Improving police performance

A new approach to measuring police efficiency

Clare Spottiswoode

Public Services Productivity Panel

Acknowledgements

This report did not just "happen". It is a collective effort, co-ordinated by a Project Team, involving many people and many organisations willing to generously contribute their advice, experience and enthusiasm. I am indebted to the people listed in Chapter 12, who were interviewed or consulted in preparing the recommendations of this study.

This study received crucial support throughout from the Home Office, HMIC and HM Treasury, who not only provided advice to the study, but more importantly seconded staff to the Project Team on either a full-time or part-time basis. Their assistance and co-operation made this study possible.

Several people kindly took the time to review the draft of the report, and provided many useful suggestions on how this inherently technical subject could be best presented.

I would like to specifically mention two teams, without whom this report would not have seen the light of day. The Project Team completed the detailed analysis to shape the many views into a coherent set of recommendations. The Economics and Resource Analysis Unit of the Home Office provided the technical spark to the study, as well as ongoing technical support to the Project Team. My sincere thanks to both teams.

However, the lynchpin to the whole project has been Michael Papesch, without whom this major project could never have been completed, and who has dedicated enormous energy and perseverance to bring the report into existence.

Clare Spottiswoode

Project Team – on either a full or part-time basis:

Michael Papesch (Team Leader)
David Black
Mike Boyles
Anne McDaid
Juanita Roche

Economics and Resource Analysis Unit – Home Office:

Richard Price (Unit Leader)
Gareth Harper
Richard Harries
Andrew Webster

Contents

Foreword by the Chief Secretary to the Treasury

Successful public services lie at the heart of a successful economy and society. At their best, they respond quickly to their customers, operate efficiently, and set the highest standards of quality and probity. In practice there have been wide variations in quality and in some areas ineffectiveness and waste have been tolerated. Modernising Government – the drive to achieve better, more responsive government and public services – means raising all services to the standards of the best and recharging our public organisations with fresh vigour, incentives and ideas.

As part of its comprehensive plan for modernisation, the Government has recruited a team of top private sector managers to the Public Services Productivity Panel – with a remit to advise on improving efficiency and productivity. The focus on productivity is linked in particular to the Public Service Agreements (PSAs), now published for all government departments, that enable the Government to track performance, improve accountability, sharpen responsiveness and inform investment. The role of the Panel is to support departments in raising standards to achieve or outperform their PSA targets, providing a source of practical ideas and new approaches.

Each member of the Panel is focussing on a different area, working with individual government departments and agencies to identify solutions that will increase productivity. In doing so, the Panel recognises that there is no monopoly of wisdom in the private sector about how to raise productivity. Instead, it is the blend of fresh ideas *and* learning from good practice in our public services that can provide the spur to improvement. The findings of each project will be published and the Government aims to draw together the overall lessons of the Panel's work during the first half of 2000.

I am grateful to Clare Spottiswoode, Associate Partner, PA Consulting, for her report that has identified a new way to measure the efficiency of the police. Clare's report is an important contribution to the Government's crime reduction strategy. She has identified a way of using existing information to make a more meaningful analysis of which police forces are doing the most to prevent, solve and reduce crime within available funding. More importantly, Clare outlines a means by which police forces can better understand their strengths and weaknesses. This will help to raise the performance of all police forces to the level of the best, within the framework of the *Best Value* initiative. Her fresh approach clearly has wide potential application across all public services.

Our challenge now is to translate these finding into clear and meaningful public benefit. Expectations of the police are rightly ambitious. By involving valuable private sector experience in our drive to modernise government, our promise is to match that ambition with excellence for the many, not the few.

Rt Hon. Andrew Smith MP
Chief Secretary to the Treasury

Glossary

Explanation of abbreviations and technical terms used in this report:

ABC	Activity Based Costing. A technique for calculating the cost of a good or service by measuring the direct labour and other costs of each activity associated with producing the good/service, as well as allocating an appropriate share of overheads.
Absolute efficiency	The maximum theoretically possible level of efficiency in producing outcomes of a given quality, for a certain level of input.
ACPO	Association of Chief Police Officers.
Allocative efficiency	Refers to the situation where, for any level of production, inputs are used in the proportion that minimises production costs, given the relative prices of those inputs.
APA	Association of Police Authorities.
Authority	Police Authority – the body responsible for maintaining an efficient and effective police force for its area.
Best Value	A statutory duty placed on local authorities (including police authorities) from 1 April 2000 to deliver services to clear standards – covering both cost and quality – by the most effective, economic and efficient means available. It involves a programme of reviews over five year periods that challenge whether, how and by whom a service should be provided.
Best Value Performance Plan	An annual plan prepared by local and police authorities setting out achievements and future plans for service reviews under *Best Value*.
BPA	Business Process Analysis. A technique that breaks down processes into a series of sub-processes, so that they can be analysed and benchmarked to find redundant steps and other sources of inefficiency.
BVPI	*Best Value* Performance Indicator. The indicators to be used to measure and track local and police authority performance under *Best Value*.
BCU	Basic Command Unit. This is the most basic operational unit for the delivery of local police services.
CSR	Comprehensive Spending Review.
DEA	Data Envelopment Analysis. A linear programming technique that measures relative efficiency within a sample based on a comparison of the observed and best practice providers.

Effectiveness	The degree to which the outputs of a service provider meet the stated objectives of that service. Effectiveness captures the need to both provide the right mix of outputs, and to provide them to an appropriate quality.
Efficiency	Refers to the situation where the maximum amount of outcomes is produced for a given level of inputs. Can also be expressed as minimising the resources required to produce a given level of outcomes. Efficiency covers allocative efficiency, technical efficiency and dynamic efficiency (the degree to which producers alter their services over time to meet changing user needs as well as technological change).
HMIC	Her Majesty's Inspectorate of Constabulary.
Inputs	The physical resources used in the production of goods or services, such as labour.
Linear programming	A technique for maximising (or minimising) an objective function, subject to a set of linear constraints on the variables.
Malmquist index	In this context, a method for identifying how much of an efficiency gain is due to one force catching up to the best ("frontier") force(s) and how much the frontier itself has improved over time.
Net Revenue Expenditure	The expenditure on employees and running expenses, less sales of assets, income from fees and charges and specific grants.
Outcomes	The impact that a set of outputs has on society as a whole – such as a lower crime rate or the maintenance of public order.
Outputs	The goods and services supplied by providers to other entities or to the public – such as crime investigation or a police patrol.
Productivity	A measure of the physical output produced from the use of a given quantity of inputs.
Regression analysis	A technique for calculating the expected value of a (dependant) variable based on the observed values of a set of other (independent) variables.
Relative efficiency	The efficiency of one producer in delivering outcomes compared to the most efficient producer.
SFA	Stochastic Frontier Analysis. A form of regression analysis in which the difference between the observed performance and the predicted performance is broken down into measurement error and inefficiency.
Technical efficiency	The efficiency of a provider in converting physical inputs (such as labour) into outcomes. The measurement of technical efficiency compares each provider's ratio of outcomes to inputs with the ratio achieved by best practice.

Executive Summary

The police service is one of the most important of all the public services. Effective policing is fundamental to a safe and secure society, and to everyone's sense of wellbeing. The police are central to the Government's programme to reduce crime and the fear of crime. It is a difficult and complex job.

Best Value is the central plank in the drive to improve police performance. A systematic measure of police efficiency – where "efficiency" is a measure of the police's performance in meeting their overarching aims and objectives for the money spent – is crucial if *Best Value* is to work effectively. There is a plethora of indicators and information about police outputs and outcomes. But, to date, it has not been possible to draw this information together to build a comprehensive or systematic measure of relative police *efficiency* in meeting their ultimate objectives of promoting safety and reducing crime, disorder and the fear of crime.

This is a serious gap in the *Best Value* strategy. This gap hits police authorities and forces the hardest because:

- they lack good measures to fulfil their *Best Value* obligations comprehensively to compare performance (including efficiency) with others.
- police authorities and forces themselves do not always know what the scope for efficiency gains is or even where they should be looking for them.
- efficiency targets should take into account the actual position of each force, and thus the gains made in recent years. The Government cannot do this if it lacks a systematic efficiency measure. So, to date, the Government has set uniform, across-the-board efficiency targets.

Resolving these issues through the provision of good comparative efficiency measures will always be problematic and controversial. Reliable comparative efficiency measures are technically hard to construct and rely on assumptions and simplifications that could always be challenged or objected to. Searching for the perfect approach would be fruitless. The task is to identify the best available approach, recognise its strengths and weaknesses, build on the strengths and establish approaches to limit the downsides.

This study takes this approach. It recommends the joint use of two of the most advanced relative efficiency measuring techniques – Stochastic Frontier Analysis and Data Envelopment Analysis. These techniques have been used for measuring the relative efficiency of regulated private sector industries, and are increasingly being used in the public sector in other countries. This report outlines how these techniques would be integrated with the existing *Best Value* outcome performance information. Used this way, these techniques would effectively benchmark police authority and force performance in meeting police objectives for the resources that each authority and force has. They would help identify which authorities and forces put their resources to the best use. As a result, they would help to spread good practice and to reduce variations in police performance.

The resulting comparative efficiency measures would fit seamlessly into the *Best Value* framework. Comparative efficiency would be measured using the key outcome measures from the existing suite of *Best Value* Performance Indicators. A programme of analysis, review and inspection is recommended to understand the raw comparative efficiency measures – and to adjust them as necessary. This will identify the key actions that police authorities and their forces need to take. Differentiated efficiency targets – to improve the level of outcomes for the available budget – can be drawn up to reduce, over time, the variations in performance.

Compared to the current across-the-board targets, these targets would better reflect the starting position of each authority and force. This approach would also mean that "efficiency" is about finding ways of improving the performance delivered for the money that each authority and force has. It is *not* about simply cutting costs, or corners, or reducing the number of services or staff.

There must be meaningful incentives and rewards for top performing forces. Options for an incentives framework are identified, based on setting differential efficiency targets that close only part (not all) of the estimated gap between authorities and forces, broad (rather than detailed) rankings, rewards for top performers, and assistance for those needing it. More importantly, developing good outcome-based efficiency measures will lessen the need for central controls over the use and deployment of police inputs. It is a virtuous circle – as the better efficiency measures allow central controls to be relaxed, managers have greater freedom to manage and overall police performance can improve even more.

The two recommended techniques must be used with care. They can only provide estimates of relative efficiency, and so the raw results will need interpretation and explanation. They are critically reliant on the quality of the data used. However, data risks would be minimised by using *audited Best Value* data and by checking, during implementation, that the best input and outcome measures are being used. And, although in general there should be a good correlation between the two approaches, when properly specified, each technique is different and each will provide slightly different results from the other. This study recognises that the techniques will present different efficiency estimates, and identifies how the different estimates can be used to better understand the drivers of each one.

Most importantly, the techniques cannot be used mechanistically or interpreted simplistically. If they are, the wrong conclusions will be drawn. Careful analysis and judgement must be applied to the results – the report outlines a detailed process to do this. Care must be taken to avoid perverse consequences – such as incentives on the police to fail to record crime or incentives to respond solely to the efficiency measures rather than the needs of the public. This is one reason why this study recommends that outcomes, not outputs, be measured. Moreover, the reform of central controls on the deployment of inputs will make it easier for the police to offer the service that the public wants.

The recommendations in this report set a direction which, when implemented, would:

- provide a systematic, comprehensive measure of relative police efficiency for the first time. This will allow differentiated efficiency (performance) targets to be set for the police.
- allow across-the-board efficiency targets to be abolished.
- reduce the administrative burdens on police authorities and forces – by getting rid of the bureaucracy surrounding the current, separate efficiency plans.
- give police managers the freedom to manage their resources better – as central controls on flexibility of input use are relaxed.
- provide a tool that assists each police authority and their force to assess its relative strengths and weaknesses in delivering outcomes for the resources it has, and thus improve efficiency over time.
- provide a means to meet fully the *comparison* requirement of *Best Value* and to sharpen the focus of each *Best Value* Performance Plan.

This study presents a strategic framework for better measuring police efficiency. It has not estimated relative police force efficiency – nor has it attempted to build the appropriate models. The next step is – through an inclusive process – to develop working, practical models for assessing relative efficiency in time to use the first audited *Best Value* data in mid-2001. To do this, the suggested input and outcome measures need to be validated, the models specified, built and peer reviewed, and tested in a pilot. While the Home Office would take the lead in this work, it is critical that these steps be undertaken in an open and transparent manner involving all stakeholders at all stages. These requirements are at the core of the implementation process recommended in this report.

Policing stakeholders consulted in the course of this study are aware of the limitations in the current police measures and have assisted this study throughout. They have given their support to pilot the recommended approach in a positive and constructive manner. This demonstrates the willingness of the police to develop meaningful performance information and to use that information to get underneath the difficult task of measuring the relative efficiency of a complex public service.

Summary of Recommendations

An overview of this report's detailed recommendations is outlined below. More specific recommendations can be found at the end of each chapter.

- *police efficiency measurement should focus on the performance of police authorities and forces in delivering outcomes that contribute to the overall aims and objectives of policing, for the available funding.*
- *police authorities and forces should be given differentiated efficiency targets for the available funding, based on the results of systematic measures of relative police efficiency. The aim should be to raise the poorer performers up to the level of the best.*
- *a small set of input and outcome measures should be used for measuring relative police efficiency. The input measures should correspond to the annual costs of police forces (including a measure for capital assets used), less uncontrollable costs. The outcome measures should be focussed around (but not limited to) a subset of the existing outcome-focussed Best Value Performance Indicators (BVPIs).*
- *the choice of appropriate outcome measures should be reviewed as part of the implementation of this report. If this review finds relevant police outcome measures outside the current BVPI suite, the BVPI suite itself should be changed.*
- *two efficiency measurement tools, Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA), should be used to provide an assessment of the relative efficiency of police forces in delivering police outcomes.*
- *SFA and DEA place weights on each outcome measure when assessing relative efficiency; DEA can allow these weights to be any value. To prevent DEA placing a high weight on a relatively low priority outcome, the weights given to each outcome measure should be constrained in DEA.*
- *the SFA and DEA analysis will each provide an estimate of the efficiency of each police force which, collectively, will give a spread of efficiency estimates. The raw SFA and DEA results should be reviewed and analysed, to understand differences in results as well as any data issues or differences in service quality influencing the results. This analysis may lead to adjustments to the raw results as necessary.*
- *the analysis of the SFA and DEA results will indicate which model is the most technically robust and performs the best. The results of the best performing model should be used to group forces into relative efficiency "bands" based on sensible demarcations of the actual results and statistical tests of the significance of each band. These bands should be used by HMIC to target its inspection and review programme to assist police authorities and forces understand the efficiency measurement results.*
- *the current efficiency plans and the associated across-the-board efficiency targets should be abolished from 1 April 2002. Differentiated efficiency targets – to improve the level of police outcomes for the funding available – should be introduced thereafter based on the banding system. Authorities and forces in the top band should be set a comparatively smaller target to further improve their efficiency. Authorities and forces in the lower bands should be set targets to close about half of the estimated gap between themselves and those in the top band.*
- *any overall efficiency target set for the police should be distributed amongst police authorities and forces according to the principles outlined immediately above.*

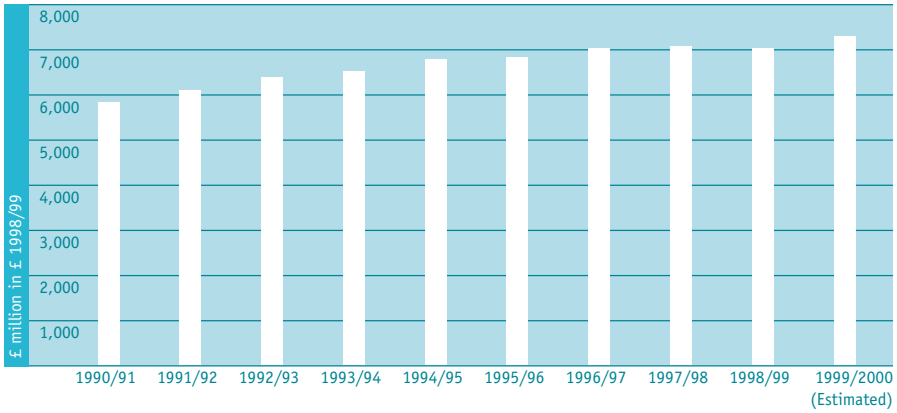
- *the Home Office should publish relative efficiency tables based on the banding system, rather than detailed force-by-force rankings.*
- *a more comprehensive performance management and incentives regime should be developed by examining options to reward top performing forces and by developing a programme to review the impact of police regulations. This work should finish no later than the proposed pilot of SFA and DEA.*
- *the Home Office, in consultation with its policing stakeholders, should review the appropriate input and outcome measures, outcome weight ranges, and SFA and DEA models with a view to first using them in mid-2001 using audited BVPI 2000/01 data. The task of specifying, building and testing the models should be contracted out to independent experts. The models should be developed using an open and transparent process in which all parties have access to all data and to independent technical advice.*
- *a pilot of the SFA and DEA techniques should be undertaken in late 2000 to test the reliability of the proposed models and to demonstrate the value of the approach.*
- *once the results of the pilot are known, HM Treasury, with assistance from the Home Office, should initiate a programme to present the results to other possible users in the public sector.*
- *the Home Office should evaluate the recommendations in this report at the end of the pilot – December 2000 – and once the techniques have been used with the first audited BVPI data – late summer 2001.*
- *the Home Office should present a report to the Ministerial Committee on Public Services and Expenditure (PSX) in early 2001 on its success in implementing the recommendations of this report. The Home Office report should be reviewed by the Public Services Productivity Panel prior to submission to PSX, with the Panel's comments also provided to PSX.*

Chapter 1 – Introduction

Improving the efficiency and performance of police forces[1] matters. The police are at the forefront of the Government's drive to reduce crime, reduce the economic costs of crime and to make people feel safer.

A large amount of resources is invested in policing each year. Just over £7.3 billion (in £1998/99) was spent on policing in 1999/2000 – an increase of just over £1.4 billion in real terms since 1990/91[2], as shown in the following graph.

Police force expenditure in real terms 1990/91 to 1999/2000



Source: Home Office. Figures pre 1995/96 are current expenditure; figures post 1995/96 are revenue expenditure (ie: includes capital financing costs met from police revenue). Figures are converted from nominal to real 1998/99 expenditures using the HM Treasury's GDP deflator.

The Government, rightly, expects improvements in the efficiency and performance of all public services, including the police. Improving police performance and efficiency is a critical element of any crime reduction strategy. All police authorities and their forces should be expected to contribute to this strategy – even the most efficient. But the focus should be to reduce any variations in performance between forces.

The Government and its policing partners have initiated a multi-faceted strategy to improve police productivity, focused around the introduction of *Best Value*. Under *Best Value*, from 1 April 2000 all police authorities must:

- develop a five-year programme of service reviews.
- summarise their findings and action plans in published annual Performance Plans that show past and projected performance.
- undergo a rigorous programme of audit and inspection.

As part of this programme, police authorities will be required to *challenge* why and how a service is provided, *compare* their own with others' performance across a range of relevant indicators, *consult* local users about the order of service reviews and the setting of performance targets, and embrace fair *competition* as a means of securing efficient and effective services. To set a wider framework for *Best Value*, overarching aims and objectives were developed for the police in 1998. Key performance targets and objectives for measuring progress against these aims have been established, and progress will be reported annually. These measures complement the wide range of performance information currently monitored by Her Majesty's Inspectorate of Constabulary (HMIC) and the Audit Commission.

1 Police authorities are accountable for maintaining an efficient and effective police force in their area. Authorities set the overall strategic direction of their local police force, and are also accountable for meeting police obligations under *Best Value*. Efficiency targets are formally handed to police authorities, who in practice, forward them to police forces. Police forces are accountable for the delivery of the strategy, and for managing day-to-day operations and resource use. Any reference in this report to the efficiency of "forces" is shorthand for the more complex set of relationships and accountabilities between authorities and forces, bearing in mind that it is the operational arm of policing – police forces – whose efficiency is being measured.

2 Part of the increase in expenditure is accounted for by growth in pension expenditure in recent years. In 1990/91, pension expenditure accounted for 7.2 percent of police spending; in 1999/2000 it is expected to be 13.8 percent.

Important as it is, the strategy to improve police performance is not limited to *Best Value*. Other critical elements include:

- the 1997/98 Comprehensive Spending Review (CSR), which highlighted the need for savings in key areas such as sickness, ill-health retirements, asset management and procurement. As a result, police authorities and forces have been required to find two percent year-on-year efficiency savings for the three years starting in April 1999, and to include the planned efficiency improvements in their annual Policing Plans.
- the adoption by the Association of Chief Police Officers (ACPO) of Activity Based Costing (ABC) in all forces, to improve the understanding of the cost of services and to guide more effective decision-making at force level. ACPO is also piloting the use of Business Process Analysis (BPA) to understand differences in processes, costs and effectiveness between forces.

A theme of the Productivity Panel's studies is the importance of bold and aspirational short and long-term targets, as part of a broader performance management framework. Well-designed efficiency targets are important vehicles for improving productivity. The focus of these targets should be improving the level of outcome performance for a given level of police funding, and to bring the performance of all forces closer or up to the level of the best. This means that the highest efficiency targets should be set for the poorest performers.

Doing this requires good and reliable comparative efficiency measures. While there are a number of frameworks that measure aspects of police performance in England and Wales, none systematically relates cost to outcomes.

This is a fundamental gap in the police efficiency and *Best Value* strategies. Without a cost-to-outcomes measure, it is impossible properly to compare police performance across forces under *Best Value* because, for example, one force's very good crime outcomes may have been achieved on a larger asset base compared to other forces.

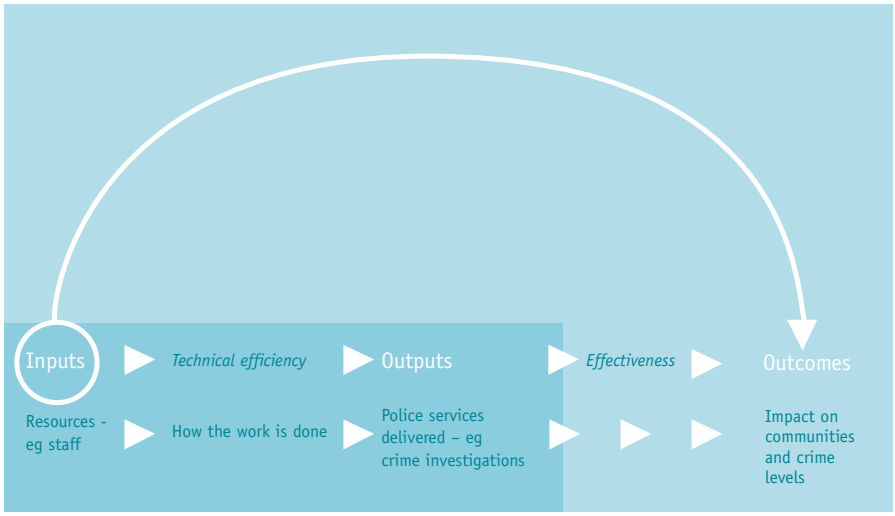
Systematically measuring relative police force efficiency – relating inputs to police outcomes – is notoriously hard to do with traditional analysis methods because of the complex and often indirect relationship between inputs, outputs and outcomes. To illustrate, a single police patrol could be simultaneously contributing to a number of outputs or outcomes – preventing crime, solving a crime, and reducing the fear of crime. Efficiency measurement also has to take into account that socio-economic and environmental factors influence the success of police work. Moreover, the data requirements to estimate efficiency can often be daunting.

But, while difficult, a systematic measure of comparative efficiency is also desperately needed. This Public Services Productivity Panel project was therefore commissioned to:

- review the various initiatives currently in place to improve understanding and measurement of the relationship between police inputs and outputs – including the development of ABC, BPA, police efficiency plans and new *Best Value* indicators.
- specify which tools and techniques, that might add to this knowledge, are likely to have most practical value, and what further steps should be taken to apply them in the police context.
- identify other areas of public sector activity that might benefit from a similar approach.

Chapter 2 – What Do We Mean by "Efficiency"?

There are many possible dimensions to the concept of "efficiency", that are outlined in the following diagram.



The diagram shows the relationship between inputs, outputs and outcomes. The typical interpretation of "efficiency" considers the relationship between inputs and outputs – shown in the shaded box. The widest concept of efficiency would consider the relationship between inputs and outcomes – the bold arrow at the top of the diagram. The widest definition of efficiency would implicitly capture two concepts:

i *technical efficiency:* ensuring that the highest level of outputs is delivered for a given level of inputs.

ii *effectiveness:* ensuring that the right outputs are delivered to meet the desired overall outcomes, and that the outputs are of good quality.

This report uses the widest concept of efficiency. This reflects the Government's overall focus on finding better ways to achieve the desired policy outcomes. Improving efficiency is about finding better ways to improve the performance of forces in meeting their overarching aims and objectives for the resources that each police force has. It need not be about cutting costs or reducing staff or shifting cost burdens onto others.

This is a crucial point. Whenever this report discusses "increasing efficiency", what it is really discussing is *how to improve police performance for the available funding*.

Making a link between inputs and outcomes assumes that such a link, in fact, exists. In reality, before a measure of input-outcome efficiency is made, it is impossible to definitively state what that relationship is – one of the purposes of doing the efficiency measurement is to actually discover what the true relationship is. But it is reasonable to assume that both direct and indirect input-outcome relationships exist – otherwise there would be no point having the police. Some police work might have only an indirect link to outcomes – crime preventative work and foot patrols might be examples. Other police activity would demonstrate quite a direct relationship on outcomes – such as targeted tactics aimed at persistent burglars or premises at high risk of burglary.

There are two approaches to measuring inputs-outcomes efficiency. The first approach is to measure "absolute efficiency". This is a measure of the theoretically possible performance of a police force; that is, if the police force used all of its resources in the best possible manner. While this would be the ideal efficiency statistic, it is also not measurable – while some forces may claim to be at the absolute cutting edge, nobody really knows what the highest theoretically possible level of performance is, or how it is changing over time. This problem is not unique to the police – it bedevils both the public and private sector.

The second approach is to measure "relative efficiency" – to compare the performance of each force with others. This needs to be done in a way that recognises that even the best relative performers should not be standing still, but improving their performance over time. This is the approach explored in this study.

Chapter 3 – The Use of Police Efficiency Measures in England and Wales

Police authorities measure and monitor police performance on a local basis, and have been doing so for years. At the national level the Home Office, the Audit Commission and HMIC monitor police performance through a framework of performance measures and (in the case of HMIC in particular) an inspection/review programme. The focus of their national frameworks is measuring police economy and effectiveness.

To date, none of the performance measures, at either the local or national level, directly relate cost to outcomes. The national performance measures that include any reference to costs are the current two percent police efficiency targets, force cost per head of population and running costs of estates. But none of these measures effectively relates inputs to outcomes, and so there is no good measure of police efficiency currently available.

Box 1: Why the Current Two Percent Police Efficiency Plans Do Not Provide a Good Measure of Efficiency

The current two percent police efficiency plans outline how each authority and their force will meet this year-on-year efficiency target. The identified efficiencies are to be invested back into front-line policing.

The current plans ask forces to identify both savings and how they will increase outputs from the saved resources. But:

- the plans only have to relate inputs to outputs for the two percent of savings; they need not relate inputs to outputs across all of a force's expenditure.
- there is no need to relate the size of the savings to the size of the gain in outputs.
- outputs are poorly defined – the guidelines refer to an increase in front-line policing or meeting or improving on Key Performance Indicators.
- the format of the efficiency plans does not encourage good or flexible management practice as there is a heavy input focus.
- there is no standard costing for calculation of efficiency savings; for example, there was a wide variance in the unit cost of a constable used.

More fundamentally, the current efficiency plans only match inputs to outputs, not to outcomes.

The efficiency plans were the only practical approach for implementing the key findings of the CSR. While they were appropriate for their time, it is now time to move on.

The absence of good efficiency measures means:

- police authorities and forces are unable to demonstrate their current level of efficiency and their previous progress towards becoming more efficient. This means that the police are poorly placed to argue for more funding.
- all authorities and forces have been given the same two percent savings target irrespective of their past record or scope for further savings.
- under *Best Value*, police authorities will be unable to compare systematically input-to-outcome performance.
- police authorities and forces find it hard to identify where efficiency improvements could be made or to set internal efficiency targets.
- the Home Office and HM Treasury do not know how efficient forces are, or which are the most efficient. As a result, these departments tend to estimate the size of possible efficiency gains on the basis of top down reviews such as the CSR, and to

propose across-the-board mechanisms such as the current efficiency plans.

This study recommends a way of filling these gaps. It also outlines how the resulting comparative efficiency measures should be used to set targets designed to raise the performance of all forces closer to the level of the best.

Recommendations:

a **a better approach to measure relative police efficiency systematically is needed. This report recommends an approach that measures efficiency by relating key outcomes to the overall cost of each force.**

b **the measurement of police efficiency should focus on the performance of police forces in delivering outcomes that contribute to the overall aims and objectives of policing, for the available funding.**

c **future police efficiency targets and measurement should be based on the results of systematic measures of relative police efficiency. The results should be used to identify differential efficiency targets at authority/force level, rather than setting across-the-board targets. The aim should be to raise the poorer performers up to the level of the best and to minimise the danger that efficiency targets penalise efficient forces.**

Chapter 4 – Input and Outcome Measures

All techniques for measuring comparative police efficiency would work best when there are a limited number of input and outcome variables relative to the number of forces being measured. In the case of the police in England and Wales, the number of forces (43) is relatively small, and discipline is required on the number of input and outcome variables to be used. This technical constraint would be overcome if efficiency measurement is undertaken at Basic Command Unit (BCU) level (see Box 2 below) or by using time-series data. However, even if BCU level data is used, a limited number of variables should be preferred, as this would facilitate the task of establishing and tracking relationships between variables, and limit the scope for data and measurement error.

Box 2: Efficiency Measurement at BCU Level

The focus of this study has been on force-level analysis because the collection of input and outcome information is, at this stage, relatively straightforward at this level. However, the Government has already signalled its intention to increasingly analyse BCU performance, for example, by piloting HMIC inspections at BCU level.

The techniques outlined later in this report can be easily adapted to analyse results at BCU level – provided the necessary input and outcome data is available. However, this would require further detailed work, especially on how to:

- capture outcome data at BCU level.
- address data comparability issues – for example, in the allocation of a force's overhead costs between BCUs. Solving these issues might possibly require central regulation of data standards.
- handle the more pronounced statistical fluctuations that would accompany measurement of performance at BCU level.
- manage the tensions that would emerge in accountabilities – as accountability for performance under *Best Value* would rest at police authority level, not at BCU level.

An initial, non-statistical, *a priori* analysis of possible input and outcome measures was undertaken as part of this study. Its sole purpose was to assess whether a relatively small number of measures could capture the essence of police inputs and outcomes. If so, that would help identify which efficiency measuring techniques could be considered subsequently. A summary of this analysis is given in the remainder of this Chapter. Details are provided in Annex A.

It is stressed that this is an exploratory analysis only. As discussed under the "Next Steps" section later in this Chapter, a more comprehensive and systematic analysis is required. This subsequent analysis should also test to see what aggregation biases – if any – arise from specifying a relatively small number of input and outcome variables.

Input Measures

The high-level police inputs are staff costs, operating costs and the consumption of capital assets (buildings, vehicles, information technology). The first two inputs can be approximated by "net revenue expenditure". However for the purposes of efficiency measurement, it is recommended that uncontrollable pension spending[3] be excluded from net revenue expenditure, as this relates to past wage structure and demographic factors entirely outside a force's control. Removing uncontrollable pension costs would avoid the possibility of efficiency measures being unduly influenced by these factors. The principle of removing cost items completely outside a force's control should extend to other similar costs, such as the National Crime Squad levy and the NCIS levy.

At present there are no consistent measures of the capital assets consumed in delivering police outcomes on an annual basis. Capturing capital usage is important

3 Police pensions are funded on a pay-as-you-go basis, and therefore almost all current pension costs result from historical legacies. The only elements of current pension costs that are under police force control are ill-health pensions and injury awards expenditure, where there is some variation in costs between forces, reflecting the varying strength of management action to control them. Information on the costs of ill-health pensions and injury awards is currently not collected, but HMIC is expecting to collect some data shortly. When this data becomes available, the costs of ill-health pension and injury awards liabilities should be *included* in the total costs used to measure the efficiency of each police force.

because some forces may have received relatively high capital investment over the years, which should have an effect on current measured efficiency. If a capital consumption measure is included as an input variable, any one force's relatively high use of capital will be taken into account.

Further analysis of how best to capture capital usage will be required during the implementation of this report.

Outcome Measures

The Government has recently developed a set of overarching aims and objectives for the police, which requires the police to:

- promote safety and reduce disorder.
- reduce crime and the fear of crime.
- contribute to delivering justice in a way that secures and maintains public confidence in the rule of law.

More recently the Government, in consultation with policing stakeholders, has developed a suite of *Best Value* Performance Indicators (BVPIs) that, amongst other things, identify a range of high level outcome performance measures that can be linked back to the overarching aims and objectives. As far as possible, the outcome measures for efficiency estimation should come from this *Best Value* suite. Selecting BVPI outcome measures will make it easier to fit the recommendations of this report within the *Best Value* framework. Using the BVPIs should also mean that efficiency estimation should not generate demands for new information sets from the police. More fundamentally, police authorities and forces have agreed that the BVPIs represent key police outcomes that they should be measured on.

That said, it is also critical that the selected outcome measures capture the essence of police outcomes and thus, implicitly or explicitly, the many dimensions to policing. This does not mean that there has to be a multitude of outcome measures. The focus of the outcome measures should be on what the police are being expected to achieve for the money they have. This is different from trying to model everything that forces do on a day-to-day basis.

While there is every reason to expect that the BVPIs would provide a good starting point for the final outcome measures, it may become necessary to alter the current BVPI measures or even look outside the suite.

Not all of the BVPIs will be necessary or appropriate for measuring comparative police efficiency. A very initial review of the BVPIs suggests that the key police outcomes could be captured in a relatively small number of indicators. These are listed in Box 3.

It cannot be stressed enough that the selection of a subset of the agreed BVPIs does not mean that somehow they are more important than the remainder. The BVPIs identified in Box 3 are likely to form the core of BVPIs that are important for measuring cost-to-outcome efficiency. The other BVPIs are important indicators of overall police authority (and force) performance, detailing aspects not covered by a strict efficiency measure, such as aspects of police quality of service, community relationships and corporate health. They will also be important in forming judgements on efficiency – as discussed in Chapter 6.

Box 3: Possible Outcome Measures using BVPI Data for Efficiency Measurement Purposes

An a *priori* review of the *Best Value* indicators suggests that the following selection of BVPIs might capture the overarching aims and objectives of the police and have a strong relationship with costs:

BV125R Recorded crime: total crimes per 1000 population.
BV125D Recorded crime: percentage detected.
BV126 Domestic burglaries per 1000 households.
BV127 Violent crimes per 1000 population.
BV128 Theft of and from motor vehicles per 1000 population.
BV129 Number of offenders dealt with for supply offences in respect of Class A drugs per 1000 population.
BV130 Public disorder incident count per 1000 population.
BV132 Number of road traffic collisions (RTC) involving death or serious injury per 1000 population.

A number of these BVPIs include both the level of recorded crime and detection rates. As reducing crime is one of the police's overall aims and objectives, the level of crime would be the better measure.

Other key police outcomes include surveyed measures of crime rates and public feelings of confidence and security. These measures are reliant on good, comprehensive survey data. This data is currently not available but could be provided with the 2001 British Crime Survey. Once this information becomes available, the following could also be included as key outcome measures:

BV120 Level of crime, using British Crime Survey.
BV121 Fear of crime, using British Crime Survey.
BV122 Feelings of public safety, using British Crime Survey.

If the survey measure is robust enough, the surveyed measure of the level of crime (BV120) might be better than, and could replace, total recorded crime (BV125).

The recommended review of the selection of outcome measures should ensure that double counting is avoided. A number of these measures capture elements covered by others. For example, the total recorded crime statistics (BV125R) will also include burglaries (BV126), violent crimes (BV127) and automobile-related theft (BV128). The final choice of outcome measures should avoid double counting of outcomes – but in a way that keeps a "level of overall crime" measure, as this captures other crimes, such as fraud, not specifically measured. Less obviously, perhaps, if crime rates drive the fear of crime, having both measures in the models could also lead to double counting.

It should be noted that some of these measures are of the form that "more-is-better" (eg: a higher detection rate is preferred over a low one) while, for others, "less-is-better" (eg: lower crime rates are better). When it comes to efficiency measurement, all of the indicators will have to work in the same direction. This will mean that, in practice, some indicators will have to be "inverted".

Socio-economic and Environmental Factors

A range of environmental and socio-economic factors will significantly affect policing costs and outcome performance. These factors describe aspects of a force's operating conditions that are outside the force's control. Box 4 discusses what some of these external factors might be – using, as examples, the factors identified in the Home Office's models for forecasting crime levels. Any analysis of comparative police authority and force efficiency will have to identify and take into account these effects, so that the results and comparisons are not distorted by differences in operating environments.

Box 4: What Do We Mean By "Socio-Economic and Environmental Factors"?

External factors influence police outcomes. The Home Office's models of crime rates illustrate the types of things these factors cover, such as:

- number of young men.
- stock of goods available to be stolen.
- changes in consumer expenditure.

This study has not explored the extent, if any, to which these factors influence police outcomes. Nor has this study explored whether other external factors should be incorporated into the models. This must be done as the relevant models are being built.

An alternative to placing socio-economic and environmental factors explicitly into the models is to compare the raw results using HMIC's "family of forces" methodology. This approach identifies key socio-economic and environmental factors facing forces,

and groups forces facing similar conditions into "families" that could be compared with each other – but not with members of other families without first adjusting the outcome or cost data.

While not ruling out the family of forces approach, environmental factors can be explicitly factored into efficiency-measuring models. This should be the preferred approach because this offers the potential benefit of quantifying the impact of these factors. But if the models are not able to adequately control for external factors, the family of forces approach could be used in addition.

Data Quality

Any efficiency measure is only as good as the data that it is using. Historically, crime recording practices have varied significantly between police forces – for example, in the recording of apparent crimes, the categorisation of crimes and in clear-up statistics. As a result, cross-force data consistency has been a constant issue.

The Audit Commission will be auditing the *systems* used by forces in collecting BVPI data. This promises a greater level of consistency than has historically been the case. HMIC also inspect data collection systems as a routine part of its inspection process. But, even with sampling, neither the Audit Commission nor HMIC can audit the day-to-day data recording and categorisation practice for every BVPI. The Audit Commission will not be giving an assurance that each BVPI measure is completely accurate. So while the BVPI data should be better, it will not necessarily be perfect.

These data risks must be acknowledged from the outset, and steps taken to address them. In the first instance, this means not always accepting the efficiency estimates at face value, but doing some analysis to understand and validate them. But the data risks are not a reason not to measure police efficiency. The real answer to data problems is to improve the data. *Best Value* should provide an incentive to do that – forces will be progressively upgrading their systems if they have not done so already – although the recommendations of this study should give that work even greater urgency.

Another data risk may arise if the analysis of socio-economic and environmental factors relies on dated data, such as census data, which could be up to ten years old. Using the most current information, as soon as it becomes available, and cautiously interpreting the raw efficiency measures are the only ways to address the problem of dated data.

Next Steps

BVPIs are in their early stages of development. For a number of the BVPIs, some issues about how they will be defined and measured have yet to be resolved. The development of BVPIs needs to continue alongside the analysis of their applicability and completeness for comparative efficiency measurement.

As BVPIs appear to capture the key outcomes required to fulfil the overarching police aims and objectives, BVPIs should be the starting point for an appropriate set of outcome measures. However the Home Office should be prepared to move beyond the current set of BVPIs if necessary, provided that doing so would not lead to a *net* increase[4] in the statistical information required from the police. The principles guiding this analysis should include:

- the outcome measures should relate back to the overarching aims and objectives for the police.

4 That is, any new data requirement that proves necessary for efficiency measurement should be offset by a reduction in data requirements elsewhere.

- the outcome measures should relate to the level of costs being used in the efficiency analysis.
- the outcome measures should relate to outcomes that police forces are able to influence.

If the further analysis of outcome measures finds relevant police outcome measures outside the current BVPI suite, this should lead to changes in the BVPI suite itself. BVPIs should reflect the key outcomes that the police are expected to achieve. If the suite is currently inappropriate, it should be changed.

This analysis should also map and quantify the socio-economic and environmental factors that influence BVPI outcomes. These steps are taken into account in the recommended implementation plan in Chapter 9 of this report.

Recommendations:

a a small set of input and outcome measures, agreed between the Home Office and its stakeholders, should be used for measuring relative police force efficiency.

b the input measures should be based around:

i "net revenue expenditure" less annual uncontrollable expenditure on police pensions and other uncontrollable costs.

ii a measure of capital assets "consumed" by forces in producing the outcomes.

c the selected outcome measures should meet the following principles:

i the outcome measures should relate back to the overarching aims and objectives for the police.

ii the outcome measures should relate to the level of costs being used in the efficiency analysis.

iii the measures should relate to outcomes that police forces are able to influence.

d before implementing the recommended relative efficiency measuring techniques, further analysis should be completed to ensure that:

i starting with (but not limited to) the agreed BVPIs, all outcome measures meeting the principles in (c) are identified and captured in the analysis.

ii all measurable socio-economic and environmental factors having a significant impact on outcome efficiency are identified and the effect quantified.

iii appropriate measures for the consumption of capital assets, and uncontrollable aspects of pension and other costs, are identified.

e if the further analysis of outcome measures finds relevant police outcome measures outside the current BVPI suite, the BVPI suite itself should be changed.

f the development of the input and outcome measures should take into account the possible extension of comparative efficiency measurement to BCU level at some point in the future.

Chapter 5 – A Better Measure of Police Efficiency

There is no flawless method for unambiguously measuring relative police efficiency. All methods necessarily rely on simplifying assumptions and approximations to be able to model the complexity of what happens in the real world. All techniques are therefore open to criticism that they are too simplistic and fail accurately to capture all real world nuances. Such criticisms – although undoubtedly correct in their own way – should not rule out the use of techniques that provide insights into, and estimates of, police efficiency. The task is to find the best possible approach to measuring relative police efficiency while recognising the limitations that any techniques have.

The most straightforward approach to measuring comparative police efficiency would be a simple efficiency index. In principle, this index would take the weighted sum of the relevant BVPIs, and relate this weighted sum to total cost. Box 5 provides details of how such an index might be constructed.

There are two key problems with using a simple efficiency index. First, it allows for no variation in the weight that could be assigned to each outcome for each authority and its force. But the relative importance of each outcome could legitimately vary from force to force, reflecting local circumstances and local police plans. An approach that lets the outcome weights vary is preferable. Second, such an index would implicitly assume a "linear" or straight-line relationship between inputs and outcomes; that is, if a force doubled the inputs it would get double the outcomes. This is unlikely to hold in practice. An approach that allowed for non-linear relationships between inputs and outcomes would be preferable.

This study has therefore sought alternative approaches to a simple efficiency index. But the alternatives are, in principle, extensions of the simple index. What the extensions do is to find ways around the fixed weights, and the linearity assumptions that hamper the more straightforward index.

Box 5: A Simple Efficiency Index

A simple efficiency index could take the following form:

$$\text{Relative Efficiency Index} = \frac{BVPI_1w_1+BVPI_2w_2+BVPI_3w_3+.....+BVPI_nw_n}{\text{Total cost per 1,000 population}}$$

where:

- w_i = the weight given to each outcome measure, reflecting the relative importance of each outcome. Under a simple index, each weight would need to be set centrally, based on subjective judgement. The weights need to add up to 100 percent in total.
- $BVPI_i$ = the particular *Best Value* Performance Indicator, expressed in the same direction (ie: they are all expressed so that "*more-is better*" or "*less-is-better*").
- total cost = total resources used, which is expressed in cost per 1,000 of population as the *Best Value* indicators also relate crime statistics back to population in this way.

Under this approach, the index values would be computed for each force, and the highest index value given a score of 100 percent. The index values for all the other forces could then be related back to this highest score.

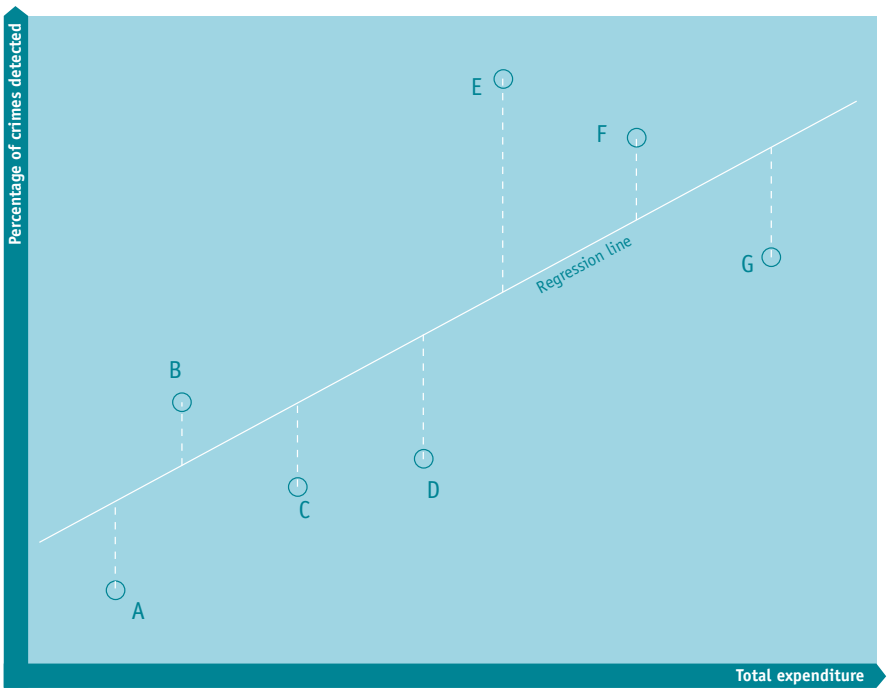
Two Techniques: Stochastic Frontier Analysis and Data Envelopment Analysis

Stochastic Frontier Analysis (SFA)

SFA is a form of "regression analysis" that perhaps can be best explained by building up a simplified, stylised diagram.

In this very simple example, there is only one outcome – percentage of crimes detected – and each force's performance on this measure is plotted against that force's total expenditure.

It is first useful to plot a simple regression line through the data. This simple regression identifies a "line of best fit" through the observed data. There will be a discrepancy between the cost predicted by the regression line, and the observed cost for each force. In simple regression analysis, all this difference will be attributed entirely to (in)efficiency.

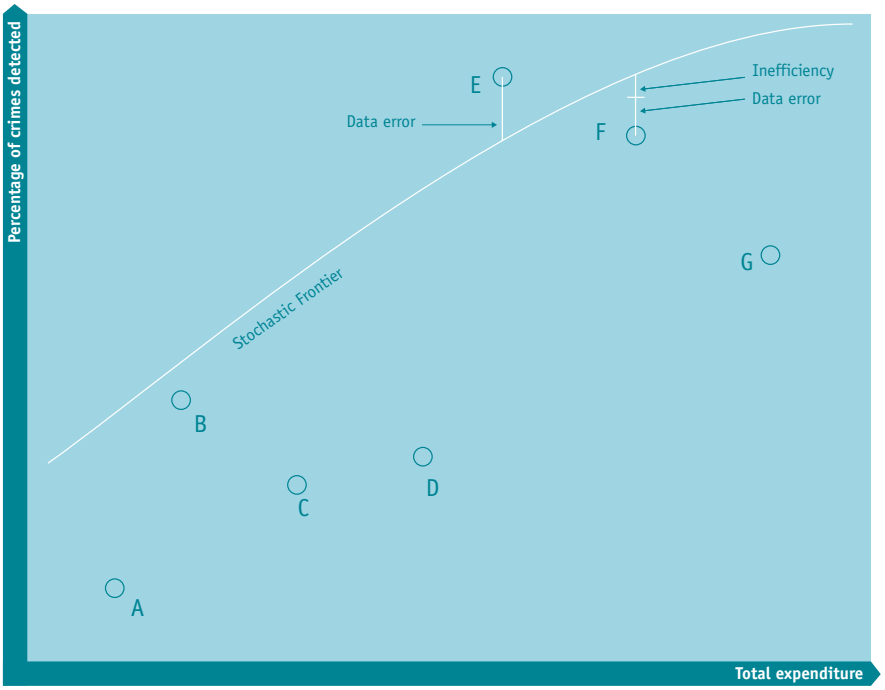


In the graph, the forces above the regression line (Forces B, E and F) would be considered to be relatively efficient (they detect more crime than expected for the expenditure incurred). In contrast, those beneath the regression line would be considered to be relatively inefficient (they detect less crime than expected for the expenditure). In ordinary regression analysis, the forces above the line could be regarded as relatively "super-efficient".

However, SFA assumes that part of the observed gap between predicted and observed performance will be data or measurement error, part will be omitted variables and only the remainder will be actual (in)efficiency. To address these problems, SFA makes an assumption about the distribution of these errors, and excludes the estimated errors from the efficiency calculation.

In effect, what SFA does is identify the most relatively efficient force – in this case it would be Force E. If there were no data or measurement error, Force E would lie exactly on the stochastic frontier. However, because the measurement of Force E's performance is not perfect it actually lies above the frontier.

The stochastic frontier identifies the predicted performance for the best force, should its actual costs vary. The other forces would be below this frontier and therefore are relatively inefficient compared to the best.



But, for forces below the frontier, SFA assumes that some of the gap between actual and predicted best performance will be data and measurement error. Based on this assumption, SFA decomposes the gap between actual and best performance into data error and estimated inefficiency. This is shown for Force F – only some of the gap between its own performance and the predicted best performance will be attributed to inefficiency.

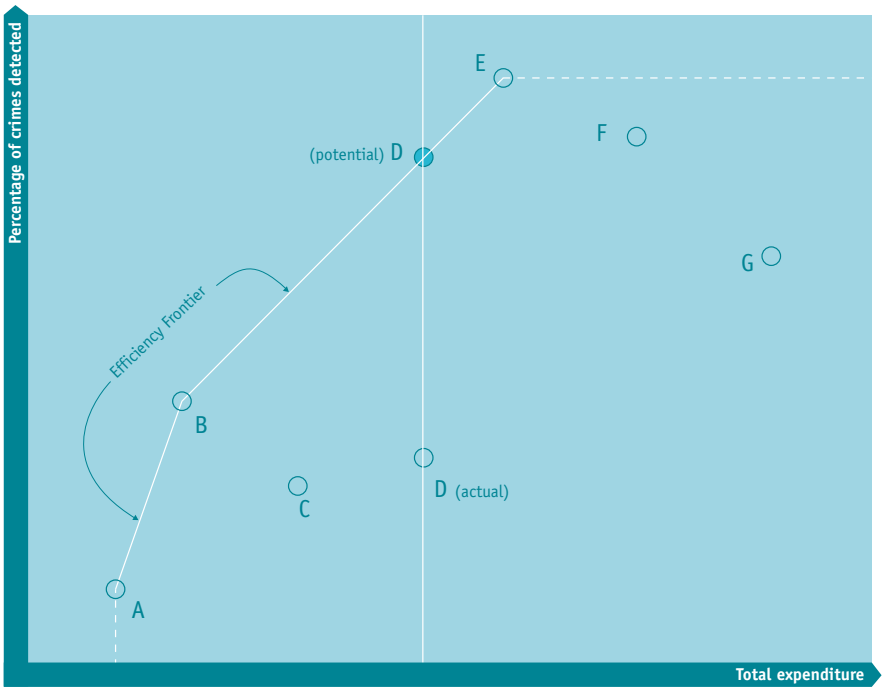
Although the diagram is shown linking cost to a single outcome, SFA can handle a large number of outcomes. If there is more than one outcome, SFA effectively calculates the weights that would best explain the observed data, and then sets the weights at those proportions. SFA therefore establishes a common set of outcome weights across all forces.

A fuller, more technical, discussion of SFA can be found in Annex C.

Data Envelopment Analysis (DEA)

In the context of measuring relative police efficiency, DEA is a form of "linear programming" that mathematically seeks to find the best level of outcome performance for each force, subject to the constraints and resources they face.

It is again best explained by looking at a simple diagram, using the earlier example that discussed SFA. As before, there is only one outcome (percentage of crimes detected) which is related to a single input (total expenditure). The performance of all forces is again plotted.



DEA will identify the forces with the outer-most observations. These forces effectively define the boundary of an "efficiency frontier". In this example, Forces A, B and E define the boundary of the "efficiency frontier" – that is, for their level of expenditure, they generate the highest detection rates of all of the forces. Force D's position is within the frontier, and cannot be more efficient than any of the forces on the frontier. It is possible to map the position that Force D could take on the frontier if it were more efficient. This potential point is also shown on the frontier. By measuring the distance that Force D is from the frontier, it is possible to estimate by how much it is relatively less efficient than the best.

Applying this methodology to the real world police environment is obviously more complicated than this very simple example. But the underlying concepts – identifying the best performing forces and comparing the efficiency of the others against the best – remain unchanged.

In this very simple example, as there is only one outcome measure, the DEA model will put a 100 percent weighting on this one outcome. If there were more than one outcome, DEA would need to determine what weight it will put on each of the outcomes. In the purest form of DEA, the weights on each outcome would vary so as to maximise the efficiency score for each authority/force. This allows any authority/force to be compared as favourably as possible with all the others. But it could also allow a very high weight to be given to a relatively low priority outcome. Box 6 discusses more fully how weights are set in DEA, as well as show how DEA is actually an extension of the very simple efficiency index discussed at the start of this Chapter and in Box 5.

A fuller, and more technical, explanation of DEA can be found alongside the discussion of SFA in Annex C.

Box 6: A Simple Efficiency Index – and its Relationship with DEA

The simple efficiency index in Box 5 took the following form:

$$\text{Relative Efficiency Index} = \frac{BVPI_1w_1+BVPI_2w_2+BVPI_3w_3+.....+BVPI_nw_n}{\text{Total cost per 1,000 population}}$$

In its most basic form, DEA also starts with exactly the same simple efficiency index for each force. But DEA then allows the weights to vary in such a way that a force's performance would be viewed in the best possible light compared to other forces. DEA does this by solving a linear programming model that identifies a set of outcome weights that gives each police force the highest possible DEA score (although the highest possible score of all the police forces is 100 percent).

Before it accepts this score for any one force – let this force be "Force A" – the DEA model checks to see what impact that set of weights would have for all the other police forces. The model will reject the solution for Force A, if Force A's set of weights would generate scores greater than 100 percent for any one else. Force A's weights are therefore constrained so that if they were applied exactly to the BVPIs of any of the other 42 police forces, none of the other forces could get a score higher than 100 percent.

Strengths and Weaknesses of Each Approach

The preceding discussion illustrates that these techniques approach the problem of efficiency measurement from complementary, but different, perspectives. As would be expected, SFA and DEA have different strengths and weaknesses based on the assumptions underpinning the respective approaches.

One distinguishing feature of DEA is that it does not depend on any prior assumption about the relationship between outcomes and costs – other than to assume that an efficiency frontier can be constructed in a "piecewise linear" manner (that is, in the above simple example of DEA, a straight line can be drawn between the furthermost observations). That said, DEA does not need a linear relationship between inputs and outcomes to work well – it can assume a non-linear relationship between them. As discussed earlier, a further strength of DEA is that it allows the weights on the outcomes to vary in such a way to show the performance of each authority/force in the best possible light.

DEA assumes no data or measurement error in the observations – it assumes that the forces forming the frontier are genuinely the most efficient. This raises the risk that incorrect outlying data points that form part of the frontier can adversely affect DEA results. Forces that are particularly large may end up forming part of the frontier just because of their size – irrespective of their efficiency – as could forces benefiting from favourable measurement errors. As DEA does not assume any particular relationship between costs and outcomes, it is not possible to complete any statistical tests on the veracity of the underlying model, nor of the results.

In contrast to DEA, SFA assumes a particular relationship (which again need not be linear) between the inputs and the outcomes. This allows statistical tests to be made on the accuracy of the assumed relationships, the significance of each variable or parameter, and the significance of the results. A further key advantage of SFA is that it does not assume that all of the gap between the observed and expected performance is due to inefficiency. The gap could be due, in part, to data error or missing variables. The main drawback with SFA is that a simple assumption must be made about the distribution of these errors. Changes to this assumption will affect the efficiency estimates (but probably not the rankings).

This discussion illustrates that the strengths of one technique tend to address the weaknesses in the other. Therefore, the challenge is to find the way to get the best of both worlds.

A Recommended Approach to Measuring Police Efficiency

This study recommends the use of both SFA and DEA to better measure police efficiency, using the processes to analyse the results outlined later in this report. Each technique would be used for each force, with the results being used as a cross-check on the results of the other.

A range of technical advice was sought on this approach. There was generally consistent advice from leading experts in the fields of economics and econometrics having practical experience in efficiency measurement. The conclusions of these experts are reproduced in Annex D. Although each of the economics and econometrics experts identified important technical issues to be addressed in formulating and implementing this approach, their general conclusions are summed up in the following excerpt from the report presented by Professor Leigh Drake and Doctor Richard Simper of Loughborough University:

"The main conclusion of this report is that we strongly endorse the use of DEA and regression analysis in tandem ... The use of both DEA and a parametric frontier technique such as ... SFA undoubtedly represents the 'state of the art' in terms of relative efficiency analysis and would represent the optimal approach to efficiency analysis across police forces."

As well as being technically "state of the art", the key benefit of both SFA and DEA is that they are able to relate inputs to outcomes – the overall measure of efficiency. Of the known efficiency measuring techniques, they are the only ones to offer this capability.

The next chapter of this report outlines how the results of each technique could be used to check the results of the other, as well as provide diagnostic information on the underlying causes of relative (in)efficiency.

There was, however, dissenting advice from other experts who were concerned that SFA and DEA were deceptively simplifying "black box" techniques. Concern was expressed that these techniques potentially obscured important value judgements and assumptions about the nature of efficiency, whereas it was important that these were made explicit, tested rigorously and laid open to critical review.

An alternate (or complementary) approach was proposed for development. In this approach, criteria would be established for the prior selection of outcome measures, and weights applied to the various outcomes delivered by the police to reflect their true value to society.

The alternate approach suggested by the other experts would be based around the simple efficiency index discussed at the start of this Chapter. The index was not further developed by this study for the reasons outlined earlier. The need for an explicit, rigorous and transparent process is acknowledged and forms the basis for the implementation process discussed in Chapter 9.

Box 7: The Use of DEA and Regression Analysis in Efficiency Measurement

Although the techniques have existed for over two decades, the use of DEA and variations of regression analysis/SFA for systematic public sector efficiency measurement is relatively new and mostly confined to the academic literature. But a couple of examples show what can be done.

The Office of Water Regulation (OFWAT) used both ordinary least squares regression and DEA in its 1994 review of water companies' costs. It found the results generally agreed:

"In most cases the results [on company efficiency in clean water operations] were similar to those of the regressions. If they were significantly better, the Director [of OFWAT] moved the company up one band [on efficiency in clean water operations]."

Researchers from York University have recently used a range of modelling techniques –including SFA and DEA – and a range of model specifications within each technique to explore the efficiency of NHS trusts and to check the performance of the various techniques. The study found that the specified models of SFA and DEA measured slightly different forms of efficiency (see Annex C) and did not always produce consistent results. But the study concluded that each technique has strengths and weaknesses, and several specifications of each approach should be used to develop ranges of efficiency rather than a single point estimate efficiency score. In this way, the two methods "can best be used in conjunction with one another".

The New South Wales Treasury has used DEA to explore the relative efficiency of a number of its agencies, including its police forces. The nature of the NSW budgeting system means that DEA has most value as an internal management tool to understand and promulgate best practice. These agencies are now requesting that other techniques – such as SFA – are used to test the sensitivity of the DEA results.

Sources:
Monopolies and Mergers Commission; "South West Water Services Ltd" (London: HMSO 1995) page 415.
van der Merwe, R; "Efficiency as a factor in Hospital Cost Variation" (Department of Economics and Related Studies, University of York, 1999).
Personal communication with the NSW Treasury.

Principles for the SFA and DEA Models

Appropriate SFA and DEA models for police inputs and outcomes will need to be developed as part of the implementation of this report. These models should meet the following principles.

A Limited Number of Input and Outcome Measures

Given the relatively small number of police forces, SFA and DEA will work best when they are used with a limited number of input and outcome variables, provided that other biases are not introduced by using only a small number of variables. Net revenue expenditure (less uncontrollable pension expenditure, and other expenses completely outside a force's control) and consumption of capital assets should be used as the input measures. A selection of BVPIs should form the starting point for the selection of a relatively small number of outcome measures.

The Weights on the Outcomes Should be Constrained

Setting weights on outputs or outcomes is not new – they exist within the current police funding formula[5]– albeit in this case the weights are on functional areas of policing rather than on outcomes. Weights are also implicit in any priority setting process – be they the national priorities indicated by the Home Secretary or the local priorities set out in local policing plans.

As noted above, the pure form of DEA allows the weights on outcomes to be chosen by the technique to show each police authority and its force in its best possible light. Depending on what an authority/force is particularly good at, this could allow implausibly high weights to be placed on some relatively minor outcomes. DEA could therefore allow an authority/force to be evaluated almost entirely on thefts involving motor vehicles, even though violent crimes might be a higher priority in the locality. In the extreme, allowing a free range of weights could allow manipulation of DEA as authorities identify the relative strengths of their forces and seek to place a very high

5 The police funding formula uses econometric models to estimate the relative demand for various police functions – such as call management, crime management and traffic management. Each of these functions is then given a percentage weight in the formula determining funding allocations.

weight on them irrespective of their overall impact on police overarching aims and objectives. At the same time, the relative importance of crimes can vary across the nation – for example, the concerns of rural authorities and forces are not always the same as those of inner city forces.

A balance must be struck between the need to have the DEA analysis based on intuitively reasonable sets of weights, and the need to reflect local priorities in policing. It must also be recognised that any process for selecting weights will rely on subjective judgement – there is no scientifically "correct" set of weights.

One approach to striking this balance is to set a national framework for ranges of weights within which police authorities could set their own local weights (or allow DEA to do it for them). The ranges could be set in a number of ways. One way is to survey police authorities and Chief Constables for their view on the relative importance of various outcomes. Another option is to look at the priorities implicit in local policing plans. A "dummy run" of the DEA analysis could also be completed, to find what weights were derived by the models and then assess whether these weights are, in fact, reasonable.

Whichever methodology is chosen, Chapter 9 of this report recommends that a representative Steering Group be established to oversee implementation decisions. The Steering Group should be charged with identifying an appropriate range of weights for each outcome measure that should be recommended to the Home Secretary who would make the final decision on the weight ranges. Local police authorities would then, if they wished, be able to narrow the nationally agreed weight limits for specific outcomes, to reflect the relative priorities in local policing plans. Alternatively, authorities could allow the DEA technique to chose the weights that would show their force in the best possible light, within the range specified by the Home Secretary.

The question is how narrow or wide the permissible range of weights should be. It is critical to bear in mind that the weights would reflect the relative importance of each outcome – say the domestic burglary *rate* – and not just the volume of work in each area. The weight would therefore reflect the relative importance of keeping the violent crime rate low. This may vary across the country – but probably not by much. A domestic burglary is likely to have more-or-less the same importance relative to a motor vehicle theft, irrespective of where in England and Wales the offence occurs. The permissible ranges of weights should be kept fairly narrow. Box 8 provides further elaboration.

Box 8: Keeping the Ranges of Weights Narrow

This study outlines how selected BVPI or other outcome measures could be used to measure relative efficiency. The BVPI outcomes currently suggested are related to crime (eg: the rate of recorded crime in each area per 1000 population) and, eventually, perception measures (fear of crime). These are substantively different from volume measures, that have frequently been used to assess police performance, and can lead to a different approach to weighting.

Rural communities have different crime patterns to urban centres – for example, there is less violent crimes in rural areas. Thus, rural authorities and forces have typically devoted proportionately fewer resources to violent crime activity.

This makes sense when considering the *volume* of work. But when a violent crime does occur in a rural location, in principle it should get much the same priority as if it occurred in an inner-city area. Moreover, the same importance should be placed on keeping the rate of crime low – no matter the locality. Focusing on *rates* of crime and *perception measures* means that a relatively uniform set of priorities can be established. This translates into a reasonably common set of weights across outcomes across the nation.

This is not to say that volumes of crime do not influence the perceived importance of crime – clearly volumes do matter. If there is a long term increase in, say, the burglary rate, this type of crime should command increasing importance. The ranges of weights should therefore be reviewed periodically to capture these changes in circumstances. But the police also need stable, longer term messages about the relative importance of various outcomes if they are to plan effectively. Outcome weights should not automatically respond to short term pressures or crime waves.

Socio-Economic and Environmental Factors Should be Identified, Quantified and Incorporated

As noted earlier, a range of environmental and socio-economic factors that significantly affect policing costs and performance in delivering outcomes should be incorporated into the SFA and DEA models.

Alternative Efficiency Measuring Techniques Considered

In addition to SFA and DEA, further techniques were evaluated as part of this study. These included Decision Conferencing, Systems Dynamics and BPA, all of which have historically been used in various contexts for assessing police efficiency or performance. These, along with good practice examples of their application, are described in more detail in Annex B. None of these techniques provide a systematic pan-force measure of comparative police efficiency. However they will have value in helping diagnose, at a force level, what in practice needs to be done to improve efficiency. This is discussed in Chapter 6.

All forces as part of their need for better management information are introducing ABC. The information from ABC is typically used to calculate the total costs of each broad function of policing. The information is used to help develop costed annual policing plans. ABC could help measure the police's *technical* efficiency, if it identified the cost of police internal units or processes and these costs were benchmarked across forces. This would be workable provided that uniform definitions and costing methodologies were employed. This is becoming increasingly possible with a standard set of coding systems being employed and common software suites introduced across forces. However, as ABC cannot cost outcomes, it is not recommended for wider relative police efficiency measurement. But ABC should continue to be implemented at a rapid pace, as it can be used as a diagnostic benchmarking tool to understand the sources of relative inefficiency. A fuller discussion of ABC is included in Annex B.

Recommendations:

a **Stochastic Frontier Analysis (SFA) and Data Envelopment Analysis (DEA) should be used to provide an assessment of the relative efficiency of each police force in delivering police outcomes. The results from each technique should be used as a cross-check on the results of the other, using the processes outlined further in this report.**

b **the weights given to each outcome measure should be constrained in the DEA model. The weights would be allowed to vary within a range for each outcome measure determined by the Home Secretary, taking full account of the advice provided by the Steering Group to be established to implement this report (see Chapter 9).**

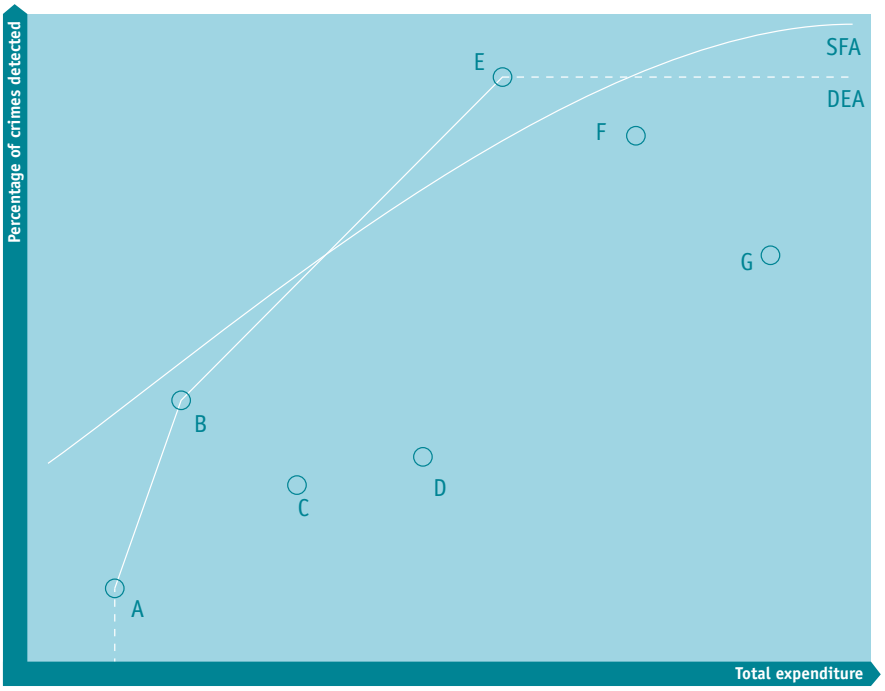
c **Activity Based Costing should continue to be implemented, in order to give police managers basic management cost information, and to assist forces carry out the benchmarking necessary to fully understand the results of the SFA and DEA analysis.**

Chapter 6 – Presentation and Analysis of the Results of SFA and DEA

The initial results of the SFA and DEA analysis will need to be carefully scrutinised and evaluated to make sure the key conclusions and findings can be corroborated. This will be particularly important in the first couple of years, and if doubt exists over the quality of some of the data.

Even though the recommended SFA and DEA framework should lead to a fairly good relationship between their results, they are different techniques, using different approaches and underlying assumptions. Each will therefore give a different estimate of relative efficiency. Therefore, between them, the techniques will provide a *spread* of comparative efficiency estimates, not a *single* relative efficiency score.

The relationship between the results of each technique can be shown in the following diagram.



This diagram shows the SFA and DEA frontiers derived in the example used in Chapter 5. It shows that, in principle, there should be a fairly good relationship between the SFA and DEA results. The observations within the frontiers (Forces C, D, F and G) are about the same distance from the respective frontiers. However, the SFA inefficiency estimates will differ from the DEA estimates because part of the gap between the observations and the SFA frontier will be assumed to be data and measurement error. This diagram also shows that DEA may identify a number of forces as being fully "efficient", (Forces A, B and E) whereas SFA will show only one force as being 100 percent efficient (in this case, Force E).

Sometimes the spread of efficiency results will arise, not because of the different approaches being used but because of the effect of one or two factors having a disproportionate impact in one model but not in the other (eg: one outcome having a

higher weighting in DEA compared to SFA). In these cases, it is important that there is a subsequent process for analysing the initial results of SFA and DEA – to get underneath the numbers and explain the key drivers and factors. *SFA and DEA can only be a guide to further diagnostic analysis – they are not a complete answer in themselves.*

This process also needs to relate the raw SFA and DEA results back to wider measures of performance, captured by (for example) the other BVPIs, to ensure that apparent efficiency has not been achieved by compromising quality or other aspects of performance.

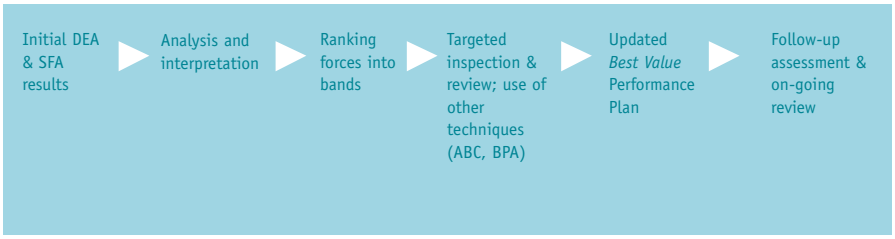
This chapter outlines the process that should be used to do this analysis, and how the resulting relative efficiency could be integrated into the wider *Best Value* process.

Understanding and Using the Results of SFA and DEA in *Best Value* Performance Planning

Best Value requires police authorities to develop *Best Value* Performance Plans, identifying the five year programme of reviews to be undertaken as a basis for improving performance. A regular series of inspections and audit – by HMIC as the *Best Value* inspectorate and the Audit Commission respectively – will take place to ensure compliance and validate outcomes. The comparative efficiency information provided by SFA and DEA will be an important element for police authorities, assisted by Chief Constables, in framing the *Best Value* Performance Plan.

Placing SFA and DEA within the *Best Value* framework suggests that analysis of the raw SFA and DEA information needs to take place through a structured process involving a number of stages summarised in Box 9.

Box 9: Stages in Analysing the Results of SFA and DEA



Analysis and Interpretation

The initial task is to understand the raw results of the analysis, to form a view of their reasonableness and reliability as well as understand what is driving them. This will involve several steps:

- first, an analysis to understand as far as possible the results of the SFA and DEA analysis. As the techniques will almost certainly offer different relative efficiency estimates, the task is to understand the differences in the scores by getting underneath the drivers for each estimate. Discrepancies may be due to the effect of one or two factors or assumptions in one model not in the other. This analysis should identify what causes the differences – which often will highlight areas for subsequent analysis and investigation – and consider which approach is generating the more reliable and robust results overall.

- second, use available diagnostic tools to interpret the high level statistics. To illustrate, one DEA software application currently being considered by the Home Office can identify:
  - a force's efficient "peers" (those forces operating in similar environments, and having a similar mix of BVPIs, but which generate better results for the available resources). The software can then diagnose how they compare on each variable. This may identify areas of best practice that can be developed in the less efficient forces.
  - which outcomes can be improved, for any given force, based on a comparison with that force's peers. It will also identify areas in which a force is performing well.
  - how many times a force has been identified as an efficient peer for other forces. This identifies key performers.
  - which variables require the most improvement across all forces. This may provide a focus for a national initiative to improve these areas.
- third, corroborate these analyses through site visits and discussion with forces on the one hand, and with other performance data on the other. This work should investigate whether there are any force-specific data issues influencing the results. It should consider the impact of extraordinary circumstances influencing the results, such as the impact of a crime wave, the impact of a serial criminal or the workings of a Crime and Disorder Partnership on the results. Finally, the analysis should ensure that apparently good or bad performance is not driven by differences in service quality (as measured by other BVPIs and other indicators).
- fourth, control for perverse incentives. Some desirable police activities – such as campaigns to increase reporting of offences – would adversely affect apparent police efficiency (by increasing the level of recorded crime) even though it would lead to a better service to the public. To prevent an efficiency measure creating perverse incentives, such factors need to be identified and taken into account in the analysis.

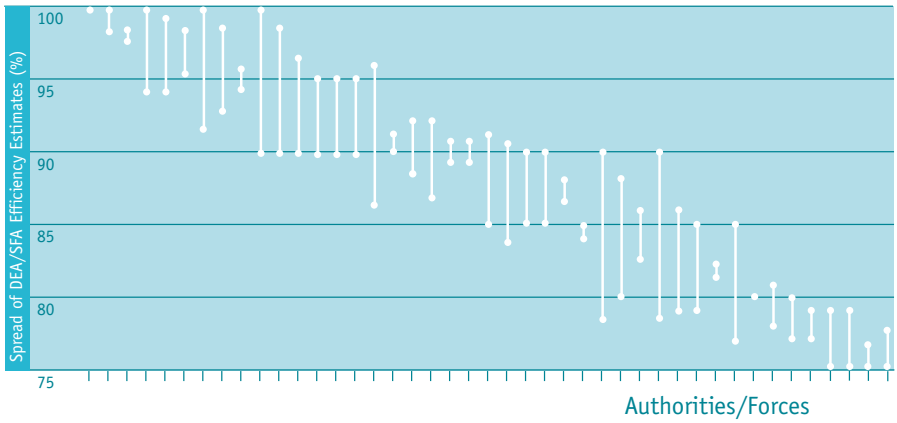
Collectively, this four-step analysis should provide a well-rounded perspective on the SFA and DEA results. There will be cases where this analysis of the factors underpinning the efficiency estimates will lead to an adjustment to the raw results.

Ranking Authorities/Forces into Bands

This initial analysis would allow authorities and their forces to be grouped in bands. An ideal situation might have them ranked into four bands, corresponding roughly to quartiles, as this would fit in best with the wider *Best Value* framework. But it may not make sense to try rigidly to determine quartiles. If there is a large group of forces within the same range of efficiency estimates, a band should be drawn around this group even if this only leads to (say) three bands being drawn.

The question is how to derive appropriate bands, given that there will be a *spread* of efficiency estimates for each authority/force. The following chart shows a hypothetical example of the SFA and DEA estimates, showing the highest and lowest estimate.

DEA and SFA Estimates



Authorities and forces could be placed into one of the bands either on the basis of the higher of the DEA or SFA estimates, or by identifying which of the SFA and DEA models is generating the more robust results, and accepting this model's results. The latter approach is recommended. While each technique is offering a valid measure of comparative efficiency, at the end of the day one is likely to perform better than the other. The best performing model should be chosen as the basis for banding authorities and forces. This would lead to more consistent, less ad-hoc, estimates of relative efficiency compared to simply choosing the best score for each.

While the boundaries of the bands should be chosen on sensible demarcations of the results, statistical tests can be completed on the bands. These tests can establish, for example, whether there is a statistically significant difference between the average of the scores of the forces in each band. These tests should be used to guide the derivation of the bands.

Part of the reason for grouping forces into authorities and bands is to target inspectorate resources into areas most needing performance improvement – as discussed in the next section. But detailed ranking within each band is not recommended. This recognises that, even with the best model specification and data collection, there is still scope for errors to occur (for example, some relevant variables might still be missing). The broad banding of the results will prevent the results from being quoted with spurious accuracy.

The initial analysis, and the process of ranking into bands, will provide pointers on the areas requiring improvement. However only a more detailed review will indicate the specific actions that should be incorporated into each police authority's *Best Value* Performance Plan to bring the poorer performers up to the level of the best. A targeted set of inspections is recommended to facilitate this process.

Targeted Inspection and Review

The grouping of authorities and their forces into bands, and the diagnostic information available from the models, provides a means for informing the targeting of HMIC/Audit Commission reviews as part of the process of assisting police authorities and Chief Constables understand the results and identify key actions.

The Local Government Act 1999 widened the role of HMIC to include the inspection of police forces for the purposes of *Best Value*. To support this role, the HMIC has designed a Diagnostic Inspection Model based on risk assessment and inspection protocols. However, the HMIC Diagnostic Model does not itself currently capture the

concept of cost-outcome efficiency. Accordingly, the HMIC Diagnostic Process needs to be supplemented if the perspective of efficiency is also to be taken into account for the purpose of informing inspections and reviews of forces.

The approach outlined below (based on three bands, for ease of explanation) is intended to complement and strengthen the current and planned HMIC process rather than replace it.

In completing their inspections and reviews, HMIC, police authorities and forces would selectively use detailed diagnostic tools such as BPA and ABC. Box 10 outlines the role that these techniques could play.

- *Top Band: Routine Inspection and Review*
  In addition to meeting HMIC's statutory obligations, a routine inspection and review would involve an analysis of the force's own performance and efficiency data to understand the indicated results. This would probably require sample inspections in a few areas. Routine inspections would pay particular attention to identifying areas of good practice, to facilitate information exchange and assist efficiency improvements in other forces. Forces in the top band can improve their own performance even further by sharing good practice with each other, since their good efficiency result is relative and not absolute.
- *Middle Band(s): Selective Inspection/Review*
  This would involve more detailed inspections in selected areas as directed by diagnostic information from the efficiency estimation techniques. As well as meeting HMIC's other obligations, it would also identify the areas of good practice in these forces for wider dissemination.
- *Lower Band: In-depth Review*
  A more detailed examination would be undertaken for those identified as the poorer performers. An in-depth inspection and review would again take as its starting point the diagnostic information from the SFA and DEA results. The in-depth inspection and review would look beyond this to understand in detail the results and to distil the lessons for improvement, for incorporation into a revised *Best Value* Performance Plan. The understanding of good practice obtained from the analysis of the top band performers, and a force's more efficient peers, would be particularly important in identifying areas for improvement.

Box 10: How Traditional Techniques Can Identify Key Actions for Improving Efficiency

As SFA and DEA can identify the top performing forces for a particular area of outcome, they will facilitate the use of traditional techniques – BPA and ABC – to identify the changes that need to occur in practice.

BPA uses a multi-level mapping methodology to break down high-level processes into sub-processes, activities and tasks. Each of these tasks can then be subjected to close scrutiny in order to determine if they are "value adding" or whether they can be rationalised, streamlined or the workflows otherwise improved. If ABC-based costs are available, the costs of each process can be benchmarked. BPA therefore facilitates effective benchmarking between forces.

Within the Police Service, the ACPO Harmonisation of Business Processes Sub-Committee is co-ordinating work to develop a national database of BPA models, making the results of the analysis comparable between forces (through standard definitions and measures). The first studies are considering probationer recruitment, crime scene management, case file preparation, custody procedures, crime information management and call handling.

If methodological and data comparability issues can be resolved, ABC could also derive unit costs of processes and organisational units through a systematic process of data collection. This will also facilitate the analysis of force costs and benchmarking.

Updating the Best Value Performance Plan

The focus of this process of analysis, diagnosis and targeted inspection and review is to provide a basis for police authorities to identify areas for improvement and action in their individual *Best Value* Performance Plans. These plans will be updated annually, although it may not always be appropriate to wait for these revised plans to start to make improvements. The Audit Commission will appoint an auditor to review these plans.

Follow-up Assessment and On-going Review

Once areas for improvement and targets have been set, an on-going process of monitoring performance will track actual changes in performance. Box 11, in Chapter 7, summarises the process whereby the DEA and SFA analysis would be repeated annually to provide this information.

Accountabilities

The focus of the targeted HMIC inspection programme is to assist police authorities and forces identify strengths and areas for improvement, as well as meet HMIC's wider statutory obligations. Determining the scope and focus of the *Best Value* Performance Plan, and its implementation, would continue to be the responsibility of police authorities, assisted by Chief Constables. Police authorities, and through them Chief Constables, who would remain accountable for the delivery of the Plan and the achievement of results.

Discussion of accountabilities raises the question of the incentives within the system to encourage high, and improving, levels of performance. These issues are discussed in the next chapter.

Recommendations:

a the raw SFA and DEA results should be analysed through detailed review and site visits in order to:

i understand differences in results.

ii identify any force-specific data issues influencing the raw results.

iii identify differences in service quality influencing the raw results.

iv identify the key factors underpinning relatively good or poor performance.

v identify and avoid any perverse incentives on the police.

and adjust the raw results as necessary.

b the analysis in (a) should also identify whether the DEA or SFA model performs the best overall. The best performing model should be used to group authorities and their forces into relative efficiency "bands" based on sensible demarcations of the actual results and statistical tests of the significance of each band. To avoid the results being presented with spurious accuracy, detailed force-by-force rankings within the bands should not be given.

c the Home Office should publish, for scrutiny, all of the judgements made and information used to:

i adjust forces' raw efficiency estimates.

ii identify which model is performing the best.

prior to the Home Office taking final decisions on these matters.

d the adjusted results of the SFA and DEA models should be used to assist police forces, aided by targeted HMIC inspection and review, to understand the efficiency measurement results. HMIC should target its inspection and review programme as follows:

Top Band:	Routine Inspection and Review
Middle Band(s):	Selective Inspection and Review
Lowest Band:	In-depth Review

The purpose of this programme is to identify good practice for dissemination to other forces and areas for improvement for incorporation into revised *Best Value* Performance Plans.

e nothing in the recommended HMIC inspection and review programme should diminish police authorities' accountability for identifying and delivering the required improvements in performance and contents of a *Best Value* Performance Plan.

Chapter 7 – Incentives

The Productivity Panel's wider work has stressed the need for an integrated performance management framework for all organisations starting with bold aspirations, flowing into appropriate short and long term targets, ownership and accountability, performance review and reinforcement/incentives. These aspects need to be integrated: there is no point having the first four elements if the last one is missing. Reinforcement and incentives have been shown to be critical to developing an effective performance culture.

A range of incentives will motivate staff to produce consistently high and improving levels of police performance. The purpose of the incentive framework sketched out in this report is to provide a balanced approach to:

- reward successful achievers. Leading and managing a police authority and a police force to improve efficiency, by increasing outcome performance for the available funding, is a challenging task. It requires innovative leadership in an uncertain and demanding environment. There must be tangible rewards for police authorities and forces that deliver efficiency gains.
- ensure that efficiency measures feed into a cycle of continuous improvement for police forces.
- ensure that action is taken to address consistently poor performance.

An initial analysis of the key elements in an incentives framework is outlined below. This framework requires further development to put effective arrangements in place. This needs to take place alongside the development of the technical aspects of SFA and DEA.

Target Setting

The first element of an incentives framework is to build, within the *Best Value* framework, appropriate targets to raise crime reduction performance by reducing variations across police forces over time.

Phasing out the Current Two Percent Efficiency Plans and Establishing Differential Outcome Performance Targets for Best Value Performance Plans

The development of robust comparative police efficiency measures based on outcomes will mean that the current, separate efficiency planning process sitting alongside *Best Value* is no longer required or desirable. The current efficiency plan process should be seen through to its conclusion, at the end of the 2001/02 year, but not be renewed. Instead, differential targets to improve outcome performance should be set from 1 April 2002, based on the results of the DEA/SFA analysis. *These targets should be expressed in terms of improving crime reduction outcomes for the given level of funding.* The targets should be about better service delivery, not cutting budgets.

An overall target should set for the police service – based, for example, on historical productivity improvements in the overall economy.[6] The distribution of this overall target amongst forces should then adopt the following methodology.

Authorities (and thus forces) in lower bands would be set targets to close the part of the gap between themselves and those in the top band. Each authority in a particular band would be given the same target. Reflecting the risk that some of the measured gap in efficiency might be due in part to modelling or data error, authorities in each band would be set targets to close the measured gap between themselves and the top band by about 50 percent in the next *Best Value* period. Authorities in the top band

6 Annual labour productivity growth in the UK was about 2.2 percent, and overall productivity grew at about 1.7 percent per annum, between 1973 and 1995. See O'Mahony, M; *Britain's Relative Productivity Performance 1950-1996: A Sectoral Analysis* (London: National Institute of Social and Economic Research 1998) Table 1.2 Panel A and Table 1.9 Panel C.

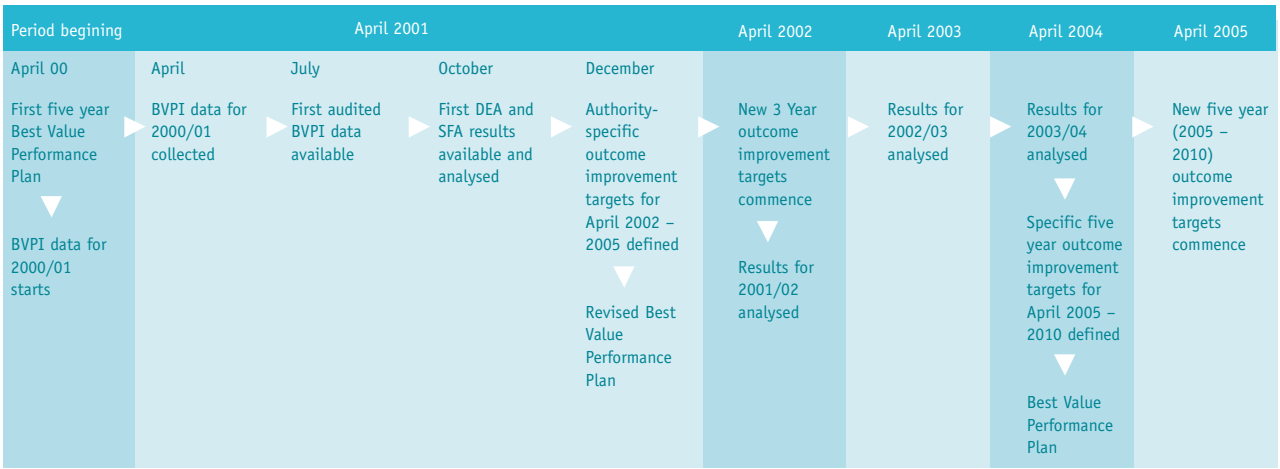
should have a smaller outcome improvement (efficiency) target. The target for top band authorities should be sufficient to make sure that the overall target for the police service is met, bearing in mind that authorities in the lower bands would be expected to generate most of the overall improvement.

Although targets would be set for the 5-year *Best Value* periods, a year-on-year track should be agreed at the outset and progress monitored against this track. To aid performance tracking, DEA can generate "*Malmquist*" indices of efficiency change for each force. For the top forces, these indices can show how much they have improved and thus how much the original efficiency frontier is moving outwards over time. For forces in the lower bands, these indices can identify whether they have become more efficient in absolute terms over time, even if they have not caught up in relative terms on the leading forces (who would also be improving).

The approach outlined above enhances *Best Value* by providing further information to set challenging targets that are tailored to the performance and funding position of each force. Police authorities would be required to identify how they propose to meet this target in their annual *Best Value* Performance Plan.

The suggested sequencing and interaction of the *Best Value* Performance Plans commencing on 1 April 2000, the 5-year *Best Value* cycle and the DEA/SFA analysis is outlined in Box 11. The first results of the SFA and DEA analysis should be available in mid-2001, which would allow new differentiated efficiency targets to apply for the three years from 1 April 2002. The *Best Value* Performance Plans would be modified at this time, to reflect the new differentiated efficiency targets. New 5-year targets would be set for the *Best Value* period from 1 April 2005, based on the most recent SFA and DEA findings then available.

Box 11: Timeframes for *Best Value* Performance Plans, the *Best Value* Cycle and the Recommended DEA/SFA Analysis



Over time, this approach would lead to the convergence in force performance, at which time attention should move to promoting general efficiency improvement (moving the SFA and DEA frontiers outwards) rather than reducing variations in performance.

Published Relative Efficiency Tables

In the interests of transparency, the Home Office should publicise the relative efficiency of forces, using the ranking bands, and the approach to ranking discussed in Chapter 6. In particular, detailed force-by-force rankings should *not* be published to avoid the risk that spurious accuracy would be attributed to the results.

Accountability for Performance

The *Best Value* framework provides that the Home Secretary has the right to intervene where there is clear evidence that a police authority is failing to meet its statutory obligations. This should include failure in relation to the efficiency targets set under *Best Value*.

Supplementing Traditional Approaches

While the foregoing will be necessary elements of a comprehensive re-inforcement regime, a wider menu of positive incentives should also be developed to reward innovation and strong performance. It is important to recognise that strong performance could represent being either one of the most efficient forces, or being one of the biggest improvers.

The following are presented as a possible menu of options that should be considered further as the DEA/SFA approach is piloted and implemented.

Rewards for top performers

Top performers should get a tangible reward – either financial or non-financial – reflecting their achievement.

While there are many approaches that could be taken, one option for financial rewards is to give top performers preferential access to discretionary funding mechanisms, such as the *Crime Fighting Fund*. This type of approach would also ensure that the extra funding available would be going to those forces that have the track record to show that they could do the most with it. An alternative approach is to focus financial rewards at teams of officers. A separate Productivity Panel report has reviewed performance incentives in other parts of the public services, with a focus on team rewards.[7] The principles outlined in that work may have application for any design of police team incentives.

7 Makinson, J; *Incentives for Change: Rewarding Performance in National Government Networks* (London: Public Services Productivity Panel, 2000)

Traditionally, concerns have been expressed that financial rewards may result in safe locations receiving extra funding, at the expense of forces who have the highest crime rates and who would have the greatest need for the funding. However, under the framework outlined above, rewards would go to the forces that do the most with the available funding, or who improve the most. These are not necessarily the safest areas – they could be the areas with the highest absolute crime rates.

Rewards for top performers do not have to be financial. Non-financial rewards to top performing forces could range from a lighter HMIC inspection regime – as suggested in Chapter 6 for forces in the top band – through to recognition awards for top performing forces.

Irrespective of the approach taken, it is critical that any financial rewards are met from specific funds allocated for the purpose, and not by top slicing existing police funding.

Funding for Investment

The detailed analysis and review of force performance may identify the need for investing in new capital assets or policing approaches to make substantial lasting performance improvements. In principle, funding should be targeted towards the best business plans based on the expected improvement in outcome performance resulting from the proposed investment, and the capability to implement them.

Freedom to Manage

The *Best Value* approach, and the approach recommended here, is based on measuring success in improving outcomes and establishing clear accountabilities for performance. These accountabilities can only be exercised if the people holding them have the ability to control the resources required to deliver.

Police operational activity is constrained by a range of regulations that centralise control of resources. These regulations inhibit police authorities' and Chief Constables' freedom to manage. Provided sufficiently robust controls over outcomes are in place, greater control over police inputs could in principle be handed over to police authorities and forces. This clearly raises much wider policy and legislative issues that would need to be carefully considered.

Next Steps

In keeping with its strategic focus, this study has not sought to develop a detailed incentives regime for the police service to supplement the existing and more traditional approaches. However it is important that new and dynamic incentives are developed alongside the technical aspects of this report.

Recommendations

a the current efficiency plans should be abolished after 1 April 2002 and differentiated outcome improvement targets introduced thereafter based on the following principles:

- **i all authorities, and thus forces, being given outcome improvement targets, including those in the top band.**
- **ii authorities and forces in the lower bands being given targets expected to close about half of the estimated gap between the average of that band and the average outcome performance of the top band by the end of the *Best Value* period. Authorities and forces in the top band should be give a smaller target.**
- **iii any overall target for improving police outcomes for the available funding should be distributed amongst authorities and forces according to the above principles upon the receipt of the first efficiency estimates in mid-2001.**

b the first differentiated outcome improvement targets should be set for a three year period from 1 April 2002 – 31 March 2005 based on the SFA and DEA results for 2000/01. From 1 April 2005, outcome improvement targets should be set for the forthcoming five year *Best Value* period, based on the DEA/SFA results from the preceding period.

c the differentiated outcome improvement targets should be consistent with other crime outcome improvement targets set for individual authorities and forces.

d the Home Office should publish relative efficiency tables based on the banding system, so that these tables do not show detailed force-by-force rankings.

e the Home Secretary should use the intervention powers available in the Police or Local Government Act (as appropriate) if police authorities (and their forces) fail substantially to deliver the outcome improvement targets set for them.

f a comprehensive performance management and incentives regime should be developed to complement the above approaches by no later than the end of 2000. The options in developing this incentives regime should include, but not be limited to, the following:

i the development of financial and non-financial performance rewards for top performing authorities and forces (the most efficient, as well as those that make the greatest outcome improvement gains each year).

ii the ability of authorities and forces to present a business case for additional funding to deliver specific performance improvements. The business case should clearly specify the promised improvement in performance, show the links between the investment and better outcome performance as well as the authority's/force's capability to deliver the improvement.

iii a programme to review existing constraints on the freedom and flexibility of forces to manage, as part of a package based on robust and enforced outcome improvement targets.

Chapter 8 – Benefits

Benefits

This study has outlined a framework for better measuring comparative police efficiency, with implementation to be co-ordinated by the Home Office over the next 18 months. The recommended techniques are the best available means of relating inputs to outcomes, while allowing some flexibility in the priority given to each outcome and explicitly capturing the impact of outside influences.

The recommendations in this report mirror the approach taken by industry regulators to the UK's regulated industries – such as water and telecommunications. Although the context is obviously different, a number of the industry regulators use similar or identical techniques to those recommended in this study. They have found that, provided the techniques are linked to a strong incentives regime, a step change in efficiency and performance is possible. Where there is a monopoly supplier – as currently exists in policing – there is no better approach.

One of the key benefits of these approaches is that, as increasingly better efficiency targets are able to be set, other central input controls and regulations can be removed. Police forces would be able to manage their resources better.

In this context, the benefits of the recommended approach include:

- a means of differentiating efficiency targets amongst police authorities and thus forces. Efficiency targets will better reflect the actual position of each force, as well as reflect the gains they have made in recent years.
- scrapping the current across-the-board efficiency targets. This will reduce the burden on police authorities and their forces – by scrapping the current, separate efficiency plans and the bureaucracy around them.
- a systematic, comprehensive measure of relative police efficiency, with clear links to the overarching aims and objectives of policing. Taking this focus means that improving efficiency is about better and more effective police services – improving overall performance – not arbitrarily cutting costs or staff numbers.
- a means for systematically identifying and promulgating good practice as part of an overall strategy to reduce variations in performance. As the techniques analyse the relative strengths and weaknesses of each force in delivering outcomes, it will direct police and HMIC scrutiny to areas of greatest value.
- a means to implement fully the *comparison* requirement of *Best Value*.

These approaches are also adaptable to meet the specific needs of the police, and to minimise the compliance and regulatory costs. For example, the recommended approach:

- uses, as far as possible, existing data – notably audited BVPI data for outcomes and net revenue expenditure for inputs. The current gaps amongst existing data are for the consumption of capital assets, controllable pension costs and (possibly) the size and impact of environmental and socio-economic variables.
- places a premium on trying to get the most reliable and robust results possible, capturing the circumstances of each force, particularly by:
  - allowing some variability in outcome weights to reflect local priorities.
  - explicitly taking into account environmental, social or demographic influences on policing outcomes.

  - recognising that police authorities and forces do not necessarily start from equal starting points – for example, in terms of past capital investment in the area.
- includes diagnostic capabilities to identify the key factors underpinning the results. As discussed in Chapter 6, this allows the targeting of police force and HMIC resources to the areas requiring greatest attention.

Risks

Although off-the-shelf computer software packages are available, these techniques are technically quite complicated to apply and rely on a range of assumptions. This relative complexity is necessary because policing is a complex business and it is important to explicitly or implicitly capture as many of the real world nuances and subtleties as possible, if the results are to be meaningful.

The recommended approach should produce reliable results only if a number of conditions are satisfied: the data is robust, the models contain the relevant variables, and the underlying assumptions of the models hold true.

As discussed in Chapter 4, data reliability and consistency have historically been issues for policing. The use of audited BVPI outcome data provides the most promising route to overcoming the historical problems, although it may not address them all. But a by-product of using these approaches is that there will be much stronger incentives to improve the underlying quality and consistency of data.

As efficiency estimation models cannot contain all the real world variables that influence police outcomes, there is always the risk that the models will be wrongly specified and estimated. The best insurance policy against this risk is an open and iterative process of model building and critique, and sensitivity testing of various specifications and assumptions, involving all key stakeholders. Such an approach is recommended in the next Chapter.

Chapter 9 – Implementation

It is important that the SFA and DEA models and data used to estimate relative police efficiency are built and operated in the most transparent manner possible. This not only minimises the risk of modelling or data error, but will also enhance the credibility of the results amongst stakeholders.

Implementation Principles

The models have yet to be constructed, and the data sets validated. These tasks should be undertaken with the full participation of the Home Office and key stakeholders. Both the models and data sets should be thoroughly tested and peer reviewed at each stage of development. This is a time-consuming task, and will take a concerted effort if the approach is to be ready to be used when the first audited BVPI data comes on stream in mid-2001. It would also be a potentially expensive task if each stakeholder independently sought technical verification of the models and the data sets.

The following implementation principles are recommended to address these risks:

- the Home Office co-ordinates the development of appropriate SFA and DEA models through an open and transparent process involving a Steering Group comprising key stakeholder groups, including the APA, ACPO, HMIC, HM Treasury and the Audit Commission. This forum should consider not only the technical modelling issues, but also the choice of input and outcome measures (as discussed in Chapter 4) and make recommendations on the appropriate weight ranges (Chapter 5).
- the Home Office contracts out the technical work of specifying, building, validating and testing the SFA and DEA models and datasets to experts in the field to make sure that the models are being built by the best expertise available.
- the suggested models and datasets are quality assured and peer reviewed by other independent experts acting for the Steering Group.
- to demonstrate the practical application of the techniques, as well as identify technical issues to be resolved in later stages of model building, the implementation process should involve a pilot in mid-2000 of the SFA and DEA approach in selected urban and rural forces. The choice of forces should be based not only on their willingness to engage in the pilot, but the consistency and reliability of their input and BVPI data at operational unit (BCU) level.
- The estimated cost of implementing this approach is £1 million over the next four years. The new demands on the Home Office, and the costs of the pilot and other stakeholder costs involved in implementation, should be met by additional funding in the 2000 Spending Review, provided that the cost could not be met from lower priority Home Office outputs.

Implementation Timetable

The target for full implementation should be mid-2001, to allow time to calculate comparative efficiency estimates using the first year of audited BVPI data and to revise the *Best Value* Performance Plans for 1 April 2002 – 31 March 2005 in light of the results and the differential outcome improvement targets. The implementation programme itself will need to include the following components:

- completing the analysis of which outcome measures are most appropriate for efficiency measurement, and a mapping of the relationships between these outcomes and socio-economic and environmental factors.

- validating cost and outcome data.
- specifying, testing, validating and peer reviewing the SFA and DEA models.
- completing the pilot.

Box 12 outlines the key milestones for implementation.

Box 12: Key Project Implementation Milestones

To be ready for full implementation by mid-2001, work will need to commence immediately. It will be particularly important that the piloting of prototype models be completed by December 2000 as this will provide a key test for whether the proposed detailed models are working well. In order to achieve this tight deadline, it is important that the following milestones are met by the dates indicated:

a. Review inputs/outcome measures by the end of June 2000.

b. Formulate model(s) specification(s) and sub-contract model development by end June 2000.

c. Complete model development by end September 2000, in preparation for the pilot.

d. Complete the pilot and evaluation by end December 2000.

The programme for 2001 should be constructed once the results of the pilot are known.

Recommendations

a the Home Office, in consultation with the APA, ACPO, HMIC, the Audit Commission, and HM Treasury (the "Steering Group"), should develop the appropriate SFA and DEA models for first use in mid-2001 using audited BVPI 2000/01 data. These models should be developed using an open and transparent process based on the following implementation principles:

i the Home Office co-ordinating the development of the models and datasets but with the technical work of specifying, building, testing and validating the models and datasets being contracted out to independent experts.

ii prior to endorsement and ratification, the draft models and datasets are peer reviewed and quality assured by further independent experts acting solely for the Steering Group as a whole.

iii the implementation programme includes a pilot of the initial SFA and DEA models with selected forces.

iv the costs of implementation (estimated to be £1 million from 2000/01), including the costs of the pilot and the costs of providing peer review for the Steering Group, should be met by additional funding in the 2000 Spending Review provided that the cost cannot be met from lower priority Home Office outputs.

b the Steering Group should also be the forum used by the Home Office to:

i agree an appropriate suite of input and outcome measures (as discussed in Chapter 4 of the report)

ii recommend to the Home Secretary an appropriate range of weights for each outcome measure in the DEA analysis (Chapter 5).

Chapter 10 – Application to the Wider Public Sector

Not only have these methodologies been used in the regulated industries and overseas (see Box 7), they are applicable more widely across the public sector in services having similar institutional features to the police service. The key necessary features are:

- having a large number of discrete and identifiable service providers relative to the number of inputs/outcomes being assessed.
- all the providers contribute to a reasonably consistent set of outcomes.
- appropriate input and outcome data being available.

Public services having these characteristics might include other emergency services, education providers, health providers, regionalised service delivery functions (such as DSS).

Once the results of the pilot are known, HM Treasury, aided by the Home Office, would be in a position to present the approach to public service agencies having these characteristics.

Recommendation

a **once the results of the pilot are known, HM Treasury, with assistance from the Home Office, should initiate a programme to present the results to other possible users in the public sector.**

Chapter 11 – Evaluation

The recommended approach represents a new and innovative way of measuring comparative police efficiency. As it is a new approach, undoubtedly a range of issues will emerge and need to be resolved as implementation proceeds. The recommended implementation provides a process for constantly reviewing and testing the models and the datasets, as well as a vehicle for handling any other issues. But a wider evaluation of the approach also needs to be undertaken.

The recommended implementation process provides two natural points to evaluate the approach. They are:

- at the end of the pilot – in December 2000. This provides the first opportunity for all participants to see SFA and DEA "in action" and to identify ways of making the approach more effective.
- once the techniques have been applied to the first set of audited BVPI data and the results of the HMIC targeted inspections are known. This will be August-September 2001.

The Home Office, in consultation with key stakeholders, should conduct full evaluations of the recommended approach at both junctures.

Furthermore, as with other Productivity Panel projects, implementation progress should be reviewed after a year – the beginning of 2001 – with the results reported to the Ministerial Committee on Public Services and Public Expenditure (PSX).

Recommendations

a the Home Office should evaluate the recommendations in this report after two implementation milestones:

i at the end of the pilot – December 2000.

ii once the techniques have been used with the first audited BVPI data – August-September 2001.

b the Home Office should present a report to PSX by February 2001 on its success in implementing the recommendations of this report. The Home Office report should be reviewed by the Public Services Productivity Panel prior to submission to PSX, and any Panel comments also provided to PSX.

Chapter 12 – People Consulted

The following people were specifically consulted during the course of this study. They were helpful in either shaping or testing the ideas and recommendations in this report, without in any way being responsible for them:

Doug Andrew	Civil Aviation Authority (CAA)
Jim Bradley	Home Office
Professor D S Broomhead	University of Manchester Institute of Science & Technology (UMIST)
Tony Burden	ACPO
Tony Butler	ACPO
Roger Carrington	NSW Treasury
Peter Chard	ACPO
Barry Coker	HMIC
Judith Cooke	Home Office
Catherine Crawford	APA
Professor John Cubbin	Department of Economics, City University, London
Peter Culham	Office of Telecommunications (OFTEL)
Professor Leigh Drake	Department of Economics, Loughborough University
Professor V Farewell	University College London (UCL)
Kate Flannery	Audit Commission
Donald Franklin	HM Treasury
John Gaughan	Audit Commission
David Gilbertson	HMIC
Fionnuala Gill	APA
Professor P A Glendinning	UMIST
Mike Goodliffe	CAA
Tyson Hepple	Home Office
Andrew Hooke	PA Consulting
Sion Jones	National Economic Research Associates (NERA)
Jim Knox	PA Consulting
Richard Kornicki	Home Office
Geoff Lewis	HM Treasury
Liz Lloyd	Number 10 Policy Unit
Tim Lunn	Home Office
Phil Maggs	NERA
Fiona McLean	HMIC
Denis O'Connor	ACPO
Fiona Pethick	Office of Water Services (OFWAT)
Steve Pilkington	ACPO
Annamaria Robb	Warwickshire Constabulary
Paul Rowlandson	HMIC
Barry Shaw	ACPO
Dr Richard Simper	Department of Economics, Loughborough University
John Simpson	OFWAT
Professor Mervyn Stone	Department of Statistics, UCL
Professor Emmanuel Thanassoulis	Aston Business School, University of Aston
Dr Chris Watkins	Royal Holloway College, London
Anne Wheeler	Metropolitan Police Consultancy Group
Brian Williamson	NERA
Ian Williamson	Home Office
Peter Winship	HMIC
Cliff Young	Home Office

Further copies of this document and copies of the supporting volume of technical annexes are available from:

The Public Enquiry Unit
HM Treasury
Parliament Street
London SW1P 3AG

Tel: 020 7270 4558
Fax: 020 7270 4574
E-mail: public.enquiries@hm-treasury.gov.uk

This document can be accessed from the Treasury's Internet site:
http://www.hm-treasury.gov.uk/pspp/index.html