

**MANAGING THE PUBLIC SERVICES**

**Series editor:** Alan Lawton, Professor of Organisational Ethics,  
Teesside Business School, University of Teesside.

*Current titles in the series:*

George Boyne *et al.*: *Evaluating Public Management Reforms*  
Paul Joyce: *Strategic Management for the Public Services*  
Alan Lawton: *Ethical Management for the Public Services*  
John Wilson (ed.): *Financial Management for the Public Services*

**Evaluating public  
management reforms**  
Principles and practice

**George A. Boyne**  
**Catherine Farrell**  
**Jennifer Law**  
**Martin Powell**  
**Richard M. Walker**

**Open University Press**  
Buckingham · Philadelphia

7

are actually adopted are accurate reflections of the theoretical principles. Third, are the reforms put into practice? Even if the reforms are theoretically appropriate, there may be a substantial gap between the policy that is formulated and the policy that is implemented. Finally, are the effects of reform consistent with theoretical expectations? This last question raises a host of methodological issues concerning the establishment of cause and effect. We show that the severity of this 'attribution problem' varies with the way that a reform is designed and put into practice.

In Chapters 4–6 we apply this framework to public management reforms in health, education and housing in the UK. Each chapter follows a common format. First, to what extent have changes occurred in the extent of competition, the number and quality of performance indicators, and the size of service delivery organizations? In other words, has reform actually occurred, and is it in the direction prescribed by public choice theory? How different are the 'pre-reform' and 'post-reform' structures and processes? Second, what have the consequences been for efficiency, responsiveness and equity? How have these variables changed? Are the changes in the direction predicted by public choice theory? To what extent can better or worse performance be attributed to changes in market structure, performance reporting or organizational size?

In Chapter 7 general conclusions are drawn from the analysis of reforms in health, education and housing. What are the problems of establishing the extent and impact of public management reforms, and what are the prospects for solving these problems? Have the reforms been a success or a failure, and is public choice theory a suitable basis for redesigning the public sector? Under what circumstances are public choice prescriptions likely to succeed or fail?

In a sense, this book is a series of case studies on three levels. First, it is a case study of a particular theoretical perspective on the need for public management reform. An investigation of alternative theoretical perspectives might yield different evidence and conclusions. Second, it contains case studies of three specific types of reform. Again, different reforms might produce different signals about success or failure. Third, there are three case studies of parts of the public sector in the UK. It is conceivable that marketization, performance indicators and changes in organizational scale have had different consequences in other areas of service provision and in other countries. For all these reasons, our conclusions cannot be sweeping or comprehensive. Nevertheless, our evaluative framework does provide a firm basis for judging whether the reforms have worked as expected, and for assessing the potential impact of such reforms in the future.

## TWO

### Criteria of evaluation

The current Labour government in the UK claims to be open-minded about methods of providing public services. Traditional dogma has, so we are told, given way to the 'modern' maxim that 'what counts is what works' (Davies *et al.* 2000; Newman 2001). But does this mantra, in itself, provide a clear basis for judging different approaches to the management of public services, or is it all style and no substance? An answer to the apparently simple question whether a reform has worked turns out to be very complex. This question comprises numerous sub-questions, all of which raise thorny conceptual and empirical issues. For example, what criteria should be used to evaluate reforms? What information is required to apply these criteria? What counts as 'good evidence' on the impact of a reform? What methods can be used to disentangle the separate impact of a reform from other influences on its intended consequences?

In this chapter and the next we develop a set of principles for evaluating the consequences of public management reforms. In the present chapter we identify criteria for evaluating reforms. Here we explore the meaning of public choice theorists' preferred measures of policy impact: efficiency and responsiveness. We also analyse different interpretations of the 'missing' criterion identified by critics of public choice theory: the equity or fairness of public service provision.

In Chapter 3 we consider fundamental questions that are seldom given much attention in the literature on public management reforms. Did reform actually take place? How different were the arrangements for public services before and after the supposed reform? Even if radical reforms were adopted, to what extent were they actually implemented? If only minor changes occur in the design of public services, and if even these are implemented only in part, then there is little reason to expect significant shifts in efficiency, responsiveness and equity. Politicians and their officials may have an interest

in 'talking up' the scale of a reform, and in emphasizing that their policies differ substantially from those of their predecessors. It is important to be sceptical about such claims, and to assess carefully whether public management reforms have really occurred.

Later on in Chapter 3 we identify the problems of establishing whether organizational performance in the public sector has improved or deteriorated, and whether this is the result of reforms. Have movements in efficiency, responsiveness and equity been positive or negative? This raises issues about the existence of a baseline for monitoring such movements (progress is difficult to judge if there is little information on the 'pre-reform' level of performance), and the availability of information on the 'post-reform' trajectory of change. We also analyse questions of cause and effect, or 'attribution'. Put simply, is it possible to establish 'beyond reasonable doubt' that any observed changes in performance can be attributed to reform rather than other variables (e.g. changes in social and economic conditions, or changes in other aspects of government policy)? It is possible that efficiency and responsiveness have improved in spite of, rather than because of, the reforms imposed on public sector organizations. Throughout Chapter 3 we deal with such issues at a fairly broad level, in order to clarify the general principles that we are proposing for the evaluation of reforms. More specific and detailed questions will be addressed in Chapters 4-6 when the evaluative principles are applied to the reform of health, housing and education.

### Criteria for assessing the impact of public management reforms

Evaluation research is bedevilled by a lack of agreed criteria for judging policy consequences (Hogwood and Gunn 1984; Pollitt 1995). This reflects wider problems in the evaluation of organizational performance (Hall 1996). There are many dimensions of success or failure, and many different stakeholder groups who place different weights on these dimensions (Connolly *et al.* 1980). For example, service users may pay most attention to service quality and quantity; staff in public organizations may be more concerned with *how* a service is provided (including their own working conditions) than with *what* is provided; and taxpayers are likely to place as much emphasis on cost as on effectiveness.

A variety of methods for establishing evaluative criteria have been proposed. One is to assess reforms against the goals of policy-makers, as expressed in legislation, government documents and other pronouncements (such as speeches by government ministers). However, these can be vague and conflicting, and may reveal only the official objectives of public management reforms. For example, the stated intention of compulsory competitive

tendering (CCT) in UK local government in the 1980s was to improve the efficiency of local service provision, but it has been argued that the unstated intentions were to break the power of local government trades unions and cut the size of the public sector workforce (Cutler and Waime 1997). Another approach is for evaluators to use their own criteria to judge the success or failure of public management reforms. In the specific context of our analysis, the implication is that the five authors of this book should develop their own list of evaluative criteria. This method, too, is problematic: it could be idiosyncratic and subjective, and would encounter all the problems of multiple dimensions of performance and multiple perceptions of their relative importance.

A third approach to the development of evaluative criteria deals directly with the latter set of problems. This 'multiple constituency' method consists of three steps: identifying all of the relevant stakeholders who are affected by a public management reform; eliciting their evaluative criteria and the relative importance that they attach to these; and attaching weights to the preferences of each stakeholder group. These steps are all very difficult and likely to be politically controversial. Which stakeholders should be included? Do they have clear preferences, and can they rank them in order of importance? What weights should be attached to the views of the various stakeholders? One approach to this last question is to use their relative power to alter the behaviour of organizations that provide public services, but how is power to be measured, and by whom?

In sum, all of these approaches to the identification of evaluative criteria pose as many problems as they solve, and are impractical in various ways. Moreover, for the purposes of the analysis in this book, they tend to miss the point of evaluation. This is because we are undertaking a *theory-driven* evaluation. Therefore, the relevant criteria are those stipulated by the theoretical perspective that is adopted, which in this case is public choice theory. The key writers in this tradition have consistently advocated two principal performance improvements that are expected to follow from the reforms which they advocate: efficiency and responsiveness (Boyne 1998a; Mueller 1989). These criteria, which are explained in more detail below, provide the basis for an 'intrinsic' evaluation of public choice reforms. In other words, improvements in efficiency and responsiveness are the appropriate criteria for judging public choice theory on its own terms, because they have been selected by the advocates of this approach to the design and delivery of public services.

Nevertheless, an intrinsic evaluation of public management reforms is too narrow. An exclusive focus on intended outcomes effectively loads the evaluative dice in favour of the theoretical perspective that has been selected. A theory can appear to be an appropriate prescription for change because it only considers a restricted range of (supposedly) positive outcomes. What about the unintended consequences of reform that are overlooked by

public choice theorists? An answer to this question requires a broadening of focus to 'extrinsic' as well as intrinsic evaluative criteria. Yet, this wider perspective must not become a loss of focus: in principle, a long 'shopping list' of extrinsic criteria could be added to the intrinsic criteria of efficiency and responsiveness. In order to retain the emphasis on theory-driven evaluation, and to make the task manageable, it is essential to consider only those unintended outcomes which themselves have a strong theoretical basis. The most sustained and coherent critique of public choice reforms is that they have an adverse effect on the equity of public service provision (e.g. Self 1993; Udehn 1996). We have, therefore, added this extrinsic evaluative criterion to the two intrinsic criteria of efficiency and responsiveness, and now proceed to consider each of these yardsticks for the assessment of public management reform.

### Efficiency

The idea that public services can be produced more efficiently has been widespread since the expansion of the modern state in the early twentieth century (Downs and Larkey 1986). The distinctive twist added by public choice theory is that public organizations are *inherently* and *massively* inefficient. The root problem is regarded as the monopoly power of public officials. Niskanen (1971) argues that the cost of services in a monopolistic public market is up to *twice* as high as in a competitive market. The trio of public choice reforms (more competition, smaller organizations, more performance information) is intended to reduce if not remove the supposed existence of this chronic waste.

But what is the meaning of 'efficiency', and how can it be measured? Economists usually distinguish between two concepts of efficiency (Jackson 1982). Technical efficiency is the ratio of service 'inputs' (e.g. spending) to 'outputs', which are the goods or services actually produced by an organization (e.g. new houses built, teaching provided in schools, operations performed in hospitals). Allocative efficiency is the match between such outputs and the preferences of the public (e.g. whether the houses built meet the needs of families who lack suitable accommodation, whether the teaching reflects the demands of parents and pupils, whether priority is given to operations for the most urgent medical complainants). This second definition of efficiency overlaps closely with the concept of responsiveness as used by public choice theorists. The concept of efficiency will, therefore, be interpreted in this book in its technical sense. Moreover, this is the meaning of the term that has become familiar to policy-makers and managers through debates on the 'three Es' of economy, efficiency and effectiveness (Tomkins 1987).

Measures of technical efficiency usually focus on 'unit costs', which can be interpreted simply as the financial resources required to produce a 'unit'

of service provision. If public choice theorists are correct, and if the reforms which they advocate were adopted and implemented, then unit costs should fall over time. In other words, there should be a significant difference between the unit costs of public services in the pre-reform and post-reform periods. This check on whether public choice reforms have worked sounds straightforward in principle, but is rather more complicated in practice. Here we focus simply on the *measurement* of unit costs – other problems, such as attribution, are discussed in Chapter 3.

First, which costs are to be included in the calculation? One option is to focus only on the direct expenditure of the organizations which provide a service that has been reformed. However, this may neglect the indirect costs of reform that are borne by other agencies. For example, the introduction of CCT in UK local government in the 1980s was associated with lower spending on services such as refuse collection that were 'contracted out' to private firms (Boyer 1998a). Yet these expenditure savings were largely achieved by redundancies amongst former local authority employees. This had direct consequences for the social security budget (through payments to the unemployed). It may also have put pressure on the budgets of the National Health Service and local social services departments (which tend to deal with problems that are partly associated with lack of income). Thus lower spending in one part of the public sector can lead to higher spending elsewhere. A comprehensive assessment of reform should attempt to take direct and indirect costs into account. More mundanely, but equally importantly, it is necessary to measure both revenue and capital costs (because an apparent saving on staff may be accompanied by the need to purchase new equipment), and to ensure that inflation is taken into account so that costs are judged on a truly comparable basis before and after reform. Not adjusting for inflation could result in a reform being judged a failure (absolute spending has risen) when it has been a success (real spending has fallen).

Second, what counts as a 'unit' of service provision? At face value, this refers to the *quantity* of the outputs of a public organization (for example, number of dwellings made fit, number of music lessons for school pupils, number of patients treated by general practitioners). However, judgements about the consequences of reform also need to consider *quality*, which refers to the standard of service provided (Walsh 1991; Reeves and Bednar 1994). Relevant questions for the services cited above would include: what facilities (e.g. double glazing, central heating) the newly renovated dwellings had; what range of music was taught, and on how many instruments; and whether patients were treated with appropriate respect and courtesy by their doctors. All these issues imply that the definition of efficiency that was given above needs to be reformulated more precisely. In particular, measures of efficiency should capture the 'cost per unit of *constant quality*'. Efficiency, in other words, comprises not two but three variables:

Table 2.1 Combinations of service cost, quantity and quality that are consistent with higher efficiency

Cost	Quantity	Quality
1 Lower	Higher	Higher
2 Lower	No change	No change
3 Lower	No change	Higher
4 Lower	Higher	No change
5 No change	No change	Higher
6 No change	Higher	No change
7 Higher	No change	% increase higher than the % increase in cost
8 Higher	% increase higher than the % increase in cost	No change
9 Higher	% increase higher than the % increase in cost	% increase higher than the % increase in cost

cost, quantity of output, and quality of output. Movements in all three need to be included in an evaluation of public management reforms. Otherwise, an improvement in efficiency may be missed (even if costs and quantity have stayed the same, quality may have improved); or an apparently successful outcome may turn out to be an illusion (costs have fallen, quantity has remained constant, but quality has fallen by even more than costs).

The separation of efficiency into its three components reveals that a variety of reform outcomes are consistent with public choice theory. The most clear-cut support for public choice reforms would be a combination of lower cost, higher service quantity and higher service quality. However, eight other outcomes also support public choice arguments to a greater or lesser extent (see Table 2.1). Similarly, the clearest contradiction of public choice claims would be a combination of higher cost and a lower quantity and quality of service outputs. Yet nine other outcomes are also inconsistent with higher efficiency (see Table 2.2). Note that the apparently 'neutral' outcome of no change in all three variables does not support public choice predictions on the impact of reform.

In sum, higher spending does not necessarily indicate lower efficiency, and lower spending does not automatically entail higher efficiency. Rather, judgements about the impact of public management reforms on efficiency must be based on information concerning the cost, quantity and quality of services. Moreover, at least nine possible combinations of these variables are consistent with improvements in efficiency, and at least ten combinations are consistent with a deterioration in efficiency. Thus careful tracking of relative movements in cost, quantity and quality is required in order to

Table 2.2 Combinations of service cost, quantity and quality that are inconsistent with higher efficiency

Cost	Quantity	Quality
1 Higher	Lower	Lower
2 Higher	No change	No change
3 Higher	Lower	No change
4 Higher	No change	Lower
5 No change	No change	No change
6 No change	No change	Lower
7 No change	Lower	No change
8 Lower	No change	No change
9 Lower	% decrease greater than the % decrease in cost	% decrease greater than the % decrease in cost
10 Lower	% decrease greater than the % decrease in cost	% decrease greater than the % decrease in cost

measure whether a public management reform has been successful on this evaluative criterion.

### Responsiveness

State intervention in areas such as education, health and housing was, in part, intended to deal with the failure of private markets to provide such services (Fraser 1984). In other words, the economic system, left to its own devices, was not believed to be able to meet public needs and demands. Public provision was deemed necessary because 'market failure' led to an insufficient quantity of services (e.g. too few schools to meet the educational needs of the economy) or to a politically unacceptable distribution of services (e.g. mansions for the rich and hovels for the poor). In short, the growth of the welfare state was intended to remedy a lack of *responsiveness* in private systems of service production and allocation.

Public choice theorists stand this conventional logic on its head. Instead of market failure, they emphasize 'government failure' (Wolf 1988). Responsiveness to public preferences is argued to be at least as weak in public as in private organizations. Government failure results from a combination of the problems of bureaucracy that were outlined in Chapter 1: monopoly power, absence of performance information, and excessive size. According to public choice theorists, bureaucrats look inwards towards their own interests rather than outwards towards the interests of the public. They are, therefore, either unaware of the needs of their supposed clients and customers, or pay too little attention to them.

In order to judge whether public management reforms have led to more or less responsiveness, it is necessary to address two main questions. First, *in what respects* should public services be responsive? High responsiveness implies that members of the public are happy with public services – but happy with which aspects of them? It would, in principle, be possible to take ‘consumer satisfaction’ as a proxy for responsiveness – but satisfaction with what? Second, *to whom* should public officials be responsive when making decisions about the design and delivery of services? The public can be divided into a variety of ‘stakeholders’. Should all of these potential stakeholders be given equal weight?

The activities and achievements of public organizations can be divided into a variety of categories (Boyne and Law 1991; Boyne 1997a). These typically fall into three broad headings: *inputs* to services (such as expenditure, staffing and equipment); *outputs* of services (such as quantity and quality, which in turn can be divided into the speed and accessibility of service delivery); and *outcomes* which relate to whether formal objectives are achieved (e.g. whether school pupils pass exams, whether houses are renovated) and whether any positive or negative side-effects occur (e.g. job creation as a result of house renovation programmes). Public organizations can also be judged on their *profit* (e.g. whether money is spent for legitimate purposes, whether any of it has been diverted for fraudulent use by politicians and officials). Furthermore, some of the activities and achievements can be combined to form important performance ratios (e.g. *efficiency* is the ratio of outputs to inputs, and *cost-effectiveness* is the ratio of outcomes to inputs, otherwise known as ‘value for money’).

Public choice theorists have emphasized responsiveness to public demands concerning a restricted subset of these dimensions of organizational performance. In particular, they have focused narrowly on expenditure inputs and technical efficiency. Their obsession with these financial aspects of responsiveness is based on their belief that the public sector is bloated, spendthrift and profligate in its use of resources. Perhaps they assume that if these problems are solved then responsiveness on the remaining dimensions of performance will also improve. However, this line of argument is, at best, implicit in public choice thinking. A more plausible interpretation is that wider aspects of responsiveness have simply been ignored by public choice theorists. Even if citizens are happy with expenditure and efficiency, this does not guarantee that their preferences are fully reflected in all aspects of public services. It is, therefore, important to expand the intrinsic evaluative criterion of responsiveness to include not only expenditure and efficiency but also outputs, outcomes, probity and cost-effectiveness.

Another problem with public choice theory is that it is unclear to whom bureaucrats should be responsive. Whose demands are to be reflected in the pattern of public services? Are all demands to be treated equally and met in full? Classical public choice arguments on bureaucracy make little

reference to these issues (e.g. Tullock 1965; Downs 1967; Niskanen 1971). Nor has more contemporary work in the field addressed such questions (see recent volumes of the ‘house journal’ of the movement, *Public Choice*). Improvements in responsiveness are regarded as urgent, but the intended beneficiaries of this improvement in performance are shadowy figures.

An alternative strand in public choice theory sheds some light on this issue. This is the literature on the behaviour of political parties rather than bureaucrats. According to Downs (1957), the ideal political system would be like a highly competitive economic market. In a contest between two major political parties, each ‘firm’ would seek to maximize its share of the vote by appealing to the ‘median voter’ in the middle of the ideological range. Thus, in a hypothetical electorate with 99 voters, a government should try to match its policies with the preferences of voter number 50 in the distribution (an equal number of 49 voters are to the left and right of this position). Any other position can be defeated by an opposition party that successfully targets the median voter.

If this argument is carried over to the behaviour of public officials who design and deliver services, then their objective should also be responsiveness to the median voter. However, this is unsatisfactory for a number of reasons. First, it implies that public services are delivered in an undifferentiated and ‘Fordist’ manner: anyone can have any service that he or she wants, provided that it is the service demanded by the median voter. Even perfect responsiveness to the demands of the median voter would leave most people to the left and right of this position deeply unsatisfied with public services. Second, the ‘party competition’ view of responsiveness focuses exclusively on the electorate, but this is a subgroup of the adult population (indeed, only around 60 per cent of the *registered* electorate actually voted in the 2001 general election in the UK). Thus the median voter and the median citizen may differ significantly in their preferences. Third, the Downs (1957) model of responsive parties deals with policy on a single dimension, usually taxation or expenditure (see Hoffman 1977; Boyne 1987). However, as noted above, responsiveness can vary over numerous dimensions that also include service outputs and outcomes. The distribution of voter preferences may differ between these, so that there are as many ‘median voters’ as there are dimensions of responsiveness (Mueller 1989). In short, if the public choice answer to the question ‘responsiveness to whom’ is ‘responsiveness to the median voter’, then this is an impractical and inappropriate basis for evaluating public management reforms.

An alternative approach is to divide the public into a variety of stakeholder groups, and to judge the level of responsiveness to their diverse preferences (Boschken 1994). These groups include direct service users and their representatives (e.g. parents of pupils at school, carers for senile relatives who require medical treatment), members of the public who use services indirectly (e.g. employers who recruit school leavers to their workforce),

and taxpayers (who may pay for services that they do not use because they opt for private provision of education, health and housing). These different constituencies are likely to have different interests and demands (Fountain 2001).

A final major stakeholder group needs to be taken into consideration: the staff who provide public services. Even if the public choice assumption of bureaucratic selfishness is rejected, the view that service users and service providers have different interests remains plausible (Wilson 1993). Different segments of staff are in turn likely to have conflicting interests (Dunleavy 1991). However, for the purposes of exposition, they will be treated here as homogeneous. Public choice theorists argue that bureaucrats place too much weight on their own interests, but presumably would accept that some attention has to be paid to staff preferences in the design and delivery of services. Otherwise, the public sector workforce could be completely demotivated and demoralized. This may not be a recipe for successful reform – high organizational performance seems unlikely to follow from low individual performance. Thus the extent to which services are matched with staff preferences is relevant to an assessment of responsiveness.

It is possible to evaluate the impact of public management reforms on responsiveness by constructing a matrix that combines stakeholder groups with dimensions of organizational activities and achievements. Figure 2.1

Organizational activities and achievements	Stakeholder groups			
	Direct service users	Indirect service users	Taxpayers	Staff
Expenditure				
Outputs				
Outcomes				
Efficiency				
Cost-effectiveness				
Probity				

Figure 2.1 Matrix for evaluating the impact of public management reforms on service responsiveness

shows a simplified version of this matrix (with only four groups and only six dimensions of service provision). Even this 'checklist' of 24 boxes, however, represents a formidable evaluative challenge: first, because of the range of information required to fill the matrix; and second, because a clear conclusion of success would require higher responsiveness in all 24 boxes, and an unreserved judgement of failure would be possible only if responsiveness were lower in all of the boxes. Anything in between these extreme (and perhaps equally unlikely) outcomes is ambiguous. A clear picture can only be produced if weights are attached to the service elements on the vertical axis, and if weights are attached to the preferences of the different stakeholder groups on the horizontal axis. Such weights could then be entered in a formula which produces a 'magic number' that reveals how responsiveness has been affected by a public management reform. However, this would be not so much a mathematical as a political process, because the weights would be deeply controversial and contentious. Nevertheless, an important first step is to begin to put numbers in the boxes and thereby provide an 'evidence base' for policy debates.

### Equity

Critics of public choice reforms have questioned whether the intrinsic evaluative criteria of efficiency and responsiveness have been met (see Self 1993; Boyne 1998a). However, their major objection to public choice propositions has focused on the extrinsic criterion of equity. Even if reforms such as more competition do lead to higher efficiency and responsiveness, they argue, these positive outcomes are outweighed by greater unfairness in the provision of public services. This is a particularly sore point because, as noted above, a major rationale for state provision of services is that private markets allocate resources inequitably. Public organizations are expected to redress the imbalance that would otherwise occur in the distribution of services such as health, housing and education. If public choice reforms undermine the achievement of equity, then this can be regarded as a fundamental flaw in their design.

Products in private markets are generally allocated on the basis of ability to pay: people with higher incomes can obtain, for example, better cars, bigger houses, and more exotic holidays. For many goods and services, this is widely regarded as fair. In other words, the criterion of equity is 'you get what you pay for'. However, in some aspects of life this criterion is regarded as politically unacceptable, and so the state intervenes to supplement or supplant the market mechanism. When this happens an alternative criterion of equity is required (Blanchard 1986). After all, if the state also allocated resources on the basis of ability to pay there would be little point in providing public services. The standard of fairness that is

normally used in the public sector is 'need' (Doyal and Gough 1991). Thus the most important criterion of equity for judging public management reforms is not 'do people get what they pay for?' but 'do they get what they need?'

An evaluation of equity requires answers to two questions that have been debated widely in the social policy literature. First, which aspects of public services are to be provided equitably? For example, should equity be judged against the expenditure inputs offered to different groups, or against service outputs? Second, what are the relevant groups in society who are to be treated equitably? If different segments of the population have different needs, then equity requires that services should be allocated in proportion to these needs.

Le Grand (1982) has developed a widely cited conceptual framework that seeks to answer the first of these questions – see also Powell (1995) for a critique of the framework. Le Grand argues that the fairness of public services can be judged in the following five ways:

- 1 *Spending.* Criticisms are frequently voiced that the distribution of public expenditure between different people and geographical areas is unfair (Boyne and Powell 1991). A problem with spending measures is that efficiency varies across public organizations, so equity in spending may not lead to equity in outputs. Furthermore, spending may vary across areas because of differences in labour costs (for example, public sector salaries are higher in London than in the North of England). Thus spending is a weak criterion of equity.
- 2 *Use.* Different groups may use the same service to varying degrees through choice. Thus perfect equity in use could only be achieved by enforcing consumption in proportion to need. This infringement of individual liberty is likely, in a democratic society, to be regarded as too high a price for the achievement of equity in use. Yet, equal use for equal need remains a reasonable objective, even if it can never be perfectly attained. Thus if public choice reforms lead to less equitable use, they can be viewed as unsuccessful.
- 3 *Access.* This is partly a question of the physical proximity of services (e.g. how far patients have to travel to a hospital) and partly a question of 'opening hours' (e.g. whether outpatient clinics are held at convenient times). Le Grand (1982: 15) argues that 'the requirement that all individuals should have equal access to a service can most easily be interpreted as implying that the costs to all individuals of using that service (per unit) should be equal'. Again, fair access is not much help in achieving equity for service users if the distribution of output quantity and quality is unfair (hospitals that are equally spread in relation to population and 'open all hours' may meet the standard of fair access, but still provide services that do not reflect different needs in different locations).

Nevertheless, if a public management reform leads to less equitable access then this can be regarded as a negative result.

- 4 *Outcome.* This is analogous in some ways to the concept of effectiveness that was discussed earlier in this chapter. Le Grand (1982: 15) argues that 'precisely what is meant by outcome will vary from service to service. For health care it could be an individual's state of health; for education, the bundle of skills with which an individual emerges from the education system, . . . for housing the conditions of individuals' dwellings.' Thus outcomes in this sense refer to whether various 'end states' (e.g. better health) have been achieved. Average improvements across the population as a whole are probably consistent with more responsiveness (people want to be more healthy), but do not logically entail more equity. Rather it is the distribution of service outcomes that counts. The success of a health reform on this criterion can be assessed by whether the distributions of morbidity and mortality become more equal. This does, of course, open up the perverse possibility that success would be achieved if otherwise healthy people became more ill and died earlier. Greater equity, in other words, can imply the redistribution of death as much as the promotion of health.
- 5 *Final income.* This refers to the combined value of all material resources at the disposal of different groups in society. According to Le Grand (1982: 14), this criterion of equity implies that 'services should be allocated in such a way as to favour the poor, so that their "final incomes" (roughly private money income plus the value of any public subsidy received in cash or kind) are brought more into line with those of the rich'. Thus any public management reform that results in a lower share of services for poor households can be categorized as a failure on this criterion of equity. This, again, could lead to the apparently perverse outcome that economic decline accompanied by redistribution to the poor would count as success. However, for people who regard the 'final income' criterion as paramount, this is presumably a price worth paying.

In sum, four of Le Grand's (1982) criteria of equity seem relevant to the evaluation of public management reforms: use, access, outcomes and final incomes. Public choice reforms are widely believed to have negative consequences for these variables (Cutler and Waite 1997). For example, more competition between service providers is claimed to lead to the neglect of 'difficult' but especially needy patients and pupils; performance indicators for schools and the 'league tables' associated with them are believed to have created an increase in the exclusion of disruptive pupils (and thereby less equity in service use); and smaller organizations supposedly find it more difficult to switch resources between prosperous and needy clients. To some extent, such arguments may be based on an inappropriate point of comparison: the post-reform level of equity is not compared with the

## Illustrative population characteristics

Criterion of equity	Income	Age	Sex	Race
Use				
Access				
Outcome				
Final income				

Figure 2.2 Matrix for evaluating the impact of public management reforms on the equity of service provision

pre-reform level, but with an 'ideal state' that never existed. This issue will be explored further in Chapters 4-6.

The second major question concerning an evaluation of equity can be dealt with more briefly: which groups in society should be included in the analysis? Le Grand (1982) notes that the major focus of studies of equity has been on social class, usually defined in relation to income. However, this is a somewhat anachronistic agenda. Debates about equity in contemporary society have expanded to include issues concerning age, gender, race, language, sexual orientation and the 'differently abled'. If Le Grand's (1982) four criteria of equity are combined with a list of different population groups, then a matrix such as that shown in Figure 2.2 can be constructed. This can then be used to establish whether public management reforms have led to a growth or decline in the equity of public service provision. As with the evaluation of responsiveness on the basis of the matrix in Figure 2.1, the only unambiguous evidence of success would be greater equity in all of the boxes in Figure 2.2. Similarly, an unconditional conclusion of failure would only be justified by a decline in equity in every box in the matrix. In between these extremes, judgements about equity can only be made if weights are attached to the different criteria and population groups.

### Conclusion

Public choice reforms can be evaluated on the basis of changes in efficiency, responsiveness and equity. In principle, a theory-driven evaluation can offer crisp and clear criteria for assessing the impact of reforms. However,

we have shown that in practice this procedure quickly becomes complex and contestable. Each of the criteria consists of numerous sub-criteria that are either not defined within public choice theory or not given relative weights. Only a handful of the many potential outcomes of public choice reforms can be unambiguously deemed a 'success' or 'failure'. Furthermore, even if precise judgements can be made on efficiency, responsiveness or equity, no theoretical guidance exists (in the public choice literature or elsewhere) for determining the relative importance of these three policy objectives. For example, are improvements in efficiency and responsiveness more valuable than a deterioration in equity? Is a decline in efficiency worth trading off for increases in responsiveness and equity? Thus even a theory-driven evaluation leads to loose ends, because it encounters political rather than technical questions.

