The Serco Institute

What Gets Measured

Contracting for Delivery

Briony Smith

22 Hand Court
London WC1V 6JF
T: + 44 (0)20 7421 6475
F: + 44 (0)20 7421 6471
E: institute@serco.com
W: www.serco.com/institute

Contents

Foreword	i
1. Why measures matter	1
2. When measurement works	9
3. New challenges, fresh approaches	14
4. Developing the right competencies	38
Endnotes	44
Selected bibliography	46
Index of case studies and figures	48

Foreword

The history of contractual performance management is a venerable one. In the English-speaking world, this discipline was largely invented by the political philosopher, Jeremy Bentham, in the late 18th century, and developed by the last of his personal secretaries, Edwin Chadwick. This early literature on performance incentives was part of a wider debate about the respective places for interest and duty in motivating individuals to serve the public good, and between them, Bentham and Chadwick developed many of the principles that lie at the heart of good public service contracting today.

There is perhaps no better illustration than Chadwick's description of the changes made to the performance conditions for the contractors transporting pauper emigrants to North America in the 1830s:

> In the first voyages of transport ships the skippers were paid per head on the numbers embarked, and the explosions of epidemic and the losses of life on some of the long voyages were terrific. By a happy and simple change of the terms of the contract, they were paid only per head on the numbers landed alive, and to the terms of this principle we adhered. On going on board one of the first pauper emigrant ships under this form of contract, I was charmed by the way in which I was received by the skipper; he was so desirous, so pressingly anxious to receive any suggestion for the improvement of the sanitary condition of the ship. Of their own accord the skippers under this form of contract engaged ship surgeons for the care of the passengers, and transferred to these surgeons, medical officers of health, the pressing responsibilities of the contract, by making their enrolments dependent solely on the number of passengers landed alive. The working of the principle was excellent. . . there were no longer any epidemic visitations upon the passengers. The perceptions and observations of points of sanitation by the ship surgeons acting under these conditions were acute, original, and highly instructive. Moreover, under these conditions, without benevolence and with the common everyday principle of interest, we did the work of uncommon, active, and enlightened benevolence. We at all events secured for every lone emigrant, who might after all be lost, at least one sincere mourner.[1]

After such a promising start, it might have been expected that contractual performance management had developed into a well-defined science today, but sadly that is not the case. In spite of the scale of public service contracting that has taken place in the UK over the past two or three decades, and the complexity of the contracts now being let, the published literature on what works well under what circumstances is remarkably small.

This brief report, which is the published output of a much larger study, has been issued in the hope of stimulating a wider debate over contractual performance management, founding a literature and (perhaps) re-launching an old professional discipline.

Gary L. Sturgess,
Executive Director,
The Serco Institute

1. Why measures matter

Political context

This is an era of performance management in the public sector. In the UK, public service reforms since 1997 have focused heavily on performance and delivery. A multitude of measures are now used to assess the performance of government departments and public services – from departmental efficiency measures to league tables and performance ratings for schools, hospitals, police forces and prisons.[2] The US has experienced a similar performance movement, culminating with legislation at federal, state and local levels, whilst the Canadian government introduced performance accounting in 1997.[3] Other governments around the world have followed.

Individual agencies and local administrations also use performance management. Well-known examples include Compstat, the system introduced to drive police reforms in New York City, and Citistat, a derivative model used in Baltimore in the US to address workforce issues. Both systems use computer generated statistics to provide timely performance information, in order to drive improvements in operational performance.

Performance contracting has been part of this broader movement and, over the past few years, contractual performance measurement has also become more intensive, and more complex. Both in the UK and overseas, governments are showing greater interest in using contracts to improve the qualitative performance of public services. A better understanding of how to measure and incentivise performance is essential to extract the maximum value from this.

Delivery potential

At the heart of any public services outsourcing project is a formal agreement – the contract – which specifies the services to be delivered by the contractor and the payments to be made by the customer in return for those services.

Most contracts include some form of performance regime, which sets out the standard of service required and the criteria against which the performance of the contract is to be measured. In most cases, the performance regime is linked to a payment mechanism that determines the final sum paid to the contractor, through a system of performance-based penalties or rewards.

Performance measures and incentives are not simply a safeguard. Their purpose is to drive service delivery by incentivising the contractor to deliver the desired results. The success of the contractual performance regime has a direct impact on the ultimate success of a service. Badly designed measures can undermine delivery, stifle innovation and provoke conflict in the contractual relationship. However, the design of the performance regime is often overlooked in the commissioning and procurement process, as negotiations concentrate in other areas, such as price.

Scope for improvement

Over a period of twelve months, the Serco Institute carried out an in-depth study of contractual performance measurement, involving more than 40 service contracts from across Serco's contract base. The purpose of this study was to try to understand the tools and mechanisms of performance management in complex service contracts: which approaches work best under which circumstances, and how successful performance measurement has been in driving service delivery. Although measures now play a key role in most projects, the field remains largely unstudied, and there is much to learn about this critical element of public service contracting.

The sample contracts for this research were drawn from a range of sectors, including local government, education, health, transport, science, defence and criminal justice. All involved soft service elements, alongside hard FM and asset management components in the majority of cases. Most of the contracts were for large, long-term PPPs (public private partnerships) or PFI projects (under the UK government's Private Finance Initiative), but the sample also included a number of more traditional outsourcing contracts.

The research took the form of interviews with staff from each of the contracts involved, combined with detailed analysis of the contractual performance regimes. This was supplemented by a review of secondary literature.

This report addresses the core findings of the research, concluding that:

- a great deal more work is needed in order to understand the impact of performance measurement in the delivery of public services under contract and the conditions under which it works well; *and that*

- although performance measurement has evolved in recent years in response to emerging market conditions, there remains significant room for further development.

Breadth of analysis

To date, much of the published research on the subject of contractual performance measurement has focused on performance measurement in construction projects.[4] However, some work has recently been done in the UK looking at operational performance and performance measurement in PFI projects, which covers the services element of the contracts, in addition to the building and hard FM elements.

In July 2003, in its document 'PFI: Meeting the Investment Challenge', the UK Treasury expressed its intention to carry out research on the 'impact of performance regimes and payments deductions on private sector responses'. A report was subsequently released, alongside the 2006 UK budget, which examined the performance of construction, hard FM and soft service elements in PFI projects.

At the same time, Partnerships UK released a separate report on the operational performance of PFI projects, on which some, but not all, of the Treasury's conclusions were based. Partnerships UK's report also included a number of questions relating to service performance.[5]

However, further research is needed to gain a thorough understanding of the particular issues involved in performance measurement and incentivisation in contracts for public services. This is particularly important for complex soft services where performance regimes are often extensive, detailed and qualitative, and conclusions about performance are thus inclined to be more subjective.

Other published research on operational performance has focused largely on individual contracts or specific sectors. For example, a number of the value for money reports produced by the National Audit Office (NAO) have examined issues of performance measurement. A report published in June 2003 on 'The Operational Performance of PFI Prisons' dealt with some aspects of performance measurement in PFI prison management contracts, amongst other issues. Reports in 2004 on two high-profile PPP projects, the Criminal Records Bureau and the London Underground PPPs, examined the use of performance incentives in those contracts. The latter report in particular undertook a detailed assessment of the types of measures and incentives used in the PPP contracts, and their success. A 2005 report on the Darent Valley Hospital examined the PFI contract in action, including the performance measurement regime.[6]

Although some interesting insights have emerged from studies such as these, there has been little comparative analysis to understand what approaches work best under different circumstances. Effective and innovative approaches to the commissioning and procurement of complex services are being used in some areas, but this is not being applied across the board. More analysis to understand why some projects excel while comparable projects struggle would enable similar innovations to succeed more often.

Access to evidence

The operational guidance available to commissioning and procurement agents, contractors and independent advisers involved in the design of contractual performance regimes tends to analyse performance measurement away from the context of actual service delivery. The Office of Government Commerce (OGC) has published all-purpose guidance for stakeholders involved in performance management; however, the value of such guidance depends on how effectively it is used.[7] One approach cannot suit all circumstances, and for commissioners to follow the model prescribed for a standard contract without understanding how it applies to the service in question is rarely effective. On the other hand, there is a role for guidance materials and exemplar used as a toolkit by knowledgeable practitioners, able to adapt them to purpose.

Unfortunately, there is a gap in understanding when it comes to performance measurement that existing guidance does not adequately address. Those involved in the procurement and delivery of public services under contract do not always understand the link between effective performance measures and successful service outcomes – or the negative implications of a poorly designed performance regime. Even when the connection between performance measures and delivery is well understood, the skills and experience necessary to design effective performance regimes are not always readily accessible. Experience in this area is not routinely shared, and good practice is rarely captured and transferred between individual projects or sectors, with the result that significant lessons are lost.

One solution to this problem would be for those involved in designing performance regimes for complex services – in both government and the private sector – to have access to an accessible body of practical evidence-based examples, to help them to understand their trade and deliver better results. However, little material of that nature is currently available.

The power of performance measurement

Many policy documents and guidance materials still treat performance measurement as a safeguard, rather than as an incentive mechanism or delivery tool. Even where the importance of the performance regime is recognised, much of the existing operational literature on the subject lacks subtlety, proposing universal solutions, which in practice are not appropriate for all contracts and services.[8]

Approaches to contractual performance measurement vary considerably across different sectors and projects. Levels of knowledge and experience, and the success with which lessons have been captured and communicated, differ from case to case. Some programmes and individual projects are using sophisticated and effective performance regimes, whilst in other areas basic mistakes are still being made.

Performance measures should be developed with clear reference to practical service considerations, and in consultation with operational experts. But, too often they are conceived in a vacuum. In many cases, operational managers interviewed for this project reported that they did not know who had negotiated and drafted the performance regime in their contract, or that they had had no direct contact with those individuals. Operational staff still do not seem to be routinely consulted about the design of the performance regime, despite their knowledge of the service and the insights they could provide. In a number of the cases examined, the performance regime was drafted by external consultants with no further connection with the project. Whilst access to external expertise can be beneficial, performance measures developed in isolation by procurement technicians, without reference to the operational service, are less likely to be effective.

The foundations for success

The success of final delivery has roots as early as the commissioning stage – the period prior to procurement when public officials decide what services they wish to procure, and what procurement and contract models they wish to employ. In spite of policies directing public officials to pursue value for money, price frequently takes precedence over quality when bids are evaluated, without consideration of the long-term impact of such an approach. Artificially low prices can result in a highly contractual approach to performance management, setting a cap on performance by restricting the scope for innovation and flexibility.

Commissioning authorities must also adopt a realistic approach to risk transfer. Experience suggests that the limitations of risk transfer are not always acknowledged. Attempting to transfer risks that suppliers are not well placed to manage can drive up service costs unnecessarily: either performance incentives will be impossible to achieve, leading to the award of disproportionate penalties that result in financial difficulties for the contract, or the supplier will be forced to negotiate wide margins to ensure that additional risk is captured in the contract price.

Contracting for delivery

The responsibilities of the procuring organisation continue after the contract is signed: competition and negotiation should always be conducted with delivery in mind.[9] However, it is not always recognised that once a contract is in operation, the procuring organisation retains a level of responsibility for operational performance, even when significant risk transfer has occurred.

Recent research on operational PFIs identified poor transition as a problem for many projects.[10] The same is true for all contracted services. Successful ongoing performance management requires a constructive attitude to problem resolution and the mindset to make changes to the performance regime and to take a flexible approach where necessary for the good of the service. Naturally, steps like this involve collaboration from all parties involved in the contractual arrangement.

The relationship between project partners also has a significant impact on operational service performance. Contracts built on strong partnership principles and clear, shared delivery objectives are better placed to manage performance problems, instigate improvements and avoid conflict during the lifetime of the service.

And the approach taken by contract monitors can be equally important. Performance measurement should be about delivery, not compliance. Only then is it possible for innovation to flourish.

Bridging the gap

Solutions to many of the challenges involved in the design of effective performance measures and incentives already exist, but could be harnessed much more effectively. The following chapters explain the impact of contractual performance measures and the associated financial and reputational incentives, they analyse some of the current challenges for performance contracting and they open discussions about some of the range of solutions available. The final chapter outlines a number of recommendations about what more can be done to improve understanding of contractual performance management, and to share the lessons.

Case Study 1: Evolution of performance measurement in individual services

Comparison of two different services within the same sector, with contrasting experiences of performance measurement.

Contracts for two distinct Services, A and B, within the same sector, have been let to external suppliers for more than a decade. However, progress and lessons learned in the delivery of Service A have not always been followed for Service B.

Service A:

There have been considerable developments in the way that performance is measured and managed in contracts for Service B:

- contracts which were originally input-based are now measured in terms of outputs;
- performance measures have changed to reflect changing policy priorities for the service;
- more flexibility is now built into contracts, because the importance of being able to respond to change is better understood;
- contractual incentives now include rewards for exceptional performance, as well as penalties for failures;
- contracts contain more user-focused measures, recognising individual needs;
- steps have been taken to reduce the measurement burden by reducing measurement overlaps and contradictions;
- specific problem areas such as outdated, ambiguous or perverse measures have been addressed.

The contracts to deliver Service A have been let under long-term PPP/PFI arrangements. A number of years ago it was recognised that the original performance regimes had become outdated. The performance requirements were re-negotiated in a formal process to bring them up to date. To avoid the need for this costly, time-consuming process to be repeated in the future, the new regimes contain clauses that allow for regular review and revisions without the need for a full contract re-fresh.

Service B:

For Service B, unlike Service A, there has not been a clear shift from input- to output-based measures. Where new performance measures have been introduced to reflect changing policy priorities, the new requirements have been specified as detailed input measures, rather than an output-based approach that would have enabled more responsive service delivery. Input measures are appropriate for some elements of this service – in some areas, innovation is not sought, whereas the government client is keen to retain adequate control of inputs to ensure consistent, predictable service delivery across all contracts for the service. However, for other elements of the service the use of inputs is restrictive.

Performance regimes in recent contracts for Service B are much more detailed overall than in previous contracts for the same service. This contrasts with a drive elsewhere in the sector to reduce the measurement burden. In addition, a new, part-automated measurement system allocates performance penalties automatically, with a mechanism for dispute and disproval after the fact – a system that has caused conflict and a significant burden on resources when used in other sectors.

Whereas best practice for Service A has been consistently built and shared between different service providers, private and public, client side staff changes for Service B have resulted in an inexperienced team with limited knowledge of the operational needs of the service, ill-placed to judge the cost and implications of proposed performance regimes. The use of external consultants to draft the newer Service B performance measures, without adequate input from operational experts, reduced the success of the regime.

2. When measurement works

In 2000, the UK government set a target for all patients seeking GP appointments to be seen by a doctor within 48 hours. An unforeseen consequence of introducing this measure was that some GP surgeries were reluctant to make advance appointments for non-urgent cases, instead asking patients to call on the day. This resulted in telephone queues at opening time and uncertainty for patients keen to fit appointments around work or school hours.

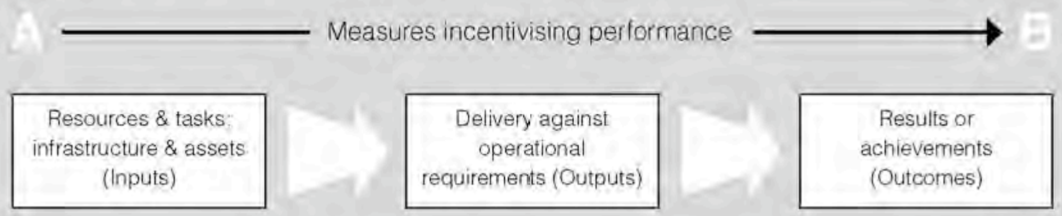
The GP waiting time target was designed to improve services for one group of patients (those seeking urgent appointments). Unwittingly, it created a ‘perverse incentive’ to limit choice for other users who did not fit into that group.[11] Targets and measures like this are not designed to cause problems. They are set in good faith, in the belief that they will help to improve the services in question. However, if they are set without adequate understanding of the impacts, they can create inappropriate incentives.

The same phenomenon also occurs in contracting when contractual performance measures are set without adequate understanding of the effect those incentives will have on the delivery of the service. An extreme example occurred in the contracts let to train operating companies (TOCs) in the rail sector. In early contracts, the penalty for cancelling a train was so high that operators would go to great lengths to avoid it, neglecting other performance measures and disrupting passenger travel by running trains late system-wide.

Understanding the service

Designing an effective performance regime requires an understanding of the service needs and the service model (the business processes linking inputs to outputs to outcomes), in order to set measures that motivate the contractor to deliver the desired results. It is about understanding what is required to get from point ‘A’ to point ‘B’, as illustrated in Figure 1:

Figure 1: Understanding the service model



The more that is understood about the different elements of a service, the easier it is to design effective measures and incentives to drive performance. Understanding of resources, tasks, infrastructure and assets makes it easier to appreciate the capabilities of the service and how it should be managed. Understanding of the operational processes involved in delivering the service makes it easier to identify what is required to make it operate efficiently and to choose measures that will incentivise the desired behaviour. An understanding of the service objectives and what can realistically be set as operating goals is essential in order to create measures with purpose and value.

But understanding the service model is just the beginning. Designing a performance regime that will drive an effective contract involves decisions about structure, incentives, weightings, the level of prescription or flexibility required to make the incentives effective, and how the regime will be managed and monitored. Each decision affects the ability of the contractor to deliver, and the motivation to do so.

Less is more

More measures and increased complexity do not necessarily lead to better service outcomes. Most of the operational expertise involved in service delivery resides with the service providers. Therefore, it is not usually desirable for commissioners to set out detailed operational arrangements in advance. Whilst it is important to understand the delivery objectives of a project, and to agree the allocation of risk and responsibility, it is usually unwise to attempt to specify over-detailed performance criteria in the contract.[12]

In output-based contracts, it can be more effective to focus measurement on a limited number of core deliverables (sometimes called key performance indicators, or 'KPIs') that reflect the key success criteria for the project, allowing the operator to decide the details of delivery as the service evolves. However, when concentrating on a small number of powerful incentives, it is particularly vital to get them right, or to ensure that they can be adjusted or exchanged if required, to avoid driving the wrong behaviour.

Many of the decisions about structure, weightings and other tools used to refine the performance regime can be simplified or avoided if the commissioner specifies outcomes, leaving detailed decisions about resources and processes to the service provider.

Financial reward is not the only incentive

The performance regime is the most important tool for driving service delivery. However, commissioners of services must also understand that providers are motivated by incentives other than the risk of financial failure against contractual targets.

The power of the client-contractor relationship should not be underestimated. Many of the most successful contracts examined in this study were based on strong partnership principles, often underpinned by a specific undertaking – documented in partnering agreements or memoranda of understanding – formalising the commitment of client and contractor to pursue shared objectives for an effective partnership and a mutual focus on delivery.

Performance is also motivated by a range of reputational, cultural and intrinsic incentives: the commercial incentive to develop and maintain a good reputation in order to win future

Figure 2: Selected conditions of effective performance measurement

When measurement works	When measurement does not work
Good understanding of the service model => measures that drive the right behaviour	Poor understanding of the service model => 'perverse measures' that drive the wrong behaviour
The right performance incentives - structure, weightings, measures and management	The wrong incentives - too many measures, wrong emphasis, ambiguity or inappropriate detail
Good partnership built on shared objectives - a climate for success	Poor partnership affected by conflict - undermines performance and stifles innovation
Good contracting - the right bid criteria, good design and effective partnering	Poor contracting - bid environment encourages gaming, contracts undermine performance and trust

business; the professional incentive to excel in the delivery of a service or craft; the personal motivation to serve well those members of the public with whom staff interact day to day.

It is possible to cultivate these incentives through good contracting – the right bid evaluation criteria, good contract design and effective partnering. It is equally possible to undermine them through poor contracting – a bid environment that encourages aggressive gaming behaviour and contracts that undermine performance and damage trust.

Case Study 2: When measures work

Comparison of performance regimes used in four light rail contracts

The Institute's study examined the performance regimes in four different light rail contracts. The service needs were broadly similar for the four contracts. However, each used a different combination of measures to drive and assess performance. The design of the performance regimes had an impact on the contractors' ability to meet the needs of service users, and on the overall success of the service.

Contract C:

The performance regime for Light Rail Contract C was designed to measure 'headways' – the gap between trains. The contractor was required to keep headways as even as possible, in order to provide a regular service for end users.

Headways are the most important measure in light rail services, which are usually timetabled by service intervals, rather than specific arrival times at stations. However, because Contract C *only* measured headways, there was an incentive for the contractor to preserve headways above all else. If a train was running late, the most sensible action to take was to allow following trains to drop back as well, to maintain regular gaps between those trains. Eventually, if a train was to run too late, it could be pulled out of service altogether.

Contract D:

The performance regime for Light Rail Contract D was designed to measure journey lengths. The contractor was required to complete each journey from start station to terminus within an agreed maximum time frame, in order to cover as much track distance as possible.

Because this regime measured *only* journey length, not headways, the contractor had an incentive to bunch the service at the front end: If a train was running late, it would not complete its journey in time to meet requirements. Holding subsequent trains back to maintain headways would prevent those trains from running on time. The most viable action contractually was to run subsequent trains as close as possible behind the late running service to cover as much distance as possible in the time allowed.

Contract E:

The performance regime for Light Rail Contract E was more comprehensive. It measured both journey lengths and service intervals (headways). However, the requirements under this contract were so prescriptive that there was limited flexibility for the contractor to cope with any problems or disruptions that might affect the everyday running of the service. Under certain circumstances, the contractor was placed in a position where any action would result in a penalty. The incentive was therefore to make operating decisions on the basis of what approach would result in the lowest penalty.

Contract F:

The performance regime for Light Rail Contract F again measured both journey lengths and service intervals (headways). However, in this case the performance requirements were more flexible – with lower risk from financial penalties, allowing the contractor to concentrate on service delivery.

The structure of the performance regime allowed lower performance results from periods of high demand, or when specific problems had occurred, to be balanced against higher performance results achieved at other times during each reporting period. Financial penalties and rewards were also smaller in this contract, compared to many other rail sector contracts, which decreased the likelihood of perverse incentives. This enabled the contractor to concentrate on the best outcomes for rail users, even under circumstances when perverse incentives in the contract meant that the most effective service solution for users would result in financial penalties for the operator. In the occasional event that the contract and service needs diverged, there was room to absorb the financial impact of the resulting performance payments.

3. New challenges, fresh approaches

The task of designing a contractual performance regime is becoming more complex. In part, this is because public services are being brought to market where the service models are little understood. Commissioners are shifting from purchasing inputs towards outputs and outcomes, and increasingly they are involving end-users in the measurement of success, creating new challenges for those designing and managing performance regimes. Complex, long-term contracts raise the challenge of ensuring flexibility in performance measurement.

Challenge 1: Designing performance regimes without a clear understanding of the service model

Recent years have seen the emergence of contracts for new government services for which there is no precedent in the public sector. When a service is entirely new, it can be difficult to design an effective performance regime before business assumptions and operating systems have been tried and tested. With inexperienced staff and new equipment and processes to take into account, it is hard to predict the capabilities of the proposed business solution. And with no established point of reference as to how to operate the service, it can be hard to determine the best incentives.

Even if the service itself is not new, when a service is outsourced for the first time, there are new issues to consider during the contracting process. Data about delivery capabilities and the condition of existing assets might not have been recorded prior to contracting out. Indeed the tendering process has often been credited with improving the amount and quality of information available about public services and how they perform.[13] In many cases, the service model will have developed incrementally over many years, and even where it was deliberately designed, knowledge and experience may have been lost over the years, a process that may be exacerbated through significant staff turnover immediately before and during the contracting process.

Lack of information about the inputs, lack of understanding about the structure of the service, or lack of experience in delivering and managing it, make it more difficult to design an effective performance regime. As contracting projects become more ambitious and the type of services delivered under contract more complex, it is becoming increasingly important to capture as much of the available information and expertise as possible at an early stage in the procurement process, to ensure the best possible input into the design of the contract, including the performance regime.

Bedding-in new contracts – some initial observations:

A number of the newer contracts studied during this research made provision for a review of performance measures after the first months or years of operation. In other cases, the contractual arrangements required financial penalties and rewards to be suspended for the initial period of operation, to allow the service model and performance measures time to bed in. No new arrangement is perfect, but such mechanisms create opportunities for measures to be tested in practice, reducing the risk of perverse incentives or other measurement problems occurring once a contract is fully operational.

Service needs also evolve over time, so increasingly contracts undergo periodic reviews throughout the operating life of a project. This helps to keep costs at a minimum if improvement or adjustment is required after the commencement of operational delivery.

In addition to these contractual mechanisms, the client-contractor relationship has an important role to play in ensuring the effective delivery of new services. Both client and contractor must be ready and willing to take a flexible approach in the early stages. The most successful partnerships are those that look beyond the words of the contract to focus on delivering service outcomes.

Case Study 3: Delivering performance for a new service - looking beyond the contract

When the UK Government first outsourced electronic monitoring services for the Home Detention Curfew (HDC) scheme in the mid-1990s, the service was completely new, so there was no public sector precedent for the service model or performance measures.[14]

The HDC scheme allows prisoners to live outside prison custody under curfew rules. Prisoners released under the scheme must sign a licence that commits them to remaining at their agreed accommodation at certain times. They are fitted with an electronic monitoring device that will initiate an alert if they break their curfew conditions. The monitoring equipment is installed and monitored by contractors.

The first electronic monitoring contracts in the UK were let on a five-year basis. However, a number of developments during the early stages of these services made contractual revisions necessary. Growing demand after the initial period of the contracts placed additional pressures on the services. At the same time, growing familiarity with the new service and delivery model meant design improvements were identified for the performance regime, including:

- Rationalisation of measures – in some instances, three measures were found to cover the same elements of the service.
- Improved balance of incentives – at the outset, all aspects of performance were subject to the same penalties for failure; these were redistributed to reflect the greater significance of some for service outcomes.
- Detail and clarity – some requirements were too prescriptive to allow adequate flexibility, and others turned out to be ambiguous.
- Adjustment of perverse incentives - some measures actually undermined the incentive to deliver.

The success of this service owed much in these early years to the attitude of client and contractor. As with most new services, it was vital for both to work together to make the changes required in order to improve delivery. It was necessary to consider what worked and what did not work in practice, and to make adjustments and allowances. In some areas the contractor delivered added value services at no charge to accommodate changes in demand; in other areas the client adopted a lenient approach to performance penalties when outside circumstances or design issues in the contract affected performance unduly.

The performance measures were subsequently refined when the contracts were re-tendered.

Challenge 2: Changing purchasing patterns require more sophisticated measurement systems

There has been a shift in recent years away from contracting for inputs towards contracting for outputs and outcomes (performance). Requirements defined as resources, tasks or processes (inputs), have given way to performance-based targets for the punctuality, reliability or technical efficiency of a service (outputs), or even the performance results, achievements or levels of customer satisfaction attributed to the service (outcomes).

Contracting for outputs or outcomes makes the contract more flexible, because it allows the service provider to decide the best way to deliver the service in order to meet client requirements, and therefore provides more scope for innovation. It also allows the commissioning agency to set higher goals for the project: rather than specifying contractual targets about tasks to complete or processes to follow, the buyer can contract for results to ensure a certain standard of delivery.

There can be no doubt about the overall advantages of a shift away from restrictive input-based contracts towards more flexible performance-based contracts. However, purchasing services at the higher level of outputs and outcomes is also more challenging. It is more difficult to set high level measures in advance without a sophisticated understanding of, and comprehensive information about the service model, in order to set realistic but challenging thresholds. The service provider must also have a suitable amount of control over the inputs in order to assume the risks of delivering the desired performance.

In addition, the shift towards performance contracting usually signifies higher client expectations, so it is likely to involve more complex service requirements and a broader range of specific demands.

Figure 3 : Examples of measures used to purchase services at different levels

Service type	Measures (grouped by level of purchase)		
	Input specification	Output specification	Outcome specification
Training/ Education	Provide staff or premises	Deliver teaching or training 'hours'	Achieve target student attainment levels
Cleaning	Provide staff or products	Complete scheduled tasks	Achieve cleanliness levels defined by industry standards/ customers
Catering	Provide meals when or where specified	Deliver meals of appropriate temperature/ portion size/ nutrition	Achieve quality targets as rated by customers/quality control

Lessons for risk management in purchasing:

There is a direct correspondence between the level at which clients purchase services and the level of performance risk (or flexibility) transferred to the contractor:

Low level purchase (input/process)	=	*low risk transfer to contractor / low delivery flexibility*
High level purchase (output/outcome)	=	*high risk transfer to contractor / high delivery flexibility*

Risk should lie with the party best able to manage it, so that, despite the advantages of performance-based contracting, there is a place for the use of input or process measures, under some circumstances:

- *Certainty:* The early stages of a project can be affected by high levels of uncertainty, usually a result of incomplete information about the service model, infrastructure quality or performance capability. The process of establishing an output-based payment mechanism to transfer unknown risks can become expensive and time consuming. Input measures are sometimes used until a service has become established and it is possible to set effective output or outcome measures, based on acquired operational knowledge. This initial period can be defined in the contract.

- *Assurance:* New clients who lack confidence in the contractor, the contracting process or the service model sometimes seek reassurance by retaining control of certain inputs. These input measures often become obsolete as the contractual relationship matures. However, where clients wish to retain the right to control certain inputs throughout the life of a service, they retain the risk in those areas.

Some risks are not transferrable. The higher the level of purchase, the more likely it is that delivery will be subject to external risks from variables that cannot be managed by either client or contractor. The more complex the service, the greater the impact from outside variables: levels of cleanliness can be controlled by a simple adjustment in approach or materials used, whereas school attainment levels are subject to a range of individual and socio-economic factors that cannot be controlled. The impact of outside variables needs to be taken into account when measures are set and assessed.

Figure 4: Risk variables and purchase patterns

Challenge 3: Changing project and market characteristics demand more flexible pricing and payment structures

The link between pricing and payment structures and the allocation of risk and responsibility between client and contractor makes this a key area of negotiation during the contracting process. Trends such as the development of markets for increasingly complex services, the shift towards performance contracting, and a drive to explore new delivery models that focus on outcomes and the user, create conditions of uncertainty that can cause difficulties for project partners trying to reach financial close.

Projects involving high levels of uncertainty are more likely to require extended negotiations over payment structures and risk allocation, which can contribute to cost increases. Alternatively, costs may rise if suppliers are forced to price this risk into their proposals. In recent years, a number of high profile projects in the defence sector have experienced significant cost escalations and delays – even cancellations. Typically, these projects were long-term, large capital risk transfer projects, involving high levels of complexity and uncertainty, so this sector offers a useful case study for the risks faced by equivalent projects in other markets.

In response to these problems, there has been a gradual movement to explore more flexible pricing structures. In 2004, a UK Public Accounts Committee (PAC) report on major defence projects suggested that *'using a variety of pricing methods should allow [the Ministry of Defence] and industry to achieve a fairer risk and reward balance appropriate to the circumstances of an individual programme'*.[15] The PAC developed a simple chart to illustrate their findings about the circumstances when different pricing models may be appropriate.

Figure 5: Alternative pricing mechanisms*

* Source: Public Accounts Committee, 2004

The PAC's analysis illustrates one response to the problem of uncertainty – the use of different pricing models. However, for many buyers this is not a practical option. Fixed price contracts are often preferred, because of the cost certainty that they offer. For this reason, practitioners have also developed ways of managing uncertainty within fixed price contract structures. One of the key ways that they do this is by including appropriate features in the payment and performance mechanism. The result is that those commissioning public services have available to them a variety of ways of securing the desired flexibility.

Options for adapting pricing and payment structures under conditions of uncertainty:

Punctuating the contract period

Contracts can be designed so as to avoid setting measures too far ahead. Options include making provision in the contract to benchmark or market test the service, or the inclusion of an appropriate clause for a review of performance measures after a specified interval.

Pricing structure

Whilst there are obvious benefits from fixed price arrangements, under the right circumstances, alternative structures offering greater flexibility may provide a better solution. Cost reimbursement models (such as 'alliancing') are the most flexible, but the customer retains high levels of financial risk. Target cost models are now becoming more common, for example, where a mid-range target cost is set with client and supplier sharing the risks of over- or under-spend. Shared financial incentives act as a cost control in this model.

The performance-payment mechanism

Within even the most inflexible pricing structure, uncertainty can be managed using flexible performance-payment structures. Case Study 4 describes a variety of performance-payment models used to accommodate a changing service model (Contract G), changing priorities (Contract H), and hard to manage delivery risks (Contract I).

However, it is almost always inappropriate to link performance criteria that are overly subjective or ambiguous to the pricing or payment structures of a contract. For PFI/PPP models in particular, the strictness of risk transfer means that this rarely offers value for money, as the cost of agreeing and managing payment conditions linked to subjective criteria is likely to outweigh the performance benefits.

Gainsharing

Gain-sharing and loss-sharing arrangements are another way of aligning client-contractor objectives, particularly where there is the opportunity for earning third party revenue.

Case Study 4: Managing uncertainty through the payment mechanism

Contract G:

Contract G is a long-term science and technology contract, recompensed under a target-cost incentive fee model. The agreement covers cost reimbursement to target level (over- and under-runs are shared equally between contractor and client), a relatively small basic fee, and a performance-related incentive fee. A target cost model is used, rather than fixed-price, because the service needs to be flexible enough to respond to scientific advances and strategic requirements as they emerge. However, the prospect of shared cost over- and under-runs provides an incentive for the contracting parties to remain within budget.

To further increase the flexibility of this contract, the performance measures change annually, in line with the business plan for the project. Thus, performance requirements and (where necessary) payment levels, can be adjusted to meet the dynamic needs of the service, but potential expenditure risks are contained.

Contract H:

Contract H is a medium-term strategic partnership to provide IT and communications support services for a government agency. The contract was let on a fixed price basis, a defined percentage of which is at risk from penalties for under-performance. The contract also incorporated service credits (to be offset against penalty points), to reward superior performance.

A particular requirement of this service is the ability to embrace new technology and service capability as it becomes available. Accordingly, the contract does not specify fixed inputs or outputs. Of 120 contractual performance indicators, only 20 are Key Performance Indicators (KPIs) attached to financial incentives. The selection of 20 may be re-negotiated during the course of the contract to reflect changing service objectives. In this way, the pricing model does not require adjustment, yet the performance requirements remain flexible.

Contract I:

Contract I is a medium-term local authority partnership. The pricing model is fixed fee, subject to penalty deductions for under performance (or bonus payments for over performance). The complexity of this contract lies in the fact that the desired performance measures are high-level social outcomes that are difficult to control, making it hard to set appropriate targets.

To manage this, the targets include wide thresholds, with no automatic penalties for below-target performance until delivery falls below a specified threshold. This allows for the effect of outside factors. In addition, the targets and thresholds are renegotiated on an annual basis, taking into account the previous year's performance, as well as any change in significant variables. The uncertainty of the hard-to-control outcomes is mitigated by the flexibility in the performance incentives.

Challenge 4: Adopting a broader approach to incentivisation

Government guidance on contractual performance incentives, particularly that dealing with PPPs and the payment mechanism, has tended to focus on the use of penalties to reduce the risk of poor performance.[16] However, whilst (effective) penalty incentives play an extremely useful role in driving contractual delivery, for some complex services, their impact can be limited. The shift towards purchasing outputs and outcomes, and a better understanding amongst some commissioners of supplier motivation, is setting the scene for a broader approach to incentivisation.

Lack of certainty: High levels of uncertainty make it more difficult to establish effective financial incentives. When the UK National Audit Office (NAO) reviewed the success of the three highly complex PPP deals to finance and manage London's Underground system, two years after the first contract was signed, and in spite of a period of 'shadow running' to review arrangements and make adjustments, they found it *'hard to determine whether [the measures were] easy or difficult to achieve, and whether they [were] sufficiently aggressive'*.[17] The link between financial incentives and risk transfer means that inappropriate incentives are often worse than no incentives at all. If penalties are too onerous, they will undermine delivery and morale, and can create perverse incentives. However, if penalties are too light, then the costs of administering the performance regime can outweigh the delivery benefits.

Lack of control: Services which involve risks that are not capable of being fully managed by the contractor (for example, outside variables affecting social outcomes, such as educational attainment or offender recidivism levels, which cannot be fully controlled within the contract period), create a particular challenge for the development of an effective performance-payment mechanism. However, the nature of such services, which are often linked to high-level social outcomes, means that it is also particularly desirable to find ways to incentivise delivery in these areas.[18]

Rewards for over-performance: Service contractors have long been disappointed by a perceived reluctance amongst commissioners to reward service providers for above target performance.[19] Although rewards or credits are more common in newer contracts, they remain relatively rare. However, as more projects come onto the market where the desired outcomes include hard-to-transfer risks (such as the high-level social outcomes described above), that are hard to specify, measure and control, the potential of reward-based incentives is becoming increasingly apparent.

Unlocking the wider potential of contractual incentives:

The case for reward-based incentives

Penalties drive compliance, whilst rewards can be used to incentivise innovation and improvement. This can be particularly useful where specific targets are hard to define in advance, and explains why rewards tend to be more common in services that deliver complex social outcomes. So, with more such services entering the market, why are rewards still so rare?

In theory, the financial implications of rewards or penalties should be the same. In practice, research suggests that contractors tend to be risk averse, making them more sensitive to the risk of penalties than the prospect of rewards.[20] This might seem to suggest that penalties are more effective than rewards. However, what it actually means is that penalties are more likely to drive compliance. Therefore, if the wrong measures are used, the risk of perverse incentives can increase under a penalty regime.

This has implications for hard to define services. Moreover, because penalties drive compliance with pre-defined criteria, if a customer is interested in incentivising the contractor to act beyond the narrow confines of the performance regime and explore innovative new ways of delivering services and/or to pursue the interests of the end-user, then a reward-based regime may be more effective.

Figure 6: The Financial Implications of Penalties & Rewards

Incentive type	Basic Payment	Available variation	Successful delivery	Performance failure
Reward	90	+10	100	90
Penalty	100	-10	100	90

However, there are political difficulties with rewards-based regimes which do not arise for penalty-based regimes. It is politically acceptable for a flagship service to suffer penalties when it does not perform. For example, it was accepted that HMP Altcourse, described in a 1999 inspection report as 'a "jewel in the crown" of the Prison Service', had accrued fines amounting to £200,000 in the preceding year for contractual non-compliance.[21] By contrast, it is much more challenging for a failing service to earn an equivalent sum in rewards, even if (as Figure 6 illustrates) the end price paid for that service would have been identical under a penalty-based regime.

A recent UK Treasury report questioned whether the performance and payment mechanisms in PPP contracts should include as standard the flexibility to make payments for 'over-performance' – performance above the level specified in the contract – or the delivery of added value services.[22] However, it would be surprising if guidance on this subject recommended discretionary rewards as standard. The difficulty lies in the political acceptability of recognising superior performance when the public expects high standards of service delivery in the first place – indeed citizens' expectations are rising all the time.

A solution to this dilemma is to use restricted mechanisms that reward over-performance only within specified limits. There are several possible models for this.

One approach is to use a system of credits or points that can be 'earned', not as direct rewards, but for offset against penalty points. Credits can usually only be earned for performance against a limited number of designated measures, where their purpose is to act as an incentive for performance in that area (perhaps where a measure relates to hard to manage risks, so the value of setting onerous penalties is limited). However, rewards earned can be balanced against performance in other areas of the contract. Payment is usually capped at the maximum level available for full performance against the contract, with no penalty failures accruing.

Another possible model involves the use of restricted direct (cash) rewards, where bonus payments can be earned, but as before only against designated measures. Under such an arrangement, the bonus payments are not usually offset against penalties earned against individual measures. Instead, total bonus rewards earned are usually offset against total penalty payments accrued. The potential to earn bonus rewards is usually capped at a specified threshold.

Harnessing non-financial incentives

Financial incentives are not the only tools used to manage contractual performance. The most successful contractual arrangements often combine financial incentives with other performance drivers, such as non-contractual partnership structures (memoranda of understanding or statements of shared principles), or reputational drivers (customer satisfaction surveys, peer review mechanisms or comparative assessment tools that rank and measure performance against equivalent establishments or services).

The range of non-financial performance management tools used in public service contracts has broadened in recent years, in line with the shift towards more complex services, models and expectations.

Commissioners need to become more aware of the potential of non-financial incentives. Partnership structures, reputational incentives (customer satisfaction, peer review, league tables) and cultural or organisational factors are all powerful performance drivers, which can be used to good effect in complex projects to manage aspects of performance that are difficult to control under a formal contractual payment mechanism.

Figure 7: Financial and non-financial incentives in public service contracts

Category	Incentive Mechanism	Structural Tools	Performance Drivers
Financial	Penalties	Points/weightings; deduction thresholds	Contractual
	Rewards	Credits/weightings	Contractual
Non-financial	Partnership	MOU; mutual objectives; review meetings; co-location	Contractual; constitutional
	Reputation	Surveys; peer review; benchmarking; balanced scorecard	Contractual; constitutional; cultural
	Professionalism	Organisational culture; training & development; reward & recognition	Constitutional; cultural

Case Study 5: Using reward incentives to help deliver social outcomes

In recent years, performance management in the UK prisons market has evolved away from the use of input measures and process compliance towards output and outcome measures. The change is testimony to an increasingly sophisticated approach, with greater focus on delivery, rather than compliance, and increased transfer of performance risk.

Until recent years, prison contracts offered no mechanism to reward over-performance, however, many contracts now incorporate a service credits system. This allows the contractor to accrue credits for offset against points awarded for performance failures. Although it is not possible to win bonus payments for over-performance, it is possible to balance success against failure.

In these contracts, service credits are used to incentivise the delivery of service outcomes in areas where many of the delivery risks are impossible for prison staff to control, and therefore difficult to transfer under the contract, but where the highest possible level of delivery is desirable because the outcomes involved hold high-level social implications.

Contractors are not usually required to take direct responsibility for recidivism rates, but there is now greater emphasis on measures that can help contribute to rehabilitation. In the area of training and education services, measures which featured in older contracts such as the number of hours spent in education or related activity are supplemented in newer contracts by outcome-based measures covering prisoners' educational and vocational attainment. Delivery of the former is incentivised by the prospect of penalty deductions, whilst achievement of the latter is encouraged by the opportunity to earn credit-rewards. It is worth noting, however, that opportunities to accrue credits are far smaller than the potential to earn penalty points.

Challenge 5: Increased emphasis on finding ways to measure value

Historically, it has been more common to assess contractual performance using quantitative measures. These are easier to set than qualitative measures, and assessment is less subjective. However, as issues of quality and customer responsiveness rise up the agenda in the public service reform debate, and as more frontline services are put out for tender, there is increasing emphasis on measuring the quality of performance.

Some qualitative assessment takes place at the bid evaluation stage. There are a number of tools available to assist with this. Multi Criteria Analysis (MCA) techniques help public authorities to evaluate non-financial criteria (which cannot easily be quantified or measured) when considering the potential impact of policy proposals, an approach that can also be used to assess non-financial criteria in bids. The UK Ministry of Defence uses a specific tool called SIBET (Soft Issues Bid Evaluation Tool) to evaluate soft issues in procurements.

Perhaps because of the difficulty associated with setting and measuring qualitative criteria, there has been no clear move to systematise qualitative assessment in operational contracts. Because it is inherently subjective, qualitative assessment is rarely relied upon as an exclusive measure of contractual performance. However, many recent contracts now include some form of qualitative assessment, ranging from customer satisfaction surveys to peer review mechanisms. There are examples from individual contracts where innovative approaches to qualitative measurement have been used as core elements of the performance regime. Lessons from these contracts could be used more widely.

Figure 8: Mechanisms of qualitative measurement

Forms of qualitative measurement	Examples
Qualitative compliance measures	• Audit against independent standards, e.g. International Organization for Standardization (ISO) standards, British Institute of Cleaning Science (BICS) cleaning standards
End-user assessment	• End-user feedback or 'satisfaction' survey • Direct comments/complaints mechanism • Mystery shoppers & other user-centric tools
Independent assessment	• Inspection by independent bodies, e.g. Office for Standards in Education (OFSTED) & Her Majesty's Chief Inspector of Prisons (HMCIP) • Peer review

Making qualitative measurement work:

For qualitative assessment to form an element of the performance payment structure, quality needs to be easily recognised and defined by objective assessment criteria, so that the performance risks can be managed.

However, many projects use qualitative assessment as a separate measurement tool in addition to the formal performance payment regime (in the form of user surveys, peer review or other tools). Without the challenge of linking to the financial structure of a project, this approach is often far easier to manage because it increases the flexibility of a contract and avoids the cost implications and time delays of negotiating complex measures. In many cases, this approach is both more appropriate and better value.

External assessment

It is important that clients also take external assessment into account when judging the quality of operational performance, and the reputation of their potential suppliers. Many significant performance drivers exist outside the structure of formal contracts. The high profile reports produced by bodies such as the Office for Standards in Education (OFSTED) or Her Majesty's Chief Inspector of Prisons (HMCIP) are powerful quality drivers because they attract such wide attention and have a significant reputational impact.

Challenge 6: Increasing involvement of end-users in service design and measurement

Some of the approaches described above, used to assess the quality of public services, involve gathering feedback from users about the service that they have received. This feeds into a wider UK government drive to increase the involvement of service users and citizens in both the design and assessment of public services.

The concept of user involvement or citizen accountability in public services has always been a strong theme for the current UK Government. The Modernising Government White Paper, published in 1999, promised *"Responsive public services: we will deliver public services to meet the needs of citizens, not the convenience of service providers."*[23] The focus on service users has since increased in intensity.

There has also been a drive by policy makers to personalise service delivery and to increase the flexibility and responsiveness of frontline services. As part of its preparations for the 2007 Comprehensive Spending Review, the UK Treasury announced its intention to use the opportunity of the review to develop the national performance management framework for public services in a way that balances the need for clear strategic goals at the national level, with the requirement for more personalised responsive service provision at the local level.[24]

Services delivered under a contractual performance regime already face the dual challenge of being flexible enough to respond to citizen needs, whilst operating within a performance management framework designed to deliver the availability and quality of service that communities need and expect. Increasingly, however, this dual commitment plays out in more direct ways.

Most customer facing services, including those delivered under contract, now seek user feedback in some form, through surveys and other mechanisms. Some contracts include one or more such mechanisms as part of the formal performance regime, in some cases linked to the payment regime. However, this usually works best for simple services, where performance is clearly visible and measurable – and there is a caveat: user expectations tend to increase as standards rise, making it increasingly difficult and costly for service providers to achieve high satisfaction scores over time.

Increasingly, users and other stakeholders are also involved in consultation during project development, or in advisory or consultation roles throughout operation.

A key challenge of user involvement is that user views are both subjective and changeable. This is compounded by rising expectations and the difficulties of involving users in the development and assessment of complex services, where outcomes are often affected by a wide range of variables and success may be difficult to define. However, the evidence shows that, with some thought and innovation, it is possible to meet these challenges and to involve users in the development and assessment of even the most complex services, where their input may actually be more important.

Meeting the challenge of user involvement:

Clients and contractors need to understand the risks and challenges of user involvement in order to develop appropriate strategies:

- At the design stage, those drafting specifications and performance measures need to be realistic about what is affordable and deliverable within the constraints of each individual project. It is not necessary to cover every element of delivery in the performance regime.

- In operation, the reputational and cost implications of rising user expectations could be addressed through flexible contractual arrangements that offer opportunities to introduce new service innovations over time, or to raise alternative revenue streams. Elements of user experience that the service provider cannot control should not form part of the contractual performance assessment.[25]

- For user assessment to be linked to payment, success or failure must be easily assessable by users, with little or no ambiguity. Separating user assessment from the formal contractual performance measures allows survey criteria to be adapted more easily as user needs evolve, which benefits most services, and more complex services in particular.

- The services that need to be most responsive to individual users are often in areas where users have least control, such as healthcare and custodial services. For different reasons, the choices for recipients of these services are limited, so that comments and complaints mechanisms are often used to give users a voice and to gather valuable information. However, users are more likely to complain than to comment, so other tools, such as surveys, usually supplement this method of assessment.

Case study 6: Involving users and stakeholders in managing service delivery – examples from three local authority contracts

Contract J:

Contract J is a strategic partnership with a local borough council to supply environmental services. The performance regime for Contract J was groundbreaking when it was first rolled out, in that the contractor's payment is largely determined by the customer satisfaction scores the service receives from regular surveys by independent consultants.

The contract uses a performance-related fee in addition to the contract core price: customer satisfaction must not fall below 60 per cent or no service fee is paid. Above 60 percent, the fee is incremental as satisfaction levels rise through agreed bands.

Customer satisfaction levels are assessed quarterly. Each quarter, independent market research consultants carry out a telephone survey amongst local residents. The survey contains questions based on the Audit Commission Best Value Performance Indicators (BVPIs), and other local performance indicators, as agreed.

The user-centric measurement regime for this service works because delivery is relatively simple to assess (main elements of the contract are refuse collection, street cleaning and parks maintenance). Users are well placed to comment on frontline performance, and their feedback helps to drive and shape delivery. At the same time, because performance is assessed against national measures, the Authority can rank the performance of this service against equivalent services in other boroughs, which provides national, as well as local quality control.

Contract K:

Contract K is for the management of a cluster of local authority leisure centres. It uses a service model designed to balance the needs of diverse stakeholders at the local and strategic levels.

There are several players in the partnership. The main leisure centre in the cluster is built on land owned by a local university. The project to build the Centre was funded by charitable grant, including lottery funding. The primary service provider is a not-for-profit Trust, for which the private sector partner acts as managing agent. The structure of the arrangement means that the company is accountable both to the Trust, governed by stakeholders including University staff and funding bodies, and the local Council. Under the terms of the arrangement, the Council retains privileged usage rights to the Centre, and the right to request special services, such as exclusive use for special events.

The contract also incorporates a variety of mechanisms to assess satisfaction levels amongst members of the public who use the centre, including 'mystery shoppers', quality assurance inspection and benchmarking. Users can also feedback directly through a comments and complaints system. Customer satisfaction ratings are displayed in the foyer of the Centre.

Contract L:

Contract L is for the delivery of education services for a local borough council, one of nine high-profile intervention arrangements designed to turn around failing LEAs. Although education services are managed by the contractor, a range of local stakeholders, including teachers, governors, parents and diocesan authorities, are directly involved in the formulation of local education policy, through a Policy Panel formed as part of the strategic partnership for education in the area. Other partners are the schools, the Council and the operating company.

The Policy Panel has advisory and decision-making powers. It advises the Council Executive and its decisions have a direct bearing on the operation of the service. It therefore provides a mechanism for involving a diverse group of stakeholders in delivery of the service. The original decision to outscurce the service also involved wide-ranging consultation, with groups including representatives from community and special interest groups.

Challenge 7: Contract performance is being assessed against a growing range of criteria

Commentators from government, industry and academia have identified the need to rationalise performance measurement to avoid creating a management and cost burden that outweighs the performance benefits. Yet, evidence from this research suggests that recent contracts have often contained a greater range and number of performance measures than older contracts.

In part this trend can be attributed to the introduction of new measures covering areas not dealt with in earlier contracts. Four principal factors have contributed to this development:

- Increased emphasis on accountability to service users has led to more detailed specification of contractual requirements. The tighter contractual approach also means that specific measures are often included to cover any overlap or ambiguity of responsibilities between different stakeholders or providers; or to build specific flexibilities into the contract, in anticipation of service needs. Most contracts are also now subject to a number of different layers of performance assessment, from contractual measures, to departmental standards and measures, to national targets.

- Increasingly, contracts are designed to ensure that services are responsive to particular user needs, taking into account individual human rights, diversity and well-being. In frontline human services more contracts now contain special measures and requirements covering specific user groups. Customer satisfaction measures also feature in more contracts, either as part of the performance regime, or as separate measurement tools.

- A growing number of contracts now include requirements that fall under the broadly defined 'sustainability agenda', referring to elements of service delivery that will contribute to projects' environmental and/or political sustainability. Examples include environmental criteria, such as waste reduction or recycling targets, or socio-political criteria, such as workforce diversity targets.

- The shift to measuring outputs and outcomes does not always see old measures replaced by new. In practice the scope of the contractual measures may simply be widened to incorporate new objectives, or in response to new policy initiatives.

Less is more

It is important that those involved in the design and development of contractual performance regimes remember that less is often more. Incentives to deliver can function effectively without the need for a specific measure covering every aspect of delivery. Professional and reputational incentives are powerful motivators for most service providers, with or without performance measures.[26]

Commissioners should recognise that a simple, delivery-focused performance regime is often more effective than a detailed schedule that is prescriptive and resource-intensive, and thus de-motivating for the operator. Early involvement of operational experts followed by regular review and adjustment of performance measures helps to avoid unnecessary proliferation and duplication and to achieve more effective incentives.

Challenge 8: Models based on financial risk-transfer have generated a more contractual approach to performance management

The use of risk transfer models such as the Private Finance Initiative (PFI) has also contributed to the trend towards increasingly detailed specification of requirements. Since the introduction of the PFI, there has been a gradual evolution across public service markets towards a more contractual and legalistic approach to procurement and contract management, even in non-PFI contracts. This arises out of the necessity in PFI and other risk transfer models to define all requirements and responsibilities tightly, to ensure that risk is clearly allocated over the life of the contract to the various partners in the project.

Risk transfer to the private sector in this type of contracting is designed to increase flexibility in delivery, through the use of output-based contracts, and to maximise the incentive for the service provider to perform and innovate. But those involved in setting and designing the contracts – and setting the performance measures – need to understand how to achieve those aims if the contracts are to achieve their full potential: how to write flexible contracts and the conditions in which innovation can flourish.

Complexities of long-term contracting

The performance regime for a PFI contract cannot easily be changed or updated after the contract has been signed without agreement from the funding banks providing the project capital. Just one small change to the performance measures could alter the risk profile for a contract significantly. Yet, PFI contracts tend to cover a long-term – some as long as 35 or 40 years – and PFI is used in a range of service sectors including human services such as prisons, schools and hospitals where changes in policy priorities and service needs over time are inevitable.

In 2006, the UK Treasury reviewed the value for money test for the inclusion of soft services in PFI projects. A key concern raised by the Treasury was the difficulty in specifying contracts for complex soft services many years in advance, as required under long-term PFI contracts.[27]

In fact, there are numerous ways to ensure that long-term contractual arrangements can adapt to meet long-term needs. As the research institution for a large service provider, the Serco Institute is particularly interested in challenges of long-term and complex contracting. It is evident from this research that these challenges are not insurmountable. There are enough examples of the successful delivery of complex services under long-term contracting arrangements to demonstrate the potential for success. These contracts yield useful lessons for future projects.

Contract structures that respond to changing service needs over the long-term

The simplest flexibility solution available to project commissioners is to use short-term contracts. Many PFI models incorporate shorter contractual arrangements for services within the broad PFI structure, usually around 5-7 years in length. Indeed, many operators of large, long-term contracts routinely sub-contract specific services such as cleaning and catering, to secure value for money and quality through regular competitions.

Another mechanism, discussed earlier in this chapter, is the inclusion of review or break clauses years or months into operation, to allow ongoing re-adjustment of the performance measures. Setting contractual commitments on performance and price only a few years in advance helps to reduce transaction costs and save time during the negotiation of a project. There is usually no need to try to predict service needs thirty years ahead, and often little advantage in doing so.

Case studies featured earlier in this report show that some projects incorporate a flexible performance-payment mechanism, where measures and targets are renewed or even replaced annually, or six-monthly. Models such as these can be useful, for example, in high technology contracts, which must be dynamic enough to respond to technical advances and changing service needs.

The problem faced by the commissioners, procurers, suppliers and users of public services is that the lessons from successful – and unsuccessful – projects, such as those examined throughout this report, are not routinely captured, analysed and used. This is discussed in the following chapter.

Figure 9: Summary of challenges and solutions in contract design

Market development	Challenge	Solutions
1. More new services entering the market without an established model for service delivery	Limited information about infrastructure and operating systems, and lack of experienced staff => performance measures difficult to set and manage	Range of contractual tools available, including: review after initial period; penalty/ reward holiday during early operation; effective joint working, with a focus on delivery not just contractual terms
2. Shift from input specification to performance contracting	More complex level of purchase - requires more information about the service model, and comes with higher expectations	Level of purchase = level of risk transfer, which should influence choice of measures There is a place for input/ process contracting: e.g. if uncertainty is high or assurance low Some risks are not transferrable
3. Movement to explore more flexible pricing and payment structures	More complex projects, shift to outcomes, and more focus on non-objective measurement criteria, such as user satisfaction => need to explore flexible pricing and payment structures	Contracts that do not set measures too far in advance Flexible pricing models, e.g. target cost with shared over/under-runs Other structural incentives, e.g. opportunities to earn third party income
4. Some commissioners leading the way with a broader approach to incentivisation	Higher level of purchase, including high-level social outcomes requires a broader range of contractual incentives, including rewards for over-performance	Political challenges of rewards can be met by using appropriate models, e.g. credit systems or restricted direct payments Potential of non-financial incentives should not be over-looked: reputational and organisational drivers

5. Greater emphasis on qualitative performance measures	More difficult to systematise, more scope for ambiguity/subjectivity	Link to payment usually most practical where quality can be defined objectively Range of options for qualitative assessment outside the payment mechanism: user survey, peer review, standards audit Non-contractual assessment can offer powerful incentives and should be taken seriously
6. Increased focus on users	User involvement easier for simple services; complex services require innovative mechanisms, bringing time, cost and management implications	Expectations need to be managed in context. No need to cover every aspect of service in the performance measures User experiences that cannot be managed should not form part of the performance regime There are benefits to separating user assessment from formal performance measures
7. Wider ranging needs: Climate of increased accountability and raised performance expectations	Broader, more civerse demands => more measures; more complex performance regimes	Contracting for outcomes should reduce the number of measures, and shift focus to delivery Harnessing non-contractual incentives and choosing the right partner helps raise performance without prescriptive measures
8. More contractual approach stemming from use of risk transfer contracting, such as the PFI model	More detailed specification; implications for flexibility and innovation in these long-term contracts	Range of flexibility mechanisms can be included in contracts and measurement systems

4. Developing the right competencies

The limitations of generic guidance

Contractual performance regimes are in many respects tighter, more sophisticated and more detailed today than ever before. However, this is not necessarily a sign that they are more effective at delivering public services.

The increasingly sophisticated mechanisms, models and measures used in many recent contracts reflect the new challenges for public service markets: as services become more complex, new contract models enter the market, and as markets mature, customer and citizen expectations continue to rise. However, as the case studies throughout this report demonstrate, different approaches to performance management work better under different circumstances. More sophisticated models are only more effective if they are right for the service in question. The problem is that lessons about what works, where and when, are not being effectively captured and applied, and access to commissioning and procurement teams with the skills required to deliver good results is not always available.

Materials that seek to capture lessons from performance contracting in the form of generic guidance are of limited practical use. For the lessons to add real value, they must be studied in their particular context, and then adapted to the specific needs and circumstances of each individual project. This assumes that those using the guidance materials possess – or can access – appropriate knowledge of the service in question, and appropriate experience in designing and managing a performance-based contractual regime. This is not always the case.

Building better systems

- Capturing and transferring best practice

  The quality of contractual performance measurement varies enormously, both within and between sectors. In some areas, performance measurement (and contract design more widely) have developed consistently over time. In others, the approach taken in individual procurements appears arbitrary.

There are many successful contracting and performance models for designers and practitioners to draw on. But numerous innovative models and instructive insights are overlooked or disregarded because knowledge is not shared effectively between different contracts and sectors. External advisers retained to advise on the design of performance measures often move on to other assignments without downloading to the client what has been learned from a project.

Recommendations:

1. Government should designate a central agency to collect evidence-based case studies that can be accessed to supplement existing guidance.

2. At the conclusion of their work on major procurements, as one of the conditions of final payment, external advisers should be required to provide a summary of the lessons learned. This should be filed with a central repository, where it can be accessed and used to inform future projects.

3. In studying the effectiveness of contracts for the delivery of public services, audit bodies should report on the utility of different performance regimes.

- Developing procurement and contract management skills

Procurement skills also vary enormously from project to project. To achieve the best outcomes from a project, commissioners and procurers must have the necessary skills to make effective decisions in designing and negotiating the performance regime, to have the confidence to step beyond generic guidance when it does not apply, and to recommend appropriate amendments to the performance regime if it is contributing to poor performance.

Once a contract is operational, the approach taken by contract managers, monitors and auditors may be as important as the performance measures themselves. No performance regime is perfect and no contract is immune to changes in the external environment that are beyond the control of the service provider. Of course, there must be consequences for failure, but it is also important that performance targets are managed constructively and intelligently.

Research by Partnerships UK into the performance of operational PFI projects suggests that there is currently too little continuity from commissioning, through procurement to the operational phase.[28] This can have a detrimental effect on performance, for two principal reasons. Firstly, lack of continuity means that valuable knowledge and understanding of the specification is lost, whilst relationships built up during the procurement and negotiation phases may be undermined. Secondly, if operational staff are involved too late, there is an increased risk that important design issues in the performance regime or the broader contract may be overlooked, leading to perverse incentives being written into the contract, that drive the service provider to deliver the wrong outcomes.

Recommendations:

4. Governments need to ensure that procurement and delivery expertise is accessible across all departments and public bodies. Experts might be seconded from one project to another, or drafted into a department or agency to train in house teams. Central governments need to play an active role in managing this process.

5. The procurement process from commissioning to delivery should be continuous. Commissioners' responsibilities do not end once the contract is signed – there may be an ongoing requirement to change or adjust the contract or performance measures. Those ultimately responsible for commissioning a project need to ensure that project partners, procuring bodies, contract managers and monitors collaborate effectively throughout the life of the project, with a shared focus on delivery.

- Access to operational knowledge and skills

To achieve successful outcomes from performance-based contracting, operational experts should be involved early in the procurement process. This is necessary to ensure that the impact of the proposed contract and performance regime on operational service delivery are clearly understood, and potential problems addressed before the service is in operation.

It is clear from this research that performance regimes are not always drafted with adequate input from operational experts, that they may not be tried, tested or fully understood before operational outset, and that, in practice, adequate processes are not always in place to enable either flexibility in interpretation or change in the formal documentation when the need arises (although there is growing awareness of the importance of this).

Measures should not be conceived in a vacuum by policy professionals or external advisers without involving individuals who understand the workings of the service in question. Whilst the increased use of external consultants, financial advisers and lawyers to inform the contracting process holds obvious benefits, there is a caveat: those lacking operational experience may not fully understand the practical implications of certain proposals. On the other hand, those with operational experience may also be inclined to mandate traditional ways of working, and projects may well benefit from the fresh insights offered by external parties. It is important to strike a balance.

The quality of information available to enable commissioners and suppliers to understand the service model is another important factor in the success of the performance regime. The more that is known about the service model (the linkages between inputs, outputs and outcomes), the quality and capabilities of inputs (particularly physical infrastructure and assets), what it will cost to deliver the service, as well as what it is required to achieve, the easier it is to design performance incentives that will deliver the desired outcomes.

The process of developing performance measures should not be rushed to meet artificial deadlines set for the procurement. Commissioners and suppliers should agree a realistic timetable for the procurement from the outset. For projects that are found to be more complex than anticipated, there is an argument for delaying final agreement on performance measures until the service has been more thoroughly tried and tested (although this will be more difficult where the supplier is assuming financial risk on a long-term asset).

Recommendations:

6. Performance measures should be drafted in consultation with operational experts who understand the impact of the proposed incentives and how the performance regime will work in practice. It is not enough to consult advisors. Procurement experts with no long-term commitment to effective outcomes do not have the same incentives as operational experts, and their knowledge of the development and negotiation process surrounding the performance measures will not carry forward into the delivery stage.

7. Public bodies should take steps where possible to conduct appropriate analysis in advance, and on an ongoing basis.

8. Commissioners must set realistic timescales at the outset of procurements, allowing for the development and negotiation of workable performance regimes. Where it is not possible to acquire this knowledge within an acceptable timescale, then contracting models must be developed that allow for learning and amendment of the performance regime over time.

- Build in flexibility

Successful public services are dynamic and adaptable, which means that performance measures also need to allow room for innovation. There must also be room for adjustment and improvements as priorities change and lessons are learned.

The need for flexibility has increasingly been recognised and incorporated in contracts through individual performance measures. However, more attention needs to be paid in this area. Where older contracts do not offer adequate scope for flexibility, more needs to be done to make room for conditions to be refreshed. When designing the performance regimes for new contracts, commissioners should not feel constrained to set inflexible service measures 30 years in advance, when requirements are bound to change. Unless there are compelling reasons associated with the delivery of integrated solutions, service contracts should not be negotiated for such long periods, and where it is necessary, then provision must be made for robust benchmarking of performance. In extreme cases (such as the PPPs for the London Underground), provision can be made for the renegotiation of terms and conditions using third party arbitrators.

It is also important to understand the impact of high-level policy decisions on front line service delivery. When new government initiatives are announced, the implications for existing services operating under a contractual performance regime are rarely taken into consideration. But if new requirements are incompatible with the existing performance regime, the service provider may become torn between conflicting delivery objectives. This needs to be recognised and managed.

Recommendations:

9. Appropriate attention should be paid to the design of contracts and performance measures for complex projects, including a full exploration of alternative models.

10. Project teams should have access to case studies and lessons learned from past projects, to gain a full understanding of the range of mechanisms available to incorporate flexibility in long-term service contracts.

11. The impact of government policy changes on operational contracts must be recognised and handled appropriately.

- Pressures of the current environment

  There is understandable pressure on procurement agents to deliver efficiency savings. In the UK, for example, at the time of writing, civil servants were under pressure to deliver efficiency gains of more than £20bn in 2007-2008, following a government review conducted in 2004 for the UK government by Sir Peter Gershon.

  In his efficiency review Gershon reminded the government of "the need to ensure that savings are not delivered at the expense of impacting on service delivery".[29] In this same vein, commissioners of public service contracts must be aware that artificially low prices can affect the service provider's capacity to deliver the required service, and prescriptive performance conditions can drive an overly contractual approach. Despite increased emphasis on quality, there is still too much emphasis on price when bids are awarded.

  There are also limits to how much risk can be transferred, particularly in long-term service contracts. Commissioners and buyers need to have realistic expectations about risk transfer, or they may drive up the cost of services unnecessarily, fail to deliver improvements in service delivery and inflict reputation damage on both government agencies and private providers.

  Ultimate ownership of a contracted service remains with the government customer. The customer specifies the service that it requires. It chooses the organisation that it wants to deliver that service. It must ensure that the desired service and standards are actually delivered, which means ensuring that the right performance regime is in place, and the right customer-side staff to manage the contract effectively once the operational phase is underway.

Recommendation:

12. Authorities and departments should be incentivised to focus on quality and long-term value for money, not short term cost-cutting measures. From commissioning through to operation, the focus should be on delivery. Departments and authorities should remember that they are contracting out provision, not affecting a wholesale hand-over. They must put appropriate structures in place to continue to manage performance, including a designated senior responsible owner (SRO) for each project, co-location, where appropriate, and willingness to develop and adapt the service as appropriate throughout its lifetime, in ways that will deliver the best possible outcomes.

Endnotes

[1] Edwin Chadwick, 'Health', *Journal of the Society of Arts*, 21 December 1877, p.70, republished in Edwin Chadwick, *Public Health, Sanitation and Its Reform*, London: Routledge/Thoemmes Press, 1997.

[2] Whilst the current government has dominated the UK 'performance movement', its origins go back further. In 1982 the Financial Management Initiative introduced performance management structures covering the efficiency and productivity of government departments, whilst the 1988 'Next Steps' report famously pointed out that although the UK civil service was dominated by policy staff, the principal role of government is to deliver services. Various performance management policies have since been introduced. See NAO: *Measuring the Performance of Government Departments*, London: The Stationery Office, 2001, p. 57, for a chronology of key UK government performance management initiatives up to 2000.

[3] The 1993 US federal Government Performance and Results Act (GPRA) required agencies to develop strategic plans, annual performance plans and performance reports detailing the effectiveness of programmes. State initiatives range from performance budgeting in Texas to 'Managing for Results' in Maryland. The Canadian initiative was introduced under a Treasury Board policy called 'Accounting for Results'.

[4] The NAO published a report in February 2003 which evaluates the success of PFI construction projects in the delivery of timely good quality assets, and the OGC has produced a procurement guide covering the different stages of project evaluation and performance monitoring in the sector. NAO: PFI: *Construction Performance,* London: The Stationery Office, 2003; OGC: *Improving Performance, Project Evaluation and Benchmarking, London, 2003.*

[5] Partnerships UK: *Operational PFI projects,* 2006; HMT: *PFI: Strengthening Long-Term Partnerships,* 2006

[6] NAO: *The Operational Performance of PFI Prisons,* 2003; NAO: *London Underground – Are the Public Private Partnerships likely to work successfully?*, 2004; NAO: *Criminal Records Bureau PPP*, 2004; NAO: *Darent Valley Hospital: The PFI Contract in Action*, 2005

[7] OGC: *How to Manage Performance*; London: Format Publishing, 2002.

[8] See, for example: HMT: *PFI meeting the investment challenge,* 2003; OGC: *Standardisation of PFI contracts Version 3*; April 2004 and *Version 4*, March 2007.

[9] For further reading on this subject, see: Ertel, Danny, *Getting Past Yes. Negotiating as if implementation mattered*, Harvard Business Review, Article No. R0411C, November 2004.

[10] Partnerships UK: *Operational PFI Projects*, 2006.

[11] The problems with the GP appointments target drew media attention when they were raised with Prime Minister Tony Blair by a BBC Question Time audience in April 2005.

[12] The report by the National Audit Office (NAO) on the Darent Valley Hospital PFI noted that leaving the detailed development of performance measures until after the contract had been let meant that potentially important risk transfer issues were still being refined at that stage. However, the report also notes that lessons have been learned as the performance measurement system for the contract has evolved over time. It is not clear how much evolution would have been possible under a detailed pre-determined regime. See: NAO, *Darent Valley Hospital: The PFI Contract in Action*, London: The Stationery Office, 2005, p.17.

[13] See for example: Walsh, K: *Competitive Tendering for Local Authority Services: Initial Experiences,* Department of the Environment, The Stationery Office, London, 1991.

[14] There had been a brief trial in 1989-1990, but the market had not been progressed due to difficulties with service design and technology.

[15] *Ministry of Defence: Major Projects Report 2003*, House of Commons Public Accounts Committee, 2004, p.13.

[16] See for example HM Treasury, *PFI: Meeting the Investment Challenge*, 2003; pp. 33-34; or Office of Government Commerce, *Standardisation of PFI Contracts Version 3*, April 2004; pp. 87-79.

[17] NAO: *London Underground: Are the Public Private Partnerships s likely to work successfully?*, 2004; pp. 24-5. The PPPs had experienced significant delays during the procurement. The NAO attributed this, at least in part to, lack of certainty affecting elements of the transaction, and to subsequent difficulties agreeing the output specifications.

[18] A brief by the Commission for Racial Equality & the Confederation of British Industry, *Public Procurement and Race Equality* (2003, p. 2), suggests the use of reward mechanisms *'where the service provider is required to undertake some sort of positive action…'*.

[19] See for example: Industry Commission: *Competitive Tendering and Contracting by Public Sector Agencies*, Melbourne: Australian Government Publishing Service, Report No. 24, January 1996, p. 331.

[20] Domberger, S. *The Contracting Organization*, 1998, Oxford University Press, p. 120.

[21] HM Chief Inspector of Prisons: Report on a Full Announced Inspection of HM Prison Altcourse, 1-10 November 1999, p. 5 and p. 108.

[22] HM Treasury: *PFI: Strengthening long-term partnerships,* 2006; p. 75.

[23] The Cabinet Office: *Modernising Government*, London: The Stationery Office, March 1999.

[24] HM Treasury: *Releasing the resources to meet the challenges ahead: value for money in the 2007 Comprehensive Spending Review*, July 2006, p. 37 and p. 42.

[25] The London Underground PPPs use 'mystery shoppers' to assess train and station ambience, but not to assess overcrowding, which is considered an external variable.

[26] One of the longest standing contracts studied for this research, for the delivery of a sensitive service of national importance, contains no financial performance incentives. Staff cite professional pride as their most powerful performance incentive.

[27] HM Treasury: *Value for Money Assessment Guidance,* London: HMSO, November 2006.

[28] Partnerships UK: Operational PFI Projects, 2006, pp. 28-29: In response to a question asking whether there was a formal handover from procurement to contract management, the report concluded 'The hypothesis testing demonstrates statistically that a formal handover process improves the operational performance of the PFI project'.

[29] Gershon, Sir Peter: Releasing Resources to the Frontline: Independent Review of Public Sector Efficiency, 2004; p. 3.

Selected Bibliography

Audit Commission, 2000. Aiming to Improve: The Principles of Performance Measurement, Abingdon: Audit Commission Publications.

Australian National Audit Office, 1999. The Use and Operation of Performance Information in Service Level Agreements, ANAO.

Bozovic, M. (Ed.), 1995. Jeremy Bentham: The Panopticon Writings, Verso.

Byatt, Sir I. and Lyons, Sir M. 2001. Role of External Review In Improving Performance, Public Services Productivity Panel.

Cabinet Office, 2001. Getting it Together: A Guide to Quality Schemes and the Delivery of Public Services, London: MBO Publications.

Chadwick, E. 1877. 'Health', Journal of the Society of Arts, 21 December 1877, p.70, republished in Chadwick E. 1997, Public Health, Sanitation and Its Reform, London: Routledge/Thoemmes Press.

Confederation of British Industry, 2003. Competition: A Catalyst for Change in the Prison Service, London: CBI.

Confederation of British Industry and Commission for Racial Equality, 2003. Public Procurement and Race Equality, London.

Constructing Excellence, 2000. Measured Partnering: The Only Way to know whether you are succeeding, M41 Case Study Project No. 20.

Department of the Environment, Transport and the Regions, 2000. Guide to Quality Schemes and Best Value, Cabinet Office Archives, UK.

Domberger, S. 1998. The Contracting Organization, Oxford University Press.

Domberger, S.; Meadowcroft, S.A. and Thompson, D.J. 1986. Competitive Tendering and Efficiency: The Case of Refuse Collection, Fiscal Studies, 7(4), pp.69-87.

Domberger, S. and Rimmer, S. 1994. Competitive Tendering and Contracting in the Public Sector: A Survey, International Journal of the Economics of Business, Vol. 1, No. 3, pp. 439-453.

Eggers, W.D. 1997. Performance-Based Contracting, How-to Guide No. 17, Reason Public Policy Institute.

Ertel, D. 2004. Getting Past Yes. Negotiating as if implementation mattered, Harvard Business Review, Article No. R0411C, November 2004.

Gershon, Sir Peter, 2004. Releasing Resources to the Frontline: Independent Review of Public Sector Efficiency, London: The Stationery Office.

Gosling, T. (Ed.), 2004. 3 Steps Forward, 2 Steps Back, Reforming PPP Policy, London: IPPR.

Hall, C. and Rimmer, S. 1994. Performance Monitoring and Public Sector Contracting, Australian Journal of Public Administration, 53(4) 453-61.

Harding, R. 1997. Private Prisons and Public Accountability, Oxford University Press.

HM Chief Inspector of Prisons, 1999. Report on a Full Announced Inspection of HM Prison Altcourse, 1-10 November 1999.

HM Treasury, 1999. How to Manage the Delivery of Long Term PFI Contracts, Treasury Taskforce Technical Note 6. London: The Stationery Office.

HM Treasury, 1999. Modernising Government, A White Paper, London: The Stationery Office.

HM Treasury, 2003. PFI: Meeting the Investment Challenge, London: The Stationery Office.

HM Treasury, 2006. PFI: Strengthening Long-Term Partnerships, London: The Stationery Office.

HM Treasury, 2006. Releasing the resources to meet the challenges ahead: value for money in the 2007 Comprehensive Spending Review.

HM Treasury, 2006. Value for Money Assessment Guidance, London: The Stationery Office.

House of Commons Public Accounts Committee, 2004. Ministry of Defence: Major Projects Report 2003, London: The Stationery Office.

Industry Commission, 1996. Competitive Tendering and Contracting by Public Sector Agencies, Melbourne: Australian Government Publishing Service, Report No. 24, January 1996.

Institute for Public Policy Research, 2001. Building Better Partnerships – The Final Report of the Commission on Public Private Partnerships, London: IPPR.

Le Grand, J. and Bartlett, W. 1993. Quasi Markets and Social Policy, London: Macmillan Press Ltd.

Lendrum, T. 2000. The Strategic Partnering Handbook: The Practitioners Guide to Partnerships and Alliances, Chapter 7 – 'Linking Measurement, Performance, Risk/Benefit and Remuneration', McGraw Hill Book Company, Australia.

Jenkins, K, Caines, K. and Jackson, A. 1988. Improving Management in Government: The Next Steps, an Efficiency Unit Report to the Prime Minister, London: The Stationery Office.

Kettl, D. 1995. 'Measuring Performance when there is no Bottom Line', in Jonathan Boston (Ed). The State Under Contract, Wellington: Bridget Williams Books Ltd, 1995, pp160-176.

Mayne, J. 2003. Reporting on Outcomes: Setting Performance Expectations and Telling Performance Stories, Office of the Auditor General of Canada.

National Audit Office, 2004. Criminal Records Bureau PPP, London: The Stationery Office.

National Audit Office, 2005. Darent Valley Hospital: The PFI Contract in Action, London: The Stationery Office.

National Audit Office, 2004. London Underground – Are the Public Private Partnerships likely to work successfully? London: The Stationery Office.

National Audit Office, 2004. London Underground PPPs: Were they good deals? London: The Stationery Office.

National Audit Office, 2001. Measuring the Performance of Government Departments, London: The Stationery Office.

National Audit Office, 2003. PFI: Construction Performance, London: The Stationery Office.

National Audit Office, 2003. The Operational Performance of PFI Prisons, London: The Stationery Office.

Office of Government Commerce, 2002. How to Manage Performance; London: Format Publishing.

Office of Government Commerce, 2003. Improving Performance, Project Evaluation and Benchmarking, London.

Office of Government Commerce, 2004. Standardisation of PFI contracts Version 3, London.

Office of Government Commerce, 2007. Standardisation of PFI contracts Version 4, London.

Partnerships UK, 2006. Operational PFI projects, London.

United States General Accounting Office, 1997. Performance-Based Organisations: Lessons from the British Next Steps Initiative, GAO.

United States General Accounting Office, 1992. Performance Measurement: An Important Tool in Managing for Results, GAO.

Walsh, K. 1991.Competitive Tendering for Local Authority Services: Initial Experiences, A report for the Department of the Environment, London: The Stationery Office.

Young-Chool Choi, 1999. The Dynamics of Public Service Contracting – The British Experience, University of Bristol: The Policy Press.

Index of case studies and figures

Case Studies

1. Evolution of performance measurement in individual services 7
2. When measures work 12
3. Delivering performance for a new service – looking beyond the contract 15
4. Managing uncertainty through the payment mechanism 21
5. Using reward incentives to help deliver social outcomes 26
6. Involving users and stakeholders – examples from three local authority contracts 31

Figures

1. Understanding the service model 10
2. Selected conditions of effective performance measurement 11
3. Examples of measures used to purchase services at different levels 17
4. Risk variables and purchase patterns 19
5. Alternative pricing mechanisms 20
6. The financial implications of penalties and rewards 24
7. Financial and non-financial incentives in public service contracts 26
8. Mechanisms of qualitative measurement 28
9. Summary of challenges and solutions in contract design 36

The Serco Institute was established in 1994 by the international services company Serco Group plc, to undertake practical research into public service contracting and the design and management of public service markets.

Publications

Research Papers

Emma Reddington, 'Good People, Good Systems', December 2004.

Megan Mathias & Emma Reddington, 'Good People, Good Systems: What Public Service Managers Say', January 2006.

Briony Smith, 'Built to Serve: The Benefits of Service-Led PPPs', January 2007.

Briony Smith, 'What Gets Measured - Contracting for Delivery', July 2007.

Policy Studies

Gary L. Sturgess, 'A Fair Field and No Favours: Competitive Neutrality in UK Public Service Markets', Policy Study 1, Serco Institute & CBI, January 2006.

Gary L. Sturgess & Briony Smith, 'Designing Public Service Markets: The Custodial Sector as a Case Study', Policy Study 2, September 2006.

Case Studies

Gerald Cranley & Megan Mathias, 'Education Walsall', Case Study 1, April 2006.

Megan Mathias & Briony Smith, 'HMYOI Ashfield - the Health Service', Case Study 2, April 2007.

Discussion Papers

Gary L. Sturgess, 'Bound for Botany Bay: Contracting for Quality in Public Services', Discussion Paper 1, October 2005

Gary L. Sturgess, 'To Gladden the Heart of Miss Nightingale: Contracting for Complexity', Discussion Paper 2, September 2006.

The Serco Institute

The Serco Institute studied over 40 service contracts from across the public sector, to understand the tools of performance measurement in complex services. The insights that emerged have important implications, not just for performance contracting, but for public sector performance management more broadly.

Public service targets are set with the intention of improving service delivery. But inadequate understanding of the impact of such measures can result in undesirable outcomes. Performance measurement now plays a key role in most areas of public service delivery. A better understanding is essential to achieve maximum value from this.

serco