

Performance Regimes Amidst Governance Complexity

Donald P. Moynihan*, **Sergio Fernandez[†]**, **Soonhee Kim[‡]**, **Kelly M. LeRoux[§]**, **Suzanne J. Piotrowski****, **Bradley E. Wright^{††}**, **Kaifeng Yang^{‡‡}**

**University of Wisconsin–Madison; [†]Indiana University–Bloomington; [‡]Syracuse University; [§]University of Illinois, Chicago; **Rutgers University–Newark; ^{††}University of North Carolina–Charlotte; ^{‡‡}Florida State University*

ABSTRACT

Much of the appeal of performance measurement is explained by its image as a simple and value-neutral way to monitor and improve government. But contemporary governance is characterized by complexity. Few public officials have the luxury of directly providing relatively simple services, the context in which performance regimes work best. Instead, they must work in the context of a disarticulated state, with policy problems that cross national boundaries and demand a multi-actor response. At the same time, traditional democratic values must be honored. This article examines the tensions between performance regimes and the complexity of modern governance, identifying implications and questions for research and practice.

INTRODUCTION

Previous Minnowbrook scholarship touched on the need to “recover” public management by relating it to organizational performance (Dilulio 1989). Since then, efforts to make government more performance oriented, via routines such as strategic planning and performance measurement, have taken on a new importance that prior Minnowbrook meetings could not address. In this article, we use the term performance regimes to refer not just to the practices of measuring and managing performance indicators but also to capture the embedded nature of these practices in almost all aspects of contemporary governance. Increasingly, performance routines are among the most important tools by which governments structure relationships, state values, and allocate resources with employees, third-party providers, and the public. The term performance regimes reflect the fact that the cumulative effect of the application of these tools is greater than the sum of their parts, resulting in a basic and fundamental change in how we govern. The clearest reflection of this change is the unprecedented pressure that public actors are under to perform, in a context where performance is defined by quantitative indicators.

This article examines the role of performance regimes in contemporary governance. Even as much of the appeal of performance measurement is explained by its image as

Address correspondence to the author at dmoynihan@lafollette.wisc.edu.

doi:10.1093/jopart/muq059

© The Author 2011. Published by Oxford University Press on behalf of the Journal of Public Administration Research and Theory, Inc. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

a simple and value-neutral way to monitor and improve government (Moynihan 2008; Radin 2006), we argue that such simplicity is actually at odds with the complexity of contemporary governance. Successful examples of performance regimes tend to be of organizations that directly provide relatively simple services. But contemporary governance is more frequently characterized by a disarticulated state, networks of providers, and a global context (Frederickson and Frederickson 2006). Consistent with Ashby's law of requisite variety, performance regimes that do not reflect the complexity of governance will fail in some fashion. They run the risk of being ignored, discarded, or manipulated. They may foster unanticipated consequences and undesirable behavior at odds with traditional values of governance.

There are both scholarly and practical implications. We need to better understand the outcomes that result from the interaction of performance regimes and their context. From a practical view, we need to design performance regimes that are consistent with basic normative choices that characterize democratic government, while acknowledging that such values are often in conflict (Rosenbloom 1983). In short, we need more carefully designed performance regimes that reflect the complexity of contemporary governance. We need to caution against the belief that performance measurement will make the messy business of governing simple.

The article is structured in three parts. First, we examine the difficulty of measuring performance. In particular, how performance measurement processes often fail to incorporate democratic values and struggle to find agreed upon indicators for network and global levels of governance. Next, we consider the effects of performance regimes. Finally, the article concludes by identifying a research agenda that should inform any new performance regime paradigm.

MEASURING PERFORMANCE

A central appeal of performance measurement was that it could overcome some of the seemingly inherent characteristics of public agencies. Performance measures could provide a bottom line to replace vague and contradictory goals and offered a means for political principals to cut through the information asymmetry enjoyed by agents. This assumed that programs were simple enough that a small number of measures could definitively represent performance. But public programs are characterized by task complexity, with multiple goals, multiple ways to measure those goals, and different perspectives among stakeholders about which choices to make (Moynihan 2008). To measure a manageable number of goals therefore implies excluding other competing goals and measures. But there are other types of complexity apart from task complexity. There is also the complexity of incorporating democratic goals and recognizing the collaborative and frequently global nature of governance. Performance measurement was designed to make bureaucracies work better, but to remain useful, it has to be adapted to models of governance that seek to incorporate the rights and views of citizens, as well as the complications of networks and global forms of coordination.

Democratic Values

Participants at previous Minnowbrook meetings worried that values such as social equity were given insufficient attention relative to efficiency and effectiveness values

(Frederickson 1980). So, what would Waldo and the scholars of the New Public Administration make of this era of governance by performance management? They would likely worry that the growing ubiquity of performance regimes has further privileged efficiency and effectiveness values. Some contemporary scholars have taken up this argument, pointing out that performance measurement systems displace attention from other democratic values (e.g., due process, equity, integrity, citizenship, and transparency). For example, Piotrowski and Rosenbloom (2002) show that performance reporting under the Government Performance and Results Act failed to incorporate available measures of transparency. Federal government performance evaluations have also been criticized for ignoring the social equity implications of programs, even those with redistributive intent (Radin 2006). Others have criticized the exclusion of citizenship indicators for programs that are likely to shape an individual's civic understanding (Wichowsky and Moynihan 2008).

Performance measurement has displaced attention from democratic values because it draws from an implicit political theory of reform that narrowly defines the goals of public organizations in terms of mission achievement and cost effectiveness (deLeon and Denhardt 2000; Piotrowski and Rosenbloom 2002). Traditional democratic values that are not obviously central to mission achievement are conceived of as something other than performance. Although it may seem that the burden of collecting data on these other values may not be worthwhile, democratic values can be justified in ways that make sense to a scholar or practitioner with a mission-oriented viewpoint. For example, many functions require citizen coproduction and, therefore, depend on values such as perceptions of fairness and due process (Wichowsky and Moynihan 2008). Such values can therefore be part of a performance logic model that lays out the links between inputs, outputs, and outcomes. Furthermore, Hatry (1999) has argued for logic models to capture side effects and unintended consequences and the displacement of democratic values fall into this category.

There are strategies by which performance measurement can align with democratic values. Some, such as Stivers (2008), worry that performance measurement has become the handmaiden of a broader approach to governance, centered on market ideals, that has reduced the capacity for communal experience of government. But we suggest that measurement can be used as a tool to reopen that shared space. By incorporating more democratic values, measurement can track how citizens experience the state. Although this may be a difficult task, Wichowsky and Moynihan (2008, 917) argue that "(t)here is no intrinsic attribute of performance measurement that prevents it from refocusing administrative attention to democratic concerns. Indeed, it could become a powerful means to do so." Some organizations already make democratic values a part of their mission—such as transparency for the U.S. National Archives and Administration and equity for the U.S. Equal Employment Opportunity Commission. And some government scorecards already attempt to measure performance in terms of fairness, participation, due process, and transparency, and there are a variety of ways to engage citizens in this process (Epstein, Coates, and Wray 2005). Additional strategies to utilize performance measurement processes to democratic ends have been proposed. Rosenbloom (2007), for example, sketches how democratic constitutional scorecards could measure the impact of reform on individual rights, constitutional integrity, transparency, and the rule of law (e.g., the value of transparency could be partially tracked through the measurement of Freedom of Information Act processes). Valid measures for different aspects of citizenship—for example, social trust, political efficacy, procedural justice, equity of treatment, and political participation—exist and

can easily be applied in measuring the impact of appropriate programs (Wichowsky and Moynihan 2008).

Performance regimes seek to establish legitimate accountability mechanisms, but they cannot achieve this goal if the practice of performance measurement excludes democratic values that provide an older and more durable basis for public legitimacy. Citizens that feel treated fairly by officials, for example, are more likely to see government as trustworthy and responsive (Wichowsky and Moynihan 2008). If we are to heed the Minnowbrook tradition of reconciling democracy and administration, there needs to be an explicit discussion of values that culminates in a coherent strategy to make democratic-based values a priority in performance management. Although we call for increased awareness of democratic values among the actual implementers of performance systems, there are also important questions for future research (see table 1). The answers to these questions are essential if we are to understand not only the current but also potential democratic implications of performance measurement.

Table 1
Key Research Questions for Performance Regimes amidst Governance Complexity

Democratic Values

- What values do or can performance information represent?
- How should goals and measures be set and by whom?
- What are useful ways to measure and disseminate democratic-based values?
- What legitimate roles should and do the public play in performance regimes?
- How should and does performance information interact with traditional forms of democratic involvement?

Collaborative Governance

- How does the dispersed nature of collaborative authority, featuring multiple actors with different goals and interests, shape performance measurement?
- Is it possible to tie performance measures to specific actors if responsibility is diffuse?
- How do we measure network performance? What are appropriate “levels” of network performance (e.g., client, network participant, network as whole, community)?
- Are different “levels” of network performance at odds with one another? Is it possible to develop aggregate and integrated measures of network performance across “levels”?
- How do process measures of network coordination contribute to network outcomes?
- How does the evolution of networks affect network outcomes?

Global Governance

- How should we define and measure the performance of country-based regulatory and economic agencies in the context of globalization?
- What are valid and useful indicators of democracy and development?
- How do cultural and institutional differences shape the application of performance regimes?
- How are performance indicators diffused, adopted, and implemented in different settings? What roles do international institutions, national and local governments, and stakeholders play in this process?
- How do power and knowledge differentials shape the application of performance regimes? What roles do and should poorer countries play in defining the terms of debate and accepting responsibility for outcomes?

Effects of Performance Regimes

- What is the likely response to the introduction of performance regimes—passive; political; purposeful; or perverse?
 - What factors encourage different forms of performance information use?
 - Do performance regimes improve performance? Does the use of performance information improve performance?
 - As performance regimes become more dominant, how do they change basic aspects of governance—such as the distribution of power, individual motivation, organizational culture, and behavior?
 - How do principals respond to gaming by agents? Can they respond effectively?
-

Collaborative Governance

The risks that democratic values will be displaced in performance regimes may be compounded when service delivery occurs in a collaborative setting, especially if key actors are primarily motivated by profit maximization, directed by measures, and do not face the same constraints as government officials. Although efficiency is a common rationale (or rhetoric) for the decision to engage in collaborative service delivery, public managers have a role in ensuring that public service values of efficiency and effectiveness are appropriately balanced with other values, such as equity, political accountability, and responsiveness (Brown, Potoski, and Van Slyke 2006). This section further examines the complexity of the collaborative setting.

The same managerial reform initiatives that demanded improved government performance also called for a shift toward more market-based governance (Donahue and Nye 2002). But there is a key tension between results-based reforms and changes that have hollowed out government: the doctrine of performance management drew from a model of organizational improvement that assumed a single direct provider rather than the complexity or collaborative nature of the disarticulated state (Frederickson and Frederickson 2006). As a result, government agencies today are expected to provide citizens with high-performing programs and services, even as much of these services are actually delivered by others via a variety of collaborative forms, that is, contracts, networks, intergovernmental agreements, and public–private ventures.

There are a number of challenges in measuring performance in a collaborative context. Not least is that scholars disagree on the definition and meaning of collaboration, leading to inconsistent measurement of the concept and wide-ranging conclusions about the inputs and processes that contribute to high-performing collaborative service arrangements (Thomson, Perry, and Miller 2009). Another challenge arises from the nature of goods and services for which government partners with external entities. Public agencies often enter into collaborative relationships in order to address complex problems that government alone cannot solve or even clearly define. As a result, public agencies are disadvantaged by the “smart buyer” problem (Kettl 1993) and experience difficulty specifying a priori service quality and quantity requirements for collaborators (Brown, Potoski, and Van Slyke 2008).

There has, thus far, been little consideration of how the dispersed nature of collaborative authority, featuring multiple actors with different goals and interests, shapes performance measurement. The need for network actors to be responsive to multiple principals often leads to role ambiguities and uncertainty on the part of actors about how to prioritize performance goals. Even though network participants are united by a common policy objective, networks actors “face the challenge of reconciling the needs of multiple stakeholders, diverse expectations, and varying organizational missions and roles, while delivering a complex public service” (Romzek 2008, 6). As group size increases, goal heterogeneity and preference divergence become increasingly likely, generating high transaction costs for the network administrator (Post 2004). The interdependence of network participants also makes it difficult for network administrators to hold any one actor accountable for unsatisfactory performance.

The complexity of the network setting is reflected in calls to measure network performance at multiple levels, even as performance at one level may conflict with another. For example, Provan and Milward (2001) call for assessing network effectiveness at the

community, network, and participant levels, whereas Mandell and Keast (2008) offer criteria for measuring performance at the environmental, organizational, and operational levels of the network. Although these criteria offer a good start, network scholars have yet to apply them in studies of network effectiveness. Thus, a central challenge for both network theorists and scholars of performance management is to test how well these conceptual criteria work at each level and explore the extent to which they can be integrated, so as to produce a refined framework for measuring network performance.

Process goals are also an important but often overlooked dimension of network performance. Mandell and Keast (2008) have convincingly argued that conventional performance indicators such as cost and client outcomes by themselves offer an insufficient assessment of network effectiveness because they overlook the relational dimensions that are of paramount importance to network performance. As these authors suggest, the “core business of a network is centered on developing relationships and processes to facilitate interactions” (Mandell and Keast 2008, 718). Another key challenge for measuring network performance, therefore, lies in specifying process measures and identifying how they contribute to network outcomes at each level.

Public administration scholarship over the last decade has yielded some important insights about how to measure the performance of collaboration, particularly in the areas of hierarchically structured contracting arrangements. Additional conceptual and theoretical development is still needed to address the challenges raised in this section, especially as they relate to measuring network performance. In other words, there is often a mismatch between our theoretical frameworks and what we know is important. Although we are best able to structure accountability and measure performance through principal agent frameworks that rely on dyads, network success often depends on less easy to measure collaborative arrangements that may feature multiple principals and agents or even the blurring of these roles, and a reliance on trust to supplement contracts. Our knowledge of network performance is also limited by a lack of time series investigations that capture the importance of iterative processes of network development to outcomes. Research that directly addresses these limitations would help to improve performance measurement as it relates to collaboration. The research questions raised by this section are summarized in table 1.

Global Governance

Perhaps the greatest single difference since the original Minnowbrook meetings is the influence of globalization. This increases the importance of connectedness, interdependency, and collaborative governance beyond the boundaries of the nation state. Earlier scholars worked in a paradigm of comparative or development administration, approaches that assumed a certain degree of independence between countries. But globalization is fuelled by interdependence and contagion between countries, typified by phenomena such as terrorism, contagious diseases, climate change, and financial crises. National systems and policies are becoming cogs of an integrated but decentralized international system. For example, the performance of financial markets in any given country cannot be understood simply by looking at the regulatory and economic agencies in that country alone. This being the case, how should we define and measure the performance of organizations such as Federal Reserve and the U.S. Treasury?

There has an explosion of comparative tables of national performance indicators (e.g., educational performance in each country), although, as discussed below, such indicators

are inevitably imperfect (Arndt 2008). But there has been much greater difficulty in finding truly transnational indicators of performance that reflects global issues (e.g., readiness for a pandemic). In this task, the challenge of how to measure collaborative governance in a domestic setting (discussed above) remains but is more complex. Global problems are more wicked than most. As they spread across national borders, the limited capacity of any single government to respond becomes apparent and the demarcation of responsibility more difficult to pinpoint. Although most examples of collaborative governance can look to one dominant central hub (Provan and Milward 2001), this is generally not the case in transnational governance, where authority is more truly diffuse. At the transnational level, efforts to coordinate are weakened by a lack of common frame of reference, concerns about sovereignty, and long-standing tensions between countries. For all these reasons, the actors involved in global governance can be expected to have a more difficult time in coming to an agreement on what it means to measure performance.

There are additional unique challenges to applying performance tools in global governance. Most pressing is the need to balance recognition of legitimate contextual differences with shared performance expectations across national boundaries. It is difficult to develop outcome measures that are directly comparable across countries given differences in institutional frameworks and cultural and political values (Lijphart 1999; Yang and Holzer 2006). The application of western performance management models in developing countries has been criticized by scholars because of the inherent differences in government capacity, civil society strength, and market maturity (Manning 2001; Schick 1998). Perhaps as a result, performance measures are rarely developed as a result of multinational collaborative efforts. Instead, international organizations have taken the lead by developing standard cross-national indicators on government effectiveness, democratic governance, and transparency (e.g., Transparency International 2007; World Bank 2007). However, the ideology and values of these organizations, and by extension their performance measures, have come under question (Keohane 2002). For example, Transparency International and the World Bank's corruption indices are quite popular, but it is unclear how political ideologies and cultural beliefs have affected the ranking. The practices that members of Parliament used in claiming expenses in the United Kingdom, or the system of campaign finance in the United States, though legal, might be considered no less corrupt than similar extra-legal practices in other settings. As Taussig (1999, 78) comments, international comparisons often reflect a dualist notion with "north as honest and industrious and repressed, the south as anything but."

Although performance indicators developed by international organizations allow comparative analysis of governance performance at the national level, they are far from perfect. In fact, some researchers have raised concerns regarding the validity of the current indicators and/or their limited view of the public administration's role and functions (Van de Walle 2008). The World Bank's Governance Indicator Project presents an interesting example. The comparability of its indicators across countries and time, the accuracy of the different data sources on which the project relies, and the consequences of its data aggregation approach have all been questioned even as these problems are not readily apparent or considered by users of the rankings (Arndt 2008; Kaufmann and Kraay 2008). The indicators themselves reflect a mixture of New Public Management and neoclassical economic doctrine (Andrews 2010). The rich countries that score highly on the indicators actually do not follow any single model, varying in political control over government, the size of the

welfare state, the degree of government control of the economy, and the centralization of authority (Andrews 2010). Thus, the indicators are used as a tool to foster the adoption of certain practices among aid-recipient countries that donor countries themselves do not consistently pursue. The perception that such values are measurable contributes to their rigidity, which in turn reinforces perceptions that they are externally imposed with little input from or room for interpretation by the recipient countries.

Indeed, much of the tensions in global performance efforts arise from a sense that external measures are being imposed, reflecting an underlying lack of agreement about the importance and meaning of basic values (Arndt 2008). For example, programs that seek to foster democracy must first define what democracy means in disparate contexts. But “democratization in any country cannot be broken down neatly and precisely into a series of quantitative bits” says Carothers (1999, 291), who has documented the difficulty that United States Agency for International Development (USAID) has had in this regard. USAID struggles to find useful measures to assess their work or find causal connections between projects and outcomes. Disagreements on the meaning of democracy can be complicated by the application of governance indicators at different levels of government. For example, governance assessments by the World Bank and Transparency International (Transparency International 2007; World Bank 2007) rank China and South Korea poorly on rule of law and control of corruption. But Seoul Metropolitan Government can point to exceptionally high rates of online citizen participation (Holzer and Kim 2007), and Guangzhou Municipal Government has been a leader in formulating its own government information disclosure rules, citizen participation law, and administrative regulations (United Nations Project Office on Governance 2008).

Overall then, the crude use of performance indicators runs the danger of reinforcing classic problems in aid: donors demanding the adoption of one-size-fits-all models that recipients begrudgingly adopt but do not implement. In recognition of this, there has been a battle on how to measure good governance within the World Bank (Radin 2006). Radin (2006) argues that top-down, government-focused approaches tend to focus on efficiency and are doomed to failure as national governments are not deeply committed to them. In contrast, performance systems that are designed to work with NGOs who are serious about change may have better prospects for success and are more likely to capture the impacts of programs on citizens.

The global and transnational settings introduce practical and scholarly questions that are similar to those raised in previous sections but more challenging (see table 1). There is the overarching problem of how to define government performance in a way that transcends political and cultural values. Standardized approaches to measurement are often rejected as imperialist, but this criticism can become an excuse by governments seeking to avoid oversight (Arndt 2008). The challenge, therefore, is to measure performance in a way that recognizes power imbalances and meaningful contextual differences. There is need for research that examines how performance indicators are diffused, adopted, and implemented in different settings and the particular roles that international institutions, national and local governments, and stakeholders play in this process. There is also the exceptionally difficult question of drawing causal connections between project activity and changes in performance indicators. Carothers (1999) points out that there is fierce debate about the effectiveness of well-studied programs in developed countries, whereas there is almost no evidence on this question in poorer countries. This brings up the question of the effects

of performance regimes, a question relevant to all levels of government and discussed in more detail below.

FROM MEASUREMENT TO MANAGEMENT: THE EFFECTS OF PERFORMANCE REGIMES

Much of this article deals with the difficulties of measuring performance. But once measures are developed, how are the data used? How do we make the jump from measurement to management? Until data are interpreted and used, the promise of performance regimes to improve performance will not be met. Indeed, says Stivers (2008, 126), “there is a real sense in which the knowledge such as the kind produced by a standardized test only becomes “true” when we experience what happens after we get the test results.”

Results-based reforms propose to change the behavior of public officials. These reforms assume that the introduction of information will lead to its purposeful use by public officials to manage their program or allocate resources in ways that improve performance. But the simplicity of this assumption fails to reflect the complexity of governance in three ways. First, it assumes that performance information use is unidimensional, when in fact performance data can be used in multiple ways that reflect the preferences of individuals and institutions. Second, performance doctrine assumes that the simple availability of information shapes behavior and pays relatively little attention to how other factors shape use. And third, there has been little consideration of the long-term unanticipated consequences of performance regimes on basic aspects of governance. This section briefly examines these three complexities.

One of the difficulties of studying the use of performance information is that we tend to think of it as a unidimensional construct, when in fact it takes various forms. In addition to purposeful uses of performance data, scholars have observed other important forms of performance information use (Moynihan 2009). Results-based reforms can result in passive reactions, where actors do the minimum required to comply with requirements to create and disseminate information without actually using this information to make programmatic changes (Radin 2006). Performance data also are used for political purposes as a means of advocacy to defend or criticize a program (Moynihan 2008). Alternatively, public agencies may clearly specify performance expectations that can create perverse incentive structures and lead to unintended outcomes, displacing the underlying or unmeasured goals of a program (Heinrich 2007). For example, performance-based contracting for welfare-to-work has been shown to produce a “performance paradox” whereby contractors are rewarded for meeting performance expectations aimed at short-term outcomes and not for helping clients to become self-sufficient (Dias and Maynard-Moody 2007).

Researchers often focus on one or two categories of performance information use and ignore others. If scholars arrive at different conclusions about performance information use, it may be because they are actually talking about different aspects of use, with different causes and effects. If we are to emerge from this muddle, we need to conceptualize and measure these different aspects of performance information use in the standardized way that we do for other multidimensional behavioral concepts, such as organizational commitment.

In addition to recognizing important distinctions in how performance data are used, research should also examine antecedents to use. On a practical level, governments devote considerable attention to creating and disseminating performance information but give little thought as to how to foster the use of this data (Moynihan 2008). On a scholarly level,

there remains “surprisingly little effort to develop a theory of performance management,” say Jennings and Haist (2006, 174), as they elucidate a series of hypotheses that identify a wide range of different possible influences. A growing empirical literature seems to confirm a lack of an overarching theory, as well as the importance of disparate factors at the individual, task, organizational, and external levels (for a summary, see Moynihan and Pandey 2010). But perhaps the most important factor is how performance is measured.

Part of the reason this article devotes so much attention to measurement is because measures direct attention and shape behavior, or to quote the worn-out adage, what gets measured gets managed. As the previous sections have discussed, in an environment where some goals and values are explicitly recognized as important and essential to the organizational mission while other values are largely ignored or denigrated as procedural barriers to performance, it is entirely predictable that organizational actors will exert effort toward what is deemed culturally appropriate. O’Toole (1997) points to the need for public managers to carefully structure accountability mechanisms and performance expectations for third-party providers in ways that help to realize the objectives of democratic governance. If contractors are focused on highly incentivized mission-based measures, it is more likely that procedural goals such as due process and equity, and other aspects of quality that are not measured, will be displaced (Soss, Fording, and Schram forthcoming). In the area of global governance, Carothers (1999) argues that the numbers crowd out more nuanced or contextual discussion of USAID programs by overseers in Washington, arguing that “efforts to assess the impact of democracy programs by using highly reductionist indicators is a deeply flawed undertaking that is consuming vast resources, producing little useful insight or knowledge, and introducing serious distortions into the designing and implementing of such aid” (Carothers 1999, 291).

Although how performance is measured influences its use other factors also clearly matter. The nature of the task will shape ease of measurement, which in turn is likely to shape use (Jennings and Haist 2006). The tendency of performance regimes to fail to acknowledge fundamental differences between public services or contexts may result in a one-size-fits-all approach that suits some agencies more than others, causing resentment among the performance losers (Radin 2006). There is evidence that the quality and availability of the data fosters use (Bourdeaux and Chitoko 2008; de Lancer Julnes and Holzer 2001; Moynihan and Ingraham 2004; Moynihan and Pandey 2010).

Individual traits, beliefs, and job characteristics (such as organizational commitment, public service motivation, experience, and role ambiguity) are likely to affect how performance data are viewed, and what forms of uses are appropriate and feasible. There is strong evidence that some organizational factors matter, particularly supportive leadership (Askim et al. 2008; Jennings and Haist 2006; Melkers and Willoughby 2005; Moynihan and Ingraham 2004) and the organizational culture (de Lancer Julnes and Holzer 2001; Moynihan and Pandey 2010). The political environment is also likely to matter, as the role and involvement of elected officials, central agencies, the public, and stakeholders will affect the ways in which data are used (Bourdeaux and Chitoko 2008; Jennings and Haist 2006). In addition, the broader institutional environment may foster multiple uses and send mixed signals to decision makers. Organizations may measure one set of goals to gain external legitimacy (Brunnson 1989), whereas relying on other data to improve internal operations.

Finally, it is clear we need to not only understand more about how performance regimes emerge but also be aware of their long-term consequences. At a most basic level,

we lack definitive evidence about whether performance regimes ultimately improve public sector capacity and outcomes. Performance regimes are intended to change the nature of governance and alter culture, behavior, and outcomes in the public sector. They empower some actors who set the terms of performance but put others under closer scrutiny. Therefore, we need to not just identify the causes of performance information use but also to systematically examine how the sustained use of performance data affects a variety of individual and organizational phenomena in governance, including factors that shape performance information use in the short run. For example, while in the short run, intrinsic motives, organizational commitment, job satisfaction, and organizational culture might shape performance information use, the long-term application of performance regimes might in turn reshape these variables. Performance goals could be used to celebrate achievement, encouraging more committed and satisfied employees. Or they could have the opposite effect, if the goals are seen as illegitimate, demanding actions inconsistent with a sense of public interest (Dias and Maynard-Moody 2007). Principals might respond to the exploitation of incomplete performance contracts by trying to write ever-more detailed contracts or they could rely instead on more relational forms of contracting, using trust and reputation as a substitute for performance monitoring (Fernandez 2009; Van Slyke 2007). The use of performance data can also affect the extent and nature of the external environment's influence on the organization, for instance, by enhancing political support from constituents and overseers (Brunnson 1989; Moynihan 2008).

Understanding how performance regimes reshape governance represents the long-term equivalent of understanding the causes of performance information use, in that it is the key test of whether performance management is making a difference. It requires us to resist the temptation to either romanticize or stereotype the past as we track how performance regimes displace or change existing bureaucratic routines. Table 1 presents some of the key questions that must be answered in this process (see also Moynihan 2009). Tracking such changes offer the best map of how governance is being transformed, and the information to judge whether these changes are for good or ill.

CONCLUSION: TOWARD A NEW PARADIGM

The widespread diffusion of performance tools is one of the most significant developments in public administration in the 20 years since the last Minnowbrook Conference was held. A number of insights emerge from this article that can inform a new paradigm for performance regimes. We do not claim that our arguments are completely novel or comprehensive. Instead our goal is to weave together strands of existing work to identify what we see as the pressing issues for scholars and practitioners of performance.

First, there is the relationship between governance complexity and performance regimes. In this article, we have argued that the increasing complexity of governance converges with the application of performance regimes in ways that will fundamentally affect both performance regimes and governance. Figure 1 reflects this argument. There are a number of related factors that make governance more complex. As these factors increase, standard approaches to performance management become less easy to apply and more likely to experience failure or negative unanticipated consequences. Policy areas characterized by task complexity, multiple and possibly contradictory goals, disagreement about goals, and uncertainty about the link between action and outcomes run at odds with

Figure 1
As Governance Complexity Increases, Performance Regimes Struggle



performance systems that are designed to generate consensus around a manageable number of indicators or clearly hold actors responsible. Under such conditions, agents may exploit information asymmetry by gaming how they respond to incentives. Where actors are shaped by different cultural lenses and institutional interests, agreements on what a performance regime should do becomes less feasible. In complex governance settings, the greater heterogeneity of influential actors is likely to result in more marked battles about the definition of performance, and the specific measures that emerge are likely to reflect hard-fought compromises rather than obvious and neutral solutions.

A second major insight is that the complexity of governance—and therefore, the difficulties encountered by performance regimes—tends to increase as we move away from traditional public bureaucracies, toward networks and global governance. We would therefore expect that traditional public organizations to be closer to the left-hand side of the complexity continuum in figure 1, networks toward the middle, and global governance toward the right-hand side. Of course, there will be exceptions to this general claim. Some traditional public organizations have more complex tasks than others. Some networks are able to foster a high level of basic agreement on goals, either through shared norms or via contracts. But governments generally turn to networks to deal with “wicked problems” that are beyond the capacity of single organizations, and global governance introduces a different set of institutional and cultural tensions that can undermine the functionality of standard performance regimes.

A third major insight for any new paradigm is to recognize the normative aspects of performance regimes. The basic questions we raise in this article—How to measure performance? How do performance regimes affect behavior?—appear simple, but are difficult to answer. In part, this is because they have strong normative implications that were not readily acknowledged by reformers. After all, the business of choosing measures is the business of choosing what government does. Performance indicators empower some and control others. Such measures are not neutral but have been used to legitimize values such as efficiency and cost effectiveness over others, such as equity and transparency. As we consider a research agenda, we need not just strong empirical research but ultimately must face normative considerations on the relationship between performance regimes and democratic governance.

The search for a new paradigm should also be mindful of the very real gaps in theory and practice identified in this article. These gaps will be difficult to fill but are necessary to address if we are to understand how performance regimes currently function, or might function, under varying levels of governance complexity. Table 1 summarizes a broad research agenda that has clear practical implications. Although the questions are organized to align with the specific section of the article in which they were discussed, the default assumption should be that questions from one section are broadly applicable to others. For example, the question of who decides measures, and the democratic implications therein, are appropriate for all levels of governance. Questions related to performance regimes for collaborative governance are critical to understanding global governance. And the effects of performance regimes are of universal interest wherever they are deployed.

Given the potential problems posed by the interplay between the complexity of governance systems and simplicity of performance measures often used to portray them, and the gaps in our knowledge, it might be easy to suggest that the pursuit of performance regimes be abandoned. Indeed Kelman (2007) portrays contemporary public administration scholarship as basically antagonistic toward efforts to improve performance. Interestingly, the scholars who met at Minnowbrook in 2008 took a different perspective. Perhaps because this generation of scholars has not experienced a time when performance tools were not in vogue, we are receptive to the opportunities that performance regimes offer in improving governance. Internally, performance regimes can help clarify vague and often symbolic missions, translate them to action, and produce the performance feedback necessary to evaluate and adjust the goals, strategies and actions at the network, organization, unit, and employee levels. Externally, performance management can provide access and transparency to government, facilitating citizen participation, and improving the accountability, equity, and responsiveness of government.

The maximization of these opportunities requires that performance tools are not viewed as simple or neutral but rather as a necessary part of an evolving and inevitably imperfect system of governance. Thus, our argument here should not be construed as an attack on performance regimes—measuring performance is almost always preferable to not measuring performance—but a recognition of the complex context in which they operate. The final major insight we propose, therefore, is the need for a contingency perspective. Performance regimes should be designed to reflect the environment in which they work rather than a one-size-fits-all model that may work perfectly well under some conditions but fail under others. Our hope is that performance regimes lose some of their negative connotations and be recognized as a legitimate approach to democratic governance. This may require a “rebranding” of performance tools in a way that reflects a broader realignment between performance regimes and the complexity of modern democracy that remains, thus far, unrealized. We view as legitimate many of the objections that have been raised but see such criticism as a necessary part of redesigning rather than dismantling performance regimes. Performance regimes are and will remain an essential aspect of contemporary governance arrangements, but their effectiveness depends on a capacity to be as nuanced and complex as the context in which they work.

REFERENCES

- Andrews, Matt. 2010. *Good government means different things in different countries*. *Governance* 23: 7–35.

- Arndt, Christiane. 2008. The politics of governance ratings. *International Public Management Journal* 11:275–97.
- Askim, Jostein, Åge Johnsen, and Knut-Andreas Christophersen. 2008. Factors behind organizational learning from benchmarking. *Journal of Public Administration Research and Theory* 18:297–320.
- Bourdeaux, Carolyn, and Grace Chikoto. 2008. Legislative influences on performance management reform. *Public Administration Review* 68:253–65.
- Brown, Trevor, Matthew Potoski, and David Van Slyke. 2006. Managing public service contracts. *Public Administration Review* 66:323–31.
- . 2008. *Contracting for complex products*. Paper presented at Minnowbrook III Conference, Lake Placid, NY, September 5–7.
- Brunnson, Nils. 1989. *The organization of hypocrisy*. Chichester, UK: John Wiley & Sons.
- Carothers, Thomas. 1999. *Aiding democracy abroad*. Washington, DC: Carnegie Endowment.
- de Lancer Julnes, Patria, and Marc Holzer. 2001. Promoting the utilization of performance measures in public organizations. *Public Administration Review* 61:693–708.
- deLeon, Linda, and Robert B. Denhardt. 2000. The political theory of reinvention. *Public Administration Review* 60:89–97.
- Dias, Janice Johnson, and Steven Maynard-Moody. 2007. For-profit welfare: Contracts, conflicts, and the performance paradox. *Journal of Public Administration Research and Theory* 17:189–211.
- Dilulio, John J. 1989. Recovering the public management variable: Lessons from schools, prisons, and armies. *Public Administration Review* 49:127–33.
- Donahue, J. D., and J. S. Nye. 2002. *Market-based governance: Supply side, demand side, upside and downside*. Washington, DC: Brookings Institution Press.
- Epstein, Paul D., Paul M. Coates, and D. Lyle. 2005. *Wray with David Swain. Results that matter*. San Francisco, CA: Jossey Bass.
- Fernandez, S. 2009. Understanding contracting performance: An empirical analysis. *Administration & Society* 41:67–100.
- Frederickson, H. George. 1980. *New public administration*. University, AL: Univ. of Alabama Press.
- Frederickson, David G., and H. George Frederickson. 2006. *Measuring the performance of the hollow state*. Washington, DC: Georgetown Univ. Press.
- Hatry, Harry P. 1999. *Performance measurement: Getting results*. Washington, DC: Urban Institute.
- Heinrich, Carolyn J. 2007. False or fitting recognition? The use of high performance bonuses in motivating organizational achievements. *Journal of Policy Analysis and Management* 26:281–304.
- Holzer, Marc, and Seang-Tae Kim. 2007. *Digital Governance in Municipalities Worldwide*. <http://unpan1.un.org/intradoc/groups/public/documents/aspa/unpan031374.pdf> (accessed October 4, 2010).
- Jennings, Edward T., and Meg Patrick Haist. 2006. Putting performance measurement in context. In *The art of governance*, ed. Patricia W. Ingraham and Laurence E. Lynn, 173–94. Washington, DC: Georgetown Univ. Press.
- Kaufmann, Daniel, and Aart Kraay. 2008. Governance indicators: Where are we and where should we be going? *The World Bank Research Observer* 23:1–30.
- Kelman, Steven. 2007. Public administration and organization studies. In *Academy of management annals*, ed. Arthur Brief and James P. Walsh, 225–67. New York: Erlbaum.
- Keohane, Robert O. 2002. *Power and governance in a partially globalized world*. London: Routledge.
- Kettl, Donald F. 1993. *Sharing power: Public governance and private markets*. Washington, DC: Brookings Institution Press.
- Lijphart, Arend. 1999. *Patterns of democracy*. New Haven, CT: Yale Univ. Press.
- Mandell, Myrna P., and Robin Keast. 2008. Evaluating the effectiveness of interorganizational relations through networks. *Public Management Review* 10:716–31.
- Manning, Nick. 2001. The legacy of the new public management in developing countries. *International Review of Administrative Sciences* 67:297–312.
- Melkers, Julia, and Katherine Willoughby. 2005. Models of performance-measurement use in local governments. *Public Administration Review* 65:180–90.
- Moynihan, Donald P. 2008. *The dynamics of performance management: Constructing information and reform*. Washington, DC: Georgetown Univ. Press.

- . 2009. Through a glass darkly: Understanding the effects of performance regimes. *Public Performance & Management Review* 32:586–98.
- Moynihan, Donald P., and Patricia W. Ingraham. 2004. Integrative leadership in the public sector: A model of performance information use. *Administration & Society* 36:427–53.
- Moynihan, Donald P., and Sanjay K. Pandey. 2010. The big question for performance management: Why do managers use performance information? *Journal of Public Administration Research and Theory* 20:849–66.
- O'Toole, Laurence J. 1997. The implications for democracy in a networked bureaucratic world. *Journal of Public Administration Research and Theory* 7:443–59.
- Piotrowski, Suzanne J., and David Rosenbloom. 2002. Nonmission-based values in results oriented public management. *Public Administration Review* 62:643–57.
- Post, Stephanie S. 2004. Metropolitan area governance and institutional collective action. In *Metropolitan governance: Conflict, competition, and cooperation*, ed. Richard C. Feiock, 67–94. Washington, DC: Georgetown Univ. Press.
- Provan, Keith G., and H. Brinton Milward. 2001. Do networks really work? A framework for evaluating public sector organizational networks. *Public Administration Review* 61:414–23.
- Radin, Beryl. 2006. *Challenging the performance movement: accountability, complexity and democratic values*. Washington, DC: Georgetown Univ. Press.
- Romzek, Barbara. 2008. *The tangled web of accountability in contracting networks*. Presented at the Kettering Foundation Symposium on Accountability, May 22–23, Dayton, OH.
- Rosenbloom, David H. 1983. Public administration theory and the separation of powers. *Public Administration Review* 43:219–27.
- . 2007. Reinventing administrative prescriptions: The case for democratic-constitutional impact statements and scorecards. *Public Administration Review* 67:28–39.
- Schick, Allen. 1998. Why most developing countries should not try New Zealand's reforms. *The World Bank Research Observer* 13:123–31.
- Stivers, Camilla. 2008. *Governance in dark times: Practical philosophy for public service*. Washington, DC: Georgetown Univ. Press.
- Soss, Joe, Richard Fording, and Sanford Schram. Forthcoming. The organization of discipline: From performance management to perversity and punishment. *Journal of Public Administration Research and Theory*.
- Taussig, Michael. 1999. *Defacement: Public secrecy and the labor of the negative*. Stanford, CA: Stanford Univ. Press.
- Thomson, Ann Marie., James L. Perry, and Theodore K. Miller. 2009. Conceptualizing and measuring collaboration. *Journal of Public Administration Theory and Research* 19:23–56.
- Transparency International. 2007. Global Corruption Report 2007. <http://www.transparency.org/publications/publications> (Retrieved November, 2007).
- United Nations Project Office on Governance. 2008. *A management capacity framework for local governments to strengthen transparency in local governance in Asia*. Project report to the United Nations Department of Economic and Social Affairs.
- Van de Walle, Steven. 2008. Comparing the performance of national public sectors: Conceptual problems. *International Journal of Productivity and Performance Management* 57:329–38.
- Van Slyke, David M. 2007. Agents or stewards: Using stewardship theory to understand the government-nonprofit social service contracting relationship. *Journal of Public Administration Research and Theory* 17:157–87.
- Wichowsky, Amber, and Donald P. Moynihan. 2008. Measuring how administration shapes citizenship: A policy feedback perspective on performance management. *Public Administration Review* 68: 908–20.
- World Bank. 2007. Governance matters 2007: Worldwide governance indicators 1996–2006. <http://info.worldbank.org/governance/wgi2007/> (accessed August 2007).
- Yang, Kaifeng, and Marc Holzer. 2006. The performance-trust link: Implications for performance measurement. *Public Administration Review* 66:114–26.