Organizational Control Systems and Pay-for-Performance in the Public Service

Bruno S. Frey
University of Warwick, UK and CREMA (Center for Research in Economics, Management, and the Arts), Switzerland

Fabian Homberg
Bournemouth University, UK

Margit Osterloh
University of Zurich, Switzerland and CREMA (Center for Research in Economics, Management, and the Arts), Switzerland

Abstract
Under certain conditions, output related performance measurement and pay-for-performance produce negative outcomes. We argue that in public service, these negative effects are stronger than in the private sector. We combine Behavioural Economics and Management Control Theory to determine under which conditions this is the case. We suggest as alternatives to the dominant output related pay-for-performance systems selection and socialization, exploratory use of output performance measures, and awards.

Keywords
organization control, organizational forms, public administration and organizations, public service motivation

Introduction
Control and performance measurement systems are critical to how organizations function (Scott, 1992; Meyer & Gupta, 1994; Meyer & Evans, 2003). An abundant literature highlights the link between the design of control systems and the organization’s task environment (Ouchi, 1977, 1979; Turner & Makhija, 2006; Cardinal, 2001; Kirsch, 1996). However, this literature concentrates on private sector firms. Until now it has largely ignored the public sector. There seems to be a general...
disinterest for public sector organizations by management scholars (Mahoney, McGahan, & Pitelis, 2009) though insights derived from the analysis of public sector organizations could be of relevance for management in general (Benz & Frey, 2007).

Today, in the wake of New Public Management, output related performance measures and rewards become commonplace in public sector organizations in many countries (see OECD, 2003; Lah & Perry, 2008; Perry, Engbers, & Jun, 2009). They are major characteristics of public sector reforms (McNulty & Ferlie, 2004; Varone & Giauque, 2001; Mascarenhas, 1993; Sherwood & Wechsler, 1986; Siegel, 1987). These characteristics are assumed to raise public servants’ motivation and enhance service quality. Indeed, praise for output related pay-for-performance as a symbol of modern quality management is common in New Public Management (e.g. Moynihan & Pandey, 2010). As a consequence, such systems diffuse into many public areas, such as university education (Meyer & Evans, 2003), general teaching (Wragg, Haynes, Wragg, & Chamberlin, 2004; Thomas & Davies, 2005), the healthcare sector (Finn, Currie, & Martin, 2010; Mueller, Sillince, Harvey, & Howorth, 2004) or internal revenue agencies (Bertelli, 2006).

However, many authors point to the difficulties inherent to the introduction of such control systems in general (e.g. Frey, 2007; Lindenberg, 2001; Lindenberg & Foss, 2011; Falk & Kosfeld, 2006; Bowles, 2008; Bowles & Polania-Reyes, 2012) as well as in the public sector (e.g. Milkovich & Wigdor, 1991; Perry, 1986; Ingraham, 1993; Kellough & Lu, 1993; Perry, Mesch, & Paarlberg, 2006). In this vein, it has been argued that ‘reforms in the public sector often make little impact on the product and services provided, or the impact they make is quite different from what was intended. … the practical effects are equally uncertain’ (Brunsson & Sahlin-Andersson, 2000, pp. 730–1). A recent review investigating 57 studies on pay-for-performance in the public sector concluded ‘performance-related pay continues to be adopted but persistently fails to deliver its promise’ (Perry et al., 2009, p. 46). Nevertheless, the introduction of these practices has met with remarkably little opposition (Hyndman & Eden, 2001) though a host of empirical evidence is available to back both the early and recent criticisms of pay-for-performance systems (Cox, 2000). Consequently, as Dubnick (2005, p. 378) argues, ‘the popular acceptance of various NPM reform initiatives … was based in good part on the unchallenged rhetoric of “greater accountability will mean improved performance”’.

First, one major problem is the lack of an adequate theoretical underpinning of New Public Management. This problem, addressed early (e.g. by Perry, 1986), is still existing today. Second, there is no systematic review on the efficacy of pay-for-performance initiatives in the public sector. A third problem is that, to our knowledge, there is little research focusing on public sector organizations that investigates the alignment of tasks to the control-and-reward system.

We intend to fill this gap by making three major contributions. First, we ask what might be an adequate theoretical underpinning of control systems in New Public Management. We examine whether the principal-agent theory, usually taken as a theoretical background for New Public Management, really fits the main characteristics of public service. We follow Larkin, Pierce and Gino (2012) who criticize principal-agent theory’s failure to consider elements of intrinsic motivation and other psychological factors at work which are of high importance in the public sector. As we find in general that principal-agent theory is not an adequate theoretical basis for the public sector, we look for a more adequate theoretical basis. We suggest Behavioural Economics as a more useful approach.1 We argue that this approach better copes with the special challenges of public service than principal-agent theory, though it does not fully grasp the rich psychological reality in the public service. Second, we offer a systematic review of the main issues surrounding the debate on the efficacy of pay-for-performance in the public sector. In order to achieve this goal, we apply Management Control Theory and Behavioural Economics to determine the conditions that are appropriate for different control-and-reward systems. We argue that
output control, which according to New Public Management should be applied in the first place, is not adequate for most tasks in the public sector. Therefore output control has to be substituted by other forms of control. Third, we provide ideas for a future research agenda concerning incentives to motivate public servants in cases when output related performance measurement and pay-for-performance should be avoided. Overall, our contributions intend to facilitate a better understanding of public sector control systems and enable a nuanced design of incentive mechanisms in public sector organizations.

The next section contrasts different views of public service behaviour, in particular, New Public Management and Behavioural Economics. The third section applies the insights gained from Behavioural Economics to public sector organizations. In the fourth section, we review Management Control Theory to determine more closely under which conditions pay-for-performance in public sector organizations is applicable. In the fifth section we consider alternative measures which could improve performance in cases when pay-for-performance fails, e.g. the explorative use of performance measurements, (self-) selection and awards. The last section provides our conclusions.

Two Views of Public Service Behaviour

New Public Management

New Public Management is characterized by a strong emphasis on output performance measurement and by the introduction of pay-for-performance according to output indicators. Its theoretical basis is standard economics, in particular, the principal-agent view (Jensen & Meckling, 1976; Kaboolian, 1998; Arellano-Gault, 2000), as proposed by Jensen and Murphy (1990). These theories build strongly on the model of the self-interested *homo oeconomicus*. They accept as a matter of course that ‘cash compensation should be structured to provide big rewards for outstanding performance and meaningful penalties for poor performance’ (Jensen & Murphy, 1990, p. 141). High-powered monetary incentives are assumed to align the interests of the agent and the principal. Following the theoretical price effect, the higher the price the more effort is exerted. According to this view extrinsic incentives in the form of monetary rewards motivate individuals’ additional or marginal efforts. These incentives satisfy personal needs in an instrumental way because money usually serves as a means to an end – that is, buying food, purchasing a house, paying for a vacation – and not an end in itself. In contrast, intrinsic incentives satisfy needs directly or in a non-instrumental way if the activity or its outcome is valued for its own sake.

There is some empirical support for principal-agent theory and its assumptions in the private sector. For example, a study by Lazear (2000) shows that in the Safelite Auto Glass Company the introduction of piecework pay raised productivity significantly. Based on the number of glass units assembled per worker per day, productivity increased by 36%, including an incentive effect of 20% and a self-selection effect of 16%, whereas salary costs only rose by 9%. This finding confirms that pay-for-performance can raise effort and can attract highly talented workers. However, the application of principal-agent theory within public service might be misleading because the work of public officials differs from work on an assembly line. There are three reasons why this is the case.

First, there exists ample empirical evidence that private and public employees differ, in particular with respect to their motivation (Bhagwat et al., 2004; Brewer & Selden, 1998; Heath, 1999). It is argued that in the public sector prosocial motivation plays a bigger role than in the private sector, due to different selection effects and job designs (Crewson, 1997; Festre & Garrouste, 2008;
Houston, 2000, 2006; Rainey, 1982). Therefore theories based on self-interest do not suffice to analyse motivation within the public sector (e.g. Vandenabeele & Hondeghem, 2005). Behavioural Economics provides an extension to the principal-agent theory that encompasses public service motivation more adequately (see below).

Second, in the public sector – in contrast to the private sector – there is no principal as a residual claimant (Alchian & Demsetz, 1972) who supervises the agents in exchange for earning the residual gain (Burgess & Ratto, 2003; Francois, 2000). In the public sector it is not clear to whom the service providers are responsible, for example to politicians as representatives of the citizens, to user groups, or to professional peers following professional norms (Bryer, 2007). Research has shown that employees are likely to engage in resistance behaviours if such professional norms and professional autonomy is undermined by the introduction of private sector control practices (Doolin, 2002). There is also usually no market price for the outputs of public service and very often competition between different service providers does not exist. Nonetheless, quasi-competition to a considerable extent has been introduced in public service (LeGrand, 2007, 2010).

Third, in public service the goals are often characterized by high ambiguity. Therefore incomplete contracts prevail. An example is the police force. Policing is the classic textbook example of a situation in which explicit contracting for results is not feasible (e.g. Besanko, Dranove, Shanley, & Schaefer, 2006, p. 479). If, nevertheless, police forces are subjected to performance contracts that introduce financial rewards for meeting a number of quantiative performance targets (like number of fines, number of suspects per type of crime brought to prosecution) – as was the case in the Netherlands – it is doubtful that the contracts will contribute to increased crime reduction (Spekle & Verbeeten, 2009). In addition, means-end relations are often not understood (Bevan & Hood, 2006; Kravchuk & Schack, 1996) and the tasks are highly interdependent. It is often the case that the desired outcome is influenced by uncontrollable events or by decisions taken elsewhere. This means that the actors do not know or do not control the production function that transforms efforts into results (e.g. Gibbons, 1998). As a consequence, available performance measures are incomplete and incentives are often not aligned with the overarching goals but have distorting effects (Kerr, 1975). Such effects are obvious also in the private sector. As the recent financial market crisis shows, the prospect of huge salaries has turned some managers from ‘legends’ (Hegele & Kieser, 2001) into ‘crooks’ (Osterloh & Frey, 2004). However, distorting effects in the public sector are more salient than in the private sector because there is more leeway for self-serving activities, due to the ambiguity of the goals and the absence of market prices. We will discuss below in detail whether more adequate control systems than output control exist in order to deal with high ambiguity of goals and means-end relationships in the public sector.

As a consequence, we conclude that the principal agent view which usually is seen as the theoretical basis of New Public Management has serious flaws. It cannot be taken for granted that output measures linked to high-powered incentives always provide an adequate motivation within public sector organizations (see e.g. LeGrand, 2003; Moynihan, 2010; Arellano-Gault & Lepore, 2011).

The view of Behavioural Economics

Behavioural Economics suggests that the sources of an individual’s motivation are much broader than encompassed in the standard economic view of principal-agent theory (Rabin, 1998; Mullainathan & Thaler, 2001). According to Tomer (2007) it is a less narrow, less intolerant and
less mechanistic school of thought. It takes intrinsic motivation into account (e.g. Frey, 1997; Osterloh & Frey, 2000; Meier, 2006). As mentioned, intrinsic motivation is based on the satisfaction an individual derives from involvement in an activity without external rewards. The distinction of intrinsic and extrinsic motivation serves the purpose of this paper because it gives a clear-cut distinction between extrinsic (instrumental) and intrinsic (non-instrumental) motivation. The ideal intrinsic incentive system resides in the work content itself, which must be satisfactory and fulfilling for the employees.

There are two kinds of intrinsic motivation (Lindenberg, 2001). Enjoyment-based intrinsic motivation refers to a satisfying flow of activity. Examples are playing a game or solving an interesting puzzle. In each case, pleasure is derived from the activity itself which may provide a ‘flow experience’ (Csikzentmihalyi, 1975). Obligation-based intrinsic motivation, which is particularly relevant to public service, refers to an activity with the goal to act appropriately. When individuals are driven by obligation-based intrinsic motivation they follow norms for their own sake. In particular they take the well-being of others into account without expecting a reward. The welfare of the community enters into the preferences of the individuals. While the standard economic model of human behaviour – the *homo oeconomicus* – is based on the assumption of self-interested, extrinsically-motivated individuals, a growing body of empirical evidence indicates that many people are prepared to contribute voluntarily to the community of which they feel to be a part (e.g. Frey & Jegen, 2001; Frey & Meier, 2004; Bowles & Gintis, 2004; Fehr & Fischbacher, 2002).

Under certain conditions, extrinsic incentives can crowd out intrinsic motivation (Bénabou & Tirole, 2003; Falk & Kosfeld, 2006; Frey, 1997; Lindenberg, 2001; Osterloh & Frey, 2000; Weibel, 2007), an effect called ‘the corruption effect of extrinsic motivation’ (Deci, 1975; Deci & Flaste, 1995; Deci, Koestner, & Ryan, 1999). Frey (1997) introduced it into microeconomics as the crowding-out effect. There are several theoretical underpinnings of the crowding-out effect (see e.g. Lepper & Greene, 1988; Bowles, 2008; Festre & Garrouste, 2008). The most cited one is self-determination theory (Deci & Ryan, 2000a, 2000b; Gagné & Deci, 2005).

According to this theory, self-determination suffers with the introduction of external controls, such as monetary incentives. As a result, individuals shift their locus of causality from the inside to the outside. Consequently, they only fulfil their duty when external incentives exist. Related to self-determination theory is the ‘theory of over-justification’ (Kruglanski, Schwartz, Maides, & Hamel, 1978) or the ‘goal framing theory’ (Lindenberg 2001; Lindenberg & Foss, 2011). Substituting external incentives for intrinsically motivating rewards may shift the attention from the task or from normative goals to the reward.

For the crowding-out effect to become apparent, three conditions are necessary. First, intrinsic motivation must have been present before crowding-out occurs (Calder & Staw, 1975). Second, the recipient must perceive the reward as controlling (Deci, 1975; Frey & Jegen, 2001). According to self-determination theory (Deci & Ryan, 2000a, 2000b), an incentive may have positive effects on intrinsic motivation when it is perceived as supportive (see also Lawler, 1969). Third, the price effect (i.e. motivation induced by extrinsic rewards such as money) does not compensate for the decline in intrinsic motivation (Gneezy & Rustichini, 2000; Weibel, Rost, & Osterloh, 2010).

Psychologists and economists have empirically analysed the crowding-out effect in hundreds of laboratory experiments (e.g. Cameron, Banko, & Pierce, 2001; Deci et al., 1999; Weibel et al., 2010), as well as in econometric studies of real-life events (e.g., Frey, Oberholzer-Gee, & Bohnet, 1997; Gneezy & Rustichini, 2000). Frey and Jegen (2001) summarize the empirical evidence. LeGrand (2003) and Bertelli (2006) apply it to the public sector.
It is important to note that symbolic gifts do not crowd out intrinsic motivation (Rheinberg, 2008). When a superior acknowledges an employee’s extraordinary effort with a symbolic gift, for example, a bunch of flowers or a non-monetary extrinsic reward such as an award, an order or a title (Frey, 2007; Swiss, 2005), it raises the employee’s intrinsic motivation. Such gifts make employees aware that intrinsic motivation is appreciated. The more unexpected the reward, the stronger the effect will be (Rheinberg, 2008). If the employee feels that the superior’s gesture serves only an instrumental purpose, intrinsic motivation is impaired. Intentions play a highly relevant part in the crowding-out effect (Tyler, 1999; Bowles & Polania-Reyes, 2012).

The Behavioural Economics approach overcomes some of the limitations of standard economic theory but has two limitations. First, the implications of Behavioural Economics for the design of institutions or governance systems are limited. Due to its grounding in psychological research and its heavy reliance on laboratory experiments, it explores mainly individual and group behaviour but not institutional arrangements beyond that level. Practical implications on the meso or macro level of institutional design are often drawn in a speculative way. Second, Behavioural Economics – though claiming that it is based in empirics – still follows the axiomatic approach of standard economics. It starts with a highly stylized model and puts only few of its axiomatic assumptions into question (Camerer & Loewenstein, 2004). This procedure on the one hand helps to keep the comprehensiveness and elegance of the economic model. On the other hand, this model takes psychological reality into account to a limited extent (Berg & Gigerenzer, 2010; Tomer, 2007).

To overcome the weaknesses of Behavioural Economics we combine this approach with insights from the literature on Public Service Motivation as well as Management Control Theory.

**Application of Behavioural Economics in Public Service**

The concept of intrinsic and extrinsic motivation has been applied within the context of public sector employment referred to as ‘public service motivation’ or PSM. It is characterized by an individual’s predisposition to contribute to public goods with the purpose of doing well for others and society (Perry, 1997; Perry & Hondeghem, 2008).

There is an intense discussion about what motivates public sector employees (Moynihan & Pandey, 2007; Perry et al., 2008). A number of studies support the notion that public employees are motivated less by monetary rewards than their private counterparts (Brewer & Selden, 1998; Houston, 2006). Qualitative and quantitative studies support the public service motivation hypothesis: when compared with employees in the private sector, public servants have a higher degree of intrinsic motivation, they have a greater interest in altruistic activities and socially desirable outcomes, and they rate intrinsic incentives higher than extrinsic ones (e.g. Crewson, 1997; Guyot, 1962; Houston, 2000; Kilpatrick, Cummings, & Jennings, 1964; Rainey, 1982; Warner et al., 1963). Recently, evidence for the validity of the public service motivation concept beyond the US has been provided (Kim, 2009; Ritz, 2009). Empirical evidence supports the view that performance measurement systems driven by NPM reforms are ill-designed for triggering public service motivation (Monyihan & Pandey, 2010). Economists have also become interested in the concept of public service motivation (e.g. Georgellis and Tabvuma (2010) provide evidence of the stability of public service motivation, Francois (2000) and Francois and Vlassopoulos (2008) underline the benefits of not using strong incentives in the presence of public service motivation).

These empirical findings have not remained unchallenged. For example, Gabris and Simo (1995) do not find empirical evidence for a special public service motivation and suggest abandoning the concept. Also Lyons, Duxbury, and Higgins (2006, p. 614) do not find evidence for a lower preference of extrinsic rewards of public sector employees, but they argue that this might be due to
sample selection. Another strand of research investigates the issue of person-organization fit (Lewis & Frank, 2002; Christensen & Wright, 2011). There are consistent findings that higher levels of individual PSM make public sector employment more attractive to an individual. However, high levels of PSM are not found to be a predictor of public sector employment (Lewis & Frank, 2002). Other findings, for example, that of Tschirhart, Reed, Freeman and Anker (2008), support the idea that individuals having a preference for helping others (which is part of PSM) will self-select into the public sector. But they do not find an effect of salary preferences in relation to sector preference. Earlier work by Wright and Pandey (2008) provides evidence for the mediating effects of value congruence on job satisfaction. As a summary of this discussion we agree with Christensen and Wright (2011, p. 738) ‘that simply linking PSM and employment sector may be insufficient to determine person-organization fit’.

In the public sector also crowding-out effects of intrinsic motivation have been confirmed, e.g. with the US Internal Revenue Service (Bertelli, 2006). It also has been shown to affect individuals with high levels of public service motivation when switching from the public to the private sector (Georgellis, Iossa, & Tabvuma, 2011). But in general the evidence is blurry. There also exist studies showing that financial incentives induce public servants to increase their work effort (e.g. LeGrand, 2007, 2010; Andersen & Pallesen, 2008) except when strong professional norms are present (Andersen & Blegvad, 2006).

These inconclusive findings show that public servants may possess varying degrees of public service motivation (LeGrand, 2003, chapter 3). Also, with certain tasks the price-effect (increase of performance induced by extrinsic rewards) may compensate the crowding-out effect, as has been experimentally shown in fields other than public service (Gneezy & Rustichini, 2000; Weibel et al., 2010). Therefore it is useful to differentiate task characteristics in order to find out which kinds of tasks in the public service are more or less susceptible to crowding-out effects induced by pay for performance.

Management Control Theory

In the following section we apply Management Control Theory to public sector organizations in order to specify the antecedents under which pay-for-performance can be expected to deliver its promise (according to New Public Management) or under which conditions a negative outcome might take place (according to Behavioural Economics).

Management Control Theory distinguishes different kinds of control and shows to which kinds of tasks they are applicable (Eisenhardt, 1985; Kirsch, 1996; Kirsch, Ko, & Haney, 2010; Ouchi, 1977; Simons, 1995; Turner & Makhija, 2006). Three control mechanisms are distinguished: output control, process control, and input control. Although all organizations use a combination of these control mechanisms, two aspects are decisive for employing the optimal control mechanism. These aspects are: (1) the nature of the task, particularly the measurability of outputs; and (2) the knowledge of the supervisor about the cause-effect relations (Thompson, 1967), or the transformation process (Ouchi, 1977), and the appropriate rules to be applied (Turner & Makhija, 2006). Figure 1 illustrates the relationship between these aspects and the optimal control mode.

Output control

Output control does not specify the processes or type of behaviour that produces the desired outputs but measures countable output indicators. Output control is most appropriate when processes or cause-effect relationships are difficult to specify but outputs are easy to measure. This is the
reason why non-experts who do not understand the cause-effect relationship prefer output indicators, and many journalists, politicians, and other external observers are keen on output control. Output indicators seem to provide easy, understandable quality data.

However, there are some preconditions of output control that often fall short in the case of providing public services. These circumstances create an ‘illusion of control’ (Rosanas & Velilla, 2005, p. 87). First, the knowledge relating to the output must be clear-cut, stable, and not subject to change (Snell, 1992). Second, outputs must be observable and attributable (Eisenhardt, 1985) and not characterized by intensive interdependencies between the various actors involved (e.g. different agencies, private companies, and citizens).

However, in public service, desired outputs are often ambiguous (Burgess & Ratto, 2003) and the tasks are highly interdependent so that the outputs are hard to attribute to individuals. Public servants normally provide complex services such as ‘good health’, ‘good education’, or ‘important knowledge’ (Plant, 2003; Perry, 1986). In these cases, it is difficult to identify clear goals and performance measures (Rainey, Backoff, & Levine, 1976; Cutler & Waine, 2005). An example is the US Foreign Service: Its goal is to improve international relations. In this case, outputs are not measurable and specific advice cannot be given on how to ‘transform manpower into good foreign relations’ (Ouchi, 1977, p. 98). The task is not contractible (Masten, 2006; Spekle & Verbeeten, 2009). This is often true even at the lower bounds of hierarchies, e.g. with police officers, benefits assessors, or fire fighters. As Lipsky (1980) asserts, street-level bureaucrats, not top public officials, put policies into effect. Superiors have difficulties in observing their decisions. Street-level bureaucrats and other public officials have to equate efficiency with equity without having the clear-cut framework of, for example, an insurance salesperson (Burgess & Ratto, 2003).

If output based control and reward systems are applied although the task does not display the prerequisites for output control, dysfunctional effects may arise. Individuals who are not intrinsically motivated will have a strong incentive to respond to those indicators that are easy to measure, that is, the quantifiable performance-related aspects of a task. Data that is not easy to measure is disregarded, although it might be crucial to fulfilling the task. The reliance on quantitative criteria
to govern work behaviour neglects the more important qualitative aspects of public services (see also Emery & Giauque, 2005). LeGrand (2010, p. 63), for example, mentions ambulances that concentrated on dealing with emergencies a short distance away so as to meet the goal to respond within eight minutes. They ‘hit the target and miss the point’ (see also van Bockel & Noordegraaf, 2006; Hood, 2006, p. 516). This effect is also well known as the ‘goal-displacement-effect’ (Merton, 1940; Perrin, 1998) or the ‘multiple-tasking effect’ (Ethiray & Levinthal, 2009; Holmstrom & Milgrom, 1991; Kerr, 1975). There is much evidence of this effect (Staw & Boettger, 1990; Gilliland & Landis, 1992; Schweitzer, Ordonez, & Douma, 2004; Ordonez, Schweitzer, Galinsky, & Bazerman, 2009). For example, Fehr and Schmidt (2004) show that output-dependent financial incentives lead to the neglect of non-contractible tasks. An example in public service is teachers responding with ‘teaching to the test’ when they are assessed according to quotas of students who pass a certain exam (Amrein & Berliner, 2002; Heilig & Darling-Hammond, 2008; Nichols, Glass, & Berliner, 2006). One step further goes ‘cream skimming’ and ‘gaming the system’ (van Thiel & Leeuw, 2002). Dalrymple (2001) presents an illustrative example. French police officers decided not to investigate a robbery. The robbery would have increased their district’s crime rate that in turn would have cancelled the officers’ end-of-year bonuses. Other examples are chronically ill patients excluded from healthcare, teachers responding to evaluations by excluding bad pupils from tests (for empirical evidence in the US see Figlio & Getzler, 2002) or putting lower quality students in special classes that are not included in the measurement sample (Corley & Gioia, 2000). In the academic field, an example is the ‘slicing strategy’ whereby scholars divide their research results into as many papers as possible to enlarge their publication list (Butler, 2003). These effects contribute to what is called the ‘performance paradox’, namely the fact that performance measures have the tendency to ultimately lose their ability to discriminate between good, average, and bad performance (Meyer & Gupta, 1994; Meyer, 2005). This explains evidence suggesting that in spite of more sophisticated tools for output measurement performance has not improved (e.g. Marsden & Belfield, 2006).

**Process control**

Process control is an alternative to output control. The preconditions are that evaluators: (a) have the appropriate knowledge of cause-effect relationships or the transformation process of inputs into outputs, and (b) have a shared understanding of the rules obtained. The external controller must possess such knowledge in order to make a qualified judgement. Process control and peer control are comparable because peers know about cause-effect relationships and have an understanding of the rules. For example, the peer-review system in academia controls for the appropriate application of methods and consistency of arguments (see Osterloh & Frey, 2010, for a critical perspective on peer-reviews).

Process control is indeed of special importance in many areas of the public sector because it serves to provide equal treatment of all citizens according to the rule of the law as well as transparency to the public in order to participate in the political process. According to Weber (1978), this is the qualifying characteristic of a well-functioning administrative system, which he calls bureaucracy. However, this view has also met with criticism due to its inflexibility and rigidity (e.g. Burns & Stalker, 1961; Mintzberg, 1979; Woodward, 1958; for a summary see Meyer, 1995). Such bureaucracy is inadequate in a rapidly changing environment as well as for many ambiguous tasks, for example, in matters relating to health, social, and educational services. In these areas, those in charge must apply the rules in such a way as to prevent arbitrariness, but, at the same time, they must employ a sense of proportion and power of judgement. Thus, a key challenge for public sector
managers is to balance the need for process control to guarantee equity and impartiality while at the same time imposing a control mechanism that furthers adaptation to changing environments and ambiguous tasks. Under these circumstances, a combination of process control and a form of control known as input control in Management Control Theory is necessary.

**Input control**

Input control assesses whether individuals or groups that have been selected for employment fit to the organization and have internalized norms and professional standards, i.e. are dedicated intrinsically to their task. According to Ouchi (1977), who refers to this form of control as ‘clan control’, this is the control form necessary for complex and ambiguous tasks. With input control, selection and socialization take place by and inside professional groups. Internalization of norms and values is induced by senior colleagues behaving as desired by the institution. Senior colleagues function as role models for younger co-workers. Additionally, institutionalized rituals align the behaviour of junior colleagues. This is the case with professions characterized by a low degree of observable and attributable outputs and processes. Examples in the private sector are medical doctors and executive search companies (Zehnder, 2001). In public service these kinds of tasks are of special importance due to the fact that complex and ambiguous tasks exist on all levels ranging from life-tenured judges (e.g. Benz & Frey, 2007; Posner, 2010) to streetworkers, policemen and clerks (Lipsky, 1980). In all these cases careful selection and socialization processes have to be applied.

The aim of input control is twofold. First, it should make sure that civil servants have internalized norms that limit their self-interests as long as they accept their legitimacy (Tyler, 1999; Tyler & Blader, 2000). If such norms were not present, economic incentives would gain prevalence. Exemplary empirical evidence shows that in fact in public service organizations strong professional norms override the effects of money (Bøgh-Andersen, 2009). Also Brewer and Selden (1998) in their study on whistle-blowers show that high performers in the public sector were willing to set their private interests aside for the public good.

Second, input control should make sure that there is a congruence between the characteristics of individuals and the characteristics of the public service organization beyond public service motivation. This fit takes place if the employee and the organization share ‘fundamental characteristics’ (Kristof, 1996) in terms of attitudes, skills and values. The literature suggests that person-organization fit has a positive impact on performance, job satisfaction, commitment, and turnover (Kristof-Brown, Zimmerman, & Johnson, 2005). Also in the public sector there is a growing stream of research linking public service motivation to person-organization fit. Even with high levels of public service motivation there may exist big differences in common values, e.g. between employees in defence and education. Following these ideas, Bright (2007, 2008) establishes person-organization fit as a mediator between public service motivation and performance. Carpenter, Doverspike, and Miguel (2012) replicate these findings in an experimental setting. These results underpin the importance of careful input control to avoid a mismatch between individuals and public organizations even if public service motivation is high.

Input control has three disadvantages. First, it is in danger of being submitted to groupthink (Janis, 1972) and cronyism (Osterloh, 2010). The only way to overcome this danger is to foster diversity within the relevant peer group. Empirical support for this argument was found in a private sector environment: Rost and Osterloh (2010) show empirically that, during the recent financial market crisis in the Swiss banking industry, companies performed better with heterogeneous boards than with homogeneous boards.
Second, input control that relies on the norms of professional peers is susceptible to paternalism, e.g. a low accountability to politicians as representatives of the citizens, because they think that they know best (LeGrand, 2007). They may legitimize their paternalism with the great information asymmetry between them and non-experts. But on the other hand, paternalism will serve to consolidate the information asymmetry. To prevent such insulation of professional experts from the recipients of public service LeGrand (2010) suggests giving the citizens voice and choice. Referenda or other forms of direct democracy put voice and choice into practice (Frey, 1994; Frey, Kucher, & Stutzer, 2001).

Third, input control is more demanding than output or process control. Motivation, values, attitudes, and norms are not observable as outputs or processes are. Therefore recruitment policies and human resource development is a major task in public service. As a consequence it makes sense that in public service organizations internal recruitment and promotion plays a bigger role compared to external recruitment, because internal recruitment allows for more intense screening of a person-organization fit.

Notwithstanding the three limitations mentioned it can be argued that input control is the most appropriate control system when tasks are ambiguous, complex and interdependent. In these cases, output control could lead to dysfunctional reactions like gaming. Process control could result in inflexibility and rigidity. Both problems could decrease perceptions of legitimacy of the public service among citizens.

Pay-for-Performance and Control Forms

Management Control Theory clarifies which kinds of control exist and at the same time clarifies under which condition pay-for-performance makes sense. In reality, there is always a mix of the different control forms and, accordingly, a mix of different remuneration systems is applied.

Output control can be used in public administration primarily for simple and well-defined tasks, such as waste management, garbage collection services or building and infrastructure maintenance. In these cases, output control can be linked to pay-for-performance as principal-agent theory suggests. First, with such tasks, the goal-displacement effect is low because the criteria for the desired output are unambiguous. Second, in simple tasks an intrinsic motivation might exist, but it is not as necessary as is the case for complex tasks. These insights are confirmed empirically in a vignette study (Weibel et al., 2010) as well as in the field. Spekle and Verbeeten (2009) show in a study of 101 public sector organizations that overall performance will decrease if performance measures are used to determine incentives and sanctions for difficult to specify tasks.

However, this is not to say that output control has generally to be avoided with complex tasks. As empirical evidence in R&D shows, output control can enhance innovation when it is linked to public recognition as long as it is not linked to pay-for-performance (Cardinal, 2001). In this case output control may be perceived as supportive (see also Andersen & Pallesen, 2008, for research institutions) or as explorative use (see below).

Process control is appropriate if rules are unambiguously clear. If there is no room for interpretation on behalf of the controlled and controlling people then pay-for-performance may be linked to the correct fulfilment of rules. This is the case with routine-based tasks. Some examples are the processing of requests to authorities, such as applications for permits of stay, building permits, or applications for unemployment pay. However, if there is some discretion required in interpreting the rules, administrators must exercise utmost care in granting monetary rewards. Because monetary incentives cannot enforce a sensible interpretation of the rules, negative side effects might occur. Again, this is not to say that process control has to be avoided with complex
tasks as long as it is perceived as supportive and not linked to pay-for-performance (Cardinal, 2001; Weibel, 2007).

Input control has to be applied when tasks are complex, ambiguous, or the environment is rapidly changing. In these cases output and process control linked to pay-for-performance is not appropriate. Under these circumstances, monetary incentives create goal displacement effects. Even more important is that monetary incentives tend to undermine community-oriented intrinsic motivation (Frey & Oberholzer-Gee, 1997). Money signals that the relationship to the organization is purely market driven (Burgess & Ratto, 2003) instead of developing trust in the so-called ‘public service ethos’, which is prominent in public services across different nations (Horton, 2008). Voluntary compliance with ethical standards and norms must take place, even when it is personally disadvantageous for the employee (Lindenberg & Foss, 2011).

Control Systems and Incentive Schemes: A Research Agenda

We have argued that for a large number of tasks in public service – notably, the most important tasks for the functioning of public administration – output and process control linked to pay for performance should be avoided. Further research is needed to examine theoretically and empirically which measures are likely to induce intrinsic motivation in employees and simultaneously are able to reflect the work performance adequately. We outline some alternatives that can be applied when performance measurement linked to pay-for-performance is problematic. First, positive selection and self-selection effects might be an alternative to output and process control. Second, performance measurement may be used in an exploratory way to evaluate the appropriateness of current objectives and policy assumptions and to enable the voice of the citizens. A precondition is that it is not linked to pay (Spekle & Verbeeten, 2009). Third, procedural fairness should be taken into account. Fourth, awards can supplement a fixed-pay scheme.

Positive selection effects

Under pay-for-performance schemes, individuals who pay strong attention to extrinsic rewards (e.g. money) will be selected more frequently than others (Lazear & Shaw, 2007). Empirical studies provide some evidence for this effect also in the public sector, though sometimes indirectly. For example, a large scale study on the public service by Alonso and Lewis (2001) shows that public servants who believed that pay and promotion are dependent on performance in fact achieved higher performance ratings for both high and low PSM individuals. However, in a later work Lewis and Frank (2002, p. 398) found that individuals who pay high attention to extrinsic rewards and feel attracted to public sector employment were less likely to actually work for government. This finding suggests a selection effect insofar as government salaries do not retain workers who place special emphasis on high pay.

Direct selection effects are suggested by authors like Crewson (1997), Brewer and Selden (1998) and Perry and Hondeghem (2008). These authors argue that people pursuing a public career do so because of their intrinsic motivation (e.g. to promote the public good) or to comply with professional standards. When a public sector organization successfully identifies such candidates by careful selection, socialization, and training, there is good reason to trust that they will continuously perform favourably.

While there are manifold examples of organizations that use the selection approach successfully (Cardinal, 2001; Cardinal, Sitkin, & Long, 2004; Ouchi, 1977), there is a clear need for
research with regard to personnel selection systems in public service organizations. What tools can be used to identify those candidates who are committed to the public good, who are intrinsically motivated to a sufficient extent and who fit into the organization’s characteristics? For example, Ben-Ner and Ren (2008) suggest the use of the current employees’ personal social networks as a suitable tool to identify candidates who display the desired motivational profiles. Another suggestion is to identify suitable candidates by a stronger involvement of public sector managers in graduate programs offered by universities. This helps to observe candidates during a longer period and to meet the requirements of person-organization fit (Steijn, 2008; Taylor, 2008; Wright & Pandey, 2008, 2011).

**Exploratory use of output performance measures**

Performance measures can fulfil an exploratory role, if the data is used to trigger double loop learning and experimentation (Spekle & Verbeeten, 2009). It can lead to a constructive debate about the development of new policies, strategies and future developments and to strengthening the responsiveness to the citizens who are non-experts. This approach also contributes to insights on learning in public sector organizations which, according to a recent review, is an under-researched field (Rashman, Withers, & Hartley, 2009).

Further, an exploratory use of performance measures would be preferred in an environment which cannot rely on clear-cut targets, which is the case for a large number of public service tasks. In such an environment the NPM concept fails because it cannot use performance measures in an explanatory way due to the danger of goal-displacement and gaming effects. These effects can be avoided if no monetary rewards are attached to the exploratory use of performance measures.

The demand for an explanatory use of performance measurement was indirectly shown by Townley, Cooper, and Oakes (2003). The authors investigated the introduction of performance measurement through NPM reforms in Canadian public agencies. Such reforms were initially welcomed by public sector managers because they hoped that this would trigger discussion in order to allow for more informed judgements. This situation can be interpreted as a demand for explorative use of performance indicators. However, disillusionment took over when it turned out that ‘central government agencies were interested in simple standardized information’ (Townley et al., 2003, p. 1058) and not in an explanatory function. Performance indicators were not used for ‘bringing to light the justification by which actions … are pursued’ (Townley et al., 2003, p. 1045).

More research is needed to confirm the effects of an exploratory use of performance measures. Nonetheless, the approach seems promising and is more suitable for public sector organizations than the concepts provided by New Public Management.

**Procedural fairness**

Procedural fairness is important for several reasons. First, people tend to accept unfavourable decisions when a fair and transparent procedure produces them. Many studies have found that, in contrast to the equity theory of Adams (1963), people accept a perceived unfair distribution of the cake as long as the process that has led to the distribution is perceived as fair (Tyler, 1999; Tyler & Blader, 2000). Second, it increases an individual’s propensity to contribute to the common good (Kim & Mauborgne, 1998). Third, rules contrary to an individual’s interest are usually followed when they are implemented in a fair way (Greenberg, 1994).
Key elements of procedural fairness are dignity, participation, respect, and neutrality (Greenberg, 1996). Fixed-pay schemes for authorities support the element of neutrality. This is another important reason why pay-for-performance should be avoided for such positions (Frey & Osterloh, 2005; Moynihan, 2008; Tyler, 1999; Tyler & Blader, 2000).

Procedural fairness is a universal and low-cost measure to improve performance when output or process control is not feasible. Thus, future research should investigate the links between procedural fairness and the public service motivation concept in more detail. More empirical data is needed about the criteria that are perceived as procedurally fair in cases of big differences in public service salaries.

Awards

Awards are able to send signals of recognition to the employee, avoiding the detrimental effects of motivation crowding triggered by monetary rewards. At the same time they direct the motivation of employees to goals that are appreciated by the citizens. Public sector organizations may issue awards at the individual, group, or organizational level.

Examples of individual awards are medals, titles, or orders. At the group level, awards can signal an affiliation to a special group of people. For example, soldiers distinguish themselves via ‘mission badges’ that are issued for any mission in which the individual took part. Another example is a trophy issued to a team (comparable to soccer championships). At the organizational level, innovation or quality awards, such as the Beacon Scheme in the UK (Hartley & Downe, 2007), can be presented. Public sector organizations can issue awards during special ceremonies and use awards as official certificates in order to make high-performing individuals more visible.

Awards have many features that are different from monetary incentives and provide distinct benefits (Frey, 2007; Frey & Neckermann, 2008). First, they have low (material) costs but high value to the recipient because they are highly visible in contrast to individual compensation, which often is a secret. Second, they avoid the goal displacement or multiple-task effect because they represent recognition of overall performance. Ex-ante criteria for the awards usually are not clearly specified; thus, they do not provide incentives for employees to manipulate the criteria necessary to win an award (Frey & Neckermann, 2008; Kosfeld & Neckermann, 2011). Third, awards do not crowd out intrinsic motivation for two reasons. They strengthen loyalty to the organization as long as they are consistent with fairness concerns. They are perceived as supporting and not as controlling since ex-ante criteria are not specified. Fourth, they are unlikely to generate negative effects on non-recipients because awards recognize extra-role behaviour and thus behaviour that contributes to the common good (Neckermann, Cueni, & Frey, 2010).

To summarize, awards provide positive feedback and social recognition (Frey 2007; Frey & Neckermann, 2008). These characteristics significantly contribute to intrinsic motivation (Deci & Ryan, 2000a, 2000b) and to an attention to normative goals (Lindenberg & Foss, 2011) avoiding the disadvantages of output and process control. However, research on the effects of awards is still in its infancy, in particular with respect to the public sector.

Discussion and Conclusion

Our first contribution consists in providing a systematic review of the debate concerning public service motivation and pay-for-performance. We find that the results are inconclusive, indicating that the theoretical underpinning of New Public Management is insufficient.
Our second contribution consists in contributing to a more adequate theoretical underpinning of new governance approaches in the public sector. Using Behavioural Economics, we preserve the advantages of principal agent theory which consist in its theoretical comprehensiveness and its ability to investigate individual variables and their modification systematically. At the same time we take advantage of the fact that Behavioural Economics is increasingly based on more nuanced empirical foundations. It modifies the narrow assumptions of principal-agent theory step by step, taking into account intrinsic and extrinsic motivation. On the one hand, Behavioural Economics keeps the theoretical comprehensiveness of principal-agent theory. On the other hand, there is a limited capability to take into account the rich psychological reality of public service organizations. Confronting Behavioural Economics with the insights of the public service motivation literature leads to the consequence that we have to distinguish between different tasks in public service in order to meet the requirements of person-organization fit and to evaluate the applicability of pay-for-performance schemes. To do this we introduce Management Control Theory, which helps us to clarify which kinds of control exist and under which conditions output, process or input control is applicable. As a consequence we suggest a novel and comprehensive theoretical basis to integrate the literature of public service motivation with New Public Management.

Our third contribution consists in suggestions as to how to increase public service motivation in cases when output and process control is not applicable, including selection, exploratory use of output measures, procedural justice and awards. These suggestions are in themselves explorative and need more theoretical and empirical research to make them applicable.

To summarize: ‘In recent years public sector employers have invested in a more appealing image of the public sector by improving employment conditions … but it may well be that they have missed the essential motives for employees to work in the public sector’ (Leisink, 2004, p. 11). This paper confronts the dominant principal agent view in the New Public Management discourse on the one hand and Behavioural Economics, the theory of public service motivation and Management Control Theory on the other hand in order to gain richer insights into the foundations of control and performance measurement systems of public service organizations.

Funding
This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Notes
1. We refer to Behavioural Economics for ease of understanding. However, strictly speaking, we prefer the expression Psychological Economics over Behavioural Economics for two reasons. First, economists had already examined human behaviour before this new field emerged. Second, Simon (1985) pointed out that the term ‘behavioural’ was misleading because it could be confounded with the behaviourist approach in psychology.
2. Gagne and Deci (2005) recently differentiate between controlled and autonomous motivation; they consider intrinsic motivation to be a part of autonomous motivation.
3. Locke and Latham (2009) in a rejoinder provide counterevidence to Ordonez et al. (2009). However, they disregard that output control might well work for simple but not for complex tasks within an organization.

References


**Author biographies**

Bruno S. Frey was born in Basle, Switzerland, in 1941. He was Professor of Economics at the University of Constance from 1970–77 and at the University of Zurich from 1977–2012. He has been Distinguished
Professor of Behavioural Science at the Warwick Business School at the University of Warwick, UK, since 2010 and is Professor at Zeppelin University Friedrichshafen, Germany. He is Research Director of CREMA (Center for Research in Economics, Management, and the Arts). He received honorary doctorates in economics from the universities of St. Gallen (Switzerland, 1998), Goeteborg (Sweden, 1998), the Free University of Brussels (Belgium, 2009), the University of Aix-en-Provence/Marseille (France, 2010), and the University of Innsbruck (Austria, 2011).

Fabian Homberg is Senior Lecturer at Bournemouth University. He holds a doctorate from the University of Zurich. His current research interests are motivation and incentives in private and public sector organizations, top management team diversity, and decision-making biases.

Margit Osterloh is Professor (Em.) for Business Administration and Management of Technology and Innovation at the University of Zürich, Switzerland, and full Professor of Management Science at Warwick Business School, University of Warwick. She is Research Director of CREMA (Center for Research in Economics, Management, and the Arts). Margit Osterloh holds a honorary doctorate from Leuphana University Lüneburg. She was a member of Deutscher Wissenschaftsrat (German Council for Science and Humanities) until 2011 and a board member of three companies in Switzerland and Germany. Her main research areas are organization design, corporate governance, research governance, knowledge management, and gender issues.