly. Above 20 the game becomes a prisoner's dilemma.

But where do the professional norms come from in the first place? There are basically two individualist stories about this. One says, the professional norms may have emerged because people have worked together a long time and developed expectations which allow them to trust each other to collaborate. This *can* work even for a prisoners' dilemma situation. Individual agencies may get to know each other

and develop a reputation for being collaborative. Then they may continue to collaborate, even if the one-off individual payoff from suddenly turning uncooperative arises: you can only cheat once without losing your reputation, so you don't do it. The long term expected benefits of collaboration can outweigh the one-off 'temptation benefits' if reputation affects the other party's behaviour, even in a prisoner's dilemma game.

The other story about professional norms is much less orthodox in economists' terms. It says that people do not have only one kind of rationality. They may have both an instrumental rationality and other perfectly reasonable ways of deciding: one name for another way is 'expressive rationality' (Hargreaves Heap 1988). This means that what one does may be expressive of the kind of person one wishes to be or to present oneself as. Actions may be expressions of a sense of self, an assertion of self worth.

The attraction of this kind of line of thought is that it allows a way of thinking about social norms which is somehow more satisfactory than the very thin philosophical basis of most economics. Economics takes the split between action ('agency' in the philosophical sense) and structure as more or less complete; game theory treats rules of the game (e.g. 'both players choose simultaneously and cannot make binding agreements') as structure and strategies as choices or actions. But in the real world, and in experimental situations too, people are social beings from the start: there is no asocial starting point. Behaviour draws on shared meanings which are constitutive of both the people and the situation, even though actions change them.

There is nothing unusual in game theory about the assurance game having two equilibria. Many games have multiple equilibria, and it may be hard to see why one or other is chosen. The answer seems necessarily to lie in the world outside the game; or culture that forms part of the calculation. Shared meanings, shared cultures, help people choose one outcome rather than another: there has been a certain amount of thinking in game theory about the way in which such 'focal points' from common culture solve games (Mehta 1993).

So how can this be applied to the internal workings of government? To explore this, I want to turn to a more specific reflection on health services.

Modelling health services decentralisation

The standard model of health services decentralisation is to provide a clinic, a district, a hospital, with funds - increasingly called contracts, not budgets - and to agree outcomes, targets, performance measures of some kind for those units, on behalf of patients and potential patients. (I come back in the next section to other sources of

funding.) The format of the contract is likely to be 'block' - so much for offering a given type of treatment to those requiring it (as in accident and emergency, or fairly predictable surgery); 'cost and volume' - so much up to a maximum, then payment per case over that; or cost per case - on various possible payments bases. These are however notional categories into which it is quite hard to fit real contracts.

The two models above treat these contracts quite differently. The first model views the contracts as transactions among self-interested parties and asks, what are the incentives? The second model asks, to what extent are these *transactions*? In other words, do we understand - do we wish the participants to understand - these relations as exchange? It is in the nature of our understanding of exchange in the twentieth century's economy that we see it as something-for-something: in this case, cash for performance. We also invest 'exchange' with meanings about the ethics of strategic and arms-length behaviour: it is acceptable to behave instrumentally in exchange, but not in relation to - whom? Parents? Children? Vulnerable patients? Those *funding* vulnerable patients?

To construct the funding relation as a contract therefore invests it with new meaning. This may be good or bad, depending on the prior meanings and the shared understandings of the change. The danger perceived in British debate has been particularly in areas where service provision has relied on a collaborative culture (people taking responsibility for the outcome of team working), commitment to patients (e.g. unpaid overtime), and resistance to temptations to an easy life (not dodging responsibilities) - despite the manifest inequalities of the work place. If you then change the deal by, for example, bringing in managers to make decisions previously seen as professional decisions, managers furthermore who are well paid, have a more openly cynical culture and are denigratory of the skills of the lower paid, then you can break people's expectations of each other - and perhaps end up with a more expensive and less productive equilibrium.

While much of the political debate in Britain has been about the relations between clinicians and managers, there is also a debate about the provider/ funder relations. These have been newly invested with contractual or transactional status, and as a result have become more oppositional. This can clearly have good effects: demanding information about what is happening to all this public money seems a good thing. But there are acute difficulties in specifying performance measures which do not provide perverse incentives for behaviour. Perverse incentives - to keep people too short a time in hospital, to find ways to count patients twice - can induce cynicism and instrumental behaviour. And instrumental behaviour in turn can raise costs.

There are signs that this lesson is being learned, and here there is one overlap between the models. New contracts are increasingly being written in Britain in more incomplete form: not, here is what happens if demand is greater than expectations, or certain targets are not met, but, these are the circumstances which trigger a meeting to discuss the problems. The first model would specify such 'relational contracts' as likely to emerge, but say little about their content. The second model would focus attention on the relationships involved: how is the sharing of risks and problems sustained?

It is also possible to apply these models to the clinician/ patient/ citizen relationship, with the patient as 'principal'. The first model however has little to say about public provision which is free at the point of use: it does not really perceive a contract where no money changes hands. It does not credit the patient as an active partner. The second type of model is sometimes used to develop the idea of 'co-production' (Ostrom 1996): the notion that public services are in fact a joint product between the user and the clinician (in the health services case). This seems a very promising line of thought, though it has not been developed empirically in any depth (so far as I know) except for the case of urban services provision and maintenance. One could then interpret the top left equilibrium in the figure above as a collaborative relation between clinic and population which maximised the public health effects of the resources; the bottom right is a more distanced and instrumental level of provision and behaviour: not disastrous, but much less effective. (One could of course have much worse circumstances in the low level equilibrium, with very asymmetrical payoffs.) Instead of pursuing the analysis of public sector decentralisation, however, I want tentatively to apply these ideas to private charging and privatisation of provision.

Modelling privatisation

The introduction of mixed funding into health services which were previously public-dominated always involves forms of privatisation of provision as well as payment. And formal changes are often preceded by informal shifts towards payment which change perceptions of the relations among participants. In this case, the comparison between Tanzania and Kerala offers some suggestive ideas.

Kerala: relatively benign privatisation?

Kerala in South India is a state with, as is well known, a very creditable health record. The mortality statistics stand out in Indian and international terms: Kerala has a much lower rate of infant mortality and much longer life expectancy than would be

expected from its low income levels, and despite relatively high levels of morbidity (illness). This undoubted success is widely attributed to good primary health care, and to high levels of female education and involvement in the primary health care system - so that the primary care interacts effectively with good child rearing practices which include effective use of clinics and advice. Kerala has large numbers of doctors per head of the population and a good geographical spread of clinics, although some groups including people in tribal areas and coastal villages are disadvantaged.

Over the last fifteen years however, Kerala's health care system, which was public-dominated, has been under great fiscal pressure. Although India has only recently been going through adjustment and liberalisation at national level, the financial squeeze at the state level has been growing, as has the diversion of available public funds from mainstream service programmes into centrally mandated semi-permanent relief programmes. This squeeze has been associated in Kerala with two visible trends, to be investigated in more detail: a tendency to protect wages and salaries at the expense of supplies and maintenance, and a rise in informal payment for services, in the public sector; and a rapid increase in private sector provision. A quite substantial proportion of Kerala's primary and secondary health care is now privately provided.[2]

What is interesting and justifies considerably more investigation, is that the health effects of this shift do not seem to have been negative. There is no evidence in the indicators so far of a decline in health status or a rise in mortality. This seems so far, in other words, to have been a relatively benign privatisation which has not polarised access to care.

There are a number of possible, hypothetical explanations. The traditional one (following for example the line of thought in many World Bank documents) would be to expect that public services which remain have focussed on the poor, with those who are better off using private facilities. But this is an unsatisfactory hypothesis on a number of counts. The state has a low income population, so there is no sufficiently large distinct middle class to turn wholly to private medicine. And private provision exists in rural as well as urban areas in this densely populated country. Anecdote fastens on the dual effects of high expectations of decent treatment on the one hand, and a large number of competing doctors on the other, both embedded in local

---

[2]I owe some background information on the Keralan health care sector to discussions with Dr D. Narayana, of the Centre for Development Studies in Thiruvananthapuram, Kerala. He is not however responsible in any way for my discussion here. All errors of fact and interpretation are my own.

society, which together make low quality, casual or exclusionary treatment by doctors much less likely.

In other words it is at least a reasonable hypothesis that the strengths of the previous health care system have produced a pattern of *informal regulation* of private medicine in Kerala which limits exclusion and instrumental behaviour. By 'informal regulation' I mean a mix of local public pressure, expectations, competition, information and perhaps publicity which together imply that the system functions in a 'co-operative equilibrium' rather than a more polarised and instrumental one.

Suppose, just for the sake of argument, that this is what is happening. (The evidence is not yet available to support or undermine the hypothesis.) Then it is very hard to model it - for example to specify research and policy questions in detail - without bringing in ideas of culture and meaning - in other words, without turning to the second framework above. The natural way to start seems to be to investigate the meanings patients and local publics give to the relations with clinics and doctors. To what extent is it perceived as a transaction or a collaborative relationship? How does the form of payment influence that? Are there ways in which the 'transaction' element of the relationship is minimised? (There are quite a few anecdotes about that, for example, about payment being socially played down, not discussed directly as part of the consultation.) What do people expect to happen if someone cannot pay?

One could then use these kinds of questions to frame an understanding of the answers to more specifically transactional/ contractual questions. What is the payment system and how does it share risk between patient and clinic or doctor? Is there cross-subsidy among patients? What are fee levels and how are they changing? One would like to be able to offer a (plausibly well supported) explanation of fee levels, distribution and changes which included culture and expectations as explanatory elements.

Suppose for a moment that one could develop an analysis of reasonably stable informal regulation on the basis of such research. Then it would have some fairly distinct implications for policy. One might be that the contractual elements should only be 'hardened' - fee structures made more formal or more closely related to provision - after careful thought about the cultural/ informal-regulatory implications. Policy should address the system as a whole, and balance the gains and losses of informal regulation and cross-subsidy against more formal systems of accountability. The Keralan detail in that discussion is hypothesis based on anecdote and some background evidence. It is intended as a relevant peg on which to hang the argument

that informal/ cultural issues matter in health reform. Health care systems are deeply affected by culture, and very path-dependent.

Paths through privatisation in health care: some Tanzanian questions

'Path-dependency' is used in a variety of different intellectual contexts, but always with the basic meanings that the starting point, and initial direction, of a system strongly influences its continuing characteristics. The health services literature tends to emphasise that health care is very culturally embedded, and hence health care systems cannot be easily transferred from one country to another. Therefore, when there is a major change in a system - such as privatisation - there are, first, only so many ways it can happen; and second and perhaps more important, once it is launched on a particular path then that deeply influences subsequent opportunities.

The concept of informal regulation helps to think about this. Different patterns of privatisation have different patterns of both formal and informal regulation. For example, it is commonly accepted, at least among health economists, that the single most destructive private health care financing system in the long run is private individual insurance for limited types of hospital care. This is because of its second-round effects on the health care system (Bennett and Tangcharoensathien 1994). It tends to raise costs sharply; to produce lots of perverse incentives for further cost escalation; to raise medical incomes; to focus hospital provision on those types of care funded; to undermine primary care provision and gate-keeping; and to create access to medical care as an intensely divisive social issue. Perhaps worst of all, it creates a political/ financial lobby against socialising health care risk and uncertainty, since it is focussed on narrowly defining risk categories. One way of summarising that list would be to say, private individual insurance is very unlikely to be complementary to any more inclusive system. Its patterns of informal regulation are more likely to be destructive of parallel public or non-profit systems - and this is partly because of its embedded notion of health care as a consumption good.

At low levels of income per head, no health system is egalitarian - and there are limits even in Scandinavia to the achievement of equality of access and treatment. But the notion of informal regulation suggests that it is possible to create paths to a more equitable system by building in (or not losing) all the possible mechanisms available, sucha as social openness, cross-subsidy, risk pooling, cost ceilings and rights. And by ensuring that institutions created are in principle generalisable if incomes rise sufficiently. What those mechanisms and institutions are can be suggested, but the 'path' depends very strongly on culture and context - that was the point of the Kerala story.

So can one apply this way of thinking to health sector reform in Tanzania? What follows are a set of tentative arguments for discussion, which might form the basis for research.[3] Tanzanian health sector reform proposals combine decentralisation - which has meant that the public funding for health care below regional hospital level now goes to local governments - with 'cost-sharing' and privatisation (URT 1994, 1996, World Bank 1996). Donor funding is very important for preventative and primary care, and while some of it goes to the sector via the government budget, some is spent directly by donors. Mission hospitals form a substantial element of hospital and rural dispensary provision, and draw on donated funds; some are designated district hospitals, funded by the government to provide district hospital services (Mujinja and Mabala 1992, Mujinja 1995). Non-designated voluntary hospitals have long charged fees; user fees have now been instituted in the government and designated hospital sector (Mushi and Abayo 1995, Mujinja 1996). Entry of private for-profit providers has been liberalised, and there has been a sharp increase in commercial provision as a result, concentrated particularly in Dar es Salaam and in the dispensary sector (Munishi 1995)

These large changes, following a history of declining quality of provision and falling morale in the sector, pose a considerable public policy challenge. If one takes path-dependency seriously, then it follows that the next few years will set up some of the initial conditions, in terms of investment and behaviour in the independent sector(s), and also in terms of changing public health spending behaviour, which will bring consequences for many years to come. That in turn implies that policy matters, since policy may influence those initial conditions.

Let me illustrate different ways of thinking about the policy problem, drawing on two important aspects of the reforms. The first aspect is the growth of the independent sector. There are various patterns of private provision and private health care market which might emerge, and which these will be is as yet not determined. But what seems certain is that diversity will increase; that both commercial and diverse non-profit provision will co-exist in primary and secondary care; that commercial investment will rise sharply in the early stages of reform; and that donors will continue to fund some preventative and community health care provision including direct delivery via local NGOs. There is also likely to be a range of proposals for private and non-profit insurance systems, some of them provider-based.

---

[3]Some of these arguments have been developed in discussions with Dr Paula Tibandebage, of the Economic and Social Research Foundation, Dar es Salaam, Tanzania. Responsibility for the use of those discussions here lies entirely with the author.

All of these developments raise alternative possible roads through privatisation. Insurance may be more or less individual; more or less employer-based; wholly voluntary or compulsory for some categories. Private providers may be more or less interested in quality, more or less focussed on profit-seeking. Competition may emerge which is price-based or quality-based in different segments of the market. Perhaps the greatest danger, illustrated by 'polarised' Latin American systems (Frenk 1993), is the solidification of low quality - and not really very cheap - private provision for the poor and middle-income population, and of very expensive private provision for the well-off generating very high medical incomes. Once solidified, the former is hard to break, and the latter impossible ever to generalise (a point illustrated by the failure of the recent U.S. health reforms).

Trying to prevent this kind of polarisation requires public intervention. The form of the intervention in Tanzania will be influenced by the second big set of changes: decentralisation in management of public health care funding, associated with user fees and other pressures to raise more funds at local level. Effective decentralisation of responsibility for management seems not to have occurred yet, since the use of funds is largely still mandated centrally, and there are also many vertical programmes in health care to which a decentralised system will have to relate. There is also very limited understanding of the reforms at local level as yet, and very limited capacity to manage a decentralised budget - let alone to regulate at local level a mixed system in which independent providers are becoming increasingly influential.

How then should we think about public policy development in these circumstances? There seem to be two possibilities. One is to think of the commercial/ independent providers and public provision as effectively two independent sectors, and hence to see the public policy problem as largely one (and this is large enough) of providing directly some basic primary and preventative care, leaving the rest to an emerging independent sector for which some basic regulatory rules (such as licensing of practitioners) also need to be set centrally. That, it would not be too far out to suggest, is the World Bank/ IDRC project approach (World Bank 1996, TEHIP 1995, 1996).

The alternative is to think of public policy as seeking to use the limited public resources as leverage to influence the direction and effectiveness of formal and informal regulation in the health care sector. This involves a different vision of regulation: thinking of regulation not simply as formal rule-setting, but as a system of both written rules and cultural expectations, enforced by a mixture of legal threat and broad social acceptance. Formal rules such as licensing arrangements are not (of course) necessarily observed: rules are followed in some contexts and not others.

Formal regulation indeed only works when it enjoys a large measure of prior acceptance - and preferably a positive lobby in its favour.

Behind this image of regulation is an assumption that people operate on a mixture of motives - we might call them materialism and professionalism. The aim of regulators is then to try to reinforce the professionalism (the 'expressive rationality' referred to above) and to get the materialism operating in its proper place.What might this mean in practice? Here are some examples. The general health economics literature suggests that health care systems work best - and most cheaply - when they have effective 'gate keeping' through a primary care system. Suppose then that one aim of regulation was, over time, to develop the private dispensary network in the urban areas - growing fast through private investment - into such a system? What would be involved in achieving that? It cannot be done through close formal regulation of each provider: that would mean far too many resources, though regulation might include unsignalled inspections with a threat of withdrawal of a license. Involving the public seems one relevant approach. Can the urban population be turned into much more actively critical users of the services: can expectations be raised and people encouraged actively to complain? Can public/private alliances be constructed with good dispensaries to establish norms for treatment and charging, which can then be effectively publicised? Can the sector be made to some extent self-policing - perhaps through a professional association structure - on the basis that good dispensaries will make a reasonable income, and have an interest in driving out undercutting by low quality providers?

As a second example, suppose that one effect of reform - as in Britain - is to reduce the information available for public policy. And that a second effect is to threaten training provision, since the private providers may not think it their job to train. Both of these effects are quite likely, since information has strategic uses, and it is in the interest of the commercial sector to 'poach' staff. In these circumstances, the government may need to take a negotiating approach to the use of its own funds and legal powers. One way to gain access to information is through explicit joint ventures, where government funding - recognised to be limited - gains specific commitments to training, openness, treatment of those unable to pay, and collaborative behaviour over time. While the current designated district hospitals have a contractual agreement and mixed funding (Mujinja 1995), the current contract may be chiefly seen as a constraint (e.g. on charging and the use of mission funds); it may be that a better contract could be designed, explicitly aimed at getting the best public benefits available from mixed funding - and generating regulatory information.

More generally, this approach suggests that reform should aim to reduce - or not to create - perverse incentives - and to minimise the reworking of relations within the system as transactions. Transactions are expensive, and encourage self-seeking behaviour. They are bound to increase substantially with liberalisation, but there is a choice over how much they increase and the form they take. It is widely believed in Britain that three problems with the National Health Service reforms are high transactions costs, perverse incentives to massage data (expensively) because performance targets are set on throughput indicators, and an over-emphasis on material incentives which has angered and offended staff. The specific problems will be different in Tanzania, but these kinds of dangers are unlikely to be absent.

Most generally, this approach suggests that the best way to see a developing mixed health care system is as itself a joint venture where the boundaries between sectors are deliberately 'blurred'. Health care privatisation tends to be investment-led: most of the new facilities and equipment will be in the independent sector. But the government retains some revenue funding and legal (and political) powers which it can use to influence the system very strongly. Regulation needs to be three-way: government, providers and public, orchestrated by the government but also perhaps by providers' associations and by campaigns and lobbies for patients. For example, the government can use strong sanctions against bad providers if it works closely with good providers to develop acceptable regulation. On insurance, since providers inevitably engage in price discrimination, they strongly influence access, and hence should be involved in all experiments with prepayment and mutual risk pooling. Too sharp a funding/ providing distinction is probably undesirable in contradistinction to the 'conventional wisdom' model. If the public are involved in local health funding schemes, can the same organisations be used to scrutinise the quality of care received? Can non-material incentives - especially professional and public recognition - be increasingly built into the system alongside reasonable payment for work done? Can unpaid sharing of facilities and information be positively encouraged?

Conclusion: supporting benevolence?

'in any economy, the financial structure is continually reshaped by the efforts of the spending units to break out of the confines of existing financial arrangements.' (Gurley and Shaw 1960 p.50)

Once a public financial management system becomes decentralised, the 'spending units' take on a dynamic of their own, especially if enjoined to go out and seek additional funding. This rather obvious statement is however still not embedded in

the public finance literature, which tends to see public finance as rather passive in economic terms, with the exception of the fiscal balance. The new economics of regulation of social provision - still really to be created as distinct from the utilities sector work - needs to come to grips with the scope for influencing the very diverse behaviour of decentralised and autonomous providers. This paper is intended as one small contribution to this development. Its main focus is to argue that 'regulation' needs to be understood very broadly in health care, to encompass both formal and informal regulation, and that research needs to address culture as well as incentives.

Once one sees regulation in this way, the policy/management divide inherent in the new public management starts to look particularly problematic, since the system can only be effective if providers take ethical and practical responsibility for the quality of the system: the (government) regulators cannot do it for them. Policy needs to reinforce desirable behaviour at the provider level, and to recognise that government needs to open up to outside pressure and informal regulatory behaviour from other members of the system. James Meade remarks somewhere that economic systems work best if they 'economise on benevolence' by working with the grain of self-interest where possible. In health care, this is important, but so is a recognition that a large portion of 'benevolence' is the only way for the system to work effectively. Mutually reinforcing ethical behaviour is harder to create than to sustain once created - that was the point being made by the assurance game. The challenge for policy in a liberalising health care system is to try to encourage ethical and professional ways of working and institutions which will sustain them, piecemeal, but with a sense of direction over time. Hence the emphasis in this paper on path dependency.

References

Barrow, M. (1996) 'Public services and the theory of regulation' Policy and Politics Vol. 24 No.3 pp.263-276

Bennett, R. (1990) 'Decentralisation, local governments and markets: towards a post-welfare agenda' in Bennett, R. (ed.) Decentralisation, Local Government and Markets Clarendon Press, Oxford

Bennett, S. and V. Tangcharoensathien (1994) 'A shrinking state? politics, economics and private health care in Thailand' Public Administration and Development Vol. 14 p.1-17

Collard, D.(1978) *Altruism and Economy: a study in non-selfish economics* Martin Robertson

Ellwood, S. (1996) *Cost-based Pricing in the NHS Internal Market* Chartered Institute of Management Accountants

Frenk, J. (1993) 'The public/private mix and human resources for health' Health Policy and Planning Vol. 8 No.4 pp.315-326

Gurley, J.G. and Shaw, E. (1960) *Money in a Theory of Finance* Brookings

Hargreaves Heap, S. (1988) *Rationality in Economics*

Mackintosh, M. 'Competition and contracting in selective social provisioning' European Journal of Development Research Vol. 7 No. 1 pp. 26-52

Mehta, J. (1993) 'Meaning in the context of bargaining games: narratives in opposition' in Henderson, W. Dudley-Evans, T. and R. Backhouse (eds.) Economics and Language Routledge

Mujinja, P.G.M. (1996) 'Service Provision in District Designated and District Hospitals in Tanzania: Comparison of Quality, Performance and Unit Costs: a research report' mimeo, Institute of Public Health, Muhimbili University College of Health Sciences, Dar es Salaam February

Mujinja, P.G.M. (1995) 'National study: the relationship between the government and mission hospitals in Tanzania: a research report' mimeo Institute of Public Health, Muhimbili University College of Health Sciences, Dar es Salaam November

Mujinja, P.G.M. and Mabala, R. (1992) 'Charging for Services in Non-governmental Health Facilities in Tanzania' Bamako Initiative Technical Report Series No.7, New York UNICEF September

Munishi, G.K. (1995) 'A need to plan for the growth of private health care quality and equity in Tanzania' Paper presented at an Executive Policy Seminar on Health Sector Reform, Dar es Salaam, March

Mushi, R. and Abayo, A. (1995) 'Reflections of the cost recovery programme: initial observations' Paper presented at an Executive Policy Seminar on Health Sector Reform, Dar es Salaam, March

Ostrom, E. (1996) 'Crossing the great divide: coproduction, synergy and development' World Development Vol. 2 No. 6 pp. 1973-1088

Sen, A.K. (1973) *On Economic Inequality* Clarendon Press Oxford

TEHIP (Tanzania Essential Health Interventions Project) (1996) 'Guidelines for preparation of core research protocols' mimeo Dar es Salaam July

TEHIP (1995) 'The Essential Health Interventions Project (EHIP)' Background Paper mimeo Dar es Salaam October

Tirole, J. (1994) 'The internal organization of government' Oxford Economic Papers Vol. 46 pp. 1-29.

URT (United Republic of Tanzania) (1996) 'Health Sector Reform Plan of Action (1996-1999)' Ministry of Health, Dar es Salaam May

URT (United Republic of Tanzania) (1994) 'Proposals forHealth Sector Reform' Ministry of Health, Dar es Salaam December

World Bank (1996) 'Health Policy in Eastern Africa: a structured approach to resource allocation' Vols I-IV Eastern Africa Dept., Africa Region, April.

Wuyts, M. (1996) 'Foreign Aid, Structural Adjustment, and Public Management the Mozambican Experience' Development Policy and Practice Working Paper No. 36.