
Negative Political Feedback: An Examination of the Problem of Modelling Political Responses in Public Sector Effectiveness Auditing

Negative
Political
Feedback

5

Peter Smith
University of York, UK

Two principal arguments are advanced for seeking to reduce the size of the public sector (Tomkins, 1987). The first is that it enables central and local government to reduce tax burdens, thereby allowing workers and corporations to retain a larger proportion of the profits from their endeavours and so, it is said, improving the supply side of the economy. This argument has generated intense debate, and is not considered further here. The second argument advanced for reducing the size of the public sector is that public sector organizations are usually monopoly providers of their services, and are not subject to the usual discipline of markets. In a competitive market, a company's failure to produce the goods that the community demands at reasonable prices should result in the loss of markets to competitors, and therefore render the company unviable. By contrast, management in public sector enterprises are not subject to the disciplines of product markets or financial markets and, therefore, it might be argued, do not have as much incentive as their counterparts in the private sector to maximize the efficiency with which they deliver services. The inefficiencies this gives rise to may be allocative (in the sense that an undesirable mix of services is provided) or managerial (in the sense that the services provided are not run efficiently).

The United Kingdom government has introduced a number of measures in an attempt to enhance the efficiency of management in the public sector. The principal focus has been on stimulating competition. Initiatives include the introduction of compulsory competitive tendering in local government (UK Government, 1988) and, in the National Health Service, the proposals to introduce an "internal market" in health care (Department of Health, 1989). Both of these developments seek to introduce pseudo-markets into sectors previously provided by monopolistic public sector organizations such as local government.

This article is based on a paper presented on 6 April 1990 to the annual conference of the British Accounting Association at the University of Dundee. Thanks are due to Jane Broadbent of the School of Management and Economic Studies, University of Sheffield, and to the referees for their helpful comments.

*Accounting Auditing &
Accountability Journal*, Vol. 5
No. 1, 1992, pp. 5-20. © MCB
University Press. 0951-3574

An alternative, less direct method of stimulating competition in the not-for-profit sector is to enhance the provision of comparative information about the performance of public sector organizations. Major developments in the United Kingdom include the requirement of local authorities to publish comparative data in their annual reports (Department of the Environment, 1981), the release of a huge volume of statistics on health service performance in computer readable form (Department of Health and Social Security, 1985), and the publication of University performance indicators (Committee of Vice Chancellors and Principals and University Grants Committee, 1987).

The release of such information offers the possibility of judging the performance of not-for-profit organizations, and therefore potentially enhances accountability between those who fund the public sector (the electorate) and those who run it (management). However, in spite of the improved flow of information, there is a problem in encouraging citizens to scrutinize the performance of public sector organizations. Unlike the private sector, where investors in a company have available an immediate sanction — selling the securities they hold — if unfavourable information is released, the electorate have very few sanctions to exercise if a public sector body is shown to be inefficient. The only control mechanisms are migration (a costly and often infeasible option) and the electoral system, which offers very infrequent and imperfect opportunities to exercise control.

As a result, governments have found it necessary to set up independent auditing bodies to scrutinize the performance of public sector organizations. In the United Kingdom the National Audit Office undertakes this function for central government expenditure. And in 1982 the Audit Commission was set up with a remit to examine the economy, efficiency and effectiveness of local authorities in England and Wales. Under the proposals to reform the National Health Service, the Audit Commission has assumed a similar role in respect of health authorities (Tristem, 1989).

The distinctions between economy, efficiency and effectiveness are in many respects artificial: they all refer to various aspects of economic efficiency. However, accountants have found it helpful to differentiate between them. The most basic concept is *economy*, which refers to the cost at which input resources are purchased. It is loosely equivalent to the economist's concept of input price efficiency. In an accounting framework, *efficiency* refers to the relationship between inputs consumed and outputs produced, regardless of the value of those outputs to society. It is equivalent to the economist's managerial or technical efficiency. The most ambitious aspect of the Audit Commission's remit is the examination of *effectiveness*, the extent to which the organization is achieving society's objectives, given the resources at its disposal. As Tomkins explains, it is not at all clear how the concept of effectiveness can be made operational, and the interested reader is referred to his discussion (Tomkins, 1987, chapter 3). The major debate to date concerning effectiveness audit has been the difficulty of developing satisfactory measures of the outcome of an organization's intervention. The controversy surrounding the development of

“quality adjusted life years” in the health sector illustrates the problems inherent in measuring outcome (Gudex and Kind, 1988).

In this article, however, it is assumed that such measures can be developed and a case study is presented for one of the few public sector services for which widely accepted outcome measures are available — the maternity services. The article examines how these data might in practice be used by an organization such as the Audit Commission to monitor effectiveness. Implicit in the idea of the performance indicator is the need to identify a feasible production function for the service under scrutiny, and thereby infer feasible cost functions. The performance of individual management teams can then be assessed with reference to these “ideal” functions. It is, of course, rarely possible to determine the feasible production function without reference to the existing performance of all management teams. That is, value-for-money auditors must rely on comparative data, or performance indicators, in order to infer best practice, given current technology. This is not the only reason for publishing performance indicators. For example, it is also presumed that their publication will stimulate competition, as public sector organizations seek to emulate best practice amongst their peers (Smith, 1988). However I do not consider such uses here.

In addressing effectiveness, performance indicators therefore serve two purposes. Firstly, they provide the basic data required to model relationships between inputs and outcomes. And secondly they can be used to assess the performance of individual decision-making units. To make operational the view of effectiveness auditing developed above it is usual to employ numerical models of the conventional input-output sort, formalized by Farrell (1957), and developed in Smith and Mayston (1987). These assume that there are certain underlying *needs* in the community which give rise to demand for services. For example, in the health sector these needs are reflected in the magnitude of the population, its profile of morbidity, and more general social circumstances. Needs measures are often referred to as environmental variables. The consequent demand for services is met by the deployment of *resources* of various sorts, principally capital and manpower. Finally, the consequent services yield a certain *outcome*, as measured by performance indicators. In the health services such outcome might ideally be measured in terms of the health of the community.

Thus in this conventional model causality is unambiguously from needs and resources (both of which can be considered as inputs) to outcome (the impact of the organization on society). Innumerable studies have been undertaken using this model. For example, the Department of Education and Science (1984) undertook an analysis of the effectiveness of local education authorities in which the inputs were the socioeconomic characteristics of the areas and various measures of resource use, and the outcomes were various measures of examination success. Stepwise multiple regression techniques were then used to select and estimate appropriate models, with examination success as the dependent variable, and the various needs and resources inputs as potential explanatory variables. This analysis yielded general models of examination performance, based on the stochastic averaging procedure implicit in ordinary least-squares regression analysis. The regression line in a sense represents

“average” performance amongst local education authorities. The performance of an individual education authority was then assessed in relation to this model by examining the magnitude of its residual. A large negative residual was taken to indicate notably poor performance, well below the typical level indicated by the regression equation. Conversely, authorities with large positive residuals were taken to represent best current practice, achieving results in excess of the norm, after adjusting for the authority’s environment and the resources it consumed.

The purpose of this study is to elucidate serious shortcomings in the methodology outlined above, brought about by the failure to recognize that political responses to poor outcomes are potentially important. In particular, to take the DES example cited above, a natural response of a local authority to poor examination results might be to commit more resources to the education service in order to attempt to rectify the matter. Thus not only is there an expected positive correlation between the level of resources committed to education and examination results: it is equally plausible to propose a negative relationship between examination results and resources, in the sense that poor examination scores might generate pressure for increased resources. This phenomenon might be called “negative political feedback”, and is in general not considered in most value-for-money audits. The concept is illustrated with a case study from the maternity services in England.

The organization of the article is as follows. The next section introduces the National Health Service performance indicator package. There follows a section in which a statistical model of the effectiveness of maternity services performance is developed. The concept of negative political feedback is then introduced, and incorporated into the conventional statistical model. The article ends with some thoughts on whether effectiveness auditing is ever likely to be worthwhile.

National Health Service Performance Indicators

The National Health Service (NHS) has at long last started to put its tradition of massive data collection to some constructive purpose. After a little prodding from the Health Services Management Centre at the University of Birmingham (and the House of Commons Public Accounts Committee — see Harley, 1988, and Carter, 1989) the Department of Health has now made serious efforts to disseminate the huge volume of data it requires from health authorities in a form that might be useful to auditors, managers and users of the service. The turning point was the issue in 1985 of a range of over 400 indicators in a form suitable for analysis by Lotus® software, together with a rudimentary, though flexible, analysis package. Since then an expert system has been made available by the Department of Health (DHSS, 1987). Though too late to be considered here, a new expanded system has been issued which takes account of the enlarged dataset now collected by health authorities.

The ready availability of so much hitherto inaccessible data might on the face of it be an unalloyed benefit. Certainly it is said to help the central government in its performance review process (Bowen and Payling, 1987). The high visibility given to the performance indicators might also improve the quality of the data,

as authorities come to realize that the information they provide is to be subjected to detailed scrutiny. And researchers throughout the country have seen their productivity leap now that they do not have to wade through reams of manual returns. However, the principal aim of the package was to “. . . help[managers] to assess the efficiency of the services for which they are responsible” (DHSS, 1983). Early evidence suggests that managers are not making very extensive use of the package (Jenkins *et al.*, 1988). Nevertheless, the interest of the Audit Commission in health authorities is likely to stimulate interest in performance indicators.

In this study a single specialty — the maternity services — is examined. The principal reason for choosing the maternity services is that, because they serve a homogeneous group of clients and have a long-established tradition of outcome measurement, they offer the best hope of yielding a manageable model of effectiveness. Indicators of resource provision and process variables form the vast bulk of the NHS performance indicator package, and the maternity services are adequately covered in this respect. However, unusually for the package, the maternity service indicators contain two types of variable almost completely absent in other specialties, namely epidemiological and outcome indicators. The epidemiological variables, equivalent to the needs variables noted in the preceding section, indicate the underlying health of the community, and refer to various aspects of low-weight births. The relevant outcome variables are various neonatal and perinatal mortality rates, well established indicators of maternity services performance, accepted for many years at national and international level. Thus the reason for choosing the maternity services as the subject for analysis becomes apparent. The data for these services offer the best prospect of yielding meaningful insights into effectiveness.

In order to assess effectiveness we first need to build a suitable production function, and therefore require estimates of needs, resources and outcome. Clearly in the maternity services the needs of the population are determined first by the number of pregnancies arising in the area. Almost all the indicators in the NHS package are in the form of ratios. They therefore make implicit allowance for the size of authority and thereby permit valid comparison between areas. Of course the use of such ratios presumes that no economies of scale exist. There is also a plethora of evidence to suggest that the health and social circumstances of mothers have a profound impact on the need for services. In particular, the proportion of low-weight births is known to be a sensitive determinant of adverse outcome. However, the low-birth-weight indicator is itself a function of both social circumstances and the local antenatal services. This study therefore uses the ten socioeconomic indices from the 1981 Census made available by the Inter-Authority Comparisons and Consultancy at Birmingham University as additional, more fundamental indicators of needs.

From the point of view of measuring effectiveness, the most useful measure of resources consumed would be in terms of costs. However, the allocation of overheads and staff time is always an arbitrary exercise, and the NHS costing returns are notoriously difficult to interpret. Perhaps because of this, no specialty costs are presented in the NHS performance indicator package. Instead, it is

necessary to use physical measures of resources. There are available measures of nursing and medical staff, although these might give a misleading picture. Some of the medical manpower resources will be devoted to gynaecology services, which do not form part of this study. And physicians (and nurses) from other specialties not included in the cited indicators might contribute to the maternity services: for example, paediatricians who specialize in neonatal care.

Only one relatively trivial aspect of capital provision (delivery rooms) is directly covered by the package. More general measures of capital stock are not available. The other major omission is any reference to the large number of support services, such as cleaners, porters, ambulances and the pathology laboratories, as well as overheads such as repair and maintenance, heating, and central administration. However, bed usage, measured as inpatients per bed per annum (or throughput) is a reasonable independent indicator of bed provision, and can therefore be used as a proxy (albeit highly imperfect) for non-staff inpatient costs. It should nevertheless be treated with caution, because like many process variables it can be affected by aspects of clinical practice such as readmissions.

So far as outcome is concerned, neonatal and perinatal mortality rates are important measures of the success of local maternity services in minimizing the grief and suffering associated with the loss of an infant. However, they offer only a limited perspective. For example, early neonatal mortality rates refer only to deaths within seven days of birth. It may be that the increased use of special-care baby facilities would reduce this index without changing the ultimate outcome. The best services will be those that not only minimize mortality, but also maximize the prospective quality of life for mothers and infants, given the resources committed to the services. However, there are no measures of handicapped survival. Moreover, at a local level, neonatal mortality rates are subject to substantial stochastic variation from year to year, so I take a three year average for the neonatal and stillbirth mortality rates which will form the study's outcome measures. Performance indicators from 1984/85 are used throughout. The list of variables used in this study is given in Table I.

Analysing the Data

Recall that the purpose of the indicators is to help managers and auditors in an individual authority. Clearly it is inappropriate simply to compare that authority with all others, regardless of their circumstances. For example, the problems of service delivery may vary considerably between impoverished inner cities, remote rural areas and affluent shire areas. In the first instance, therefore, some means is required to allow for different environmental circumstances.

One approach might therefore be to search for a collection of authorities that are in some sense "similar" to the focus of attention. This could be done using cluster analysis, which uses rather arbitrary statistical methods to split the entire population of authorities into homogeneous groups (Smith and Stewart, 1983a). The data used to perform the analysis might be the needs variables above, yielding a group of authorities with problems similar to those encountered by the authority under scrutiny. An alternative, slightly more satisfactory approach

<p>A. <i>Needs indicators</i></p> <ul style="list-style-type: none"> Percentage of population aged over 64 Percentage of elderly living alone Percentage of population aged under five Percentage of families with one parent Percentage of workers unskilled Percentage unemployed Percentage of households in poor housing Percentage of households in overcrowded housing Percentage of households with no car Percentage of households from ethnic minorities <p>B. <i>Resource indicators</i></p> <ul style="list-style-type: none"> Throughput per obstetric bed in District Health Authority (DHA) Delivery rooms per obstetric bed in DHA Gynaecology/obstetrics consultants per 1,000 deliveries in DHA Gynaecology/obstetrics senior doctors per 1,000 deliveries in DHA Nurses in maternity services per 100 deliveries to DHA residents Midwives as percentage of all nurses in maternity services Trained nurses as percentage of all nurses in maternity services <p>C. <i>Intermediate outcome indicators</i></p> <ul style="list-style-type: none"> Births <2,500 gm per 1,000 live births to residents of DHA Births <1,500 gm per 1,000 live births to residents of DHA <p>D. <i>Outcome indicators</i></p> <ul style="list-style-type: none"> Neonatal mortality rate (mothers resident in DHA) Still births in DHA per 1,000 deliveries in DHA

Table I.
Maternity Service
Performance Indicators

to identifying a homogeneous subgroup of authorities is to search for the authority's statistical "nearest neighbours" using the same data set (Smith and Stewart, 1983b). Whichever method is adopted, the performance indicators referring to resource use and outcome for the chosen subgroup of authorities can then be examined in some detail to determine the relative effectiveness of the authority of interest.

However, such methods suffer because the analyst is left without a coherent model of the services under scrutiny, and so may have to resort to piecemeal examination of the resource use and outcome data without any clear idea of how they relate to each other. The NHS manual is rather vague on how to proceed. Having identified the relevant indicators, the user is encouraged to scrutinize what are referred to as "first-line" indicators, which show the user whether the authority is devoting a reasonable mix of resources to the service, and is achieving satisfactory output, usually measured in terms of patient throughput. If the authority (or hospital) being examined exhibits "unusual" performance then second-line indicators can be examined in an attempt to gain an understanding of why the authority's behaviour is apparently aberrant. Using these principles, the expert system developed by the DHSS offers a systematic method of exploring the performance indicator system (Bowen and Payling, 1987).

However, reflecting the philosophy in the manual, this concentrates on outlying unsatisfactory behaviour, and offers no significant help to the median, or even to the mediocre. Similarly, a DHSS guidebook (DHSS, 1988) is highly selective in the performance indicators for the maternity services it chooses to discuss: the neonatal mortality rate, the low-birth-weight rate, and nurse staffing in special care baby units.

As Williams (1985) argues, it would seem desirable to go beyond this superficial view, and I shall therefore attempt to build more formal models of the determinants of outcome in the maternity services. As shown in Smith and Hagard (1982), such a model is likely to be complex. However it is possible to abstract from this complexity a manageable if crude model of the maternity services, as shown in Figure 1. Clinical intervention is thought to have an impact at the antenatal stage on the proportion of deliveries with low birth weight. Low birth weight in turn, combined with the socioeconomic factors, affects ultimate outcome in terms of perinatal mortality.

In order to make this model operational, I first deploy a naive regression technique, of the sort used in many value-for-money studies. If it is believed that outcome is determined by the local environment (or needs) and the local resources, it might seem sensible to specify a regression equation of the sort:

$$y_i = \alpha + \sum_j \beta_j x_{ij} + \sum_k \gamma_k z_{ik} + \epsilon_i \quad (1)$$

where

- y_i is the outcome observed in authority i ;
- x_{ij} are resources of type j in authority i ;
- z_{ik} are socioeconomic circumstances of type k in authority i , including low-birth-weight variables;
- $\alpha, \beta_j, \gamma_k$ are parameters to be estimated;
- ϵ_i is the usual error term.

This is the type of model used by the Department of Education and Science (1984) and, as described above, the residuals are said to indicate the relative effectiveness of the organization in securing society's health objectives. The terms included as potential explanatory variables in this specification include the three measures of resource use, the ten socioeconomic variables, and

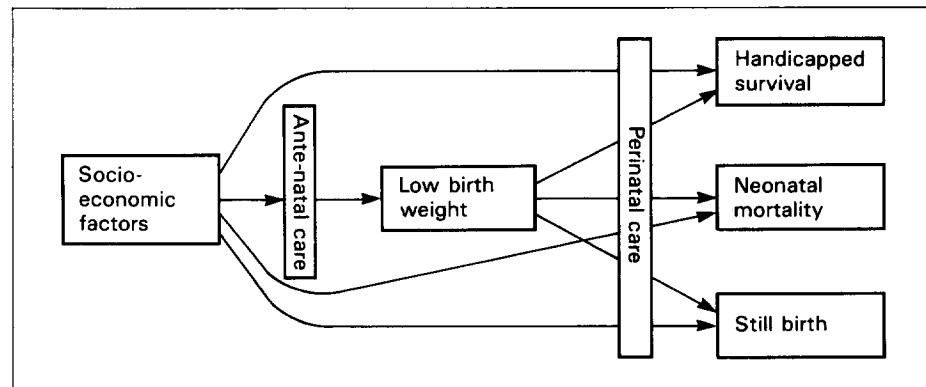


Figure 1.
Schematic Model of
Adverse Outcome in
Maternity Services

the low-birth-weight indicator. The chosen model is determined by stepwise techniques. Two such models are shown in Table II for the two principal outcome measures, stillbirth and neonatal mortality rates. The most notable features are the highly significant coefficients associated with socioeconomic variables (especially, as expected, low birth rate) and the absence of resource variables in the selected models.

One problem associated with these naive models is that, although they are intended to indicate effectiveness, they only indicate the typical level of effectiveness amongst the observed units, with divergences from the chosen model distributed either side of the regression line. It might instead be desirable to model the "ideal" level of outcome for a given set of inputs. This can be done using stochastic frontier techniques, which incorporate an "inefficiency" residual η_i as well as the usual random error ϵ_i into equation (1). The distribution of η_i is constrained to take only a positive sign (when outputs y_i are "bads" as in this application) (Schmidt and Lovell, 1979). Use of this refinement does not usually alter the ranking of an individual authority to any great extent, but might produce more realistic estimates of improvements that can be achieved.

Negative Political Feedback

The superficial implication of the models estimated so far is that resource provision is immaterial to health outcome in the maternity services! In particular, the provision of extra doctors and nurses appears to offer no returns that can be detected in the crude outcome measures. This conclusion appears to contradict common sense, as well as the Royal College of Physicians (1988), and must cast doubt on the model specification.

In seeking to find an explanation for this phenomenon it is important to note that the mortality rates used as outcome measures in this study have been

	Neonatal mortality	Still births
Low birth weight (% <2.5 kg)	0.141** (0.041)	—
Unemployed (%)	0.077** (0.022)	0.144** (0.022)
Single parent families (%)	—	-0.429* (0.138)
Poor housing (%)	—	0.115* (0.044)
Mobility (%)	-0.121* (0.041)	—
Constant	4.560** (0.507)	4.704** (0.231)
<i>n</i>	188	188
<i>r</i> ²	0.220	0.252
* Significant at 5 per cent level; ** significant at 0.5 per cent level.		

Table II.
Results of Stepwise
Regressions of
Outcome Measures on
Performance and
Socioeconomic
Indicators (Standard
Errors in Brackets)

published for many years. As a result local politicians, managers and (to a lesser extent) citizens will have been aware of the performance of local maternity services. A major weakness in the models described so far is that they fail to take account of possible reaction to the published indicators. For example, if an area suffers from high perinatal mortality rates, there is likely to have been strong local political pressure to improve matters. So if an authority is performing badly on one of the mortality measures, it is likely to make a response. The obvious response is to devote more resources to the maternity services. Consequently, instead of the "common sense" negative correlation between resources and mortality rates, we might also expect to see a *positive* correlation between mortality rates and resources, as local management seeks to improve local performance indicators. Underlying this model is the belief that performance indicators such as the perinatal mortality rates have traditionally been given a high profile by politicians and health care professionals. Comparatively poor performance is therefore likely to generate strong pressure to increase resources, even if those resources might be better allocated elsewhere. Thus, if this model is sustained, a well established performance indicator system will stimulate convergence towards some median behaviour in terms of outcome.

Using the terminology of cybernetics, this phenomenon is therefore a form of negative feedback, in the sense that it is a natural correcting mechanism for abnormal behaviour (Beer, 1966). The mechanisms by which such feedback is sustained are the various forms of formal and informal political process which influence the delivery of public services. So it seems natural to call the phenomenon "negative political feedback".

To the mathematically inclined, it can be represented in algebraic form by simultaneous equations of the sort:

$$y_i = \alpha + \sum_j \beta_j x_{ij} + \sum_k \gamma_k z_{ik} + \epsilon_i \quad (2a)$$

$$x_{ij} = \alpha' + \delta y_i + \sum_{j \neq i} \beta'_j x_{ij} + \sum_k \gamma'_k z_{ik} + \epsilon'_i \quad (\text{for each } j) \quad (2b)$$

Equation (2a) represents the traditional "technological" relationship, with resources x_{ij} (in conjunction with uncontrollable environmental factors z_{ij}) influencing outcome y_i . Equation (2b) introduces the inverse "political" relationship, in which outcome y_i influences resource provision x_{ij} . If relationships of this second sort hold, the naive model (1) is mis-specified, and it is clearly possible that a simple regression based on this faulty specification will detect no relationship between resources and outcome.

Econometricians are accustomed to modelling the more subtle feedback systems indicated by (2a) and (2b), and, by way of illustration, some results are estimated for the following specific system:

$$\text{NNM}_i = \alpha + \beta_1 \text{DOC}_i + \beta_2 \text{NRS}_i + \beta_3 \text{TPT}_i + \gamma \text{LBW}_i + \sum_k \gamma_k z_{ik} + \epsilon_i \quad (3a)$$

$$\text{DOC}_i = \alpha' + \delta' \text{NNM}_i + \beta'_2 \text{NRS}_i + \beta'_3 \text{TPT}_i + \gamma' \text{LBW}_i + \sum_k \gamma'_k z_{ik} + \epsilon'_i \quad (3b)$$

$$NRS_i = \alpha'' + \delta'' NNM_i + \beta_1'' DOC_i + \beta_3'' TPT_i + \gamma'' LBW_i + \sum_k \gamma_k'' z_{ik} + \epsilon_i'' \quad (3c)$$

$$TPT_i = \alpha^1 + \delta_1^1 NNM_i + \beta_1^1 DOC_i + \beta_2^1 NRS_i + \gamma^1 LBW_i + \sum_k \gamma_k^1 z_{ik} + \epsilon_i^1 \quad (3d)$$

$$LBW_i = \alpha^0 + \beta_1^0 DOC_i + \beta_2^0 NRS_i + \sum_k \gamma_k^0 z_{ik} + \epsilon_i^0 \quad (3e)$$

This model starts with the simple relationship developed in equation (1), with equation (3a) assuming that neonatal mortality (NNM) is dependent on the provision of senior doctors (DOC), trained nurses (NRS), throughput (TPT) as a proxy for capital provision, and the incidence of low-weight births (LBW), as well as more general socioeconomic factors. Equations (3b), (3c) and (3d) acknowledge the possibility that the provision of resources may in turn depend on the prevailing neonatal mortality rate. Equation (3e) models low birth weight as an endogenous variable. The model is estimated by two-stage least squares, using the socioeconomic variables as instruments in the first-stage regressions.

The results are shown in Table III. They are generally unsatisfactory. None

Explanatory variable	Dependent variable				
	NNM	LBW	DOC	NRS	TPT
NNM	—	—	-97.6 (249.3)	53.6 (61.0)	3.58 (3.06)
LBW	0.75 (2.20)	—	68.8 (446.8)	-15.8 (27.9)	-3.60 (2.33)
DOC	-0.01 (0.05)	-0.003 (0.004)	—	—	—
NRS	0.00 (0.03)	0.007 (0.004)	—	—	—
TPT	-0.10 (0.34)	—	-9.3 (53.6)	10.4 (6.5)	—
Under five (%)	—	—	-2.0 (237.5)	-52.7* (22.8)	—
Single parents (%)	—	—	—	—	6.73* (2.71)
Unskilled (%)	—	—	—	17.3 (9.4)	—
Unemployed (%)	-0.01 (0.55)	—	—	—	—
Overcrowding (%)	—	0.240** (0.047)	0.6 (99.0)	—	—
Mobility (%)	0.05 (0.43)	—	4.9 (14.7)	15.3 (9.7)	—
Constant	5.43 (2.89)	5.047** (0.980)	545.7 (946.1)	-211.4 (337.4)	45.01** (8.16)

* Significant at 5 per cent level; ** significant at 0.5 per cent level.

Table III.
Two-stage Least
Squares Estimates of
Neonatal Mortality
Output Model

of the tested explanatory variables is significant in the neonatal mortality model, and only a few of the socioeconomic variables remain significant in the resource models.

Discussion

There are at least six factors contributing to differences in performance between authorities:

16

- (1) *Different responsibilities*: some authorities might be responsible for functions not required of others, such as regional specialties or clinical education.
- (2) *Different objectives*: even if they have identical responsibilities, authorities might be able to assign different priorities to various services.
- (3) *Different needs*: localities experience wide variations in population age structures, socioeconomic patterns, and more prosaic determinants of need, such as local transport facilities. These variations may result in different levels and types of service provision.
- (4) *Different modes of delivering services*: localities face different costs for manpower and materials, and have different capital endowments, so may choose to deliver services in different ways in the pursuit of improved performance.
- (5) *Different accounting procedures*: these might result from inadvertent differences in interpretation of data definitions, or deliberate decisions to represent the authority in the most favourable light. The possibility of simple data error is an ever-present possibility.
- (6) *Differences in the economy, efficiency and effectiveness with which services are delivered*: these represent the principal target of most performance indicator packages.

As can be seen from this list, the diligent performance indicator analyst faces a formidable task if he or she is to disentangle this confusion of determinants and arrive at reliable estimates relating to item (6). The case study reported here has described the traditional way in which statistical techniques have been used to address the problem, and has illustrated some of the pitfalls that might arise.

Clearly, in the light of this complex model of variation, the crude presentation or citation of performance indicators is to be deplored. For example, no meaningful interpretation can be placed on the unadjusted ranking of an authority on an indicator. Cluster analysis and nearest-neighbour analysis represented two modest approaches to adjusting for factors (1) and (3). However it should be borne in mind that there remain three other sources of variation as well as various aspects of inefficiency. The regression analysis presented here represents an attempt to model the simultaneous effects of at least some of these factors.

There nevertheless remain problems with the quantitative approach described here. The incompleteness of the coverage of available performance indicators will always compromise effectiveness audit. There is a profound danger that

excessive reliance on quantifiable aspects of performance will compromise achievement of less readily quantifiable aspects of service provision. Similarly, the crude models presented here, though widely advocated as methods of scrutinizing public sector performance, are only able to address one objective (or outcome) at a time. They do not readily lend themselves to analysis of the multiple objectives that characterize public sector services. The measures of resource use that enter the model as explanatory variables contribute to a whole spectrum of outcomes, and not just the one being modelled. Some authors have advocated modelling multiple objectives with careful use of data envelopment analysis (Charnes and Cooper, 1980). However, DEA is principally interested in considerations of managerial efficiency. Its usefulness in measuring effectiveness is limited as it concentrates on determining whether management is obtaining a satisfactory level of outputs from the resources at its disposal, and does not make any judgement as to whether the mix of outputs produced is appropriate to society's needs.

These issues are discussed at greater length elsewhere (Smith, 1990). The main purpose of this paper has been to draw attention to a further drawback of the quantitative approach to modelling effectiveness — the phenomenon of negative political feedback. The argument developed above shows that there are likely to be difficulties when the impact of past performance on resource provision is taken into account. Such feedback from performance measurement to behaviour is likely to become increasingly prevalent as more reliance is placed on performance indicators. Underlying this assertion is the belief that local managers and politicians are highly sensitive to equity issues. If an area has traditionally poor performance in a particular service, it is very difficult to explain to citizens that it is not cost-effective to commit resources in an attempt to improve matters. Instead pressure is likely to arise to give a high priority to the service, and to devote more resources to it, even if the resources do not yield benefits in terms of improvements in the very performance indicators that stimulated the political need to improve the service.

This gives rise to problems because simple statistical approaches such as regression analysis will not yield satisfactory models of production functions. In particular, there will be a tendency for the importance of resource variables to be understated in such models. This may lead the incautious analyst to conclude that the level of resource provision does not make an important contribution to outcome. The policy implications of such a conclusion — that resources should be reduced in areas with high levels of resource provision — might be disastrous for services in those areas. As a result, it is imperative that the analyst makes every effort to develop well-specified models of performance.

The fundamental difficulty is that performance indicators are used both retrospectively and prospectively. In building models of production possibilities, and assessing individual performance, the auditor concentrates on their retrospective use. The politician, however, is likely to use the same indicators prospectively in setting targets for management. Such targets will usually be formulated with reference to the existing performance of comparable organizations,

and so the political process will generate a tendency to converge to “typical” levels of performance. Such is the mechanism of negative feedback.

If the prospective effectiveness auditor is daunted by the foregoing discussion, then that is as it should be. Performance assessment in the public sector is a complex process, and no amount of presentational skill can reduce that complexity. This article has addressed by far the most tractable aspect of health care — the maternity service — in which clients are relatively homogeneous and easily identified, outcome can be measured to some extent, and in which there even exist some epidemiological data. Nevertheless, the study has not addressed costs. In accordance with the performance indicator philosophy, only linear relationships have been modelled, although many health systems are manifestly non-linear. And only the crudest outcome measures have been examined. Midwives and doctors working in the maternity services, and the mothers who use them, would be astonished to learn that the auditor has been forced to use mortality rates as proxies for outcome. To the majority of users, aspects such as comfort, reassurance, convenience, education, and attention to relatively minor clinical complications are the principal yardsticks by which they judge the service. Given this list of well documented difficulties, it is not surprising that very few convincing studies of effectiveness have been published.

This study has sought to bring to light a further shortcoming of quantitative effectiveness audit. It has noted the importance of taking careful account of political responses in inferring organizational effectiveness from performance data. The discussion has suggested that those responsible for running public services are likely to be deeply concerned with equity as well as effectiveness, and that such concern is likely to be heightened following widespread dissemination of performance indicators. If this is the case then it is unlikely that statistical tools will offer satisfactory insights into organizational effectiveness.

It is important therefore to recognize that the private sector paradigm of organizational effectiveness may be inappropriate in the public sector. There are many ways, apart from audit, of promoting effectiveness in the public sector, although these have been largely ignored in the United Kingdom in recent years. For example, one approach is to nurture and promote the sense of commitment that many employees in the public sector often feel towards the ideals of the organization for which they work. Such commitment can result in employees being willing to accept relatively low pay, and to carry out duties beyond those for which they are formally contracted. This has traditionally been a source of low costs and high productivity in the NHS in particular.

Another possible means of promoting effective delivery of services is to widen the involvement of the citizens of an area in their local public services. This might entail enhancements to the democratic process which enable the electorate to feel that their views are capable of influencing policy. If a true “market” for consumers of local services can thereby be created, the need for effectiveness audit will disappear, as an informed electorate will not tolerate poor management. However, there is no sign of any such developments in the United Kingdom at present. As a result the effectiveness auditor is likely to be, at best, a very poor substitute. Indeed this study has shown that, at worst, the influence of the auditor might be distinctly malign.

References

- Beer, S. (1966), *Decision and Control*, Wiley, London.
- Bowen, B. and Payling, L. (1987), "Expert Systems for Performance Review", *Journal of the Operational Research Society*, Vol. 38, pp. 929-34.
- Carter, N. (1989), "Performance Indicators: 'Backseat Driving' or 'Hands Off' Control?", *Policy and Politics*, Vol. 17, pp. 131-8.
- Charnes, A. and Cooper, W.W. (1980), "Auditing and Accounting for Program Efficiency in Not-for-profit Enterprises", *Accounting, Organizations and Society*, Vol. 5, pp. 87-107.
- Committee of Vice-Chancellors and Principals and University Grants Committee (1987), *University Management Statistics and Performance Indicators*, CVCP, London.
- Department of Education and Science (1984), "School Standards and Spending: a Further Appreciation", *Statistical Bulletin 13/84*, DES, London.
- Department of the Environment (1981), *Local Authority Annual Reports*, DoE, London.
- Department of Health (1989), *Working for Patients*, DoH, London.
- Department of Health and Social Security (DHSS) (1983), *Performance Indicators: National Summary for 1981*, DHSS, London.
- Department of Health and Social Security (1985), *Performance Indicators for the National Health Services, Guidance for Users*, DHSS, London.
- Department of Health and Social Security (1987), *The Performance Indicator Analyst*, version 1.0, DHSS, London.
- Department of Health and Social Security (1988), *Comparing Health Authorities*, DHSS, London.
- Farrell, M.J. (1957), "The Measurement of Productive Efficiency", *Journal of the Royal Statistical Society, Series A*, Vol. 120, pp. 253-81.
- Gudex, C. and Kind, P. (1988), "The QALY Toolkit", *Discussion Paper 38*, Centre for Health Economics, University of York.
- Harley, M.J. (1988), "Performance Indicators in the Management Process", *Health Services Management Research*, Vol. 1, 29-42.
- Jenkins, L., Bardsley, M., Coles, J. and Wickings, I. (1988), *How Did We Do? The Use of Performance Indicators in the National Health Service*, CASPE Research, London.
- National Audit Office (1990), *Maternity Services*, HMSO, London.
- Ripington, T. (1988), "The National Health Service — An Annual Report" in Jackson, P. and Terry, F. (Eds), *Public Domain 1988*, Public Finance Foundation, London.
- Royal College of Physicians of London (1988), *Medical Care of the Newborn in England and Wales*, The College, London.
- Schmidt, P. and Lovell, C.A.N. (1979), "Estimating Technical and Allocative Efficiency Relative to Stochastic Production and Cost Frontiers", *Journal of Econometrics*, Vol. 9, pp. 343-66.
- Smith, P. (1987), "Performance Indicators in the National Health Service", *Public Money*, Vol. 6 No. 4, pp. 35-9.
- Smith P. (1988), "Assessing Competition among Local Authorities in England and Wales", *Financial Accountability and Management*, Vol. 4, pp. 235-51.
- Smith, P. (1990), "The Use of Performance Indicators in the Public Sector", *Journal of the Royal Statistical Society, Series A*, Vol. 153, pp. 53-72.
- Smith, P. and Hagar, S. (1982), "Planning for a Single Specialty in a Health District", *Journal of the Operational Research Society*, Vol. 33, pp. 29-39.
- Smith, P. and Mayston, D. (1987), "Measuring Efficiency in the Public Sector", *Omega: The International Journal of Management Science*, Vol. 15, pp. 181-9.
- Smith, P. and Stewart, J. (1983a), "Cluster Analysis", *Public Finance and Accountancy*, March, pp. 18-19.
-

AAAJ
5,1

Smith, P. and Stewart, J. (1983b), 'A Simple Method for Comparing Performance', *Public Finance and Accountancy*, April, pp. 36-7.

Tomkins, C. (1987), *Achieving Economy, Efficiency and Effectiveness in the Public Sector*, Kogan Page, London.

Tristem, R. (1989), 'Auditing the Health Service', *Public Finance and Accountancy*, 3 November, pp. 17-18.

United Kingdom Government (1988), *Local Government Act*, HMSO, London.

20

Williams, A. (1985), 'Performance Measurement in the Public Sector. Paving the Road to Hell?', *Arthur Young Lecture 7*, Department of Accountancy, University of Glasgow.