

Donald P. Moynihan
University of Wisconsin–Madison
Stéphane Lavertu
The Ohio State University

Does Involvement in Performance Management Routines Encourage Performance Information Use? Evaluating GPRA and PART

Donald P. Moynihan is professor in the La Follette School of Public Affairs at the University of Wisconsin–Madison. His research examines performance management, homeland security, election administration, and employee behavior. He is a fellow of the National Academy of Public Administration, the author of *The Dynamics of Performance Management* (Georgetown University Press, 2008), and the winner of the best book award from the Academy of Management's Public and Nonprofit Division in 2009.
E-mail: dmoynihan@lafollette.wisc.edu

Stéphane Lavertu is assistant professor in the John Glenn School of Public Affairs at The Ohio State University. His research examines the design and policy-making authority of federal agencies, the implementation of performance management reforms, the impact of deadlines on agency rulemaking, the role of federal advisory committees in executive branch policy making, and the impact of school choice on K–12 education.
E-mail: lavertu.1@osu.edu

The Government Performance and Results Act (GPRA) of 1993 and the George W. Bush administration's Program Assessment Rating Tool (PART) established new routines that were intended to foster performance management practices. Using data from two surveys, the authors find that the involvement of agency employees in GPRA processes and PART reviews generally had little direct effect on performance information use once other factors are accounted for. The main exception is that managerial involvement in GPRA processes and PART reviews is associated with the use of performance data to refine measures and goals. This reflects the limits of government-wide reform efforts that depend on difficult-to-observe bureaucratic behavior. The authors also find that a series of organizational factors—leadership commitment to results, learning routines led by supervisors, the motivational nature of the task, and the ability to link measures to actions—are positive predictors of performance information use.

This article offers some theoretical and empirical insights into the relationship between results-based reforms and managerial use of performance data. Performance information use is not just an explicit goal of policymakers. It also has been characterized as “the big question” for the performance management movement (Moynihan and Pandey 2010) and the best way to judge whether this movement has been successful (Van de Walle and Van Dooren 2008).

This article offers some theoretical and empirical insights into the relationship between results-based reforms and managerial use of performance data.

Organization theory suggests that behavioral change among employees can be fostered by altering their routines (Adler, Goldoftas, Levine 1999; Levitt and March 1988), especially in public sector environments, which generally offer low extrinsic benefits to accept change. To identify the causal mechanisms by which administrative reforms matter, we need to identify how they create or alter organizational routines. GPRA and PART, in different ways, both established organizational routines of data collection, dissemination, and review. We examine whether involvement in these routines is correlated with a variety of forms of performance information use, while accounting for other factors that might influence performance information use. These issues remain relevant because current reform efforts such as the GPRA Modernization Act continue to be premised on the notion that establishing performance management routines is central to promoting performance information use. The key difference from the earlier reforms is the type of routines employed.

There is something of an irony (or a tragedy, depending on your point of view) in the progress of performance management in the U.S. federal government. Reforms promise to make public managers more performance oriented, subsequently are criticized for failing to do so, and then are succeeded by a new wave of reform that promises to do the same, and similarly struggle. The Government Performance and Results Act (GPRA) was passed in 1993, at least in part, to foster performance information use. It was subsequently criticized by the George W. Bush and Barack Obama administrations and by Congress for failing in this task. The Bush administration also set itself the goal of fostering performance information through the Program Assessment Rating Tool (PART). It, too, failed, said the Obama administration, which abandoned PART. Both the Obama White House and Congress, by way of the GPRA Modernization Act of 2010, have explicitly set the goal of greater performance information use. Recent history does not confer optimism.

We estimate a series of models using data from surveys of federal employees administered by the U.S. Government Accountability Office (GAO) in 2000 and 2007. These surveys have influenced policy judgments about the progress of federal performance management (OMB 2001, 27; U.S. Senate 2010, 11–12), but previous analysis of these data has been limited

to GAO discussion of descriptive statistics (see, e.g., GAO 2008). Dull (2009) used earlier GAO data to offer an insightful study of the effects of leadership commitment. Our data set provides some additional analytical leverage over that used by Dull, as it coincides with the Bush-era PART, includes agency controls, and allows us to examine other organizational and political variables not previously tested.

The first section of the article examines the development of the GPRA and PART, highlighting the consistent desire to employ these reforms to increase performance information use. The second section examines the types of routines that GPRA and PART established, and the third section looks at relevant organizational and environmental factors that might also affect performance information use. In the fourth section, we describe the nature of the data and our statistical method. Finally, we review and discuss the results, which indicate that the involvement of agency employees in GPRA processes and PART reviews generally has not fostered greater performance information use once other factors are accounted for. The main exception is that managerial involvement in GPRA and PART is associated with the use of performance data to further refine performance measures and goals, which may be a reflection of the limits of government-wide reform efforts to induce performance management. We also find that a series of organizational factors—leadership commitment to results, the existence of learning routines led by supervisors, the motivational nature of task, and the ability to infer meaningful actions from measures—are all positive predictors of performance information use.

The Goal of Performance Information Use in Federal Reforms

GPRA underpins the contemporary framework for performance management in the U.S. federal government (Breul 2007). The designers of GPRA learned from the failures of previous results-based management reforms (Joyce 2003). Unlike previous executive branch initiatives, GPRA was set in statute and enjoyed bipartisan support. Despite some lofty rhetoric and goals, GPRA had relatively modest requirements, asking only that agencies routinely create strategic goals and disseminate performance data.

One goal of GPRA was to foster performance information use among federal managers. The preamble to the act notes that one of its purposes was to “help Federal managers improve service delivery, by requiring that they plan for meeting program objectives and by providing them with information about program results and service quality.” The Senate Committee on Government Affairs (1993) report on the legislation was even more explicit:

[U]se of performance measurement is still the exception rather than the norm in American government organizations ... The Committee realizes that, as valuable as the goal-setting and performance reporting under S.20 [GPRA] will be to Congress in its policymaking and oversight roles, its greatest value will likely be to the program managers

themselves ... Managers will use this information throughout the year to plan their activities and guide their subordinates. The Committee believes that S.20 has the potential to be a powerful tool for strengthening governmental management.

If GPRA established a bipartisan statutory framework for performance management, then the Bush administration’s President’s Management Agenda built on it (Breul 2007; OMB 2001). The Bush administration characterized GPRA as a well-intentioned but ultimately ineffective tool for performance management (Dull 2006, 17). The President’s Management Agenda stated, “After eight years of experience [since passage of GPRA], progress toward the use of performance information for program management has been discouraging ... Performance measures are insufficiently used to monitor and reward staff, or to hold program managers accountable” (OMB 2001, 27). The Office of Management and Budget (OMB) under President Bush sought to remedy this problem by creating PART (Joyce 2011). The OMB used PART to systematically evaluate federal programs. Specifically, it used PART to grade federal programs on a scale from effective to ineffective according to four criteria (program purpose and design, strategic planning, program management, and program results/accountability), weighting scores on those categories to assign programs an overall score (Breul 2007). PART evaluations were conducted in waves between 2003 and 2008 until nearly all programs were evaluated.

Congress remained supportive of GPRA, indicated by the bipartisan passage of the GPRA Modernization Act of 2010. This new version of GPRA pushed agencies to provide performance data on a quarterly rather than an annual basis and to use those data or else risk closer OMB and congressional oversight (Kamensky 2011). The revised version of GPRA also incorporated some aspects of the performance management framework that the Bush administration had established by executive order, such as creating the posts of chief operating officer and performance improvement officer in all agencies.

The new legislation represents an implicit criticism of the limitations of the original design of GPRA. It suggests that government agencies need to provide more timely data. It also suggests that agencies need to more actively use those data. The act requires quarterly reviews of data by the chief operating officer, performance improvement officer, and other key managers and calls for assigning responsibility for specific goals to individual managers. The new law empowers the OMB to review agency performance updates and identify unmet goals. Agencies must submit performance improvement plans to the OMB, and then to Congress, if goals continue to be classified as unmet. The Senate report on the new version of GPRA is explicit in its explanation that these new practices are “aimed at increasing the use of performance information to improve performance and results. ... Agencies are collecting a significant amount of information, but are not consistently using that information to improve their management and results” (2010, 11–12).

The Obama administration has echoed Congress's concerns about agencies' failure to use performance data, and it has prioritized performance information use as a key benchmark for performance management. One of three primary performance strategies articulated by the Obama administration is to "use performance information to lead, learn, and improve outcomes" (OMB 2011, 73). The Obama administration explicitly criticized GPRA and PART for failing in this regard:

The ultimate test of an effective performance management system is whether it is used, not the number of goals and measures produced. Federal performance management efforts have not fared well on this test. The Government Performance and Results Act of 1993 (GPRA) and the Performance Assessment Rating Tool (PART) reviews increased the production of measurements in many agencies, resulting in the availability of better measures than previously existed; however, these initial successes have not lead [sic] to increased use. (OMB 2011, 73)

A clear assumption of these reforms, and the performance management movement more generally, is that performance information use is a worthy goal, an unambiguous good. Reformers assume that reforms will lead to what has been labeled as a *purposeful* performance information use, that is, use that is likely to result in efficiency and effectiveness gains. But there are at least three other broad types of use that we might expect (Moynihan 2009). Managers may adopt a *passive* approach, doing the minimum required to comply with the procedural requirements of performance systems but not using data (Radin 2006). They may engage in *perverse* performance information use, using data in ways that are detrimental to goals even if some actual performance measures increase (e.g., as a result of cream-skimming or goal displacement) (Barnow and Heinrich 2010). Or managers might engage in *political* performance information use, employing data to advocate for the legitimacy and resource needs of a program (Moynihan 2008).

In this article, we focus on passive and purposeful forms of performance information use, largely because this is the type of use solicited in the secondary data set that we examine. Measures of purposeful use relate to whether managers use performance data in program management, problem solving, and employee management. Measures of passive use reflect whether managers simply use data to refine measures and goals in order to satisfy the procedural requirements of GPRA and PART.¹

Using Meta-Routines to Create a Community of Performance Managers

If both GPRA and PART were intended to foster performance information use, how were they to do so? Theories of organizational learning begin with a standard assumption in much of organization theory: that employee behavior is shaped by organizational routines. The importance of routines is especially high in organizational settings in which there are no strong extrinsic incentives toward action, such as the public sector. Employees look to a "logic of appropriateness" to guide their action, and a central reference point is the

organizational routines that they are asked to devote their time to (Levitt and March 1988). Shaping behavior, therefore, requires altering the basic routines that employees engage in.

If we are to take reforms such as GPRA and PART seriously, we need to understand the nature of the organizational routines that they establish and the way in which they involve employees and then identify the plausible causal links between routines and behavior. Routines can be studied as forms of behaviors, rules, or general dispositions (Becker 2005), but to qualify as routines, they must reflect regularities of interaction. GPRA and PART clearly fit into the tradition of examining routines in terms of rules that are intended to alter behavior. This is perhaps the easiest methodological way to study routines as long as rules are formalized, but the question of how such rules alter determine behavior is an empirical one (Becker 2005, 820).

GPRA and PART also qualify as meta-routines, that is, routines that are intended to alter existing routines or create new ones, usually in the name of regularizing the idiosyncrasies of the creative process to achieve improvement (Adler, Goldoftas, and Levine 1999). Large-scale organizational reforms intended to create new ways of acting, such as total quality management (Hackman and Wageman 1995), qualify as such routines. While GPRA and PART did not occur with the same regularity as daily routines, they were intended to alter such routines by engendering commitment to performance. Rather than depend on synchronicity or creative individuals to foster performance information use, they seek to regularize this use. Both reforms created routines of data creation, dissemination, and review involving large numbers of federal employees. By doing so, these reforms might be expected to build a community of performance information users.

We are interested in whether involvement in GPRA and PART routines is associated with greater performance information use. The type of managers most likely to have experienced such involvement is similar for both reforms: agency managers who are responsible for performance measurement, planning, and evaluation. But because any manager involved in goal setting for a program could be involved, the pool of possible participants was large. Involvement means slightly different things for each reform, which we examine in turn.

GPRA established routines whereby agency officials created performance measures in the hope that this would foster a desire to use the data that emerged. The Act obliged agencies to consult with congressional members and stakeholders in the strategic planning process and to give Congress annual performance reports. GPRA routines were continuous, predictable, and repetitive, with annual requirements to develop performance plans and reports and a three-to-five-year schedule for updating strategic plans. In all, 53 percent of the managers surveyed by the GAO reported that they had been involved in GPRA processes, either in terms of developing measures, assessing the quality of measures, or analyzing whether specific performance goals or broader strategic goals were being met (see table 1 for the specific wording of the question).

Table 1 Data Description

Variable	Activity Description	2000		2007	
		N [range]	Mean (SD)	N [range]	Mean (SD)
DEPENDENT VARIABLES: INFORMATION USE					
Variables capture the extent to which respondents reported using performance information for a particular set of activities. Responses are 0 = to no extent; 1 = to a small extent; 2 = to a moderate extent; 3 = to a great extent; 4 = to a very great extent.					
Priorities	Setting program priorities	2,089	2.51 (1.12)	2,591	2.66 (1.05)
Resources	Allocating resources	2,096	2.44 (1.09)	2,543	2.62 (10.6)
Problems	Identifying program problems to be addressed	—	—	2,627	2.71 (1.04)
Correction	Taking corrective action to solve program problems	—	—	2,631	2.70 (1.06)
Processes	Adopting new program approaches or changing work processes	2,113	2.42 (1.08)	2,625	2.58 (1.06)
Measures	Refining program performance measures	2,049	2.29 (1.11)	2,519	2.46 (1.11)
Goals	Setting new or revising existing performance goals	2,074	2.44 (1.09)	2,534	2.59 (1.10)
Expectations	Setting individual job expectations for the government employees that the respondent manages or supervises	2,081	2.40 (1.11)	2,568	2.70 (1.03)
Rewards	Rewarding government employees that the respondent manages or supervises	2,074	2.39 (1.16)	2,556	2.66 (1.06)
Predictor Variables					
REFORM ROUTINES					
PART	Whether (1) or not (0) respondent reported any involvement in "PART-related activities (This includes any involvement in preparing for, participating in, or responding to the results of any PART assessment.)"	—	—	2,937 [0,1]	0.31 (0.46)
GPR	Whether (1) or not (0) respondent reported any involvement during the past three years in "any of the following GPR-related activities: (Developing ways to measure whether program performance goals are being achieved; Gathering and analyzing data to measure whether programs are meeting their specific performance goals; Using measures for program performance goals to determine if the agency's strategic goals are being achieved; Assessing the quality of data used in measuring performance)."	2,507 [0,1]	0.53 (0.50)	—	—
RESPONDENT CHARACTERISTICS					
SES	Whether (1) or not (0) respondent was a member of the Senior Executive Service (SES) "or equivalent"	2,507 [0,1]	0.18 (0.39)	2,937 [0,1]	0.20 (0.40)
Supervisor Yrs	Number of years (from four ranges) respondent reported serving as a supervisor	2,478 [1-4]	2.71 (1.02)	2,891 [1-4]	2.49 (1.13)
ORGANIZATIONAL FACTORS					
Leadership commitment to results	Extent to which respondents agreed that his or her "agency's top leadership demonstrates a strong commitment to achieving results."	2,492	3.43 (1.18)	2,911	3.83 (1.16)
Measures link to action	Extent to which respondents indicated having difficulty "determining how to use performance information to improve the program." (reversed)	2,469	2.59 (1.22)	2,903	2.69 (1.26)
Measurement difficulty Cronbach's alpha = .82	Index that averages responses to four items inquiring about the extent to which difficulty creating performance measures hinders information collection and use: "Difficulty determining meaningful measures; Different parties are using different definitions to measure performance; Difficulty obtaining valid or reliable data; Difficulty obtaining data in time to be useful"	2,449	2.90 (1.07)	2,853	2.81 (1.08)
Discretion	"Agency managers/supervisors at my level have the decision making authority they need to help the agency accomplish its strategic goals."	2,497	3.02 (1.12)	2,922	3.16 (1.14)
Political conflict	Extent to which respondents indicates difficulty "resolving conflicting interests of stakeholders, either internal or external."	2,471	2.61 (1.43)	2,906	2.53 (1.38)
Learning routine (2000)	"The individual I report to periodically reviews with me the results or outcomes of the program(s)/operation(s)/project(s) that I am responsible for."	2,475	2.26 (1.19)	—	—
Task is motivational	"It is easy to motivate employees to be more results-oriented in the program(s)/operation(s)/project(s) I am responsible for."	2,476	1.57 (1.17)	2,900	1.87 (1.23)
Learning routine (2007) Cronbach's alpha = .74	Index that averages responses to three statements: "Managers and supervisors at my level pay attention to their agency's use of performance information in management decision making; The individual I report to periodically reviews with me the results or outcomes of the program(s)/operation(s)/project(s) that I am responsible for; Agency managers/supervisors at my level effectively communicate performance information on a routine basis."			2,780 [0-5]	3.20 (.92)
OMB attention				2,913	2.16 (1.92)
Congressional attention				2,907	1.80 (1.73)
Auditor attention	Extent to which respondents believe that the stakeholders identified "pay attention to their agency's use of performance information in management decision making."			2,914	2.21 (1.87)

The Bush administration judged that agencies needed a stronger central actor to oversee federal performance management. The OMB had been given a limited role in GPRA and did not feature prominently in President Clinton's reinventing government initiative. But it was given a primary role in designing the Bush management agenda (OMB 2001) and PART. While GPRA pushed agencies to consult with a wide array of actors, the PART process was essentially a dialogue between the OMB budget examiners and agency representatives (Dull 2006; Joyce 2011; Moynihan 2008). Agencies shouldered the burden of proof in justifying their performance, but the PART process gave the OMB the final decision on whether a program was deemed effective or not. While 1,015 PART evaluations took place over a five-year period, the review process was experienced by fewer agency actors than the GPRA process, as it was a program-by-program analysis. Just over 31 percent of the managers surveyed responded that they had been involved in PART-related activities, including any involvement in preparing for, participating in, or responding to the results of any PART assessment. Even though PART was experienced by fewer managers and with less frequency than GPRA, the involvement of the OMB and the connection to the budget process gave agency officials strong incentives to take PART seriously (GAO 2005).

While there are real differences in the nature of the reforms, the causal logic of how each reform would foster performance information use was similar. Even though the PART review process established more rigorous and adversarial routines by introducing a role for the OMB, both reforms were premised on the notion that performance data needed to be created, shared, and reviewed by a political principal. For GPRA, the principal was Congress. Indeed, Rosenbloom argues that GPRA represents a "quantum leap" (2000, 43) in inculcating congressional control of agencies. For PART, the principal was the OMB and, by extension, the president. Relative to the fragmented decision processes of Congress, the OMB could behave as a unitary actor able to make clear judgments on performance.

By requiring data dissemination, GPRA established a basis by which data transparency, and its ensuing effects on program reputation, could be expected to foster use. PART also employed the basic logic that transparency matters, but it sought to make public assessments of performance more understandable, by boiling down program assessments to an effective-ineffective scale, and more visible, by attaching them to the president's budget proposals and placing them on a central government Web site.

The main difference between GPRA and PART was not so much in the causal mechanism through which they would promote performance information use, but in the vigor with which this mechanism would be employed. Both reflected a belief that formal government-wide routines would create a community of performance information users. This article seeks to test this logic, as represented by the following broad hypothesis:

Hypothesis 1: Managerial involvement in formal performance routines will promote performance information use.

The main difference between GPRA and PART was not so much in the causal mechanism through which they would promote performance information use, but in the vigor with which this mechanism would be employed.

Policy-makers have criticized GPRA and PART for having failed to encourage use, but they also largely accept that the underlying causal theory of these reforms is sound. With a little more vigorous implementation—such as the more timely data collection and quarterly assessments required by the GPRA Modernization Act—the holy grail of widespread performance information use is believed to be achievable. The current critique of GPRA and PART seen among policy-makers is not directed at the implicit theory of the reforms, but what sort of routines will be effective in fostering information use.

Scholarly assessments of GPRA and PART are mixed. Case analyses indicate that the effects of the GPRA on performance information use were quite limited (Radin 2006). Dull (2009) finds that involvement with the GPRA had a positive impact on performance information use in a 2000 survey of managers, but not in a 1997 survey. PART did influence budget formulation within the White House (Gilmour and Lewis 2006), and qualitative analyses suggest that PART was taken at least somewhat seriously by agency actors, though it often was viewed as a burdensome imposition by the OMB (Moynihan 2008). However, PART did not appear to significantly influence budget decisions in Congress (Heinrich 2012). Congressional staff tended not to use PART data, preferring their own sources of information (Frisco and Stalebrink 2008; Moynihan 2008; Stalebrink and Frisco 2011). Prior research does not systematically examine the relative effects of GPRA and PART on performance information use across a wide range of federal employees, as we do here.

Organizational and Environmental Factors

In addition to formal reform initiatives, what other factors might affect performance information use? Moynihan and Pandey (2010) point to a range of factors: individual beliefs, job and organizational attributes, and external political influences. The GAO survey items do not enable us to account for all of them. In particular, there are few items that enable us to account for managers' beliefs, and high correlations between responses to some survey items limit the number of factors that we can account for simultaneously. Nevertheless, the data allow for the estimation of the impacts of a range of relevant organizational and environmental variables, including some that are particularly relevant to recent efforts to foster performance information use. Here we focus on agency leadership, learning routines, the capacity to link measures to actions, measurement difficulty, and the motivational nature of task. Table 1 provides full measurement details.

There is general agreement that supportive leadership fosters performance information use (Melkers and Willoughby 2005; Moynihan and Ingraham 2004). Leadership can play an indirect role by putting in place the conditions helpful to performance information use. One study finds that that transformational leadership is associated with higher goal clarity and developmental culture, which, in turn, is associated with higher use (Moynihan, Pandey, and Wright 2012). Most pertinent for the purposes of this analysis, Dull (2009), using 1997 and 2000 GAO data, finds that perceived leadership commitment is significantly and positively associated with use. We include

this same measure—the extent to which respondents agree that their “agency’s top leadership demonstrates a strong commitment to achieving results”—in our analysis.

One criticism of performance management systems is that while they have excelled at creating organizational routines for data collection and dissemination, they have been less successful at creating routines for the use of these data (Moynihan 2008). Organizational learning theory argues that learning is not just a function of organizational culture, but can be actively encouraged through structured practices or routines (Lipshitz, Popper, and Oz 1996). In performance management, learning routines can come in different forms. For example, the “stat” movement relies on senior managers or elected officials routinely querying managers in a data-rich environment (De Haven-Smith and Jenne 2006). Askim, Johnsen, and Christophersen (2008) and Ammons and Rivenbark (2008) describe how benchmarking processes facilitate the use of performance data in local governments in Norway and the United States, respectively. Such routines are distinct from the types of routines that GPRA and PART established in that they are directly focused on considering how to use the data. Here, we examine whether managers participate in routines involving performance information, focusing particularly on the role of supervisors in these routines. Note that the measures are slightly different in the pooled 2000 and 2007 models and in the 2007 model (see table 1).

The nature of some programs makes them less amenable to using performance data in decision making. We examine three ways in which this might matter. First, it is more difficult to measure performance for some programs than others (Radin 2006). To capture this, we include an index that reflects respondents’ perceptions about the difficulty of determining meaningful measures, disagreement over measurement definition, and the difficulty obtaining valid, reliable, and timely performance data. Dull (2009) finds that a similar measure is associated with lower performance information use. Second, we examine the difficulty that respondents perceive in relating actions to program results. In some cases, there may be plenty of measures, but causal inference is difficult. Radin (2006) argues that production functions, in which there are clear links between action and outcomes, are most likely to benefit from performance data, in part because it is relatively easy to make causal inferences from performance data. Our measure indicates the extent to which respondents reported difficulty “determining how to use performance information to improve the program.” Third, we examine the impact of the motivational nature of task. This is based on responses to this statement: “It is easy to motivate employees to be more results-oriented in the program(s)/operation(s)/project(s) I am responsible for.” The item is a broad measure, and it does not tell us why the task setting is motivational. For example, it could be because of the nature of employees, or employee–managerial relationships, in the group assigned with the task. It may reflect the mission valence of the task itself, that is, the attractiveness or salience of an organization’s purpose or social contribution (Rainey and Steinbauer 1999). If employees believe that performance measures can facilitate greater achievement of the mission they care about, they may be more inclined to use those measures. Regardless of why the task is motivating, it remains an important factor to control for.

In addition to agency controls, we control for a number of other factors. We control for managerial discretion, which previous research suggests predicts performance information use (Moynihan and Pandey 2010). We also control for respondents’ perceptions of the organizational political environment. We control for perceived political conflict, which Dull (2009) finds is a negative predictor of performance information use in models estimated using the 1997 GAO survey data, but not for models employing the 2000 data. We also control for stakeholder interest in performance management, with measures of perceived attention paid to performance information use by key governmental stakeholders, including the OMB, congressional committees, and audit agencies such as the GAO itself.

Data and Methods

The GAO administered surveys in 1997, 2000, 2003, and 2007 to collect data on the implementation of performance management reforms in federal agencies. It administered the surveys to a random, nationwide sample of mid- and upper-level federal managers in the agencies covered by the Chief Financial Officers Act of 1990, and, in 2000 and 2007, it oversampled managers from certain agencies to facilitate agency comparisons. This enabled the GAO to provide us with agency identifiers, which can be used to control for unobserved differences between agencies that might affect the dependent variable and bias the results. Additionally, these two waves of the survey included a set of targeted questions about GPRA and PART. Thus, our analysis focuses on the 2000 and 2007 survey data. The survey response rate was 70 percent in both 2000 and 2007, ranging between 55 percent and 84 percent in the 29 agencies identified in the 2008 survey (GAO 2008, 26) and between 54 percent and 76 percent across 28 agencies surveyed in 2000 (GAO 2000, 5).

Table 1 summarizes the measures that we employ (and includes descriptive statistics). Rather than describe them all, we focus on the manner in which we categorize the dependent variables and the key measures for GPRA and PART. We report multiple measures of performance information use to gauge whether GPRA and PART involvement have differential effects across different types of use. We categorize the types of use as performance measurement (which reflects a passive form of use) and program management, problem solving, and employee management (which reflect more purposeful uses). Our measures of GPRA and PART involvement are dichotomous variables, based on a question asking whether the respondent had been involved in any GPRA- or PART-related activities. If respondents indicated any sort of involvement, they were coded with a 1, while those not involved were coded with a 0.²

We employ ordered probit regression models because of the categorical nature of the dependent variables. All models include agency indicators to capture agency fixed effects. The models also include control variables to account for employee perceptions regarding factors such as leadership commitment and hindrances to performance measurement. While agency fixed effects should control for some variation in these perceptions, the experience of individuals across programs and supervisors still may vary a great deal within agencies.

Results and Discussion

The variables measuring managerial involvement in GPRA and PART behave quite similarly in the pooled model specified in

Table 2 Comparing the Impact of Involvement in PART and GPRA Routines on Perceptions of Performance Information Use

	Passive Use		Purposeful Use				
	Performance Measurement		Program Management		Employee Management		
	Measures	Goals	Processes	Priorities	Resources	Expectations	Rewards
PART	0.30*** (0.06)	0.26*** (0.06)	0.12^ (0.06)	0.07 (0.06)	0.01 (0.06)	-0.05 (0.06)	-0.07 (0.06)
GPRA	0.25*** (0.06)	0.20** (0.06)	0.16* (0.06)	0.11^ (0.07)	0.04 (0.07)	-0.08 (0.07)	-0.03 (0.07)
2007	0.09 (0.06)	0.07 (0.06)	0.10^ (0.06)	0.05 (0.06)	0.07 (0.06)	0.22*** (0.06)	0.20** (0.06)
SES	0.14** (0.05)	0.13* (0.05)	0.02 (0.05)	0.04 (0.05)	0.01 (0.05)	-0.00 (0.05)	-0.01 (0.05)
Supervisor years	0.02 (0.02)	0.03 (0.02)	-0.01 (0.02)	0.02 (0.02)	0.01 (0.02)	-0.02 (0.02)	0.02 (0.02)
Leadership commitment to results	0.15*** (0.03)	0.18*** (0.03)	0.15*** (0.03)	0.14*** (0.03)	0.14*** (0.03)	0.10*** (0.03)	0.11*** (0.03)
Learning routine	0.08** (0.02)	0.08** (0.02)	0.11*** (0.02)	0.08*** (0.02)	0.08*** (0.02)	0.14*** (0.02)	0.12*** (0.02)
Task is motivational	0.19*** (0.02)	0.16*** (0.02)	0.17*** (0.02)	0.20*** (0.02)	0.22*** (0.02)	0.16*** (0.02)	0.19*** (0.02)
Measures link to action	0.12*** (0.02)	0.12*** (0.02)	0.15*** (0.02)	0.13*** (0.02)	0.13*** (0.02)	0.10*** (0.02)	0.11*** (0.02)
Measurement problems	-0.09** (0.03)	-0.08** (0.03)	-0.12*** (0.03)	-0.14*** (0.03)	-0.14*** (0.02)	-0.16*** (0.03)	-0.16*** (0.03)
Discretion	0.12*** (0.02)	0.14*** (0.02)	0.11*** (0.02)	0.12*** (0.02)	0.15*** (0.02)	0.09** (0.02)	0.08** (0.02)
Political conflict	0.01 (0.02)	0.00 (0.02)	0.05* (0.02)	0.10*** (0.02)	0.09*** (0.02)	0.06* (0.02)	0.04* (0.02)
N	2,814	2,823	2,880	2,789	2,832	2,841	2,837
Wald chi ²	657.55***	667.95***	627.46***	617.34***	644.31***	566.02***	588.37***
Pseudo R ²	0.09	0.10	0.10	0.09	0.10	0.08	0.08

Ordered probit regressions estimating the extent to which respondents reported using performance information for different functions. All models include indicator variables for 15 of the 16 agencies identified by both surveys. Coefficients for agency fixed effects and cut points are not reported because of space constraints. Significance levels are based on two-tailed z-tests or chi-square tests: *** $p < .001$; ** $p < .01$; * $p < .05$; ^ $p < .10$ (^ $p < .05$ for a one-tailed test).

table 2 in terms of coefficient sign, size, and statistical significance. This suggests that the effects of these reforms have been similar, with limited effects on performance information use. In table 2, involvement in GPRA is significantly and positively associated with four of the seven measures of use, and involvement in PART is positively associated with three of the seven measures of use. Table 3, which presents the results of models that employ only the 2007 data and includes variables capturing additional political factors, indicates that PART involvement is positively and significantly related to three of the nine measures of use.

Cumulatively, what do the models tell us about the impacts of involvement in GPRA and PART? Involvement in GPRA predicts the use of data for process changes and, to a lesser extent, for choosing program priorities. PART involvement also positively predicts using performance information to inform process changes (at the $p > .05$ level for a one-tailed test). But the areas in which we can have the greatest confidence about the impact of GPRA and PART involvement are those having to do with refining performance measures and program goals. The coefficients for variables capturing GPRA and PART involvement are much larger and achieve robust levels of statistical significance for these performance measurement functions, which represent relatively passive forms of performance information use.

These findings cause us to modify the broad hypothesis that performance routines promote performance information use and to adopt a more nuanced view. Such reforms appear to most directly encourage passive forms of performance information use, reflecting

further investment in improving performance measurement rather than using the data for actual program, resource, or employee management decisions.

There are two likely reasons why GPRA and PART succeeded in encouraging performance measurement but largely failed to encourage other types of performance information use. The institutional design of GPRA and PART were highly focused on goal and data creation. The basic purpose of GPRA was to create measures and goals that did not previously exist. While Bush administration officials characterized PART as being different, it too created a process that forced agencies to set more ambitious goals and generate better measures of performance. Indeed, agency interactions with the OMB surrounding PART reviews focused on what constituted acceptable goals and measures (Moynihan 2008). A GAO evaluation found that the majority of OMB recommendations made to agencies through PART had to do with “performance assessments, such as developing outcome measures and/or goals, and improving data collection” (GAO 2005, 22). The design of PART was such that if the OMB determined that a program lacked clear goals and measures, that program would receive poor scores on three of the four sections of the questionnaire (program purpose and design, strategic planning, and program results/accountability), making it all but certain that the program would be categorized as “ineffective” or “results not demonstrated.”

A second reason why GPRA and PART are more highly correlated with passive rather than purposeful performance information use is that performance measurement functions are the only ones that

Table 3 Comparing the Impact of Involvement in PART Routines on Perceptions of Performance Information Use (2007 Data)

	Passive Use		Purposeful Use						
	Performance Measurement		Program Management			Employee Management		Problem Solving	
	Measures	Goals	Processes	Priorities	Resources	Expectations	Rewards	Identification	Correction
PART	0.29*** (0.07)	0.17* (0.07)	.13^ (.067)	0.09 (0.07)	0.04 (0.07)	-0.03 (0.07)	-0.03 (0.07)	0.05 (0.07)	0.06 (0.07)
SES	0.28*** (0.07)	0.34*** (0.07)	.020 (.073)	0.09 (0.07)	0.07 (0.08)	0.05 (0.08)	0.02 (0.08)	0.04 (0.07)	0.05 (0.08)
Supervisor years	0.05^ (0.03)	0.05^ (0.03)	.035 (.028)	0.04 (0.03)	0.02 (0.03)	-0.01 (0.03)	0.06* (0.03)	0.02 (0.03)	0.03 (0.03)
Leadership commitment to results	0.10* (0.04)	0.15*** (0.04)	.16*** (.042)	0.13*** (0.02)	0.15*** (0.04)	0.12** (0.04)	0.07^ (0.04)	0.15*** (0.04)	0.13** (0.04)
Learning routine	0.32*** (0.05)	0.29*** (0.05)	.35*** (.048)	0.41*** (0.05)	0.40*** (0.05)	0.43*** (0.05)	0.41*** (0.05)	0.40*** (0.05)	0.40*** (0.05)
Task is motivational	0.15*** (0.03)	0.12*** (0.03)	.138*** (.035)	0.13*** (0.03)	0.19*** (0.04)	0.15*** (0.03)	0.17*** (0.03)	0.08* (0.03)	0.12** (0.04)
Measures link to action	0.13*** (0.03)	0.14*** (0.03)	.148*** (.032)	0.14*** (0.04)	0.11** (0.03)	0.07* (0.03)	0.12*** (0.03)	0.19*** (0.03)	0.16*** (0.03)
Measurement problems	-0.02 (0.04)	-0.04 (0.04)	-.046 (.041)	-0.04 (0.04)	-0.02 (0.04)	-0.08* (0.04)	-0.05 (0.04)	-0.03 (0.04)	-0.02 (0.04)
Discretion	0.07* (0.04)	0.13*** (0.04)	.019 (.036)	0.06 (0.03)	0.06 (0.04)	0.01 (0.04)	0.04 (0.04)	0.08* (0.04)	0.06^ (0.04)
Political conflict	-0.02 (0.03)	-0.01 (0.03)	.028 (.031)	0.06^ (0.03)	0.03 (0.03)	0.01 (0.03)	0.01 (0.03)	0.08* (0.03)	0.03 (0.03)
OMB attention	0.00 (0.03)	0.02 (0.03)	-.041 (.030)	-0.03 (0.03)	-0.05^ (0.03)	-0.05^ (0.03)	-0.05^ (0.03)	-0.04 (0.03)	-0.08** (0.03)
Congressional attention	-0.04 (0.03)	-0.03 (0.03)	.014 (.029)	0.01 (0.02)	0.01 (0.03)	0.00 (0.03)	0.04 (0.03)	0.02 (0.03)	0.01 (0.03)
Auditor attention	0.02 (0.03)	0.04 (0.03)	.031 (.028)	0.04 (0.03)	0.03 (0.03)	0.02 (0.03)	-0.03 (0.03)	0.03 (0.03)	0.05^ (0.03)
<i>N</i>	1,441	1,439	1,461	1,449	1,422	1,432	1,429	1,460	1,462
Wald chi ²	463.47***	510.14***	417.85***	446.22***	469.53***	395.32***	402.67***	452.10***	448.99***
Pseudo <i>R</i> ²	0.13	0.14	.12	0.13	0.13	0.12	0.12	0.13	0.13

All models are limited to 2007 data and include indicator variables for 28 of the 29 agencies identified. Coefficients for agency fixed effects and cut points are not reported due to space constraints. Significance levels are based on two-tailed z-tests or chi-square tests: *** $p < .001$; ** $p < .01$; * $p < .05$; ^ $p < .10$.

political principals could easily monitor. Behavior that is easily observable can be directed and, as a result, becomes the minimum required to demonstrate implementation. Congressional staffers or the OMB might not possess in-depth management programmatic knowledge and thus struggled to assess, and therefore enforce, managerial use of performance information in difficult-to-monitor contexts such as resource allocation, problem solving, and employee management. But congressional committees could easily examine whether performance plans included goals that they requested, and OMB officials kept track of whether the recommendations they made about new measures were followed (GAO 2005).

Table 2 reveals that all of the organizational variables generally are positively associated with all types of performance information use in models estimated using pooled 2000 and 2007 data. But, as illustrated in table 3, only a few of these variables are consistently significant when the analysis is limited to 2007 data. Sensitivity tests suggest that it is the inclusion of additional agency controls in the 2007 model, rather than the smaller *N*, that results in the lower number of significant variables. Four variables are consistently significant in both sets of models (leadership commitment, learning routines, task is motivational, and measures link to action), and we focus our discussion on these.

While our results suggest a limited impact of involvement in GPRA and PART routines on performance information use, this does not mean that routines do not matter. Some of the variables hint at ways that other types of routines may foster performance information use

and align encouragingly with recent reform efforts. In particular, the models point to the consistent importance of learning routines, where managers engage in a regular examination of performance data with peers and supervisors (Moynihan 2008). This is relevant to current practice. The “stat” approach to performance management involves embedding such routines, but effort generally has been focused on engaging senior managers in the process. The GPRA Modernization Act follows this approach by demanding that high-level agency managers engage in quarterly meetings on performance data. The findings here offer partial support for this approach. Given that GPRA has established routines of data creation and dissemination, the fostering of routines of data use is a logical next step. But the findings emphasize the particular importance of learning routines at lower levels, between managers, and with their supervisors. The GPRA Modernization Act’s high-level quarterly review of data is a different sort of routine—it is more formalized, with higher stakes that may result in perverse forms of performance information use (De Haven-Smith and Jenne 2006). On the other hand, such high-level routines may lead, in turn, to the adoption of less formal learning routines at lower levels. A case study of the progenitor of the “stat” approach, CompStat, found just such an effect (Chetkovich 2000).

The finding that perceived leadership commitment to agency performance encourages performance information use confirms a previous result (Dull 2009) and supports claims that the GAO has made based on these data (GAO 2000, 2008). It is also relevant to current practice: the GPRA Modernization Act formalized the Obama

administration's requirements that all agency leaders to commit to a small number of "high-priority goals" for which they would be held accountable. The logic of this approach is that the sheer volume of performance data produced by GPRA and PART made it difficult for leaders to commit to any goals. By limiting their attention to a handful of visible goals, the hope is to engender leadership commitment to achieving them, which, in turn, should trickle down to agency managers. The risk of this approach is that agencies pursue many goals, and a laser-like focus on a small number of high-priority goals may displace attention from others.

The abilities to link mission to measures and to make causal inferences are also consistently significant predictors of performance information use. This reflects the fact that while all programs may be awash with data, whether the data are insightful or not may vary quite a bit. Do the data clearly indicate what the appropriate action is? The answer to this question depends a good deal on the nature of the program. But it may also depend on the person or team asking the question, as characteristics such as experience, knowledge, craft, and ingenuity may make some work groups and individuals better able to interpret data. A key step for using data is for groups of organizational actors to engage in learning routines, addressing basic questions about cause and effect, and how data can inform future action.

The ability to motivate is also a significant predictor of performance information use. We find that managers who reported that it is easy to motivate employees to be results-oriented in their work setting also reported greater use of performance data. The measure does not tell us what it is about the setting that makes motivation easier, but, as with other factors, this might be subject to organizational influence. To the extent that mission, vision, goals, and measures can be presented in a way that is appealing, this may increase motivation to use data in decision making.

Finally, the models limited to 2007, the results of which appear in table 3, enable us to examine the impact of stakeholder attention on performance management. The only variable that approaches statistical significance is the one that captures perceived levels of OMB attention. It is a negative predictor of information use for seven of the nine management activities, and significant at $p > .1$ (for a two-tailed test) or better for four of these tests. The prominent impact of the OMB is perhaps not surprising given the role that it played in designing and implementing PART. What is most interesting is that the coefficient is consistently negative, except in models focused on passive performance information use. This result aligns with the more general finding that the impact of OMB-led PART during this period was to promote changes to performance goals and measures. Additional attention from the OMB actually seems negatively related to other forms of use, such as problem solving and employee management. It may be that in attending to the demands of PART, agency managers were less focused on using data for functions that the OMB could not observe. It may also be the case that the result reflects endogeneity in the model, that is, managers who did not use performance data in the first place were more likely to invite or perceive OMB attention.

Limitations in Assessing the Impact of the GPRA and PART

This section briefly describes sensitivity analyses and the limitations of our approach. Given the time and effort that has been devoted to

GPRA and PART, it is important to be clear about what the results do and do not tell us and to consider the risk of a Type II error—that is, that our theoretical and statistical approach led us to underestimate the real substantive effects of GPRA and PART involvement on performance information use. Indeed, it is important to note that some model specifications resulted in more positive estimated effects of GPRA and PART involvement. If one estimates the model in table 2 without agency fixed effects, for example, the coefficient for the effect of PART involvement on performance information use to improve processes and select program priorities reaches higher levels of statistical significance.³ And if we test the impact of PART and GPRA involvement while controlling only for agency and respondent level control variables (years as supervisor and Senior Executive Service status), the results would indicate that involvement in these routines is a significant and positive predictor for all measures of use, except for employee management. In other words, the variables representing GPRA and PART involvement have some correlation with performance information use, but that correlation is no longer significant once appropriate controls are included.

One possibility is that we have overcontrolled for other factors, to the point that we introduce multicollinearity in the model. However, tests do not indicate that excessive collinearity explains the limited impact of GPRA and PART involvement. We may also underestimate the effects of GPRA and PART reforms because we treat involvement as a direct cause of performance information use, when involvement may spur use through other factors for which we control in the statistical models. For example, it seems reasonable to expect that GPRA and PART have had a long-term positive impact on whether leadership is committed to performance management (GAO 2008) and that these reforms have reduced some of the measurement difficulties that programs face. Testing whether routines such as GPRA and PART work primarily through indirect influences is a different question from the one we ask here, but it worth pursuing with an alternative empirical approach, such as structural equation modeling.

Another potential mechanism by which GPRA and PART shape employee performance practices is by slowly reshaping norms and practices across time. For example, by gradually making more and better data available, these reforms might promote greater performance information use. Studies have shown that bureaucratic perceptions of data availability are associated with use—that is, bureaucrats who perceive that there is much performance data available are also more likely to report using that data (Bourdeaux and Chikoto 2008; Moynihan and Pandey 2010). However, there has been a judgment by policymakers that the increased availability of data over time has not led to use. The results of our model in table 2 suggest that this view is generally correct. A positive result for the 2007 variable implies higher performance information use in 2007 relative to 2000, controlling for other factors. This variable is significant for only two of the seven measures of performance information use (managing employee expectations and rewards) and at statistically marginal levels for one more (managing processes). Comparisons of descriptive statistics over an even longer time period offer a more negative view. The GAO compared the percentages of federal managers in 1997 and 2007 who reported having different types of performance measures to a great or very great extent. There was a clear and statistically significant increase over time. However, for

measures of performance information use, this pattern does not hold. In aggregate, federal managers who lived through the implementation of GPRA and PART did not report using performance data at a greater rate than their counterparts 10 years earlier, leading the GAO to conclude, “We have found that despite having more performance measures, the extent to which managers make use of this information to improve performance has remained relatively unchanged” (GAO 2008, 5).

Conclusion: Different Routines for Different Goals

Our analysis suggests that managerial involvement in GPRA and PART has had a direct impact on relatively few aspects of performance information use. This finding generally supports the claims of policymakers that GPRA and PART have not fulfilled their potential (GAO 2008; OMB 2001, 2011; U.S. Senate 2011). The results further suggest that despite their very real differences, government-wide performance reforms such as GPRA and PART have been most effective in encouraging passive forms of performance information use, that is, in directing employees to follow requirements to create and improve performance information. The routines that GPRA and PART created were centered on data creation, dissemination, and review and provided a mechanism to observe whether managers were creating or improving performance data according to the principal’s wishes.

The findings tell us something about the limits of any formal government-wide performance requirements to alter the discretionary behavior of individual managers when such behavior is difficult to monitor. It may be unrealistic to expect more from such reforms, as political principals lack in-depth knowledge of how programs are run. Congress and the president are left monitoring whether agencies are creating performance information, but managers retain significant discretion in their use of performance information.

Despite the difficulties experienced by GPRA and PART, current performance management policy continues to show faith in the assumption that altering organizational routines will foster bureaucratic performance information use. The central difference is that policymakers are now experimenting with a new set of routines. These new routines are focused less on data creation and dissemination and, it is hoped, will more directly encourage purposeful use. The quarterly performance reviews demanded by the GPRA Modernization Act establish internal organizational routines of data use, while the mandate that agencies must annually respond to OMB assessments of unmet goals institutes an intraorganizational routine. The Obama administration’s effort to commit leaders to a limited set of performance goals reflects a belief that such commitment is necessary to create the environment for such routines to prosper.

These new routines face problems, some the same and some different from those of GPRA and PART. The information asymmetry problem between principal and agent has not gone away. While the new routines seek a higher level of engagement from principals,

this engagement is costly, and the commitment of busy agency leaders and senior managers may erode over time. On the other hand, the new routines may succeed too well, engendering perverse forms of performance information use.

It is worth recalling the observation made by James R. Thompson (1999) about another federal reform, the reinventing government movement. Top-down reforms always struggle, but they have the best opportunities for success when they amplify or encourage preexisting patterns of behavior. The same is likely to be true of the success of performance management routines. In organizations in which there is already some interest in performance management, reforms to encourage use will be more likely to succeed because they will face little opposition or link to preexisting operational routines. The capacity of performance management meta-routines such as GPRA or PART to actually encourage purposeful performance information use may depend on their capacity to better link to and alter existing routines.

Our analysis suggests that managerial involvement in GPRA and PART has had a direct impact on relatively few aspects of performance information use.

Notes

1. This does not imply, however, that other types of use are not occurring among our sample. Most empirical work does not attempt to classify the type of use, but the implicit classification is purposeful. One recent article offers two advances on existing measurement of the concept by developing distinct scales for political and purposeful use and demonstrating that very broad responses about performance information use (e.g., I use information to make decisions) tends to align strongly with more specific measures of purposeful use (Moynihan, Pandey, and Wright, forthcoming).
2. The PART measure, but not the GPRA measure, features five points of variation on the level of involvement, but we chose to use a dummy variable in order to model GPRA and PART in equivalent fashion. If we use the ordinal PART scale, the scores for PART involvement become significant for some of the dependent variables in table 3 (*correction*) and significant at $p > .1$ (two-tailed) for others (*problems* and *priorities*). But with the exception of the *measures* dependent variable, the coefficient size for PART involvement is a good deal smaller than other significant independent variables (we used the listcoef program from Stata to make valid coefficient comparisons).
3. This change in significance is accompanied by a significant increase in N size because the GAO did not provide agency IDs for many organizations in their 2000 survey, and because there are not matches between some of the agency IDs in the 2000 and 2007 survey, which are therefore excluded from our analysis in table 2.

References

- Adler, Paul S., Barbara Goldoftas, and David I. Levine. 1999. Flexibility versus Efficiency? A Case Study of Model Changeovers in the Toyota Production System. *Organization Science* 10(1): 43–68.
- Ammons, David N., and William C. Rivenbark. 2008. Factors Influencing the Use of Performance Data to Improve Municipal Services: Evidence from the North Carolina Benchmarking Project. *Public Administration Review* 68(2): 304–18.
- Askim, Jostein, Åge Johnsen, and Knut-Andreas Christophersen. 2008. Factors behind Organizational Learning from Benchmarking: Experiences from Norwegian Municipal Benchmarking Networks. *Journal of Public Administration Research and Theory* 18(2): 297–320.

- Barnow, Burt S., and Carolyn J. Heinrich. 2010. One Standard Fits All? The Pros and Cons of Performance Standard Adjustments. *Public Administration Review* 70(1): 60–71.
- Becker, Markus C. 2005. A Framework for Applying Organizational Routines in Empirical Research: Linking Antecedents, Characteristics, and Performance Outcomes of Recurrent Interaction Patterns. *Industrial and Corporate Change* 14(5): 817–46.
- Bourdeaux, Carolyn, and Grace Chikoto. 2008. Legislative Influences on Performance Information Use. *Public Administration Review* 68(2): 253–65.
- Breul, Jonathan D. 2007. Three Bush Administration Management Reform Initiatives: The President's Management Agenda, Freedom to Manage Legislative Proposals, and the Program Assessment Rating Tool. *Public Administration Review* 67(1): 21–26.
- Chetkovich, Carol A. 2000. The NYPD Takes on Crime in New York City. Case Study no. CR16-00-1558.3, Harvard University, John F. Kennedy School of Government.
- De Haven-Smith, Lance, and Kenneth C. Jenne. 2006. Management by Inquiry: A Discursive Accountability System for Large Organizations. *Public Administration Review* 66 (1): 64–76.
- Dull, Matthew. 2006. Why PART? The Institutional Politics of Presidential Budget Reform. *Journal of Public Administration Research and Theory* 16(2): 187–215.
- . 2009. Results-Model Reform Leadership: Questions of Credible Commitment. *Journal of Public Administration Research and Theory* 19(2): 255–84.
- Frisco, Velda, and Odd J. Stalebrink. 2008. Congressional Use of the Program Assessment Rating Tool. *Public Budgeting and Finance* 28(2): 1–19.
- Gilmour, John B., and David E. Lewis. 2006. Assessing Performance Assessment for Budgeting: The Influence of Politics, Performance, and Program Size in FY 2005. *Journal of Public Administration Research and Theory* 16(2): 169–86.
- Hackman, J. Richard, and Ruth Wageman. 1995. Total Quality Management: Empirical, Conceptual, and Practical Issues. *Administrative Science Quarterly* 40(2): 309–43.
- Heinrich, Carolyn J. 2012. How Credible Is the Evidence, and Does It Matter? An Analysis of the Program Assessment Rating Tool. *Public Administration Review* 72(1): 123–34.
- Joyce, Philip G. 2003. *Linking Performance and Budgeting: Opportunities in the Federal Budgeting Process*. Washington, DC: IBM Center for the Business of Government.
- . 2011. The Obama Administration and PBB: Building on the Legacy of Federal Performance-Informed Budgeting. *Public Administration Review* 71(3): 356–67.
- Kamensky, John M. 2011. The Obama Performance Approach: A Mid-Term Snapshot. *Public Performance and Management Review* 35(1): 132–47.
- Levitt, Barbara, and James G. March. 1988. Organizational Learning. *Annual Review of Sociology* 14: 319–40.
- Lipshitz, Raanan, Micha Popper, and Sasson Oz. 1996. Building Learning Organizations: The Design and Implementation of Organizational Learning Mechanisms. *Journal of Applied Behavioral Science* 32(3): 292–305.
- Melkers, Julia, and Katherine Willoughby. 2005. Models of Performance-Measurement Use in Local Governments: Understanding Budgeting, Communication, and Lasting Effects. *Public Administration Review* 65(2): 180–90.
- Moynihan, Donald P. 2008. *The Dynamics of Performance Management: Constructing Information and Reform*. Washington, DC: Georgetown University Press.
- . 2009. Through a Glass Darkly: Understanding the Effects of Performance Regimes. *Public Performance and Management Review* 32(4): 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(4): 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(4): 849–66.
- Moynihan, Donald P., Bradley E. Wright, and Sanjay K. Pandey. 2012. Setting the Table: How Transformational Leadership Fosters Performance Information Use. *Journal of Public Administration Research and Theory* 22(1): 143–64.
- Moynihan, Donald P., Sanjay K. Pandey, and Bradley E. Wright. Forthcoming. Prosocial Values and Performance Management Theory: Linking Perceived Social Impact and Performance Information Use. *Governance*.
- Radin, Beryl A. 2006. *Challenging the Performance Movement: Accountability, Complexity, and Democratic Values*. Washington, DC: Georgetown University Press.
- Rainey, Hal G., and Paula Steinbauer. 1999. Galloping Elephants: Developing Elements of a Theory of Effective Government Organizations. *Journal of Public Administration Research and Theory* 9(1): 1–32.
- Rosenbloom, David H. 2000. Retrofitting the Administrative State to the Constitution: Congress and the Judiciary's Twentieth-Century Progress. *Public Administration Review* 60(1): 39–46.
- Stalebrink, Odd, and Velda Frisco. 2011. PART in Retrospect: An Examination of Legislator's Attitudes toward PART. *Public Budgeting and Finance* 31(2): 1–21.
- Thompson, James R. 1999. Designing Administrative Reform That Works: The Example of the Reinvention Lab Program. *Public Administration Review* 59(4): 283–93.
- U.S. Government Accountability Office (GAO). 2000. *Managing for Results: Federal Managers' Views Show Need for Ensuring Top Leadership Skills*. Washington, DC: U.S. Government Printing Office. GAO-01-127.
- . 2005. *Performance Budgeting: PART Focuses Attention on Program Performance, but More Can Be Done to Engage Congress*. Washington, DC: U.S. Government Printing Office. GAO-06-28.
- . 2008. *Government Performance: Lessons Learned for the Next Administration on Using Performance Information to Improve Results*. Washington, DC: U.S. Government Printing Office. GAO-08-1026T.
- U.S. Office of Management and Budget (OMB). 2001. *The President's Management Agenda*. Washington, DC: U.S. Government Printing Office.
- . 2011. *The President's Budget for Fiscal Year 2012: Analytical Perspectives*. Washington, DC: U.S. Government Printing Office.
- U.S. Senate. Committee on Governmental Affairs. 1993. Report on the Government Performance and Results Act of 1993. <http://www.whitehouse.gov/omb/mgmt-gpra/gprptm> [accessed April 9, 2012].
- U.S. Senate. Committee on Homeland Security and Governmental Affairs. 2010. Report on the GPRA Modernization Act of 2010. <http://www.gpo.gov/fdsys/pkg/CRPT-111srpt372/html/CRPT-111srpt372.htm> [accessed April 9, 2012].
- Van de Walle, Steven, and Wouter Van Dooren, eds. 2008. *Performance Information in the Public Sector: How It Is Used*. Houndmills, UK: Palgrave Macmillan.