University of Bristol

CMPO

Market and Public Organisation

Stakeholder Pensions

Issue 3
June 2000

From April 2001 individuals will be able to purchase stakeholder pensions. Stakeholder pensions are a government designed financial product which private sector financial institutions will supply. They are low cost, flexible, defined contribution pension schemes. In this article **Ian Tonks** *examines these new financial products and argues that they are unlikely to meet the government's objective of providing a satisfactory pension for the less well-off.*

Over the last fifteen years pensions policy in the UK has shifted from an emphasis on unfunded pay-as-you-go schemes to funded private sector solutions. In a funded scheme it is the individual's responsibility to build up a fund throughout their working life, and then convert the terminal fund at retirement into an annuity, which will form the pension. There is evidence that low income individuals are excluded from current private sector schemes[1], and stakeholder pensions represent the Labour government's attempt to ensure that individuals whose earnings are less than the national average are provided for. However the flexibility that is built into stakeholders, in contrast to the rigid structure of personal pensions, is likely to mean that low or medium income individuals will have insufficient savings to fund a decent pension on retirement.

Contents

From April 2001 individuals will be able to purchase a new pension product: stakeholder pension plans. These schemes represent the Labour government's attempt to ensure that medium income individuals have sufficient savings to fund a pension on retirement. Stakeholder pensions are rare in being a financial product designed by the government and supplied by private sector financial institutions. They are low cost, flexible, defined contribution pension schemes, into which individuals make contributions throughout their lives, and are able to convert the terminal fund into an annuity at retirement, which will then form the pension.

Background

The introduction of the State Earnings-Related Pension Scheme in the seventies and personal pensions in the eighties represented substantial policy innovations in the pensions area. Continued concern about demographic trends

[1] See footnote 2

has meant that pensions policy remained on the political agenda throughout the 1990s. A number of policy reports have emphasised the need for additional private sector pension provision in some form.[2]

In December 1998 the Government published its Green Paper "Partnership in Pensions". Its main proposals and those of subsequent consultation papers are:

- Basic state pension will increase in line with inflation (and is therefore likely to fall in relation to average earnings).

- Means tested Minimum Income Guarantee will be introduced though Income Support, and all pensioners with full working record will receive a pension of at least MIG when they retire.

- SERPS will be replaced with a flat rate State Second Pension, which guarantees to provide those on incomes of £9,000 per year or below twice the amount that would be given by SERPS. Carers and the disabled will be given credits towards a State Second Pension. All funded pensions - occupational, personal or stakeholder schemes - may opt out of the State Second Pension scheme.

- Stakeholder pensions will be introduced, which are open to all but targeted at the middle income group (between £9,000 - £18,500 per annum), who will be encouraged to opt out of State Second Pension. Stakeholder pensions will be provided under two alternative governance structures: by affinity groups established under trust law with a board of trustees, and by a contract with an authorised scheme manager who will typically be a financial institution. Stakeholder schemes will be run on a defined contributions basis and employers (with 5 or more employees) without an occupational scheme will be required to identify a stakeholder scheme and facilitate access to it for their employees. A designated stakeholder pension must have low costs, flexibility in contributions and transfers, and transparency, comparable with the CAT standards (cost, access and terms) for ISAs. All stakeholder schemes must report regularly on the relative performance of the scheme to each member.

- Tax relief will be given on contributions into stakeholder pensions of up to £3,600 per annum. This tax relief will also be applied to personal pensions and defined contribution occupational schemes.

- The Financial Services Agency will regulate the sale of stakeholder schemes and the provision of advice, and Occupational Pensions Regulatory Authority will regulate the operation of stakeholder schemes.

a designated stakeholder pension must have low costs, flexibility in contributions and transfers, and transparency

Stakeholder Provisions

A stakeholder pension is a tax-efficient defined contribution savings scheme that must satisfy minimum standards specified by the Government. To be designated a stakeholder pension the scheme must have a single charging structure which is limited to no more than one per cent of the value of the fund per annum. This "simple and transparent" charging structure should ensure straightforward comparison between schemes, and is a reaction to the much criticised highly complex charging structures of personal pensions. Charging on the basis of fund value rather than contributions means that, unlike personal pensions, there will be no front-loading, in fact the reverse, since by requiring the charge to be a percentage of fund value when the fund is small, early in the life of the fund, the charges will be small. The standards also specify that the minimum contribution must be no higher than £20, there must be no minimum frequency of contributions, and that transfers into or out of stakeholder schemes should incur no additional charges, so that a stakeholder pension contributor may be able to stop and start contributions at any time and move between schemes ensuring flexibility in contribution patterns.

With any defined contribution scheme the size of the ultimate pension is a function of the value of the pension fund at retirement, which ultimately depends on the size of contributions during working life. According to the 60:40:20:10:5:2 rule, to generate a pension of 60% of salary if contributions are over 40 years and pensioner lives for 20 years requires a contribution rate of 10% assuming return of 5% and growth in earnings of 2%. For example saving 10% of average earnings (£20,000) per annum over 40 years will build up fund of about £340,000 which can be converted into an annuity to generate pension of about £12,000 per annum (£230 per week). But someone on earnings of £12,000 per annum, who makes

[2] The Retirement Income Inquiry (1996) refer to *Assured Pensions*, Office of Fair Trading (1997) identify *Designated Personal Pensions*, Consumers Association (1997) suggest *a Personal Retirement Account*, and these studies culminated in the Government's Green Paper (1998) which labels these schemes as *Stakeholder* Pensions.

relatively infrequent contributions averaging say £30 per quarter (only 1% of their earnings per year), will only generate a fund worth around £20,000, which will generate a pension of just £14 per week.

It is anticipated that there will be perhaps a dozen stakeholder schemes. The investment policy of any single stakeholder scheme will be determined by the trustees or stakeholder manager (financial institution). The trustees or stakeholder manager will be required to set out the scheme's investment principles, and provide a "default investment option" characterising the scheme. Within any scheme there then may be the possibility of allowing individuals some choice of investment policies over and above the default option. As with personal pensions, it is proposed that the investment vehicles of a stakeholder scheme could be a "with-profits endowment scheme" or a "pooled pension investment" (ppi) like unit-linked schemes where contributions are used to buy units whose value is linked to a specific investment fund. Traditionally "with-profits" schemes have been the popular investment vehicle for life assurance policies operated by mutual insurance companies. Under a "with-profits" policy the purchaser is guaranteed some terminal fund, but in additional bonuses may be added during the life of the policy, depending upon the profits generated by the mutual insurance company. The notion here is that the policyholders in a mutual insurance company are the "residual claimants". With the increasing trend to demutualisation of insurance companies it is no longer the case that policy holders are the residual claimants and it is likely that "with profits" schemes will decline in importance, although it is proposed that profits in these schemes are ring-fenced to prevent shareholders from benefiting from the contributions of the members.

The costless transfer between schemes should ensure that funds will flow to those stakeholder schemes offering the best returns, or lowest charges. This will put pressure on stakeholder scheme providers to ensure that their investment policies cannot be improved. In turn this will probably encourage providers to adopt similar investment policies, which together with the low fund management fees, is likely to result in stakeholder schemes investing in "tracker funds"[3].

There are a number of issues related to this type of investment policy when adopted by a large section of the market. One concern is that if providers "herd-in" on a particular investment style, the link between stock prices and fundamentals will be broken, which may have a distortionary effect on the supply of capital to some firms. For example the supply of capital to smaller less liquid stocks not in a recognised stock market index may dry up. An intriguing consequence of this move into passive fund management might be that the returns to active fund management increases, ensuring an equilibrium amount of active fund management in the market. The Treasury has expressed concern that pension funds are too short-termist and invest insufficient funds in venture capital. However the emphasis on tracker funds will only exacerbate this problem.

An additional concern is that a supplier who undertakes a more individual investment strategy which performs poorly, and if this poor performance is transparent, may take more risks with the funds under management to restore the fund's performance. The regulatory framework needs to ensure that this kind of moral hazard does not arise.

Although the Government Green Paper is critical of the high and complex costs, inflexibility and mis-selling of personal pensions, they are not directly affected by the proposals. However after the introduction of stakeholder pensions it is likely that personal pensions will become uncompetitive so that anyone opening a new pension will choose a stakeholder over a personal pension. This raises the issue of whether existing personal pension holders would be better off closing their personal pensions and opening new stakeholder schemes, or whether there would be requirement that personal pension schemes be allowed to transfer into stakeholders.

One of the reasons for the high cost of personal pensions is that they are designed for specific individuals, with the supplier providing information and advice, the cost of which needs to be recouped. Stakeholder pensions are intended to be much simpler, transparent products, so that less advice and information is needed. It is proposed that a prospective purchaser would examine simple decision trees to decide whether to purchase a stakeholder pension. In addition employers would be required to designate a stakeholder scheme. A difficult decision to be faced by current contributors to personal pension schemes will be whether to close the personal scheme and open a new stakeholder, or whether to continue with the personal pension.

after the introduction of stakeholder pensions it is likely that personal pensions will become uncompetitive

tracker fund investments may have a distortionary effect on the supply of capital to firms

[3] As recommended by the Office of Fair Trading Report (1997)

viewing stakeholder pensions as benchmarked personal pensions, their introduction is to be welcomed

Stakeholder flexibility may not impose the financial discipline necessary to generate a decent pension

Stakeholder pensions look like cheaper versions of personal pensions, except that they are constituted within a trust structure. Indeed the government has recognised[4] that the characteristics of a stakeholder scheme operated by a firm, and a Group Personal Pension Plan would effectively be identical. Consequently stakeholder and personal pensions may be substitutes and, although adding an additional pensions vehicle appears to make the choice of instruments more complex, their introduction should ultimately simplify the range of pension products since stakeholder pensions should price personal pensions out of the market.

The extent of competition between personal and stakeholder pensions will depend on the relative tax-efficiency of the alternative schemes. The Green Paper and subsequent consultation documents[5] have suggested that all money purchase schemes will be treated similarly and will obtain the same tax relief, limited to annual contributions of £3,600 irrespective of earnings. Given the similarity of tax treatment it is difficult to envisage circumstances in which personal pensions would be preferred to stakeholder schemes.

On the other hand there are other savings schemes in existence which would also be attractive tax-efficient savings vehicles. In April 1999 the government introduced Individual Savings Accounts (ISAs) to replace PEPs and TESSAs. These are retail savings products which allow for combinations of tax free fund growth and income and do not have the restriction that they have to be taken as a pension. Individuals can put aside a lump sum or save up to £416.66 per month subject to the total contributions in any tax year not exceeding £5,000 (£7,000 in 1999/2000 and 2000/2001). In contrast, contributions into stakeholder pensions will be made out of pre-tax income with a limit up to £3,600 per year, however the ultimate pension will be taxed (subject to a tax-free lump sum). Hence ISAs and Stakeholder pensions offer two alternative tax-efficient savings schemes, one that is taxed on entry with tax free income, and one that has a subsidy on entry, but whose ultimate pension income is taxed. The better scheme for a specific individual will depend on a host of individual circumstances.

In conclusion, viewing stakeholder pensions as benchmarked personal pensions, their introduction is to be welcomed. The high cost and complexity of personal pensions is well known, and by ensuring the tax treatment of all personal pensions is similar, stakeholder pensions will probably serve to make existing personal pension schemes uncompetitive. However stakeholder pensions are not without problems. It is not clear that stakeholders will ensure that the intended target group of medium-income earners will be provided with reasonable pension. Individuals in this income range may simply have insufficient income to spend on pension contributions. The flexibility, which is a hallmark of stakeholder pensions, will not impose the financial discipline necessary for regular contributions, and the pension fund may be just too small at retirement to generate a decent pension. If the government believes that low and medium income individuals do not have adequate pension provision, they should bite the bullet and consider mandatory contributions.

Other concerns are that stakeholders may appear such good value that employers offering defined benefit pension schemes (described by the Green Paper as "one of the great welfare success stories of this century") may transfer occupational schemes into stakeholder schemes. Further, the low fund management costs imposed by stakeholders will probably result in providers investing in tracker funds, with subsequent distortionary effects on the supply of capital to industry.

[4] Consultation Brief No. 5, Paragraph 39.
[5] Consultation Brief No. 6.

Performance Incentives in Government:
Evidence from a US Job Training Programme for the Economically Disadvantaged

Do government bureaucrats respond to performance incentives and can government incentive-designers design performance incentives appropriately? **Gerald Marschke** *summarises research he has conducted on the Job Training Partnership Act in the US. It shows that training agencies boost their*

performance, and their awards, without providing higher quality services. The research emphasises the difficulty government incentive designers have choosing suitable performance measures and tuning and managing the performance measures once in place.

Many policy-makers and analysts see government inefficiency as a management problem and seek remedies among the management practices they observe in the private sector. One management device attracting attention in both the U.S. and U.K. in recent years is the output-driven performance incentive. Its advocates hope that for many types of government bureaux a set of quantifiable objectives can be identified and linked to explicit performance measures backed by cash incentives. Cash-backed performance measures communicate to bureaucrats a clear objective and act to commit the government to reward them for progress toward this objective. By focusing on objectives through performance measures rather than on bureaucratic inputs through monitoring, supervision, and rules, such systems encourage local bureaucrats to use their initiative and their informational advantage over the central bureaucracy to achieve more efficient outcomes. This argument is at the core of the Government Performance and Results Act (GPRA), the most recent of a long line of high-profile (U.S.) government reform campaigns that espouse performance-based accountability systems of various sorts.

The intuitive appeal of this argument, however, contrasts with the absence of empirical evidence in its support. The key questions that need to be addressed are "Do government bureaucrats respond to performance incentives?" and "Can government incentive-designers design performance incentives appropriately?". In the last five years some evidence has accumulated on the experience of the bureaucracy created under the Job Training Partnership Act (JTPA). JTPA is one of the U.S. federal government's longest-running experiments with explicit, financially backed performance incentives. Evidence from JTPA emphasises the difficulty government incentive designers will have choosing suitable performance measures, setting accounting rules for the aggregation of performance data, and tuning and managing the performance measures once in place.

The Job Training Partnership Act

JTPA created what is presently one of the largest U.S. employment and training programmes serving the economically disadvantaged. Its current annual budget is approximately $5 billion and it serves nearly one million people annually. Programme participants receive many different kinds of services, including vocational classroom training (to become, for example, nursing assistants, office managers, computer programmers, and security guards); on-the-job training; basic or remedial education (*e.g.*, in reading and writing skills); and job search assistance, which offers resume writing and interviewing workshops, as well as employment referrals.

Performance (Mis-)Measurement

At the conclusion of each fiscal year, the JTPA incentive system pays training agencies for performance accumulated over the year. The training agency wins a financial award if its performance, measured in terms of the employment outcomes of programme participants, exceeds a numerical standard. Training agencies use their performance awards in the subsequent year to expand their training services. The performance outcomes associated with the measures are computed as the average outcome over all participants terminated (that is, officially removed from the programme's rolls) over the course of the year. For example, until 1992 an important performance measure in JTPA was the employment rate at termination. A training agency's employment rate at termination for fiscal year 1990 was computed as the fraction of enrolees who terminated during fiscal year 1990 who were employed on the date of their termination. At the beginning of the next year, the slate is wiped clean, and performance measurement begins anew.

Can government design performance incentives appropriately?

The performance measures in JTPA are intended by Congress to measure the training agency's success in developing participants' human capital, the stock of knowledge and skills that a participant can "rent out" in the labour market. Because direct measures of human capital value-added are unavailable, the U.S. Department of Labor (DOL), the part of the federal government charged with overseeing JTPA, resorts to using proxies for human capital value-added. At the heart of the JTPA incentive system is a set of performance measures based on the labour market outcomes of enrolees at or shortly after training. These labour market measures are snapshots of the enrolee's employment status, wage, and/or earnings at or shortly after the end of their training spells. The

performance measures quantify employment *levels*, while the objective of JTPA is to produce *changes* in the employability of the poor. In addition, in the first decade of JTPA, the training agency's cost per enrolee was included among the performance measures.

Performance incentive systems encourage the right kind of behavioural responses only if performance measures can successfully distinguish high from low productivity. Recent, careful statistical studies have found little evidence of any relationship between the short-run, employment outcome-based measures of JTPA and the objective of JTPA, measured in terms of the estimated net impact of training on the earnings of JTPA participants. The study estimates earnings impacts from a one-time experimental evaluation of JTPA conducted in the late 1980s. Net earnings impacts are the earnings impacts minus the costs of training through JTPA. For all performance measures used in JTPA there is at most a weak, positive statistical relationship between the measure and net earnings impacts. For some measures, in fact, higher performance is negatively correlated with lower net earnings impacts. That is, JTPA is using the performance measure incorrectly, and as a result, rewarding training centre bureaucrats for performing poorly.

JTPA is using the performance measure incorrectly and rewarding bureaucrats for performing poorly

How bureaucrats manipulate the performance incentives

The difficulty with setting performance incentives is that one is constructing a game between those setting the indicators and the bureaucrats responding to them. A recent study documents how training agency bureaucrats game JTPA's performance incentives, casting additional doubt on the effectiveness of the incentive system. In the first decade of JTPA, a key determinant of a training agency's award was its employment rate at termination. Because unemployed participants who were terminated counted against the training agency, the training agency had an incentive to put off terminating unemployed participants, even after their training concluded. Thus the training agency almost always arrived at the end of the fiscal year with an inventory of idle, unemployed participants on its books. At the end of the year, the training agency would then decide which fiscal year to terminate the unemployed participants. The advantage to the training agency of terminating them in the present year was that it got them out of the way. If they were not terminated in the present year, they would have to be terminated in the subsequent year, reducing the subsequent year's performance. The following termination strategy maximises its award stream across fiscal years. If at the end of the year the training centre finds itself either comfortably above or hopelessly below its standard, it could increase its odds of winning an award in the next fiscal year without jeopardising its chances in the current year by terminating most or all of its inventory. If the training centre found itself above but close to the standard, it could increase its award in the present year by postponing termination until the following year. The study finds that training agencies cooked the books in this way, thereby boosting their performance, and their awards, without providing higher quality services, or providing services more efficiently. In addition, the study finds evidence that this kind of gaming behaviour consumed programme resources.

for performance incentive systems to be effective they must be continuously adapted

Have recent reforms of the performance measures helped?

While the evidence suggests that performance measures based on employment levels imperfectly capture the value-added of job training, we cannot on the basis of the evidence claim that performance incentives do not increase the efficiency of job training under JTPA. We cannot claim this because we cannot compare the behaviour of JTPA bureaucrats in the present incentivized environment to the behaviour of JTPA bureaucrats in an unincentivized environment, because that unincentivized environment has never existed. If we could observe the counterfactual, we might find that the incentives may produce beneficial effects that more than balance out the gaming responses. Nevertheless, recent DOL-initiated reforms in the construction of JTPA's performance measures shed light on the DOL's ability to appropriately design performance incentives.

Anecdotal case studies of the private sector suggest that some firms devote substantial resources to developing, maintaining and improving systems for evaluating and rewarding performance. It appears that for performance incentive systems to be effective they must continuously be adapted to address dysfunctional responses and unforeseen consequences changes in the organisation's external environment, and changes in the organisation's strategy. Performance incentive systems evolve as managers and organisations try to make measurement, evaluation, and incentives more effective at motivating workers.

As in these private sector case studies, the JTPA incentive-designers have monitored the effectiveness of the programme's incentives,

sometimes changing them when it was believed they were not achieving the programme's ends. The two most striking changes occurred in how the DOL measured training agency performance. In the early 1990s the DOL (1) moved away from termination-based measures, toward performance measured three months *after* termination, and (2) eliminated of the cost measure for determining incentive awards. Both changes occurred in response to a number of government studies of the efficacy of JTPA's performance incentives. One set of studies seemed to show that termination-based performance measures, with their emphasis on the enrolee's employment state on the last day of training, were inducing training agencies to emphasise "quick fixes", that is, job placement-oriented services that had no long-term impact on enrolees' skills. In response, the DOL formulated a number of performance measures based on employment outcomes three months after the enrolee's association. The DOL introduced follow-up measures to "[promote] effective service to participants and [assist] them to achieve long-term economic independence.''

In the first years of JTPA (through 1987), the DOL required states to use a cost measure. Cost measures rewarded training agencies that kept the average cost of training an enrolee low. Throughout this period, the kinds of training offered tended to run only five months, on average. The brevity of the kinds of training offered alarmed many policy-makers and analysts. Government studies of the link between cost measures and short, low-intensity services eventually led the DOL to drop the cost measure from the set of performance measures in use.

A recent, careful study of the effects of these performance reforms has shown that they have produced mixed results. The switch from termination-based performance measurement to performance measurement three months after training ends appeared to encourage training agencies to offer the kinds of intensive training that raise the long-term earnings abilities of JTPA enrolees. Nevertheless, the gains from this reform were offset by the elimination of the cost measure. Apparently the cost measure had been discouraging training agencies from offering classroom vocational training, because it is one of the more expensive kinds of training. After the cost measure was removed, training agencies offered more classroom vocational training. Earnings impacts subsequently fell because classroom vocational training produces the smallest earnings impacts of the main kinds of training offered.

Conclusions

JTPA's experience with performance incentives offers at least four lessons for the general application of performance incentives in government. First, government bureaucrats are indeed motivated by financially backed performance incentives, even when the award is not in the form of salary increase or bonus. This offers the hope that if the incentives are properly designed bureaucrats can be motivated to work more efficiently. Second, incentive-designers may have difficulty finding measures of performance that reflect the true productivity of bureaucrats.

Third, a bureaucrat's discretion over key aspects of how her performance is measured may allow her to manipulate her performance outcomes. Performance incentives will only increase efficiency if they systematically punish low performers and reward high performers. In the presence of gaming, however, we can no longer be sure that bureaucrats with high outcomes are truly more productive than bureaucrats with low outcomes. Moreover, the JTPA experience shows that these gaming strategies may also consume resources that would otherwise be used in pursuit of the programme's objectives.

Finally, the design and implementation of the incentive system in a government organisation will be assigned to government officials in the organisation. Because of the nature of government output, any gains to the organisation from improving the design of the performance incentives will be difficult to measure and therefore reward. Government incentive-designers, such as those in the DOL, are likely to have a smaller personal stake in the success of these measures than their private sector counterparts. In so far as incentive design and management is difficult and require expertise and organisation-specific knowledge, these officials will be difficult to monitor and may not take the proper care to get them right, and when the performance incentives are found to be lacking in some way, to fix them. This is the case in JTPA: the DOL has picked performance measures that bear little resemblance to the objectives of the programme. Moreover, its recent reforms have produced no improvement in programme efficiency. Some reforms have reduced programme efficiency.

any gains from improving incentive design will be difficult to measure

Incentives in the Public Sector:
A Survey of the Evidence

In this article **Simon Burgess** *and* **Paul Metcalfe** *summarise the findings from their survey of the sizeable economic literature relating to the provision of incentives within organisations, with a focus on the lessons for the public sector. They also introduce a web-based version of the survey that is searchable and which will be updated to reflect new research.*

Introduction

The introduction of incentive pay in the public sector is currently an issue of major public debate. The case of performance-related pay (PRP) for teachers is perhaps the most controversial at present. The new comprehensive pay strategy for the NHS will replace automatic service-based pay increments with performance–related pay progression and a recent government report has proposed team performance bonuses for the Employment Service, the Benefits Agency, Customs & Excise and the Inland Revenue[1]. Much of the impetus comes from the Government's White Paper on "Modernising Government" (March 1999), which proposed "taking a more creative approach to financial and other incentives for public service staff" (HMG, 1999).

Understanding the economics behind the appropriate choice of pay structure within organisations is essential to be able to evaluate these reforms. Recent advances in economic theory have shed light on many aspects of organisational design and made it possible now to discuss the suitability of incentive schemes in different environments. A body of evidence has been accumulated testing the theory and enabling judgements to be made about the suitability of particular incentive schemes for particular environments. In our survey[2] we selectively review this literature, focussing on the evidence for the public sector; we do not exclude papers relating to the private sector, but we largely avoid the bulk of the literature which has concentrated on the incentives faced by the CEOs of private companies.

workers work harder and produce more output when incentivised to do so

The survey itself examines the impact of PRP on the organisation and issues influencing the choice of pay scheme between different types of organisation. This article summarises the main results and introduces the searchable version of the survey on the CMPO website.

Main Results

There has been a great deal of theoretical work in recent times on the appropriate structure of incentives in organisations. There has been less empirical work, and little of this covers individuals other than CEOs. There is, in particular, a substantial gap in the empirical evidence relating to people whose work is hard to measure, and whose pay is often determined by their superiors. Furthermore, there is very little evidence relating to the provision of, or effects of, incentives in the public sector.

The survey pulls together some strands of this literature with relevance to issues in the public sector. The main findings for both private and public sectors in general are as follows:

- Workers do react in significant ways to incentive schemes. The evidence suggests that, in general, workers do work harder and produce more output when they are incentivised to do so. Workers respond to financial incentives. Note that the productivity gains from the introduction of a PRP scheme come from two sources. First, workers are motivated to greater effort by the higher rewards available; this is the channel usually considered in an evaluation of PRP. Second, the higher pay attracts better quality workers. This has also been shown to have a significant effect on output.

- Workers react in sophisticated ways, manipulating the quality or timing of what they do. These are generally responses that the organisation neither intended nor wanted. One

[1] The report was written by John Makinson for the Public Services Productivity Panel: "Incentives for change: rewarding performance in national government", HM Treasury (2000). This and other aspects of public sector performance pay are discussed in "Pay in the public services 1999/2000" from Incomes Data Services Ltd.

[2] Burgess and Metcalfe (1999) "Incentives in Organisations: A Selective Overview of the Literature with Application to the Public Sector", CMPO Discussion Paper 99/016.

example is if the workers' performance is measured solely in terms of quantity, they may reduce the quality of output in pursuit of higher quantity. Alternatively, if the timing of output can be manipulated, it can be chosen to fall in the most profitable period for the worker. In 1975 Steven Kerr wrote a paper entitled "On the folly of rewarding A while hoping for B"[3] This paper contains many other examples of organisations not achieving what they intended from their (ill-designed) PRP schemes. This suggests that incentive pay schemes need to be carefully designed to reduce the scope for them to produce unwanted results. It also shows that there are circumstances where PRP will be inappropriate.

- Theoretical work has proposed a set of factors that may influence whether any particular organisation would find it optimal to use incentive pay. There is some evidence to support these hypotheses. This evidence comes in two forms. First, detailed studies on particular firms can show that they are able to choose their method of payment in line with the sophisticated ways suggested by theory. For example, in the trade-off between raising the volume of output and maintaining acceptable quality, one study showed that a business facing a variety of different operating environments selected more or less high-powered incentive schemes for each environment as predicted by the theory. Second, a broad-based cross-sectional study has shown that the pattern of existence of different types of schemes is roughly in accord with theory relating to measurement and multi-tasking. Where a worker has many tasks to perform or where output is difficult to measure, objectively assessed performance related pay is observed less frequently and subjectively assessed bonus payments are observed more frequently.

Some findings that relate specifically to public sector issues:

- We know that some public sector workers are motivated by more than just their own income. Case workers in a job training scheme in the United States systematically took on the hardest-to-place workers even though their narrow financial interest was better served by selecting more employable workers. Of course, we do not know that private sector workers are not so motivated.

- In the UK incentive schemes are less common in the public than in the private sector. This may be because the nature of the work makes it optimal not to have incentive pay. Or it may be because the public sector has until recently successfully resisted such schemes and efficiency would be raised by having incentive pay. Comparing across sectors and job types (and while noting that differences in the pattern of existence of incentive schemes between the public and private sectors are not easy to interpret) the evidence indicates that there are inefficiently few schemes in the public sector.

the evidence indicates that there are inefficiently few schemes in the public sector

Unanswered questions

While research in this area has started to produce a basis of evidence for judging policy, there remain a number of unanswered questions. We need research on these issues plus in general simply more weight of evidence. In particular, we need more evidence on non-CEOs and non-manual jobs. Some of the unanswered questions are as follows: we have little evidence relating to incentives for those whose pay is determined subjectively by their superiors; we do not understand the inter-connections between incentive schemes and the job security typically enjoyed by public sector workers; we do not understand the trade-off between incentive schemes and the need for probity by public servants; we do not have a fully worked out view of the link between incentive schemes and the lack of competition for the output of many public sector workers; finally, we do not have a very good understanding, either theoretically or empirically, of the role of incentives for teams.

we do not have a very good understanding of the role of incentives for teams

Web version of the paper

The full literature survey (see footnote 2 for details) is available in hard copy as CMPO discussion paper 99/016. We have also put a searchable version on the web at http://www.bris.ac.uk/cmpo/incentives/incentindex.htm. This document contains the full text of the review, including 'go to' section headings and 'clickable' papers. Selecting a paper takes you to the Appendix, which gives details of around 100 papers in this field. The details include the research question addressed, the data used and the main findings. It also includes the full reference. The papers in the Appendix can also be accessed directly, organised by author name or topic.

We intend this to be a dynamic review, with new papers added as we find them or are sent them. The main text will be updated annually to reflect new research in the field. Our aim is to maintain this survey as an up-to-date resource, providing a concise summary of the evidence on the use of incentives in organisations, particularly with reference to the public sector.

[3] *Kerr, S. (1975) "On the Folly of Rewarding A, While Hoping for B.", Academy of Management Journal 18, pp.769-83*

Measuring NHS R&D:
Avoiding Own Goals

The Department of Health is seeking routine ways to measure the outputs from the £400m a year it devotes to funding Research and Development (R&D). Publications are often used as an indicator of research outputs. On their own they do not, however, capture those aspects of R&D that are most important and they might introduce incentives with undesirable consequences. In this article ***Bronwyn Croxson*** *illustrates the problems likely to arise if publications are used as the sole measure of R&D output by describing the likely effects of monitoring (and rewarding) individual football players according to the number of goals they score.*

Introduction

Monitoring and measuring research and development (R&D) is high on the Department of Health's agenda. Performance measures for R&D must be chosen wisely to minimise costs to NHS organisations but also to capture that which is valuable. Publications are often used as an indicator of research outputs - on their own they do not, however, capture those aspects of R&D that are most important and they might introduce incentives with undesirable consequences. The problems likely to arise if publications were used as the sole measure of R&D output can be illustrated by examining the likely effects of monitoring (and rewarding) individual football players according to the number of goals they score. It is likely that they would stop co-operating with other members of the team, possibly reducing the total number of goals scored, and they would forget defence, possibly preventing the team from winning. This is an important insight for the NHS as it tries to find a way of measuring the value of R&D.

Measuring the value of health and biomedical R&D

There is growing interest in estimating the value of R&D in all sectors, including in health and biomedical research. This interest results from two important trends. First, it results from recognition by policy makers that R&D can produce a range of socially and economically valuable outcomes[1]. Research is recognised as an investment, with returns that should be maximised. These returns include not only financial gains, but also returns in the form of contributions to knowledge, enhanced research capacity, improved decision-making and, in the context of health-related R&D, health gains[2].

investment in R&D produces valuable outcomes

Secondly, the growing interest in valuing R&D reflects a desire to make public agencies more directly accountable for how they use tax funds[3]. This is manifest in an increasing focus on measuring and monitoring the "value" of government-funded R&D. This trend is evident in the Department of Health, which each year spends about £400m on R&D. The department is developing a performance management structure, designed to hold accountable the individuals and organisations who spend these funds. The objectives of the performance management structure are:

- to ensure funding is optimally allocated between R&D and patient services;
- to ensure optimal allocation of funds between different types of R&D;
- to get information to assist in the governance of R&D.

Most of the Department of Health's research and development funds go to NHS hospital Trusts, to support the costs of conducting R&D funded by research charities (such as the Wellcome Trust) and the research councils

[1] Kay J, Money from Knowledge. Science and Public Affairs. 1999, April, 12-13

[2] Buxton M and Hanney S, How can payback from health services research be assessed? Journal of Health Services Research and Policy. 1996.1, 35-43.
Salter AJ and Martin BR, The Economic Benefits of Publicly Funded Basic Research: A Critical Review. SPRU Electronic Working Paper, www.sussex.ac.uk/spru, 1999, 34.

[3] NHS Executive, The new NHS Modern and Dependable: A National Framework for Assessing Performance.1999, catalogue number 97FP0148

(mainly the Medical Research Council). These bodies are also increasingly monitoring and measuring the activities of the individuals and organisations they fund.

Performance indicators

Performance indicators often form an essential part of performance management structures. Performance indicators are chosen as direct measures of, or proxies for, valued outcomes. For example, publications are often used as a performance indicator for R&D, since they can be a proxy for contribution to knowledge and, sometimes, for contribution to decision-making.

In general, a performance indicator can be assessed using the following criteria:

- Does it measure what is valuable?
- Is it feasible to check that the results are accurate?
- Does it have acceptable costs?
- What are the incentives it sets up and is behaviour likely to respond to these incentives?

Designing performance indicators that will meet these criteria when measuring and monitoring health-related R&D presents particular challenges. Health R&D comprises a wide spectrum of different types of activity, including laboratory-based science, clinical trials, clinical observation, health services research, statistics and so on. These have a wide variety of valued outcomes that are often intangible; their direct outputs are not necessarily comparable and may not be able to be precisely attributed to particular funders; organisations (such as NHS Trusts) are often held accountable for the activities of individuals they employ but over whom they may have little direct influence; and collaboration across organisational or disciplinary boundaries is often an essential part of the R&D process. The pros and cons of using publications as a performance indicator, given these challenges, is illustrated in the next section by examining the pros and cons of using goals scored to measure footballers' performance.

A game of two halves …

The type of problem described above is faced in many contexts, whenever managers have to try to find a low-cost measure giving accurate information about employee performance which, if used to reward employees, also creates the right incentives. It is a problem faced by football managers, who seek a way of measuring the value of individual players, so that they know who to buy and how to stimulate good performance. The problem faced by football managers resembles that faced by research managers, in the sense that both have an obvious, easily measurable indicator: goals in the case of football managers and publications in the case of research managers. Football managers do not usually use goals as a measure of individual value, for reasons which give insight into factors that should be taken into account when designing R&D indicators.

individual footballers are not valued according to goals scored

Consider, for example, two Manchester United players: Andy Cole and Jaap Stam. By the beginning of May 2000, Cole had scored a total of 103 goals for United and Stam had scored one. Does this make Cole a more valuable player than Stam?

This is obviously not the case. Goals fall foul of the first of the criteria listed above: they do not measure all valuable activity. Jaap Stam is a defender. He very occasionally scores goals, but usually plays a different role to Andy Cole, who is a striker. Players occupying different positions score goals at different rates and under differing circumstances, just as laboratory-based scientists, clinicians, and social scientists publish at different rates in different types of journal with different audiences.

Goals have the advantage of being verifiable and, because they can be observed, it is not costly to collect information about who has scored. (These are the second and third of the criteria listed above.) Alternative measures are, by comparison, more costly to use and are not based on objective information. They often rely on subjective judgement. Using subjective judgement is, however, the only way to get information about some important dimensions of value. Take, for example, the performance of goalkeepers. We want to measure the number of goals he or she *prevents*, but it is impossible to agree an objective measure of what constitutes a 'saveable' goal.

outcomes from R&D are hard to measure using conventional methods

Similar issues arise in the context of measuring R&D. Like goals, information about publications meets the criteria of being verifiable and relatively low cost. This is very attractive in the context of NHS R&D, where hard-pressed managers and policy makers have to try and collect information from hard-pressed staff, in an environment governed by pressure to reduce management costs. Publications provide an objective, accessible performance indicator, which can often be collected without needing compliance from busy, sometimes alienated, staff for whom R&D is not their main business. Alternative methods, such as expert review, user surveys and case studies require substantive

resources and/or greater input from researchers. They do, however, also provide information on other dimensions of the outcomes of R&D, particularly whether R&D has contributed to research capacity and to the efficient, equitable delivery of health gains.

The final criteria against which performance measures have to be assessed is their impact on incentives. Are they consistent with desirable behaviour? Goals are desirable for football clubs, but ultimately the club does not want to maximise goals, it wants to *win*. Using goals scored as a measure of individual players' value may directly undermine that objective, as it creates incentives that might undermine the team's ability to win. We do not want goalkeepers forward, trying to score goals: we want them in defence, stopping the other team's goals and thereby helping the team to win. The NHS is similarly not interested in publications *per se*, but rather in R&D that promotes the efficacy, efficiency and equity of health care and which is disseminated appropriately. Under some circumstances measures based on publications can be carefully used so that they are consistent with this. Recent work at the Wellcome Policy Unit, for example, suggests that we could use publications cited in clinical guidelines to gain insight into the types of R&D that are affecting clinical practice[4]. Crude measures based on counts of the number of publications or citations are, however, unlikely to give correct incentives.

Rewarding individuals for publications may create perverse incentives

It is also important to consider the way a measure affects incentives for collaboration or co-operation. Rewarding individual football players for goals may create a disincentive for co-operation and may even reduce the total number of goals scored. Goals usually rely on the efforts of the whole team, rather than on the just the boot of the last player to touch the ball. So too, the NHS should be wary of rewarding individuals in a way that introduces disincentives to collaboration - either between individuals, between disciplines or between organisations.

Conclusion

The NHS needs a measure of R&D that measures what is actually valuable, is accurate, has acceptable costs, and does not set up perverse incentives. A vital part of getting the incentives right is ensuring that what is most valuable is captured by the measure. Publications can certainly form a useful part of a performance measurement system, but they must be used with care if they are to measure what is valuable to the NHS, including the impact of R&D on health gain. Any measure based on publications must recognise that clinically relevant research and health services research may be published at a different rate and in different types of journals than other types of science. Moreover, these types of R&D may make their most valuable contribution in ways other than formal publications. Publications should be used in conjunction with other indicators, such as expert review, user surveys and case studies. Just as football managers rely on a variety of different sources of information, including scouts, so too the NHS R&D system should use expert review to gain additional information about vitally important, difficult to measure, dimensions of R&D activity.

An edited version of this article appeared in the Health Services Journal, 6 Jan 2000.

[4] Grant J, Cottrell R, Cluzea F and Fawcett G, Evaluating the 'payback' on biomedical research by characterising papers cited on clinical guidelines. British Medical Journal, forthcoming.

The Centre for Market and Public Organisation was founded in September 1998 with the objective of furthering our understanding of an appropriate design for the reform of a wide range of activities on the 'boundaries of the state'. The Centre aims to develop research, further debate and inform policy relating to the public sector and the recently privatised entities. Further information can be obtained from:

The Centre for Market and Public Organisation, Department of Economics, University of Bristol
Mary Paley Building, 12 Priory Road, Bristol BS8 1TN
Tel: 0117 954 6943 Fax: 0117 954 6997 Email: cmpo-office@bristol.ac.uk ***Website:*** www.bris.ac.uk/cmpo/