

Robert M.

La Follette School of Public Affairs

at the University of Wisconsin-Madison

Working Paper Series

La Follette School Working Paper No. 2017-001

<http://www.lafollette.wisc.edu/research/publications>

Making Sense of Performance Regimes: Rebalancing External Accountability and Internal Learning

Mads Leth Jakobsen

Aarhus University

mads@ps.au.dk

Martin Bækgaard

Aarhus University

Donald Moynihan

Director and Professor, La Follette School of Public Affairs, University of Wisconsin-Madison

Nina van Loon

Aarhus University

March 2017

Abstract

How to account for the ubiquitous presence of public sector performance regimes given evidence that such regimes have failed to achieve their promises? We argue that this paradox perseveres partly because the dominant doctrinal approach – justifying what we label as the external accountability regime – responds to a real need for political account giving, but also partly because alternative frameworks that can satisfy that need have been slow to emerge. We describe a fundamental mismatch between external accountability regimes and the basic characteristics of the public sector. As a result, performance regimes are often experienced as externally imposed standards that encourage passivity, gaming, and evasion, and will therefore never be able to achieve performance gains that depend on purposeful professional engagement. Rather than simply criticizing external accountability regimes, we offer an alternative framework better informed by the empirical study of performance regimes. The proposed alternative internal learning regime makes the case for extensive professional involvement in the development and interpretation of goals. This type of regime offers not just a framework to inform the design of performance regimes, but also a prospective research agenda to address how performance regimes affect motivation, behavior, and public sector outcomes.

Introduction

The turn to performance measurement practices in the governance of public services is consequential enough that scholars of public organizations have proposed the term “performance regimes” to “capture the embedded nature of these practices in almost all aspects of contemporary governance” (Moynihan et al., 2011: 141). This trend has not only outlasted New Public Management, but also appears to escalate once in place (Pollitt et al., 2010). A doctrine of performance management (Moynihan, 2008) was elevated to such a sacrosanct status that its theoretical and empirical foundations have rarely been questioned in public debates (Dubnick, 2005; Pollitt et al., 2010).

Performance regimes based on external accountability have, however, created remarkable contradictions in contemporary governance. First, while such regimes are premised on greater managerial autonomy, resistance has risen among public managers and especially frontline staff as they are disempowered and exposed to increased external control (Jakobsen & Mortensen 2016; Soss, Fording, & Schram, 2011). Second, such regimes are based on the doctrinal belief that external political accountability increases performance. This, however, contrasts with the growing evidence that external accountability itself generates large transaction costs that may not be offset by gains in efficiency (Bischoff & Blaeschke, 2016; Dubnick, 2005) and unintended effects such as goal displacement and gaming. It is far from certain that the net impact of performance management has been positive (Heinrich & Marschke 2010; Pollitt, 2013: 358). A recent meta-analysis of public organizations shows only a small positive average effect on performance (Gerrish, 2016). This paper seeks to make sense of these paradoxes while identifying an agenda for research into performance regimes.

We do so by offering a framework grounded in theory and evidence on performance regimes in complex public services. While administrative doctrines such as the performance management doctrine consists of prescriptions based on claims of cause-effect relationships (Hood & Jackson, 1991: 13), our framework consists of a set of assumptions, specified scope conditions, and empirical expectations. To arrive at this framework, we conceptualize performance regimes, divide them into subtypes, and scrutinize them in the context of three stylized facts about the context of public services:

1. *Political demand for account giving.* Public organizations must account for performance (Dubnick, 2005), and performance metrics are politically viewed as a legitimate and necessary way of doing so (Moynihan et al., 2011).
2. *Task complexity.* Task complexity (Campbell, 1988) makes both monitoring and incentivizing more difficult (Dixit, 2002) while also making capacity, professional norms, and autonomous motivation more important for performance.
3. *Motivation.* Autonomous motivation is enhanced by participation in goal setting but crowded out by externally imposed goals and measures (Weibel, Rost, & Osterloh, 2010).

These stylized facts are clearly broad simplifications but at least have the virtue of being explicit claims about the state of the world (whereas claims in administrative doctrines are often implicit, making them harder to rebut) and hint at scope conditions about how broadly our claims should be applied. We also assume that political principals view regimes primarily as mechanisms to provide quality services to the public that will be rewarded with positive election outcomes. This assumption seems straightforward enough (see e.g, Gailmard & Patty, 2013), but other theories of bureaucratic politics suggest that principals prefer regime control over performance (consistent with Moe, 1989). Even if a preference for control holds, there is still a value in illustrating the costs that such a preference raises.

We examine an external accountability (EA) regime, which is the dominant performance regime type of contemporary reforms. In this regime, externally set goals are tied to incentives (Heinrich & Marschke, 2010; Pollitt et al., 2010). We contrast the EA regime with a professional regime that provides autonomy in terms of goals and does not link goals to incentives. We offer a hybrid alternative that falls between these two regimes in the form of an internal learning (IL) regime, which combines political-bureaucratic performance goals with extensive professional involvement in the development and interpretation of goals. The currently dominant EA regime aligns with the need for account giving through performance metrics but not with task complexity or autonomous motivation, which are particularly prevalent in welfare policy areas such as health care provision and public education. On the other hand, the IL regime is better aligned with high task complexity and autonomous motivation.

By developing types of performance regimes and relating them to the needs of practical public governance, one contribution of this paper is to answer the call to public management scholars to address important questions such as how to design effective governance regimes (Kettl, 2015). We are also motivated by the conviction that it is not enough for public administration scholarship to simply critique the flaws in administrative doctrines. As long as policymakers need coherent organizing frameworks to guide choices, criticism by itself does not answer the question about what they should do instead. Our approach follows a “design science” premise (Barzelay & Thompson, 2010) that scholars ought to apply what we have learned from decades of reforms to inform the future design of public organizations. Although there is a growing and impressive literature on performance information use (Baekgaard & Serritzlew, 2016; Kroll, 2015), it only touches on broader questions of the design of performance regimes. We aim to incorporate these micro-level insights with the design of macro-level performance regimes as

understanding the motives and behavior of individuals working in the public sector is essential to designing good systems (Simon, 1947).

In the context of the important question of how to design public performance systems, our analysis suggests that the contradictions of performance regimes begin to make sense when interpreted as the result of using EA regimes in situations that are better suited to IL regimes. The mismatch between regime type and a complex public context is in our framework expected to result in disempowered and less motivated employees, weaker performance, and negative unintended effects.

While there is ample scholarly criticism of the EA regime, this has not cohered into a viable alternative. Without such an alternative, the performance management doctrine justifying the EA regime – even if badly frayed – will continue to set the terms for both research and practice (Dubnick, 2005; Pollitt et al., 2010). Several scholars (Pollitt, 2013: 346; Talbot, 2010: 121) have called for research that more deeply interrogates the underlying logics and effects of performance regimes. There is also a real-world demand for such an alternative. For example, in the US, policymakers are beginning to rethink the strong incentives of the Race to the Top and No Child Left Behind (Executive Office of the President, 2015). In Denmark, regional governments are experimenting with alternative governance regimes for hospitals, altering or abandoning external financial incentives (Søgaard et al, 2014). Even early proponents of the performance doctrine, such as the World Bank, are searching for new ways of understanding performance management (Moynihan & Beazley, 2016). Our approach is not wholly novel; indeed, it builds on previous examinations of long-standing tensions within the public debate on performance management in relation to control and incentives on the one hand and freedom and autonomy on the other (Aucoin, 1990: 126; Kettl, 1997; Mashaw 1983; Pollitt, 2013: 358).

Therefore, another contribution of our paper is to offer an alternative theoretical framework to EA regimes and the doctrine of performance management that reflects the complexity of public functions. In doing so, we are wary of making excessively broad claims characteristic of administrative doctrines (Hood & Jackson, 1991) that, at their worst, consider evidence selectively and do little to establish the scope domain of claims. However, doctrines are consequential (Barzelay, 2001; Moynihan, 2008), and scholars ought to offer theory- and evidence-based frameworks that counterbalance the less carefully developed doctrines that policymakers turn to.

The framework proposed in this paper should allow both scholars and practitioners to move beyond the ordinary knowledge and rhetoric typically driving changes in dominant administrative doctrines (Hood & Jackson, 1991: 7-11) towards a more scientifically grounded design of performance regimes. The framework builds on and combines many existing strands of research that provide concepts, hypotheses, and empirical knowledge on performance regimes. We do not provide a new theory at odds with these strands of research but instead seek to combine them into a broader theoretical framework. This framework should constitute an alternative to the performance doctrine and outlines research questions and empirical questions that can be settled based on empirical evidence.

The paper is divided into four main sections. In the first section, we present and analyze EA regimes and the theoretical and empirical critique of such regimes. Next, we discuss traditional alternatives to performance regimes, followed by the conceptualization of the IL regime. Throughout the paper, regimes are analyzed within the domain of complex public tasks. We end by discussing how our analysis provides insights into the paradoxes of performance regimes while offering avenues for future empirical research.

The EA Regime and its Challenges

Performance regimes consist of specific combinations of external actors with the opportunity to steer the performance of public organizations and the specific means by which they do so (Talbot, 2010). In our conceptualization of the EA regime, these external actors are political and bureaucratic superiors that use the means of external goal setting and incentives. More specifically, the EA regime has two defining features (Boyne, 2012, 209; Heckman, Heinrich, & Smith, 2011, 1):

1. *Goals are externally determined* by political-bureaucratic actors and operationalized in performance measures to track activities.
2. *High-powered external incentives* connect the achievement of goals to rewards and punishment, primarily through budget allocations.

Drawing – even if loosely – on principal-agent theory, the core assumption of the performance management doctrine is that the EA regime increases performance by imposing goals, collecting information on goal fulfillment and tying incentives to goal fulfilment. Hence, the EA regime encourages organizations to respond to the goals of their principals. As such, the EA regime directs attention to specific goals and allows political and bureaucratic actors to draw out hidden information. For example, in the field of health care, they allow politicians and managers to observe the opaque medical profession by showing the differences in costs of specific treatments between care providers. Consistent with these claims, some studies have shown performance improvements after implementing an EA regime with externally set goals connected to incentives (Bevan & Wilson, 2013; Kelman & Friedman, 2009).

The doctrine of performance management also implies a trade of managerial autonomy in return for external influence over performance goals and measures (Behn, 2002; Moynihan, 2008). Managers receive more autonomy to manage, while in return, the political and bureaucratic superiors hold them accountable for performance. EA regimes, however, face some important challenges in relation to the other two assumptions regarding the delivery of complex public tasks: task complexity and autonomous motivation.

Task Complexity

The complexity of a task increases with the amount of information that a rational decision-maker must take into consideration when solving the task (Campbell, 1988; Latham & Yukl, 1975; March, Simon, & Guetzkow, 1993). In complex public services such as health care provision, this includes the following:

- a. *Multiple performance dimensions* (Boyne, 2002) that generate multiple goals, such as the quality of service, accessibility, efficiency, equity, and patient satisfaction
- b. *Multiple paths to reach the goals* through different treatments and services
- c. *Tradeoffs between goals (and paths)* such as between quality and accessibility
- d. *Effect uncertainty* regarding how different choices in, for instance, health treatments will affect specific patients as the capacity and needs of patients even with similar diagnoses can differ substantially.

Many complex services are hence delivered by professionals that have had years of training and experience building (tacit) knowledge on how to handle each particular case (Freidson, 2001; Noordegraaf, 2007). This contrasts with simpler services such as garbage collection, where the goals – trash collected – and means to the collection – garbage trucks making rounds – are simpler and easier to comprehend, plan, and observe.

While the EA regime is well suited for simpler tasks, many public services have some level of task complexity that challenges the EA regime and its supportive doctrine of performance management. The overall challenge of high complexity to the EA regime relates to the key doctrinal assumption that “measures capture performance”, which states that it is possible to generate measures that form valid and comprehensive indicators of performance (Hood, 2006). With complex tasks, this is highly unlikely. Instead, many complex tasks in public service can only be measured through incomplete and imperfect measures that only cover some of the organization’s performance goals. This challenge helps to better understand the paradoxes faced by EA regimes.

First, when only some dimensions of performance are measured and rewarded, organizational actors will have an incentive to direct attention to the aspects of performance to which they are held accountable and neglect unmeasured aspects (Blau & Meyer, 1971; Dixit, 2002; Holmstrom & Milgrom, 1991). This will lead to goal displacement (Merton, 1940). Indeed, there is substantial empirical evidence of displacement towards measured versus unmeasured aspects of performance (Ordóñez et al., 2009). Kelman and Friedman (2009) show that an EA-like regime improved not only emergency waiting times but also the general quality of care. However, this improvement may have been made at a cost of other aspects (Pollitt, 2000), such as investments in training of employees or preventive care.

Second, if the measure of a particular performance goal only represents one aspect of the goal, it creates incentives for undesired selection effects (Bevan & Hood, 2006). For example, organizations may focus attention only on the subset of clients most likely to lead to reward and ignore or systematically exclude those unlikely to generate a reward. Practical examples include schools focusing only on students near the passing grade for a test while neglecting others (Figlio & Rouse, 2006), job trainers failing to provide adequate help to candidates deemed unlikely to receive

offers (Heinrich & Marschke, 2010; Koning & Heinrich, 2013; Soss, Fording, & Schram, 2011), and hospitals creaming profitable or dumping unprofitable patients (Cots et al., 2011).

Third, performance data is likely to be subject to gaming and manipulation by the agents that it is intended to hold accountable. In the area of education policy, for example, the link between performance measures and rewards has incentivized cheating by teachers on behalf of their students (Jacob & Levitt, 2003). A particular form of gaming is to manipulate the measurement system via strategic reclassifications. Schools might reclassify a higher proportion of students as disabled in order to exclude them from testing evaluations (Figlio & Getzler, 2006). Health providers may systematically reclassify patients as belonging to a category of services that provide higher reimbursement rates (Kerpershoek, Groenleer, & de Bruijn, 2016). Gaming and manipulation can occur when agents have some control over the collection of performance data or where agents can manipulate work processes to give a misleading impression of success. Manipulation and gaming is hence a particular challenge for complex tasks. For example, since doctors register their own diagnoses and treatment outcomes, there is a high reliance on their capability and willingness to provide valid information. When agents do not identify with a regime – which is likely with externally imposed regimes – there is a larger risk that employees will not implement it well (Tummers & Van de Walle, 2012).

The corruption of data also becomes more likely when the incentives to perform are strong. Campbell's Law, articulated by Donald Campbell (1979: 85), notes how measures collected for social science research corrupts when used for allocation of resources: "The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor." Hence, measures connected to external incentives are more likely to be subject to gaming and manipulation. Some incentives may have more severe impacts than others.

For example, Bevan and Wilson (2013) show that naming and shaming may be beneficial for performance. Although this is an external incentive, it is a milder form than financial incentives that directly threaten the continuation of the service provision. Whereas financial incentives speak to the extrinsic motivations, naming and shaming incentivizes the internal status and ego motives through introjection (Deci & Ryan, 2000).

Fourth, even when no goal displacement, selection effects, and gaming occurs, external performance-based systems often fail to match the dynamic nature of complex services (Heinrich & Marschke, 2010). As it is impossible to capture all possibilities in rules and guidelines, the EA regime can lag behind in reality.

The above challenges are not particular for public sector organizations but are a general problem for organizations facing complex tasks. As Lazear and Gibbs point out in their introduction to personnel economics (2015), the use of strong incentives becomes problematic when tasks are interdependent, there are multiple goals and tasks, agents have little control of the outcome, and there are ample opportunities for gaming. A recent example of misaligned EA regimes in the private sector was the failings of the Wells Fargo scandal. Here, employees met highly incentivized performance goals by providing unrequested services for which they charged unknowing customers. The massive scale and routine nature of the fraud pointed to a regime failure rather than the misdeeds of a handful of “bad apples” (*The New Yorker*, September 12, 2016).

Employee Motivation

The performance management doctrine underlying the EA regime makes assumptions regarding employee motivation. Motivation is a combination of energy and direction, and self-determination theory points to the importance of whether motivation is determined autonomously and arises from within (Deci & Ryan, 2000). The more controlled or extrinsic the motivation, the more factors

external to the person, such as salary and status, drive behavior. Research on public service providers has shown that autonomous motives play an important role for employees. In particular, intrinsic motivation – emanating from enjoyment in doing the task in itself – and public service motivation – identifying with and feeling an internal obligation to contribute to society – seem to drive employee behavior, and especially for highly professionalized employees working on public tasks (Moynihan, 2010; Perry, Mesh, & Paarlberg, 2006; Perry & Wise, 1990).

Historically, complex services such as education and health care have benefitted from service providers deriving motivation from the very act of performing the task or from status and honor, professional norms, and the altruistic benefits of serving others. Those with a higher public service ethos are also more likely to enter into public work, less likely to exit, and more committed when in public employment (Perry, Mesch, & Paarlberg, 2006). By contrast, EA regimes are built on the doctrinal belief that agents have a controlled motivation, which means that they can achieve externally imposed goals via extrinsic rewards and sanctions.

The mismatch between the assumptions of EA regimes and evidence on motivation in public settings creates a number of problems. First, several aspects of EA regimes may be a poor fit for highly autonomously motivated individuals. The EA regime is likely to foster goal ambiguity between individual preferences and externally imposed goals (Pandey & Wright, 2006), especially if productivity is prioritized over other goals. This has been observed in different public settings but primarily in health, education, and welfare (Brodkin, 2011; Dias & Maynard-Moody, 2007; Heckman, Heinrich, & Smith, 2011; Langbein, 2010; Ryan & Weinstein, 2009; Soss, Fording, & Schram, 2011; Weibel, Rost, & Osterloh, 2010).

Second, when the EA regime applies extrinsic incentives to push performance, it creates risks of “crowding out” intrinsic motivation (Deci, Koestner, & Ryan, 1999; Gneezy &

Rustichini, 2000) and, in some cases, also public service motivation (Jacobsen, Hvitved, & Andersen, 2014). A classic example of this mechanism is teachers who feel pressured to pursue an extrinsically imposed goal of improving test scores at the expense of the goal of improving more general student outcomes (Ryan & Weinstein, 2009). The teachers become more cynical and less autonomously motivated in their jobs, scaling back unmeasured effort and focusing only on teaching to the test. Crowding out processes may also occur in the selection of public employees: More intrinsically motivated individuals may exit EA regimes or choose not to join them in the first place (Moynihan, 2010).

Third, the use of an EA regime for complex tasks will increase the magnitude of the negative impact of the regime on autonomous motivation. With simpler tasks where measures closely mirror tasks, crowding out is less of a concern since rewards and sanctions attached to measures should still motivate greater effort (the so called “price” effect). However, when complexity makes it difficult to truly observe agent effort, there will be a performance loss on unobserved aspects of performance to the point that the crowding out effect outweighs the price effect (Weibel, Rost, & Osterloh, 2010). Further, as intrinsic and altruistic motivations are eroded, so too are normative safeguards against gaming and manipulation (Moynihan, 2010).

Articulating and Alternative Performance Regime

The popularity of EA regimes depends partly on its responsiveness to our first stylized fact – the need for account giving – even though it is ill-suited to satisfy the second and third stylized facts – the task complexity and motivational factors present in the public sector. The empirical and theoretical evidence presented above aligns with this simple observation.

The problems with the EA regime in public settings with complex tasks propel a search for alternatives. One is to establish more dynamic EA regimes based on the assumption that the principal can modify and adjust the mixture of incentives and monitoring efforts to respond to problems arising from complexity (Cots et al., 2011; Heinrich & Marschke, 2010). For example, Courty and Marschke (2003, 2007) describe the adaptations of federal principals overseeing manipulation of job training programs by state governments.

The “dynamic principal” perspective recognizes the problems created by the EA regime but responds to these problems by developing better measures instead of addressing the challenges of complexity and motivation. Put another way, this approach assumes that principals and agents have an adversarial relationship and that accountability based on external goals and incentives are necessary to manage agents. It assumes that it is possible for principals to make more precise measures of performance than today and to recognize and remove gaming opportunities. With enough diligence on the part of principals, performance regimes can be improved and solve current problems with gaming (Heinrich & Marschke, 2010). However, such an approach still assumes extrinsically motivated actors and that it is possible to solve the problems with external measurement of highly complex services. Moreover, it assumes a limit on the ingenuity of the agent, even though the agent often possesses as much information (or more) as the principal that can be used to their advantage. A highly dynamic regime that adds and modifies goals loses value in prioritizing goals and directing behavior, which is one of the proposed gains of the EA regime. More dynamic performance principals would also require an investment in public sector administrative capacity and resources, which have so far been lacking in the application of the EA regime (Soss, Fording, & Schram, 2011), and a determined commitment on the part of the principal to invest their time in management processes.

While the dynamic principal model offers one possibility for dealing with the problems of EA regimes, we ask whether there is an alternative that better accords with our stylized facts about task complexity and motivation. In the next section, we consider in detail the possible alternatives among traditional models of governance regimes.

The Building Blocks of an Alternative

On the theoretical level, we treat performance regimes as ideal types. Ideal types are theoretical constructs that describe how specific actors, institutions, and behaviors relate to each other in an internally consistent way based on a given rationale (Stout, 2010). The EA regime and the doctrine of performance management make up an ideal type as they prescribe a specific institutional setting featuring motivational and informational assumptions that predict a positive performance effect.

One possible alternative would be a professional regime. The core idea behind the professional regime is that organizations solving highly complex tasks are best governed through the judgment and expertise of professionals held together by common norms and values (Freidson, 2001; Mintzberg, 1996; Ouchi, 1980). Professional norms and the influence of professions are hence expected to shape work practices (Exworthy & Halford, 1999; Mashaw 1983; Noordegraaf, 2007). Professionals have preferences about what good service is (Evetts, 2003; Roberts & Dietrich, 1999) as well as professional norms about appropriate behavior and service quality that counterbalance a myopic focus on economic considerations (Freidson, 2001: 127).

Our focus on the professional regime differs from the traditional contrast in the PA literature between EA and the bureaucratic regime, where governance is based on input measures such as rules, orders, and budgets (Hood & Jackson, 1991; Pollitt & Bouckaert, 2011; Weber, 1960). The bureaucratic regime is not a plausible alternative to the EA regime as it does not meet

contemporary political demand for account giving using metrics (Moynihan & Soss, 2014). Moreover, because of its reliance on rules and regulations as a steering mechanism, it is not aligned with highly complex tasks requiring case-specific judgment, knowledge, and discretion as rules are unable to capture fluid situations. Put another way, it is not well aligned with goal-based learning (Moynihan, 2005). To conceptualize a hybrid regime, we therefore draw on the EA and professional regimes as ideal types to identify and build on their rationales.

Table 1 identifies the important differences between the EA regime and the professional regime. In terms of the *substantive autonomy* to determine performance goals and measures, the professional regime allows for internally determined goals and measures tied to professional norms, whereas goals and measures are externally determined in the EA regime. The professional regime hence gives professionals a high degree of autonomy in determining how to do their work. The EA regime does imply some managerial autonomy, but in reality, the loss of substantive autonomy highly influences work practices as well. For instance, teachers have to aim for higher test scores and thus have little substantive autonomy in deciding on their outcome measure, but this goal also leads them to focus their teaching on the aspects that the test measures.

In relation to *incentives and accountability*, the professional regime has no strong external incentives. This clearly contrasts with the strong externally set incentives of the EA regime that give principals considerable influence over an organization and make its professionals accountable to them (Pollitt et al., 2010). Examples are lists of good and bad schools or publicly displaying the mortality rates of hospitals. Instead, accountability in the professional regime is achieved through internal fora such as professional associations and peer reviews, with little control from external agencies (Romzek & Dubnick, 1987). Professionals in such a regime are, however, indirectly accountable to political, societal, and client interests as the autonomy of the professionals is based on an implicit contract according to which they promise to uphold those interests in

exchange for self-determination (Noordegraaf, 2007). Still, the professional regime is often critiqued for its lack of direct public accountability (Noordegraaf, 2007).

Consequently, the *rationales or doctrinal beliefs* for how organizations improve outcomes within the two regimes also differ. The EA regime is based on the doctrinal assumption that improvement occurs when external political-bureaucratic actors set the right incentives for organizations. The organizations will then increase their effort and seek to adjust their practices to achieve a higher level of goal achievement. Organizational learning – understood as the process of “encoding inferences from history into routines that guide behavior” (Levitt & March, 1988, p. 320; March, 2010) – will have a single-loop character where inferences about practices but not underlying assumptions about the goals pursued are taken into consideration (Argyris, 1980). Nevertheless, if employees fail to identify with externally imposed goals, securing the motivation needed to facilitate genuine learning becomes less likely. It is, instead, more likely that the organization will passively collect information for account giving (Moynihan, 2005). Hence, controlled motivation drives the learning process.

In the professional regime, improvement originates from learning through professional norms and ideals, and improvements in goal fulfilment are likely to be based on the acquisition of new professional expertise (Freidson, 2001). Mintzberg (1996) proposes that for client-based complex services, learning depends on the professionalism and dedication of professional staff. The organizational professionals are, to a great extent, autonomously motivated, and they respond to incentives aimed at their norms and status. Professional values and motives therefore also play an important role in the internal learning regime.

Table 1: Types of regimes

	EA regime	IL regime	Professional regime
Performance regime	Yes	Yes	No
Goal autonomy (substantive)	Low. External, tied to political-bureaucratic goals	Medium. Lower-level goals and performance measures set by professionals	High. Goals are set by professionals, tied to professional values
Incentives	External, high-powered	Internal, aimed at norms, values, and peer status	Internal, low-powered
Accountability	External with performance focus	Internal accountability embedded within an external regulative bargain and with regular dialogue between principals and agents	Internal, tied to professional norms (with a regulative bargain)
Rationale for improvements	Improvements through external control and incentives Controlled motivation; learning about goal achievement but not goals (single-loop learning)	Improvement through learning spurred by autonomous motivation Learning on (lower-level) goals and goal achievement infused with values and experiences of experts (single- and double-loop learning)	Improvement through new professional expertise Autonomous motivation; learning on goals and goal achievement through professional norms and goals (single- and double-loop learning)

The Internal Learning Regime

Neither the EA nor the professional regime fully addresses all three stylized facts about complex public tasks, and nor does either a classic bureaucratic or a dynamic performance regime. What sort of regime, then, can accommodate the three stylized facts? To arrive at a potential alternative ideal type to the EA regime ideal type, we return to the stylized fact that the political environment does not allow public services to ignore the demand for performance-based account giving. Such accounts are usually legally required, and many professions and services have endorsed performance measures as a governing principle themselves, even if objecting to specific metrics or practices (Bouckaert & Halligan, 2007; Moynihan, 2008; Talbot, 2010: 1). Hence, we are not suggesting that the EA regime should be fully abandoned. Such a suggestion is not politically feasible, and it ignores the potential for performance measures to help public organizations.

Is a hybrid that contains some features of the EA regime but also elements of the professional regime able to overcome the problems of the two ideal type regimes? This section sketches what such a hybrid regime of EA and professional regimes would look like. We do not claim that such a hybrid regime is new from an empirical point of view. Indeed, the actual practice of performance regimes is unlikely to ever fit the ideal types that we sketch out but, instead, actively blends these approaches.¹ However, there is not a strong theoretical logic for how these hybrid forms function in the same way that there is a well-developed logic for EA regimes. As a result, the application of a hybrid regime may be misconstrued as a failure to fully implement an EA regime rather than as a managerial approach that is still goal-oriented in purpose.

¹ A variety of administrative processes could be categorized to fall into this conceptualization. For example, Mashaw (1983) portrays administrative law judges juggling adjudication metrics with some professionally-based notion of bureaucratic justice and devising learning mechanisms to bridge the two.

In practice, legal limits and political opposition may dampen the use of extrinsic incentives, or political principals may recognize the value of professional input in goal selection or the diagnosis of problems. An example of such a hybrid would be from the field of job training, in which the US Department of Labor moved away from an EA regime to focus on a new approach for goal setting that was explicitly intended to give state and local partners more control in setting goals and selecting measures and, by doing so, align beliefs and generate a sense of shared accountability (Heinrich, 2007).

Another example is seen in the application of “stat” techniques to facilitate organizational learning in public organizations, which is popular in the United States and was adopted by the GPRA Modernization Act of 2010 (Moynihan & Kroll, 2016). Here, professional input is engaged less in the setting of goals and more in the interpretation of measures to identify program problems and solutions. The approach reflects the limit of political principals – they need professionals to make sense of data – and their leverage – they can exert accountability through social routines that establish norms of accountability rather than financial incentives. The “stat” tool also evolved as it migrated across professions. It originated in policing, where a paramilitary culture allowed political principals to aggressively demand improvement in ways that bordered on humiliation – such as yelling at police officers in front of their peers – and could not succeed in professions that lacked deep hierarchical norms of obedience (Behn, 2014; Hatry & Davies, 2011). Even within policing, the unrelenting demands for continuous improvement over time was charged with eroding professional norms to the point that other policing goals were damaged, and officers made their targets by hassling citizens or by manipulating performance metrics (Eterno & Silverman, 2010). By contrast, within the US federal government, such routines were associated with a more purposeful use of performance data when they featured positive reinforcement (Moynihan & Kroll, 2016).

A non-US example comes from the effort by the Central Denmark Hospital Region to move away from an EA regime towards a regime in which hospitals wards had some financial incentives removed while they were encouraged to develop their own performance indicators (Søgaard, Kristensen, & Bech, 2015; Sørensen & Burau, 2015). The regional principals perceived the previous model to produce unintended consequences in the form of gaming and selection while hampering quality development focused on the creation of value for the patient. Indeed, this example motivated our current effort as political principals sought scholarly insight to do something different but lacked a theoretical model as to how to make such changes. We return to this example in the discussion of a research agenda.

While hybrid regimes exist in practice, and while policymakers may wish to establish such regimes, we see a lack of theory and language to describe, identify, and understand how such regimes work, reflecting a failure to conceptualize a regime that combines professional and performance features. While policymakers may already be pursuing a hybrid regime in practice, they may be doing so inadvertently, working at the intersection of countervailing political and professional pressures rather than consciously considering how to combine types of regimes. Similarly, while scholars have forwarded impressive critiques of EA regimes, they lack a conceptual language for understanding and exploring a hybrid regime. Describing the internal learning regime provides a theoretical framework that – to an admittedly limited degree – fulfills this need.

The above examples also show the tension between externally imposed goals and professional autonomy. This is a consistent theme across the critiques of the EA regime. Sociologists argue that organizations are compelled to respond perversely to maintain external legitimacy (Sauder & Espeland, 2009); psychologists suggest that externally imposed goals are harmful because they reduce autonomy and self-determination (Frey, 2000; Frey & Osterloh, 2010);

economists suggest that such goals allow a self-interested actor to engage in moral hazard to maximize a financial gain (Dixit, 2002); and political scientists note that such systems undercut traditions of loyalty between principals and agents (Hood, 2011).²

If one accepts the external locus to be the central problem, the obvious alternative becomes to give organizational actors more autonomy in determining the goals and measures of performance. This would also unleash the benefits of professionalism by removing what professionals perceive to be illegitimate constraining standards that weaken morale and encourage gaming (Kerpershoek, Groenleer, & de Bruijn, 2014). The regime would be applicable for complex tasks requiring professionalism while reducing the risk of crowding out autonomous motivation. We call such a hybrid regime an internal learning (IL) regime. Table 1 above presents the core elements of the IL regime.

The IL regime gives organizations and their professionals more *goal autonomy* in selecting goals and measures as a matter of deliberate design, thereby giving up complete political-bureaucratic control over goals. The above examples all, to varying degrees, feature political principals recognizing that professional expertise is essential if outcomes are to be achieved. The most careful critiques of principal-agent theories in public settings make a similar claim, arguing that such autonomy is crucial if employees are to be motivated (Brehm & Gates, 1999; Gailmard & Patty, 2007). Further, this expertise cannot be engaged unless professional actors play a role in establishing and interpreting performance measures, which are de facto provisions of autonomy and influence to professionals that will motivate greater effort.

² In contrast to both Talbot's (2010) and Pollitt's (2013) broader performance regime and management perspectives with multiple actors and interventions, we make the degree of externally imposed goals and incentives the key dimension on which we distinguish between different performance regimes. Another difference to prior work that has critiqued what we label the EA regime (e.g., Aucoin 1990), is that we explicitly apply a learning perspective, though note that Mashaw (1983) focused on elements of learning in his conceptualization of "internal administrative law" for disability claims.

The choice of measures should reflect democratic concerns as well as appropriate professional norms, treatments, and techniques. The more specific and the closer to professional practices goals are, the more the professionals, and not outside political-bureaucratic actors, would be the ones to design the measures. Furthermore, the regime promotes learning and capacity building based on performance information because the regime provides the professionals with discretion to pursue some of their own policy preferences, making it worthwhile for them to learn and invest in expertise (Gailmard & Patty, 2013). Using a field experiment, Andersen and Moynihan (2016) offer evidence for this claim, showing that school principals who were granted more discretion in hiring invested more effort to acquire expertise in the form of performance data and that these investments were even stronger when the performance information reflected goals that were consistent with the school principal's autonomous motivation. Such arguments for discretion also underlie the management approach of many knowledge intensive private organizations such as Google's 70-20-10 rule, according to which employees are to spend 20 percent of their working time on ideas and projects initiated without external instruction (Garvin, 2013; Steiber, 2014).

The IL regime also removes a tight link between high-powered *incentives* and performance measures. In this regime, emphasis is on performance information as a learning tool and less on performance information used to create external incentives to motivate (Greve, 2003, 2008). Weakening external incentives reduce the inclination to game in order to satisfy external goals. Goal-oriented learning and innovation will become more likely when organizational actors perceive that they work in a setting where acknowledging errors and problems will lead to a dialogue about problem solving rather than punishment.

Expanding goal autonomy also creates room for double-loop learning (Argyris, 1980; Moynihan, 2008). It is assumed that individuals have a mixed set of motives and are, to some

degree, autonomously motivated to perform and learn (Deci, Kostner, & Ryan, 1999). Thus, individuals want to learn but need the right structures and support to do so. Thus, when professionals identify more with organizational goals because they can play a role in selecting them, internal learning from performance data that captures those goals is likely to take place. Positive patterns of performance-based learning tend to emerge when organizational routines to learn are established, accepted, and incorporated as professional expertise (Moynihan & Kroll, 2016). The examples cited above feature dialogue-centered social routines as a primary basis of engaging professionals, sometimes with and sometimes without incentives.

The metaphor of dialogue recognizes not just that professional expertise is essential to learning, but that such learning arises from an iterative engagement between political principals pursuing democratic goals and professional actors providing expertise. Moynihan (2006) identifies the interdependence of dialogue, learning, and professional identity by contrasting three corrections departments implementing performance requirements. In one, the performance regime was seen as externally imposed by political principals, setting unrealistic performance expectations at odds with limited resources, resulting in a passive professional response. In another, the performance regime generated single-loop learning under conditions in which political principals supported goals also valued by corrections professionals. In the third corrections agency, double loop occurred where corrections professionals engaged in a multi-year dialogue about the purpose of a corrections agency. This learning was spurred internally by a clash between professionals with different backgrounds – some with a social science training that challenged those coming from a corrections background. The mix of professional beliefs generated conflict but eventually established a new set of organizational assumptions, which the department put into practice only through a long external dialogue with political stakeholders about the wisdom of the new approach.

The case also illustrates a deeper point about the nature of accountability for complex tasks in performance regimes. Such accountability cannot realistically be expected to arise from political principals selecting and monitoring data that they struggle to process. Instead, it depends on a dialogue between principals and the professionals who are best placed to interpret and make sense of ambiguous performance information, and such a dialogue will be most fruitful if it is collaborative. This claim is evident also in the importance that the field of evaluation places on professional knowledge in understanding and interpreting outcomes (Funnell & Rogers, 2011, Mayne, 2004). It is also present in aspects of organizational learning that emphasize the ambiguity of organizational settings (March, 2010).

Building on elements of autonomy and the removal of external incentives from the professional regime raises the issue of *performance-based accountability* in the IL regime. Compared to the EA regime, less direct control is given to principals. One implication is that different organizations may become less comparable when there is no control from principals to harmonize their measures. Our hybrid model still assumes that performance data exists and is publicly reported. Still, it would be naïve to assume constant alignment of the goals of employees and external stakeholders. An obvious point of conflict, for example, is the trade-off between cost containment, quality, and rent-seeking on the part of the profession. There is thus valid concern regarding the accountability of public service in an IL regime. However, the reduced political control under an IL regime does not necessarily mean an overall loss of accountability, given the potential tradeoff between principal control and goal achievement (Gailmard & Patty, 2007).

Professional norms and values also offer a form of accountability since they are established to demonstrate the legitimacy and efficacy of a profession to the broader public and should thus share common ground with democratic goals to provide public value. Indeed, performance-based accountability should be at least partially ensured by embedding goals within an

informal (or formal) regulative bargain (Kerpershoek, Groenleer, & de Bruijn, 2016) between the professions and the state and society. The professions are granted autonomy in exchange for regulating their own members in terms of professional conduct as well as broader, set societal goals such as cost containment (Suddaby, Cooper, & Greenwood, 2007). Such bargains are struck within a political-bureaucratic structure that allows political principals to fundamentally change the regulative bargain should the professions not live up to their end of it. Any use of substantive autonomy by the professions would thus have to take this shadow of hierarchy into account (Scharpf, 1994). Regulative bargains provide a formidable constraint that prevents the adoption of pure EA regimes, especially when professions have strong political standings. Our model adds to the existing consideration of regulative bargains by proposing that they can achieve better outcomes through ongoing routines of goal-based dialogue between political principals and professionals rather than through abrupt disruptions.

A New Empirical Research Agenda

We have outlined a framework with a potential alternative performance regime better aligned with the conditions of complex public services. Making a design science perspective compatible with social science means not just sketching an alternative regime, but also outlining a research agenda that can determine how the new framework is working and whether other alternative designs would work better. In this final section, we outline how our discussion of performance regimes motivates a new empirical research agenda that can make sense of contradictions in contemporary governance.

A key element in the public administration research agenda on performance management research should be to examine the empirical consequences of moving from an EA to an IL regime or from a professional to an IL regime . In the following, we focus on organizations

with an EA regime as a starting point because recent decades have seen public organizations move to this regime type, and as cited in our introduction, there appears to be some appetite among policymakers to move in the other direction. In situations in which there has been no such movement to EA regimes, our proposed approach may not be valid, but instead, an appropriate balance between the external accountability and internal learning regime implied by Figure 1 should be pursued.

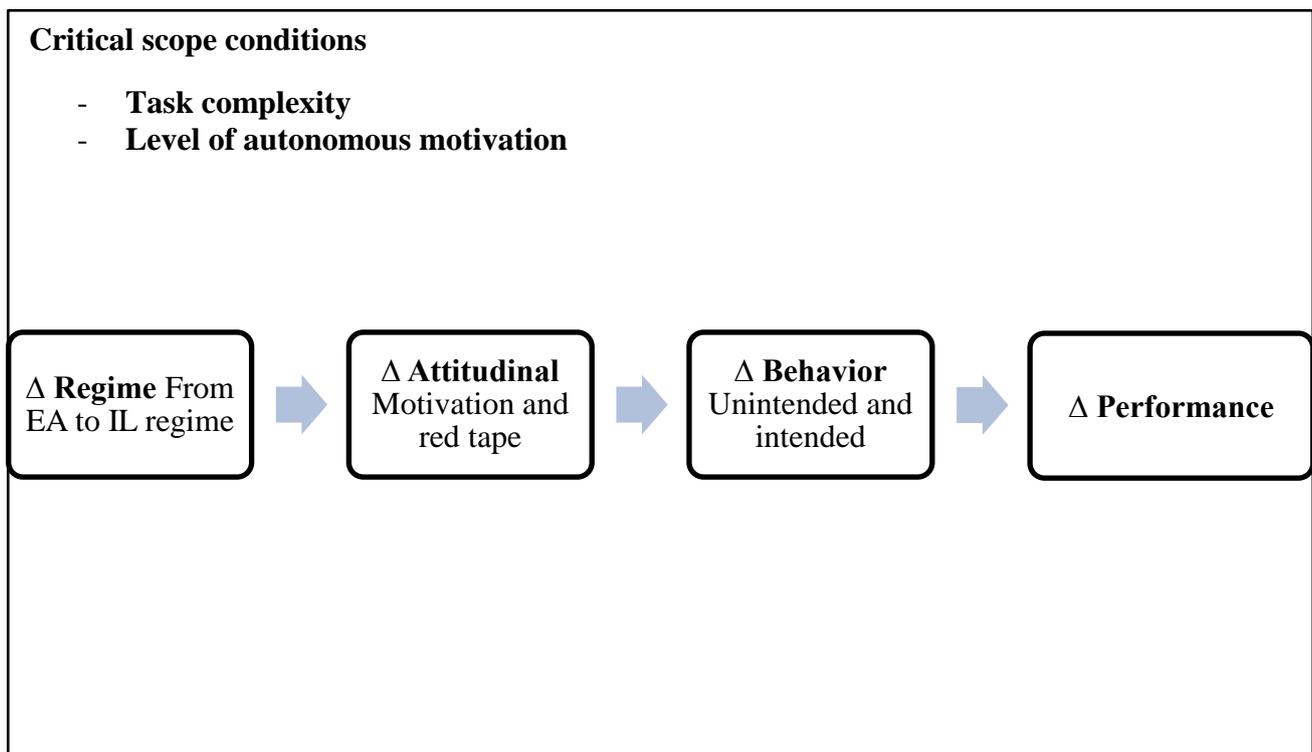
Even though our goal is a broad reconsideration of the design of performance regimes, such empirical testing requires that our framework offers clear and falsifiable expectations. We therefore offer several expectations about the effects of going from an EA to an IL regime through the removal of external incentives and increased goal autonomy. These expectations are based on the argument that the drawbacks of the pure EA regime and the professional regime can be addressed by an IL regime that gives agents some substantive autonomy in determining what performance is and how it should be measured, that lies within a regulatory bargain, and that removes any tight links between high-powered incentives and performance measures.

Figure 1 illustrates the presumed attitudinal, behavioral, and performance-related consequences of moving from an EA to an IL regime. However, not one particular performance regime is best in all contexts. How regime changes affect performance depends on two critical scope conditions. Following the argument of this paper, when there is high task complexity and high levels of autonomous motivation, moving from an EA to an IL regime is expected to have a stronger and more positive impact than in cases with little task complexity and low levels of autonomous motivation. Comparing performance regimes in different contexts can test these critical assumptions. For functions that are simple and easy to measure, where it is easy to assign responsibility and minimize the need for discretion, and where the context is stable, something close to ideal-type EA regimes would work the best. Moreover, whether performance regimes influence

behavior also depends on the implementation of the regime as the effects may be limited if organizations and employees resist the implementation (Tummers & Van de Walle, 2012).

Although a shift in regime is likely to affect several attitudes, behaviors, and outcomes, we focus here on key factors that a regime change is likely to influence.

Figure 1. Model of the Impact of Performance Regime Change



In relation to *employee attitudes*, we expect that internal regimes will increase autonomous motivation and reduce perceptions of red tape. As discussed above, studies have shown that aspects of EA regimes are likely to crowd out intrinsic and altruistic motivation (Gagné et al., 2010; Wright & Pandey, 2010). Reducing the external locus of control associated with EA regimes should reverse the negative effects on employee attitudes. The rise in substantive autonomy – having the agent

formulate goals and set targets that reflect professional norms – is likely to increase agent identification with the goals (Deci, Koestner, & Ryan, 1999), which can be expected to improve autonomous motivation. Finally, the considerable transaction costs that arise with an externally imposed performance regime may be experienced as red tape that employees have to deal with. By contrast, including employees in deciding on the indicators and data collection is likely to decrease these perceptions even if this does not reduce the actual work-load, since employees will better appreciate the instrumental values the performance regime serves (Lavertu, Lewis, & Moynihan, 2013; Van Loon et al. 2016). While we focus on employee attitudes that are hypothesized as connecting with performance, exploring actual transaction costs is also worth pursuing.

Proposition 1: A shift from an external accountability regime to an internal learning performance regime will increase autonomous motivation among employees and decrease perceptions of red tape.

We also expect changes in *intended and unintended output behaviors*. A move towards an IL regime will remove perverse incentives to goal displacement, gaming, and other forms of unintended behaviors (Kerpershoek et al., 2014). Furthermore, the shift towards the IL regime removes the disincentive for cooperation inherent in the EA regime, in which rewards and punishments typically relate to the goals of individual organizations irrespective of whether there are larger collective goals requiring cooperation between the organizations. For example, Soss, Fording, and Schram (2011) note that providers of job training services under EA failed to coordinate with one another as incentives encouraged competition rather than cooperation. Under the IL regime, actors are not rewarded for specific outputs provided and therefore do not perceive time spent in routines of dialogue as distractions from the work they are rewarded for.

Another example from the field of health care is Gawande's (2009) attempt to explain variation in costs for health services in different communities. He pointed out that hospitals that employed a mixture of fixed salary (thereby limiting extrinsic incentives) and appeal to professional values through a patient-centered culture were associated with more meetings in which specialists collectively decided on a treatment plan for the patient, which resulted in both better outcomes and lower costs than average. Here, professionals saw a team-based dialogue as a means to better pursue their professional values. By contrast, the most expensive health care communities featured strong extrinsic incentives. Hospitals and doctors were financially rewarded for each extra output, and doctors viewed meetings with other specialists as a loss of resources. Despite the higher costs, the quality of health was poorer as doctors substituted more services for an integrated treatment plan.

A regime change should therefore increase learning, especially double-loop learning through autonomous motivation and allowing professionals to pursue professional values. Autonomous motivation and autonomy have demonstrated a positive relationship with innovation and learning (Hammond et al., 2011), and these factors are expected to improve as a result of the shift towards internal learning by providing the employees with influence over the goals and increasing their sense of responsibility for attaining the goals.

Proposition 2: A shift from an external accountability regime to an internal learning performance regime will decrease gaming behavior and increase cooperation and learning.

Finally, a regime change should also affect *organizational performance*. Our theoretical framework proposes that as EA regimes are relaxed and learning values enhanced, improved behavioral outcomes should result in improved performance outcomes. However, our theory also suggests that EA regimes increase attention to the performance indicators tied to incentives. Overall, then, our theory proposes that changing regimes in the direction of internal learning accountability should

lead to improvements in some aspects of performance but may see declines in performance metrics previously tied to financial rewards.

Proposition 3: A shift from an external accountability regime to an internal learning performance regime will increase organizational performance on some aspects but may decrease performance on the previously incentivized performance measures.

Conclusion: Making Sense of Contradictions

Our approach emphasizes both broad considerations about the design of public services and social science evidence to inform those design choices. By illustrating what an IL regime would entail and by conceiving testable expectations of moving from an EA regime to an IL regime, we have provided an alternative framework to the currently dominant EA regime based on the doctrine of performance management.

This approach also allows us to make better sense of some of the contradictions of contemporary governance. One contradiction is the *discrepancy between expected and experienced autonomy*. The EA regime promises greater organizational autonomy, especially for managers. Yet, in many organizations, opposition to the EA regime is high among both public managers and their frontline employees as they feel constrained and disempowered (Moynihan & Soss, 2014). This contradiction makes much more sense when considering that the implementation of the EA regime often takes place in organizations with complex tasks and high levels of autonomous motivation among their professional staff. The external accountability rationale can conflict with internal professional and service norms and the motivation of employees. Moreover, the posited increase in managerial autonomy is ultimately not delivered. The power to determine the substance of the goals and measures to a large degree determines what happens within the organization.

Second, widespread belief in the benefits of the EA regime has driven its introduction in many public services, whereas there is *no conclusive evidence of a positive impact on performance*. This reflects the broader assessment of Hood and Dixon (2015) that the New Public Management movement of the last 30 years has not made the British central government either cheaper or better. Political principals may favor EA regimes because they imply more control and power over public service production and because of the strong rhetorical appeal of the performance management doctrine. But our model proposes that control may come at the cost of actual improvement in policy outcomes that political principals also care about (Gailmard & Patty, 2007). If the expectations developed above about the impact of changing from an EA to an IL regime are valid, the performance effects – except on explicitly measured and incentivized performance measures – would, in complex services, be stronger with an IL regime.

The contradictions in current performance regimes call for developing alternative modes of governance. The internal learning regime provides one such alternative that addresses the challenges with making performance management meaningful in complex public services. On that basis, we propose a clear prescriptive lesson for practice. For complex public services, performance regimes should be centered on learning. In such cases, governments should increase professional autonomy and be wary of high-powered incentives. This does not mean that external accountability is abandoned but that aspects of external accountability can threaten professional learning and should therefore be tempered to allow professionals to engage in a goal-based dialogue with political principals. The alternative – to continue to move to systems of external accountability – can be defended as a political prerogative but not as a model of governing consistent with evidence. The internal learning regime that we propose may turn out to be no more successful, and need to be modified, but it is time to systematically model and evaluate a different framework for managing public sector performance.

References

- Andersen, S. C. & Moynihan, D. (2016). Bureaucratic Investments in Expertise: Evidence from a Randomized Controlled Field Trial. *The Journal of Politics*, 78(4), 1032–1044.
- Argyris, C. (1980). Making the Undiscussable and Its Undiscussability Discussable. *Public Administration Review*, 40(3), 205–213.
- Aucoin, P. (1990). Administrative Reform in Public Management: Paradigms, Principles, Paradoxes and Pendulums. *Governance*, 3(2), 115–137.
- Baekgaard, M., & Serritzlew, S. (2016). Interpreting Performance Information: Motivated Reasoning or Unbiased Comprehension. *Public Administration Review*, 76(1), 73–82.
- Barzelay, M. (2001). *The New Public Management: Improving Research and Policy Dialogue*. Berkeley: University of California Press.
- Barzelay, M. & Thompson, F. (2010). Back to the Future: Making Public Administration a Design Science. *Public Administration Review*, 70(1), 295–297.
- Behn, R. D. (2002). The Psychological Barriers to Performance Management: Or Why Isn't Everyone Jumping on the Performance-Management Bandwagon? *Public Performance & Management Review*, 26(1), 5–25.
- Behn, R. D. (2014). *The PerformanceStat Potential: A Leadership Strategy for Producing Results*. Washington, DC: Brookings Institution Press.
- Bevan, G. & Hood, C. (2006). What's Measured is What Matters: Targets and Gaming in the English Public Health Care System. *Public Administration*, 84(3), 517–538.

- Bevan, G., & Wilson, D. (2013). Does “naming and shaming” work for schools and hospitals? Lessons from natural experiments following devolution in England and Wales. *Public Money & Management*, 33(4), 245–252.
- Bischoff, I. & Blaeschke, F. (2016). Performance Budgeting: Incentives and Social Waste from Window-Dressing. *Journal of Public Administration Research and Theory*, 26(2), 344–358.
- Blau, P. M., & Meyer, M. W. (1971). *Bureaucracy in modern society* (2. ed.). New York: Random House.
- Bouckaert, G. & Halligan, J. (2007). *Managing performance: International comparisons*. Routledge.
- Boyne, G. A. (2002). Theme: Local Government: Concepts and Indicators of Local Authority Performance: An Evaluation of the Statutory Frameworks in England and Wales. *Public Money and Management*, 22(2), 17–24.
- Boyne, G. A. (2012). Performance Management: Does it Work? In R. M. Walker, G. A. Boyne, & G. A. Brewer (Eds.), *Public Management and Performance: Research Directions* (pp. 207–226). Cambridge: Cambridge University Press.
- Brehm, J. & Gates, S. (1999). *Working, Shirking, and Sabotage: Bureaucratic Response to a Democratic Public* (Paperback ed.). Ann Arbor, MI: University of Michigan Press.
- Brodkin, E. Z. (2011). Policy Work: Street-Level Organizations under New Managerialism. *Journal of Public Administration Research and Theory*, 21(2), i253–i277.
- Campbell, D. J. (1988). Task Complexity: A Review and Analysis. *The Academy of Management Review*, 13(1), 40–52.
- Campbell, D T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*. 2(1): 67–90.

- Cots, F., Chiarello, P., Salvador, X., Castells, X., & Quentin, W. (2011). DRG-Based Hospital Payment: Intended and Unintended Consequences. In R. Busse, A. Geissler, W. Quentin, & M. Wiley, *Diagnosis-Related Groups in Europe* (pp. 75–92). McGraw-Hill International.
- Courty, P. & Marschke, G. (2003). Dynamics of Performance-Measurement Systems. *Oxford Review of Economic Policy*, 19(2), 268–284.
- Courty, P. & Marschke, G. (2007). Making Government Accountable: Lessons from a Federal Job Training Program. *Public Administration Review*, 67(5), 904–916.
- Deci, E. L., Koestner, R. & Ryan, R. M. (1999). A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. *Psychological Bulletin*, 125(6), 627–668.
- Deci, E. L. & Ryan, R. M. (2000). The “What” and “Why” of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, 11(4), 227–268.
- Dias, J. J. & Maynard-Moody, S. (2007). For-Profit Welfare: Contracts, Conflicts, and the Performance Paradox. *J-PART*, 17(2), 189–211.
- Dixit, A. (2002). Incentives and Organizations in the Public Sector: An Interpretative Review. *The Journal of Human Resources*, 37(4), 696–727.
- Dubnick, M. (2005). Accountability and the Promise of Performance: In Search of the Mechanisms. *Public Performance and Management Review*, 28(3), 376-417.
- Eterno, J. A. & Silverman, E. B. (2010). The NYPD’s Compstat: compare statistics or compose statistics?. *International Journal of Police Science & Management*, 12(3), 426–449.
- Evetts, J. (2003). The Sociological Analysis of Professionalism: Occupational Change in the Modern World. *International Sociology*, 18(2), 395–415.

- Executive Office of the President. (2015). Every Student Succeeds Act: A Progress Report on Elementary and Secondary Education.
https://www.whitehouse.gov/sites/whitehouse.gov/files/documents/ESSA_Progress_Report.pdf.
- Exworthy, M. & Halford, S. (1999). Professionals and Managers in a Changing Public Sector: Conflict, Compromise and Collaboration. In M. Exworthy & S. Halford, *Professionals and the New Managerialism in the Public Sector* (pp. 1–17). Buckingham, Philadelphia: Open University Press.
- Figlio, D. N., & Getzler, L. S. (2006). Accountability, Ability and Disability: Gaming the System? *Advances in Applied Microeconomics*, 14, 35–49.
- Figlio, D. N. & Rouse, C. E. (2006). Do Accountability and Voucher Threats Improve Low-Performing Schools? *Journal of Public Economics*, 90(1/2), 239–255.
- Freidson, E. (2001). *Professionalism: The Third Logic*. Chicago: University of Chicago Press.
- Frey, B. S. (2000). *Not Just for the Money: An Economic Theory of Personal Motivation*. Cheltenham: Edward Elgar.
- Frey, B. S. & Osterloh, M. (2010). Motivate People with Prizes. *Nature*, 465(17), 871.
- Funnell, S. C., & Rogers, P. J. (2011). *Purposeful program theory: effective use of theories of change and logic models*. John Wiley & Sons.
- Gagné, M., Forest, J., Gilbert, M.-H., Aubé, C., Morin, E., & Malorni, A. (2010). The Motivation at Work Scale: Validation Evidence in Two Languages. *Educational and Psychological Measurement*, 70(4), 628–646.
- Gailmard, S. & Patty, J. W. (2007). Slackers and Zealots: Civil Service, Policy Discretion, and Bureaucratic Expertise. *American Journal of Political Science*, 51(4), 873–889.

- Gailmard, S., & Patty, J. W. (2013). *Learning while governing : expertise and accountability in the executive branch*. Chicago: The University of Chicago Press.
- Garvin D. A. (2013). How Google Sold Its Engineers on Management. *Harvard Business Review*, 91(12), 74–82.
- Gawande, A. (2009). The Cost Conundrum. June 1, *The New Yorker*.
- Gerrish, E. (2016). The Impact of Performance Management on Performance in Public Organizations: A Meta-Analysis. *Public Administration Review*, 76(1), 48–66.
- Gneezy, U. & Rustichini, A. (2000). Pay Enough or Don't Pay At All. *The Quarterly Journal of Economics*, 115(3), 791–810.
- Greve, H. R. (2003). *Organizational Learning from Performance Feedback: A Behavioral Perspective on Innovation and Change*. Cambridge: Cambridge University Press.
- Greve, H. R. (2008). A Behavioral Theory of Firm Growth. *Academy of Management Journal*, 51(3), 476–494.
- Hammond, M. M., Neff, N. L., Farr, J. L., Schwall, A. R., & Zhao, X. (2011). Predictors of Individual-Level Innovation at Work: A Meta-Analysis. *Psychology of Aesthetics, Creativity, and the Arts*, 5(1), 90–105.
- Harris, D. N. (2011). Value-Added Measures and the Future of Educational Accountability. *Science*, 333(6044), 826–827.
- Hatry, H., & Davies, E. (2011). *A Guide to Data-Driven Performance Reviews*. Washington, DC: IBM Center for the Business of Government.
- Heckman, J. J., Heinrich, P., & Smith, J. (2011). Performance Standards and the Potential to Improve Government Performance. In J. J. Heckman, P. Heinrich, G. Courty, G. Marschke, & J. Smith (Eds.), *The Performance of Performance Standards* (pp. 1–14). Kalamazoo, MI: WE

Upjohn Institute.

- Heinrich, C. J. & Marschke, G. (2010). Incentives and Their Dynamics in Public Sector Performance Management Systems. *Journal of Policy Analysis and Management*, 29(1), 183–208.
- Holmstrom, B. & Milgrom, P. (1991). Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, & Organization*, 7 (Special Issue), 24–52.
- Hood, C. (2006). Gaming in Targetworld: The Targets Approach to Managing British Public Services. *Public Administration Review*, 66(4), 515–521.
- Hood, C. (2011). *The Blame Game: Spin, Bureaucracy, and Self-Preservation in Government*. Princeton, N.J.: Princeton University Press.
- Hood, C., & Dixon, R. (2015). *A government that worked better and cost less? : evaluating three decades of reform and change in UK central government* (First Edition). Oxford, United Kingdom: Oxford University Press.
- Hood, C., & Jackson, M. (1991). *Administrative argument*. Aldershot: Dartmouth.
- Jacob, B. A. & Levitt, S. D. (2003). Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating. *The Quarterly Journal of Economics*, 118(3), 843–877.
- Jacobsen, C. B., Hvitved, J., & Andersen, L. B. (2014). Command and Motivation: How the Perception of External Interventions Relates to Intrinsic Motivation and Public Service Motivation. *Public Administration*, 92(4), 790–806.
- Jakobsen, M. L. F. & Mortensen, P. B. (2016). Rules and the Doctrine of Performance Management. *Public Administration Review*, 76(2), 302–312.

- Kelman, S., & Friedman, J. N. (2009). Performance Improvement and Performance Dysfunction: An Empirical Examination of Distortionary Impacts of the Emergency Room Wait-Time Target in the English National Health Service. *Journal of Public Administration Research and Theory, 19*(4), 917–946.
- Kerpershoek, E., Groenleer, M., & de Bruijn, H. (2016). Unintended Responses to Performance Management in Dutch Hospital Care: Bringing Together the Managerial and Professional Perspectives. *Public Management Review, 18*(3), 417–436.
- Kettl, D. F. (1997). The Global Revolution in Public Management: Driving Themes, Missing Links. *Journal of Policy Analysis and Management, 16*(3), 446–462.
- Kettl, D. F. (2015). *The Transformation of Governance: Public Administration for the Twenty-First Century*. Baltimore: JHU Press.
- Koning, P. & Heinrich, C. J. (2013). Cream-Skimming, Parking and Other Intended and Unintended Effects of High-Powered, Performance-Based Contracts. *Journal of Policy Analysis and Management, 32*(3), 461–483.
- Kroll, A. (2015). Explaining the Use of Performance Information by Public Managers: A Planned-Behavior Approach. *The American Review of Public Administration, 45*(2), 201–215.
- Langbein, L. (2010). Economics, Public Service Motivation, and Pay for Performance: Complements or Substitutes? *International Public Management Journal, 13*(1), 9–23.
- Latham, G. P. & Yukl, G. A. (1975). A Review of Research on the Application of Goal Setting in Organizations. *Academy of Management Journal, 18*(4), 824–845.
- Lavertu, S., Lewis, D. E., & Moynihan, D. P. (2013). Government Reform, Political Ideology, and Administrative Burden: The Case of Performance Management in the Bush Administration. *Public Administration Review, 73*(6), 845–857.

- Lazear, E. P. & Gibbs, M., (2015). *Personnel Economics in Practice* (Third edition). Hoboken, NJ: John Wiley & Sons, Inc.
- Levitt, B. & March, J. G. (1988). Organizational Learning. *Annual Review of Sociology*, 14, 319–340.
- March, J. G. (2010). *The Ambiguities of Experience*. Ithaca: Cornell University Press.
- March, J. G., Simon, H. A., & Guetzkow, H. (1993). *Organizations* (Second edition). Cambridge, Mass: Blackwell.
- Mashaw, J. (1983). *Bureaucratic Justice: Managing Social Security Disability Claims*. New Haven: Yale University Press.
- Mayne, J. (2004). Reporting on Outcomes: Setting Performance Expectations and Telling Performance Stories. *The Canadian Journal of Program Evaluation*, 19(1), 31-60.
- Merton, R. K. (1940). Bureaucratic Structure and Personality. *Social Forces*, 18(4), 560–568.
- Mintzberg, H. (1996). Managing Government, Governing Management. *Harvard Business Review*, 74(3), 75-83.
- Moe, Terry. 1989. The politics of bureaucratic structure. In John E. Chubb and Paul E. Petersen (eds), *Can the Government Govern?* Washington D.C: The Brookings Institution.
- Moynihan, D. & Beazley, I. (2016). *Toward Next-Generation Performance Budgeting: Reflections on the Experiences of Seven Reforming Countries*. Washington DC. World Bank.
- Moynihan, D. (2005). Goal-based Learning and the Future of Performance Management. *Public Administration Review*, 65(2): 203–216.
- Moynihan, D. (2008). *The Dynamics of Performance Management: Constructing Information and*

Reform. Washington, D.C.: Georgetown University Press.

Moynihan, D. (2010). A Workforce of Cynics? The Effects of Contemporary Reform on Public Service Motivation. *International Public Management Journal*, 13(1), 24–34.

Moynihan, D., Fernandez, S., Kim, S., LeRoux, K. M., Piotrowski, S. J., Wright, B. E., & Yang, K. (2011). Performance Regimes amidst Governance Complexity. *Journal of Public Administration Research and Theory*, 21 (suppl 1), i141–i155.

Moynihan, D. & Kroll, A. (2016). Performance Management Routines That Work? An Early Assessment of the GPRA Modernization Act. *Public Administration Review*, 76(2), 314– 323.

Moynihan, D. & J. Soss. (2014). Policy Feedback and the Politics of Administration. *Public Administration Review*, 74(3), 320–332.

The New Yorker. (2016), September 12, “How Regulation Failed with Wells Fargo”.

Noordegraaf, M. (2007). From “Pure” to “Hybrid” Professionalism: Present-Day Professionalism in Ambiguous Public Domains. *Administration & Society*, 39(6), 761–785.

Ordóñez, L. D., Schweitzer, M. E., Galinsky, A. D., & Bazerman, M. H. (2009). Goals Gone Wild: The Systematic Side Effects of Overprescribing Goal Setting. *The Academy of Management Perspectives*, 23(1), 6–16.

Ouchi, W. G. (1980). Markets, Bureaucracies, and Clans. *Administrative Science Quarterly*, 25(1), 129–141.

Pandey, S. K. & Wright, B. E. (2006). Connecting the Dots in Public Management: Political Environment, Organizational Goal Ambiguity, and the Public Manager’s Role Ambiguity. *Journal of Public Administration Research and Theory*, 16(4), 511–532.

Perry, J. L. & Wise, L.R. (1990). The Motivational Bases of Public Service, *Public Administration*

Review, 75(1), 53–78.

Perry, J. L., Mesch, D., & Paarlberg, L. (2006). Motivating Employees in a New Governance Era:

The Performance Paradigm Revisited. *Public Administration Review*, 66(4), 505–514.

Pollitt, C. (2000). Is the Emperor in His Underwear? An Analysis of the Impacts of Public Management Reform. *Public Management Review*, 2(2): 181–200.

Pollitt, C. (2013). The logics of performance management. *Evaluation*, 19(4), 346–363.

Pollitt, C. & Bouckaert, G. (2011). *Public Management Reform: A Comparative Analysis: New Public Management, Governance, and the Neo-Weberian State*. Oxford: Oxford University Press.

Pollitt, C., Harrison, S., Dowswell, G., Jerak-Zuiderent, S., & Bal, R. (2010). Performance Regimes in Health Care: Institutions, Critical Junctures and the Logic of Escalation in England and the Netherlands. *Evaluation*, 16(1), 13–29.

Roberts, J. & Dietrich, M. (1999). Conceptualizing Professionalism. *American Journal of Economics and Sociology*, 58(4), 977–998.

Romzek, B. S. & Dubnick, M. J. (1987). Accountability in the Public Sector: Lessons from the Challenger Tragedy. *Public Administration Review*, 47(3), 227–238.

Ryan, R. M. & Weinstein, N. (2009). Undermining Quality Teaching and Learning: A Self-Determination Theory Perspective on High-Stakes Testing. *Theory and Research in Education*, 7(2), 224–233.

Sauder, M. & Espeland, W. N. (2009). The Discipline of Rankings: Tight Coupling and Organizational Change. *American Sociological Review*, 74(1), 63–82.

Scharpf, F. W. (1994). Games Real Actors Could Play: Positive and Negative Coordination in

- Embedded Negotiations. *Journal of Theoretical Politics*, 6(1), 27–53.
- Simon, H. A. (1947). A Comment on “The Science of Public Administration.” *Public Administration Review*, 7(3), 200–203.
- Soss, J., Fording, R. C., & Schram, S. (2011). *Disciplining the Poor: Neoliberal Paternalism and the Persistent Power of Race*. Chicago: University of Chicago Press.
- Steiber, A. (2014). *The Google Model: Managing Continuous Innovation in a Rapidly Changing World*. Cham: Springer International Publishing.
- Stout, M. (2010). Revisiting the (Lost) Art of Ideal-Typing in Public Administration. *Administrative Theory & Praxis*, 32(4), 491–519.
- Suddaby, R., Cooper, D. J., & Greenwood, R. (2007). Transnational Regulation of Professional Services: Governance Dynamics of Field Level Organizational Change. *Accounting, Organizations and Society*, 32(4/5), 333–362.
- Søgaard, R., Kristensen, S. R., & Bech, M. (2015). Incentivising Effort in Governance of Public Hospitals: Development of A Delegation-Based Alternative to Activity-Based Remuneration. *Health Policy*, 119(8), 1076–1085.
- Sørensen, M. L. & Burau, V. (2016). Why We Need to Move Beyond Diagnosis-Related Groups and How We Might Do So. *Journal of Health Services Research & Policy*, 21(1), 64–66.
- Talbot, C. (2010). *Theories of Performance: Organizational and Service Improvement in the Public Domain*. Oxford: Oxford University Press.
- Tummers, L. G. & Van de Walle, S. (2012). Explaining Health Care Professionals’ Resistance to Implement Diagnosis Related Groups: (No) Benefits for Society, Patients and Professionals. *Health Policy*, 108(2), 158–166.

Loon, N. van, Leisink, P. M., Knies, E., & Brewer, G. A. (2016). Red Tape : Developing and Validating a New Job-Centered Measure. *Public Administration Review*, 76(4), 662-673.

Weber, M. (1960). The Essentials of Bureaucratic Organization: An Ideal-Type Construct. In R. K. Merton (Ed.), *Reader in Bureaucracy* (2nd printing, pp. 18–27). Glencoe, Ill.: The Free Press.

Weibel, A., Rost, K., & Osterloh, M. (2010). Pay for Performance in the Public Sector: Benefits and (Hidden) Costs. *Journal of Public Administration Research and Theory*, 20(2), 387–412.

Wright, B. E. & Pandey, S. K. (2010). Transformational Leadership in the Public Sector: Does Structure Matter? *Journal of Public Administration Research and Theory*, 20(1), 75–90.