When economics met psychology: rethinking incentives

The generation of market failures through ineffective incentivisation schemes is a controversial area of debate. What are the assumptions of traditional behavioural economics, and how might the theory fail to explain observed phenomena in labour markets? How has economics subsequently been playing catch-up with the psychology literature in exploring what motivates individuals in practice? And how might all of this affect incentivisation in public services and not-for-profit network utilities?

Competition between ‘profit-maximising’ firms in the goods market is generally regarded as providing benefits to consumers and society in the form of lower prices and better products. Rivalry between ‘utility-maximising’ individuals in the labour market, geographical mobility, and competition between firms in attracting staff, are all also seen as desirable, reducing periods of unemployment when shocks occur, and lowering firm input costs.

However, it is also recognised that excessive rivalry, either in the goods or labour market, can have undesirable consequences where it generates externalities and induces a market failure. If two power generators compete to generate electricity, and one has lower costs simply because it has not invested in pollution-reduction measures, this can result in negative externalities (costs to society), with rivalry between the firms generating excessive pollution in equilibrium. Similarly, in the labour market, the two firms may not individually face an incentive to invest in worker training if there is too much labour mobility between them, preventing either from amortising their training costs. Given the positive externalities of training, too little training may be provided in equilibrium.

The above illustrates that the idea of excessive rivalry leading to externalities in goods or labour markets is nothing new (even though, a few decades ago, this was considered a controversial subject). However, the generation of externalities and market failures through ineffective incentivisation schemes aimed at managers and individuals is a more recent—and again controversial—area. This article looks at how economics is catching up with the psychology literature in exploring the motivation of individuals, and concludes by assessing how this might affect incentivisation in public services and not-for-profit network utilities.

What motivates people?

The manner in which economic agents are incentivised to deliver performance has been studied by both psychologists and economists. While psychologists have focused primarily on the motivation and behaviour of individuals, economists have expanded their models to encompass firm behaviour. However, until recently, economists have focused largely on ‘behaviour’, with relatively little attention paid to ‘cognition’—that is, how people think, interpret, understand and feel. This is because it was assumed that the absolute amount of utility that a person derives (i.e., how happy they are) is not strictly measurable—therefore, cognition is not an area worth exploring. Economists’ models assumed that, provided people have a given and stable set of preferences, predicting their behaviour is reasonably straightforward.

Economists also assumed that the way in which individuals (in labour markets) are motivated is analogous to the ways in which firms (in product markets) are motivated. Most models of the firm—be they perfect competition, monopoly or oligopoly—assume absolute-profit maximisation, given the demand of consumers. Similarly, competitive models of individuals making ‘labour versus leisure’ choices assume absolute-utility maximisation by an individual, subject to their budget constraints.

Thus economists became stuck in a ‘behaviourist/competitive’ paradigm. In contrast, psychologists recognised much more quickly that, ultimately, cognition interacts with the ‘triggers’ to which individuals are exposed, which in turn determines their behaviour. Arguably, therefore, a more complete economic analysis of incentives might explore each of the four key stages set out below. Figure 1 highlights the importance of understanding cognition (1) in analysing...
what might motivate people (2) to behave in a certain way (3) and thereby to deliver organisational performance (4).

It could be argued that, during the 1980s and 1990s, labour market economists began to understand this through developing theories such as efficiency wages, which recognised that individuals were different to firms and goods and that motivating individuals was not straightforward. However, it is the more recent work by labour economists on fairness and happiness that has tackled head-on the core issue of cognition. Only now does the gap between economics and psychology appear to be narrowing.

The process began with an observation that, if labour markets behaved according to the behaviourist/competitive paradigm, when firms are asked how they set wages, the response would simply be: ‘we pay the going market rate’. In this kind of world, the ‘real wage’ paid to an employee for an additional hour of work should equal their marginal productivity. This is intuitive—if marginal productivity exceeded (was below) the real wage, firms would find it more profitable to use more (less) labour input since the additional output that could be sold would be greater (less) than the extra labour cost incurred.

However, surveys undertaken in the 1980s found that managers did not behave like this—they appeared to be more concerned about a wage rate that would motivate their workers, or what might be perceived as providing for sufficient motivation and hence productivity, or at least ‘fairness’. Furthermore, competitive theories could not explain why apparently similar workers earned different amounts in different sectors. Empirical studies of inter-industry wage differentials presented evidence to support the notion that, in industries or firms in which monitoring might be expected to be more difficult, real wages were higher than those predicted by the competitive paradigm.

Two major problems were that the competitive theories assumed perfect information, and that the ‘effort per hour’ made by employees was ‘exogenous’. By contrast, efficiency wage theories suggested that monitoring might be imperfect and costly, and that workers had some discretion over how hard they worked, depending on the wage rate. High ‘efficiency wages’ thus acted as an incentive device to resolve the principal–agent problem of eliciting worker effort. This not only made sense, it seemed consistent with the above empirical findings. Several reasons were put forward for paying efficiency wages. At the hiring stage, firms with monitoring problems that paid low wages might find themselves with an ‘adverse selection’ of workers (ie, ‘pay peanuts, get monkeys’); for employees, being paid a low wage meant that those underperforming would be less concerned if they were challenged since the penalty of dismissal would be lower. Moreover, excessively low wages might also induce greater staff turnover, which, given recruitment expenses, would be costly.

While these new theories provided explanations of how wages might affect motivation, they still relied on individual workers being motivated solely by their own utility, depending only on their own pay and effort. These theories therefore took a narrow view of cognition, and were tied somewhat to the behaviourist school. Crucially, the models failed to explain other empirical facts. For example, across all occupations in certain industries, workers tended to receive better wages than in other industries, and there was a strong positive correlation between industry profits and industry wages. Thus something beyond ‘individual hedonism’ seemed to be driving worker motivation.

At the end of the 1980s, a new efficiency wage model proposed that a key motivation for performance was fairness—workers cared not only about their own wages, but how these wages compared with those of their peers. In industries in which it was considered advantageous to pay some employees highly, it was also considered fair to pay other employees well. Workers withdrew effort as their wage fell short of the perceived fair wage for ‘comparison others’. This fair wage theory—itsel influenced by equity theory in psychology and social exchange theory in sociology—tackled the
nature of cognition head-on. Models of rent-sharing further illustrated how positive correlations between industry profits and wages might be explained by behaviour more appropriately regarded as collective action or ‘joint bargaining’ by workers as part of a relationship with the employer, rather than by individual behaviour. If workers with bargaining power—however derived—see a firm making profits, they consider it fair to demand a slice. Controlling for worker characteristics, industry type and (crucially) union power, significant premia on real wages could be attributed to rent-sharing effects.\(^1\)

**Habitation, rivalry and internal motivation**

The story did not, however, end there. A fundamental assumption of the behaviourist/competitive paradigm is that, as individuals become better off through increasing real wages, society becomes ‘happier’. After all, people derive utility from goods, and having more money to spend means they can buy more, raising their utility.

New ‘mood surveys’ (and neurological scanning techniques) demonstrate that happiness is now measurable. However, while in in most developed countries incomes have risen over time, surveys reveal that reported levels of happiness have not risen significantly.\(^11\) This is at odds with the behaviourist model. Recent research reveals that there may also be three further problems with the model, all concerning cognition:

- it assumes constant preferences;
- it posits that individuals’ wants are determined only by internal forces;
- it assumes that any external incentives imposed on an individual add to their existing motivation.

In economics terms, two of these three market failures are as follows.

- **Habitation**—people become accustomed to circumstances, and adjust their wants to recent experience, giving rise to a ‘hedonic treadmill’ effect. Unlike in the behaviourist world, it is not possible to be lifted onto ‘a permanently higher plane of existence’ through buying a new car or house.\(^1\) If individuals fail to foresee their habituation, they will over-invest in work and acquiring possessions at the expense of their leisure time. Indeed, there is considerable evidence from experiments that individuals systematically under-forecast habituation.\(^13\)

- **Rivalry**—behaviourist theory also assumes that individuals care only about their own income and lifestyle, and that work effort is non-rivalrous. However, as has been discussed, people care about the income of their peers (on fairness grounds), as well as their own. Furthermore, the nature of utility is itself relative. If Mr Jones buys a new car while his neighbour, Mr Smith, does not, Mr Jones will feel happier, while Mr Smith will feel less happy than before. Mr Smith may then feel compelled to buy a new car. The net result? Neither Smith nor Jones is significantly happier than before. Thus people may over-invest in work effort to maintain differentials, compelling others to follow suit (a ‘negative externality’), resulting in a zero-sum game struggle for rank.\(^15\) There is a wide range of evidence, from experiments, cross-country comparisons, and neuroscience of these rivalry effects.\(^15\)

Under-forecasting habituation and excessive rivalry can interact to ‘ratchet up’ societal norms, while producing no net increase in happiness.\(^16\) This suggests that care is needed in developing pay and incentive structures for management and employees, because habituation affects the impact of reward structures over time, while rivalry affects the impact of reward structures across individuals.

Indeed, Layard (2003) notes that performance-related pay needs to be considered very carefully in the case of public services.\(^17\) If output is easily measured, pay can be related to an individual’s absolute performance. If it cannot, one solution adopted by firms might be to assess performance based on relative judgements across workers. However, while this can motivate those who succeed, it can demotivate those who do not, leaving the organisation no better off.\(^18\) This is particularly relevant to public services, where outputs may be complex and multidimensional, long-term, or based on teamwork.\(^19\)

The third market failure is as follows.

- **Motivational crowding-out**—economics assumes that when external incentives are introduced (extrinsic motivation), such as through performance-related pay, individuals’ internal motivation (intrinsic motivation) remains the same. However, there is a range of theoretical and empirical evidence to show that these motivations are not always additive, and that external incentives can crowd out internal motivation—another form of externality.\(^21\)

Cognitive evaluation theory in psychology shows that intrinsic motivation has two core elements:\(^21\)

- people value having the autonomy and flexibility of getting on with their job (control);
- people value being competent in doing a ‘good job’.\(^22\)

Introducing extrinsic incentives, such as incentive payments, increases external control, thereby reducing autonomy. If individuals also derive pleasure from altruism, for example, by definition, being paid for
performance reduces this more complex source of motivation.23

Thus, while behavioural theory (and indeed efficiency wage theories) suggests that, to extract performance, formal external price incentives are necessary to ‘resolve’ informational problems, cognitive evaluation theory suggests that the use of such schemes, while important, can partly destroy or crowd out the complex intrinsic motivation processes, rendering the price incentives less effective than otherwise.24 This finding is quite controversial, given that performance-related pay is increasingly relied on in a variety of private and public organisations, and regulation of utilities in both the private and public sectors relies heavily on various forms of extrinsic incentives. However, economists instrumental in designing models of extrinsic incentives now acknowledge the key role of intrinsic motivation.25

Moreover, there is evidence of crowding-out effects. For example, Deci and Ryan (1985) ran a controlled experiment on two groups of students.26 One group was paid for each correct solution, while the other was not. At the end of the allotted time, both groups were permitted to go on working. While the unpaid group continued after this period, the paid group did not. Frey and Oberholzer-Gee (1997) examined a real-life case in Switzerland of the location of a nuclear facility, and showed that offering financial compensation actually reduced people’s willingness to act on the merits of the case.27 Frey and Goette (1999) found that voluntary workers put less effort into their work if they were financially rewarded,28 and Gneezy and Rustichini (2000) showed how parents in a US town became less punctual in dropping their children off at school when financial penalties were introduced for lateness.29

Again, intrinsic motivation seems more relevant to ‘public services’ than to other institutions.30 For example, Heckman, Smith and Taber (1996) found that workers in US Job Training Partnership Act training centres were motivated to help the least well-off, even though this gave their centres a poorer placement record, and reduced the performance payments received.31 Jacob and Levitt (2003) highlight empirically that, in teaching, high-powered incentives can have both costs and benefits, and that incentives work in different ways in different institutions.32 Reeson and Tisdell (2006) conducted a number of experiments with participants, and found that the degree of crowding-out depended on initial intrinsic motivation levels, and that weak regulation and/or weak financial payments resulted in large crowding-out yet little by way of extra extrinsic incentives.33 It is therefore crucial to evaluate the importance of existing intrinsic motivation prior to introducing any external incentive scheme, and, should such a scheme be introduced, organisations should ‘pay enough or not at all’.34

Implications for incentives

Ultimately, the performance of an organisation, and the degree to which the incentives it is exposed to work, depends on the total effort made by individuals working for it. Furthermore, whereas simple behavioural models of the firm assume an owner-manager facing a clear profit motivation in selling goods, the reality can be more complex. Public sector or public-ethos-focused organisations may not act so much like ‘firms’, but rather collections of individuals motivated by both intrinsic and extrinsic incentives.

An interesting question that arises from the above discussion, therefore, is that care needs to be taken in incentivising the provision of public services via private sector participation and/or performance-related pay, especially in areas where defining and monitoring performance is difficult, and where performance depends on initial levels of intrinsic motivation. These areas could encompass, for example, healthcare provision, emergency services, education, social services, charities, and environmental policy. However, there may also be implications for utilities, particularly where these reside in the public sector.

Layard (2003) highlights that, over the past 20 years, British governments may have made:

  serious errors in their approach to the reform of public services [ , stressing] ever more the need to reward individual performance, rather than providing an adequate general level of pay and stressing the importance of the job and the promotion of professional norms and professional competence.35

From a review of various studies, Jensen and Stonecash (2005) find that, while private provision of public services and incentive contracts can deliver efficiency improvements, the evidence is mixed, and that governments making decisions between directly providing or outsourcing public services must consider the informational complexities induced by multiple and incomplete contracts, as well as the impact of introducing ‘high-powered’ incentives on intrinsic motivation.36 In the environmental field, Reeson and Tisdell (2006) note that policy design might benefit from a better understanding of existing voluntary motivations, with policy aimed at strengthening these while also providing attractive extrinsic incentives.37 Grout and Stevens (2003) note that public-spirited workers may ‘self-select’ into public sector or not-for-profit organisation jobs, in effect providing ‘donated labour’, which may be diminished if external incentives are ‘pushed too far’.38

There may also be uncomfortable lessons for network utility regulation. In the case of the utilities privatised in the UK in the late-1980s, RPI – X-style incentive
regulation was seen as a means of resolving a principal–agent problem, given the informational asymmetries between regulators and the regulated. Under this 'hands-off' approach, prices are fixed for a set period (e.g., five years), and the private firm is provided with the managerial discretion to achieve cost reductions and service improvements in order to increase its profit beyond that assumed by the regulator, while meeting the broad requirements specified. RPI – X regulation therefore relies on an absolute-profit incentive as the 'cognition core', with external incentives to modify managerial behaviour, and to generate profit and output outcomes. This encompasses incentives created to encourage rivalry with the regulator (RPI – X per se), for finance (capital market competition for equity and debt), for managerial control (threat of takeover) and—in the water and electricity distribution sectors—between firms (comparative competition).

Indeed, it could be argued that the utilities were privatised because private management was judged to be more effective at decision-making than the public sector, and because equity was better than the government at motivating management. Given the inefficiency and investment backlog of these sectors, reliance on an intrinsic public service ethic to deliver an array of poorly defined goals needed to be replaced with a clearer external profit-related motivation to deliver better-defined targets. However, as the above discussion shows, rivalry and reliance on external incentives may not generate desirable performance outcomes in all situations. In the current context, as regards organisations that were not part of the initial wave of privatisations in the late 1980s, and which may be focused more on public objectives than on profit, RPI – X regulation may be less effective and require modification.

A good example is perhaps the post-privatisation rail sector in the UK. Jeffcott et al. (2006) highlight, through interviews with rail managers and staff, how the post-privatisation performance regime contributed to a 'blame culture' across the rail industry, with trust further eroded through fragmentation and loss of expertise. The recent moves to simplify the regulatory framework and reduce fragmentation in the rail industry have arguably been aimed at reducing information problems—and excessive zero-sum rivalry—between rail companies, contractors and (what is now) Network Rail (itself a not-for-dividend company). Oxera (2006) has also highlighted that rail workers may have intrinsic motivations other than money—such as idealism and professionalism.

As noted, RPI – X relies on the existence of a profit motivation—should poor performance occur, equity provides a buffer—the dual being that equity also provides an incentive for management outperformance. Furthermore, there is little threat of takeover for public sector companies. In the absence of equity, customers and/or taxpayers may effectively become the buffer. Indeed, regulators have acknowledged that, due to the blunting of equity pressure, applying the RPI – X model is more difficult in the context of not-for-profit, public sector and/or subsidised utilities. The solutions put forward to address this have been additional direct external incentives for management.

In the case of Network Rail, additional external measures include incentives for management to improve the financial health of the business. A number of regulators have also suggested enhancing incentives through introducing performance-related pay, following the example of the Welsh Water (private sector not-for-profit) model (which has had a number of successes in achieving efficiency and performance). Oxera (2006) highlighted that, while performance-related pay could be used to improve incentives in the rail sector, this would need to be adopted 'appropriately' to avoid perverse incentives, given the presence of multiple outputs and the importance of teamwork in the sector. A scheme has also been proposed by Royal Mail (a public sector organisation), in which workers may receive 'phantom shares' in the organisation—Royal Mail hopes that this will incentivise and reward the contributions of its employees. Arguably, therefore, this scheme is intended to promote 'rent-sharing'-style incentives via profit-sharing.

As discussed above, performance-related pay may have its benefits in reducing costs or increasing service performance, but it may also generate rivalry in sectors in which cooperation and a public sector ethos are required. The discussion also illustrates that, where such external incentives are introduced (which seek to home in on encouraging good performance), care must be taken to avoid crowding out the more complex, intrinsic ones. A greater understanding is therefore required of what motivates public sector utilities, and where rivalry helps or hinders performance, in considering how the RPI – X model might be augmented to compensate for the absence of equity.

1 After adjusting for variations in skill levels, training costs and accounting for compensating differentials for risk.

2 Real wages are monetary wages adjusted for inflation, and represent the quantity of goods and services that an hour of labour earns. Marginal productivity is the additional value of output produced by this increment of labour.


11 Another analogy is that, if everyone stands up in a football stadium to get a better view, no one is better off, and in fact the fans as a whole are collectively worse off.


13 This may seem obvious to managers, as the consequences of micro-management can be reduced worker motivation, and any firm would wish to employ highly motivated and professional individuals. However, these traits tend to be ignored in the behavioural-economics tradition.

14 Another analogy is that, if everyone stands up in a football stadium to get a better view, no one is better off, and in fact the fans as a whole are collectively worse off.


16 Ibid.

17 Ibid.

18 Ibid.


27 Ibid.

28 Ibid.


34 See, for example, Gneezy, U. and Rustichini, A. (2000), 'A Fine is a Price', Journal of Legal Studies, XXX, pp. 1–18. It could be argued that, even within an economics-based game-theoretic framework, the 'rules of the game'—in so far as parents have understood them—have changed. Previously, individual parents' payoffs depended on how other parents perceived their public spiritedness (recognition), given the illegitimacy of lateness. The incorporation of explicit penalties might then legitimise lateness as a commodity with a price. Given that all parents might then understand this rule change, a new sub-game perfect equilibrium might be one in which parents are willing to be late since they feel entitled to do so, given that they pay a price for the privilege.


41 See, for example, Ofwat (2001), 'The Proposed Acquisition of Dwr Cymru Cyfyngedig by Glas Cymru Cyfyngedig: A Position Paper by Ofwat', February, revised.

42 See, for example, Ofwat (2001), 'The Proposed Acquisition of Dwr Cymru Cyfyngedig by Glas Cymru Cyfyngedig: A Position Paper by Ofwat', February, revised.

43 See, for example, Ofwat (2001), 'The Proposed Acquisition of Dwr Cymru Cyfyngedig by Glas Cymru Cyfyngedig: A Position Paper by Ofwat', February, revised.