



# EDUCATION

THE ARTS

CHILD POLICY

CIVIL JUSTICE

EDUCATION

ENERGY AND ENVIRONMENT

HEALTH AND HEALTH CARE

INTERNATIONAL AFFAIRS

NATIONAL SECURITY

POPULATION AND AGING

PUBLIC SAFETY

SCIENCE AND TECHNOLOGY

SUBSTANCE ABUSE

TERRORISM AND  
HOMELAND SECURITY

TRANSPORTATION AND  
INFRASTRUCTURE

WORKFORCE AND WORKPLACE

This PDF document was made available from [www.rand.org](http://www.rand.org) as a public service of the RAND Corporation.

[Jump down to document](#) ▼

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis.

## Support RAND

[Purchase this document](#)

[Browse Books & Publications](#)

[Make a charitable contribution](#)

## For More Information

Visit RAND at [www.rand.org](http://www.rand.org)

Explore [RAND Education](#)

View [document details](#)

## Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Unauthorized posting of RAND PDFs to a non-RAND Web site is prohibited. RAND PDFs are protected under copyright law. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see [RAND Permissions](#).

This product is part of the RAND Corporation monograph series. RAND monographs present major research findings that address the challenges facing the public and private sectors. All RAND monographs undergo rigorous peer review to ensure high standards for research quality and objectivity.

# Toward a Culture of Consequences

Performance-Based Accountability  
Systems for Public Services

---

Brian M. Stecher, Frank Camm, Cheryl L. Damberg,  
Laura S. Hamilton, Kathleen J. Mullen, Christopher Nelson,  
Paul Sorensen, Martin Wachs, Allison Yoh, Gail L. Zellman,  
with Kristin J. Leuschner



RAND EDUCATION

The research described in this report was conducted within RAND Education, a unit of the RAND Corporation, under a grant from a private philanthropic organization.

**Library of Congress Cataloging-in-Publication Data**

Toward a culture of consequences : performance-based accountability systems for public services / Brian M. Stecher ... [et al.].

p. cm.

Includes bibliographical references.

ISBN 978-0-8330-5015-1 (pbk. : alk. paper)

1. Government accountability—United States. 2. Organizational effectiveness—United States. 3. Performance—Management. I. Stecher, Brian M.

JK421.T79 2010

352.3'5—dc22

2010027552

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

**RAND**® is a registered trademark.

*Cover image courtesy Fotosearch*

© Copyright 2010 RAND Corporation

Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Copies may not be duplicated for commercial purposes. Unauthorized posting of RAND documents to a non-RAND website is prohibited. RAND documents are protected under copyright law. For information on reprint and linking permissions, please visit the RAND permissions page (<http://www.rand.org/publications/permissions.html>).

Published 2010 by the RAND Corporation

1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138

1200 South Hayes Street, Arlington, VA 22202-5050

4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665

RAND URL: <http://www.rand.org>

To order RAND documents or to obtain additional information, contact

Distribution Services: Telephone: (310) 451-7002;

Fax: (310) 451-6915; Email: [order@rand.org](mailto:order@rand.org)

## Summary

---

During the past two decades, performance-based accountability systems (PBASs), which link financial or other incentives to measured performance as a means of improving services to the public, have gained popularity in a wide range of service fields. There are many examples. In education, the No Child Left Behind Act of 2001 (NCLB) (Pub. L. 107-110) combined explicit expectations for student performance with well-aligned tests to measure achievement and strong consequences for schools that do not meet program targets. In child care, quality rating and improvement systems (QRISs) establish quality standards, measure and rate providers, and provide incentives and supports for quality improvement. In the transportation sector, cost-plus-time (A+B) contracting is used to streamline highway construction; in health care, there are more than 40 hospitals and more than 100 physician and medical group performance-based accountability (popularly dubbed pay-for-performance, or P4P) programs in place in the United States. There have also been recent efforts to create performance measures and establish consequences related to the nation's efforts to prevent, protect against, respond to, and recover from large-scale public health emergencies.

While PBASs can vary widely across sectors, they share three main components: goals (i.e., one or more long-term outcomes to be achieved), incentives (i.e., rewards or sanctions to motivate changes in individual or organizational behavior to improve performance), and measures (formal mechanisms for monitoring the delivery of services or the attainment of goals).

Today's PBASs grew out of efforts over many years and many countries to manage the private and public organizations that were growing too large to be overseen by a single manager who knew what everyone was doing. These innovative approaches focused on measuring *performance*, which was originally defined fairly narrowly. Over time, notions about what aspects of performance most mattered broadened and changed. By the 1980s, government organizations were linking performance to incentives in an effort to motivate and direct individual performance and improve organizational outcomes.

But while the use of PBASs has spread in the public sector, little is known about whether such programs are having the desired effect. Research suggests that PBASs influence provider behaviors, a first step toward achieving outcomes, but there is currently little evidence concerning the effectiveness of PBASs at achieving their performance goals, or the experiences of governments and agencies at the forefront of this trend. This monograph seeks to address the gap by examining several examples of PBASs, large and small, in a range of public service areas. This study examines nine PBASs that are drawn from five sectors: child care, education, health care, public health emergency preparedness (PHEP), and transportation (Table S.1). The cases we studied provide useful information on the formation, design, operation, and evaluation of PBASs.

The choice of cases was guided by practical as well as theoretical considerations. On the practical side, we wanted to take advantage of the expertise available at RAND, where empirical research is being conducted on a number of performance measurement and accountability systems in different service areas. On the theoretical side, we wanted to include cases in which services are provided primarily by public agencies (education, transportation), as well as sectors in which services are provided primarily by private organizations but in which the public sector has an important role in governance (child care, health care). We also wanted to include at least one instance in which measurement itself was a challenge (PHEP).

**Table S.1**  
**Cases Examined in This Study**

Sector	PBAS	Key Incentive
Child care	QRISs	Prestige associated with high rating Financial incentives
Education	NCLB	Graduated set of interventions regarding professional development, instruction, staffing, and school governance (e.g., constraints on use of funds)
	P4P	Cash bonuses, salary increases
Health care	Hospital and physician or medical group P4P programs, including quality “report cards”	Financial payments for high performance or improvement, public recognition, transparency (i.e., clarity and openness) of performance results
PHEP	CDC PHEP cooperative agreement	Withholding of federal funds for failure to meet performance benchmarks
Transportation	A+B highway construction contracting	Financial rewards or sanctions based on time to complete
	CAFE standards	Fines for failure to meet minimum average fuel-efficiency standards
	CAA ambient air pollution conformity requirements	Federal transportation funds subject to conformity with ambient air quality standards
	Transit subsidy allocation formulas	Share of state or regional funding for local transit operators

NOTE: CDC = Centers for Disease Control and Prevention. CAFE = Corporate Average Fuel Economy. CAA = Clean Air Act (Pub. L. 88-206 and its amendments).

## Research Approach

The research approach included a broad review of literature related to performance measurement and accountability, the development of an analytic framework to structure our internal discussions about research evidence in the five sectors, a 1.5-day integrative workshop that examined various features of PBASs (e.g., context in which the PBAS arose, measures, incentives, and evaluation approaches), and analysis of

sector-specific empirical results and identification of cross-sector principles. Through this process, we attempted to derive principles that might have general applicability beyond the cases we studied.

## Findings

Evidence on the effects of nine PBASs in five sectors shows that, under the right circumstances, a PBAS can be an effective strategy for improving the delivery of services to the public. Optimum circumstances include having the following:

- a goal that is widely shared
- measures that are unambiguous and easy to observe
- incentives that apply to individuals or organizations that have control over the relevant inputs and processes
- incentives that are meaningful to those being incentivized
- few competing interests or requirements
- adequate resources to design, implement, and operate the PBAS.

However, these conditions are rarely fully realized, so it is difficult to design and implement PBASs that are uniformly effective. The following sections highlight the major factors that influence PBAS development and effects in the cases we studied.

### **Decision to Adopt a Performance-Based Accountability System Is Shaped by Political, Historical, and Cultural Contexts**

In the cases we examined, the decision to adopt a PBAS was subject to multiple influences. In many sectors, the process was heavily influenced by the preferences of service providers—the very people whose behavior the PBAS sought to shape. In transportation, for instance, PBASs designed to improve local transit funding have often been strongly influenced by the local jurisdictions that are the subject of the PBASs. Given conflicts among stakeholders, it is perhaps not surprising that PBASs often proceed in spite of a lack of clear agreement on what constitutes performance and on who should be held accountable for



what. In many sectors, there is not a sufficiently strong evidence base to provide scientific guidance to would-be PBAS adopters and designers.

The creation of PBASs might be nurtured by the presence of a strong history and culture of performance measurement and accountability. In education, for instance, measurement of student performance has a long history in the United States, and standardized achievement tests are accepted as an indicator of performance for many purposes. However, such a history does not ensure the smooth adoption of a PBAS. Many PBASs, once created, exist in conflict with other PBASs and governance structures. This is especially the case in sectors with a long tradition of measurement and accountability in which service providers receive funds from multiple sources and through many funding mechanisms (e.g., transportation, health care, education).

### **Selection of Incentive Structures Has Proven Challenging**

PBAS designers face three basic design issues:

- determining whose behavior they seek to change (i.e., identifying individuals or organizations to target)
- deciding on the type and size of incentives
- measuring performance and linking these measures to the incentives they have chosen.

In the PBASs we examined, it was fairly easy in most cases to identify the individuals or organizations that are held accountable for improving service activities and reaching the PBAS goals. It has been more challenging, however, to decide which incentive structures to use to affect the desired behaviors.

Context can have a large effect on the incentive structures that PBAS designers choose. For example, when participation in a PBAS is voluntary, designers of PBASs typically use rewards rather than sanctions. We found that, when the designers of a PBAS worked within a regulatory setting (e.g., NCLB, PHEP), sanctions were more common. In contrast, designers of PBASs in which participation was voluntary—child care and A+B contracting, for example—tended to prefer rewards. The size and details of rewards vary widely across the

PBASs we studied. It is unclear how well the magnitude of rewards is correlated with the benefits of the changes that the PBAS designers seek to induce or the effort that service providers, such as doctors and teachers, must make to comply with these changes.

### **Design of Performance Measures Requires a Balance Among Competing Priorities**

The measures used to quantify performance can vary in many dimensions. PBAS designers must consider a number of competing factors when selecting and structuring measures:

- the feasibility, availability, and cost of measures
- the context within which a PBAS operates
- the alignment of measures with PBAS goals
- the degree of control of the monitored party
- resistance to manipulation by the monitored service activity
- understandability.

The selection of performance measures ultimately requires some trade-offs among these factors. PBAS designers seem to prefer measures that can be collected at low cost or that already exist outside the PBAS. To choose among potentially acceptable measures, a PBAS tends to balance two major considerations: the alignment of a measure with the PBAS's goals and the extent to which the individuals or organizations monitored by the PBAS have the ability to control the value of that measure. A natural tension arises from efforts to achieve balance between these objectives. Over time, the parties that a PBAS monitors might find ways to "game" the system, increasing their standing on a measure in ways that are not aligned with the PBAS goals. Perhaps the best-known example of such manipulation in the cases we examined is the act of "teaching to the test" in an educational setting.

Continuing vigilance and flexibility can help a PBAS manage this tension and maintain the balance between policymakers' priorities and the capabilities of the parties the PBAS monitors. Such a balance tends to be easier to achieve when the measures the PBAS uses are understandable and have been communicated to all parties.

### **Successful Implementation Must Overcome Many Potential Pitfalls**

Even a well-designed PBAS might not yield the desired results if it is not executed effectively. Our review of the literature and the nine cases identified several pitfalls that can occur on the implementation process:

- lack of PBAS experience and infrastructure
- unrealistic timelines
- complexity of the PBAS
- failure to communicate
- stakeholder resistance.

There are many strategies available to address these pitfalls. For example, when building a PBAS, exploiting the existing infrastructure, when possible, and implementing in stages can minimize both the time needed for implementation and the disruptive potential of mistakes before they can compound. Incorporating a pilot-testing phase can also head off a number of problems early. Communicating with stakeholders is also integral to the success of the PBAS, while formative monitoring can be important for identifying and correcting problems that occur during implementation.

### **Evidence of System Effectiveness Is Limited and Leads to Varying Conclusions by Sector**

In general, PBASs have not been subject to rigorous evaluation, and the evidence that does exist leads to somewhat different conclusions by sector:

- In education, it is clear that NCLB and other high-stakes testing programs with public reporting and other incentives at the school level have led to changes in teacher behavior; however, teachers seem to respond narrowly in ways that improve measured outputs (i.e., the measures) with less attention to long-term outcomes (i.e., the goals).<sup>1</sup> While student test scores have risen, there is uncer-

---

<sup>1</sup> We use the following terminology when talking about public service programs and their consequences: a program is a structured activity that transforms *inputs* into *outputs*, which are observable, measurable (e.g., blood pressure, test scores, parts per million of carbon diox-

tainty as to whether these reflect more learning or are to some degree the product of teaching to the test or other approaches to generating apparent improvement.

- In health care, relatively small financial incentives (frequently combined with public reporting) have had some modest effects in improving the quality of care delivered.
- Examples from the transportation sector suggest that large financial incentives can lead to creative solutions, as well as to lobbying to influence the demands of the PBAS regulation. The latter has been the case with the CAFE standards, which require automobile manufacturers to achieve a minimum level of fuel economy for the fleet of vehicles sold each year in the United States.
- It is too soon to judge the effectiveness of PBASs in child care and PHEP.

PBASs also have the potential to cause unintended consequences by incentivizing the wrong kind of behavior or encouraging undesirable effects. For example, in NCLB, attaching public reporting and other incentives to test scores has led to unintended behavioral changes (i.e., teaching to the test) that might be considered undesirable. In the transportation sector, some analysts have argued that CAFE standards prompted auto manufacturers to produce smaller and lighter vehicles, which, in turn, increased the number of crash-related injuries and fatalities, though this conclusion remains subject to some debate. A concern in the health-care sector is that PBASs include a narrow set of performance markers, which might increase physicians' focus on what is measured and reduce their attention to unmeasured effects. However, to date, there is an absence of empirical evidence showing such effects.

---

ide, or CO<sub>2</sub>), and easy to associate directly with the program. Ultimately, these outputs affect long-term *outcomes* that are of interest to policymakers (health, achievement, air quality). The outcomes might or might not be measurable, but it is typically difficult to draw a direct connection between the program and these outcomes. Many factors beyond the program's control or even understanding might affect the relationship between the program and the higher-level, broad outcomes relevant to policymakers. As a result, to influence behavior within a program with confidence, an accountability system must focus on measures of *outputs* that can be clearly attributed to the program.

If a PBAS does not initially meet its aims, it does not mean that a PBAS cannot be successful; it might just mean that some of the structural details require further refinement. PBASs are sufficiently complex that initial success is rare, and the need for modification should be anticipated.

## Recommendations for System Developers

We offer a number of recommendations for PBAS sponsors, designers, and other stakeholders to consider regarding PBAS design, incentives and performance measurement, implementation, and evaluation.

### Design of the Performance-Based Accountability System

Designing a PBAS is a complex undertaking, and many of the decisions that will need to be made are heavily dependent on sector-specific contextual circumstances.

**Consider the Factors That Might Hinder or Support the Success of a PBAS to See Whether Conditions Support Its Use.** The first issue is to consider whether a PBAS is the best policy approach for the policy concern at hand and whether it might be expected to succeed. From the cases examined, we identified a number of factors that tend to support a successful PBAS implementation:

- broad agreement on the nature of the problem
- broad agreement on PBAS goals
- knowledge that specific changes in inputs, structures, processes, or outputs will lead to improved outcomes
- ability of service providers, through changes in behavior, to exert significant influence on outputs and outcomes
- ability of the implementing organization to modify the incentive structure for service providers
- absence of competing programs that send conflicting signals to service providers
- political context in which it is acceptable for the PBAS to be gradually improved over time

- sufficient resources to create the PBAS and to respond to the incentives.

If a large share of these factors does not hold for the application under consideration, decisionmakers might wish to consider alternative policy options or think about ways to influence the context to create more-positive conditions for a PBAS.

**Be Sensitive to the Context for Implementation.** It is important to account for constraints and leverage opportunities presented by the context in which the PBAS will be implemented. Such considerations include the extent to which the implementing organization can alter the incentive structure faced by service providers, existing mechanisms that will affect the behavior of service providers (e.g., safety or licensing requirements) or that can be used to support the PBAS (e.g., data collection), and current knowledge of the service activity covered by the PBAS.

**Consider Applying Performance Measures and Incentives at Different Functional Levels.** If the service-delivery activities are organized hierarchically (e.g., students within classrooms within schools within districts), PBAS designers should consider the application of performance measures and incentives at different functional levels within the activity (e.g., different measures and incentives for school districts, school principals, and teachers or for hospitals, clinics, and doctors). Provided that the performance measures and incentives are structured in a complementary fashion, the results can be additive and mutually reinforcing.

**Design the PBAS So That It Can Be Monitored Over Time.** To obtain the best results over the long term, it is important to develop a plan for monitoring the PBAS, identifying shortcomings that might be limiting the effectiveness of the PBAS or leading to unintended consequences, and modifying the program as needed.

### **Incentives and Performance Measurement**

The selection of incentives and performance measures is of vital importance to the PBAS. The type and magnitude of the incentives will govern the level of effort providers expend to influence the performance

measures, while the measures will dictate the things on which the service providers should focus and what they might choose to ignore or neglect.

**Create an Incentive Structure Compatible with the Culture of the Service Activity.** Many options for incentives are available, including cash, promotions, status, recognition, increased autonomy, or access to training or other investment resources. The goal is to select options that will best influence behavior without undermining intrinsic service motivation.

**Make the Rewards or Penalties Big Enough to Matter.** The size of the incentive should outweigh the effort required by the service provider to improve on the performance measure; otherwise, service providers will simply not make the effort. However, the size of the incentives should not exceed the value obtained from improved provider behavior, since the PBAS would, by definition, not be cost-effective.

**Focus on Performance Measures That Matter.** Performance measures determine how service providers focus their efforts. To the extent possible, therefore, it makes sense to include those measures believed to have the greatest effect on the broader goals of interest.

**Create Measures That Treat Service Providers Fairly.** In certain settings, the ability of service providers to influence desired outputs might be limited. When selecting performance measures, PBAS developers should consider the degree to which service providers can influence the criteria of interest. Individuals or organizations should not be held accountable for things they do not control. In such cases, there are other options for creating performance measures that treat service providers fairly:

- Create “risk-adjusted” output measures that account for relevant social, physical, and demographic characteristics of the population served.
- Establish measures based on inputs, structure, or processes rather than on outputs or outcomes.
- Measure relative improvement rather than absolute performance.

**Avoid Measures That Focus on a Single Absolute Threshold Score.** Although the threshold approach can be intuitively appealing (in the sense that the specified score represents a quality bar that all service providers should strive to achieve), in practice, measures that focus on a single threshold can prove quite problematic. Low achievers with no realistic prospects for achieving the absolute threshold score will have no incentive to seek even modest improvements, while high achievers will have no incentive to strive for further improvement. Alternatives include use of multithreshold scores and measurement of year-over-year improvement.

### **Implementation**

It is possible to create a potentially effective design for a PBAS and then fail to implement the design successfully; thus, special attention needs to be paid to the way the PBAS is implemented.

**Implement the Program in Stages.** Because most PBASs are quite complex, it is often helpful to develop and introduce different components in sequence, modifying as needed in response to any issues that arise. For example, initial efforts and funding might focus on the development of capacity to measure and report performance, with measures and incentives rolled out over time. Pilot-testing might also be used to assess measures and other design features.

**Integrate the PBAS with Existing Performance Databases and Accounting and Personnel Systems.** A PBAS is not created in a void; rather, it must be incorporated within existing structures and systems. It is thus important to think through all of the ways in which the PBAS will need to interact with preexisting infrastructure—e.g., performance databases, accounting systems, and personnel systems. These considerations might suggest changes in the design of the PBAS or highlight ways in which the existing infrastructure needs to be modified while the PBAS is being created.

**Engage Providers, and, to the Extent Possible, Secure Their Support.** To garner the support of providers, it is helpful to develop measures that are credible (i.e., tied to outcomes about which they care), fair (i.e., that account for external circumstances beyond the control of providers), and actionable (i.e., that can be positively influenced



through appropriate actions by the service provider). A good approach is to involve providers in the process of developing the measures and incentives. While, to some degree, it can be expected that service providers will seek to weaken the targets or standards to their benefit, those responsible for implementing and overseeing the PBAS will need to judge whether lowering performance expectations would ultimately undermine the success of the PBAS.

**Ensure That Providers and Other Stakeholders Understand Measures and Incentives.** Communication is key. Particularly in cases in which there are numerous providers with varying internal support systems to enable engagement—as, for example, with health-care P4P systems and child-care quality ratings—it can be helpful to employ multiple communications channels (e.g., email, website, conference presentations) as appropriate.

**Plan for the Likelihood That Certain Measures Will “Top Out.”** As service providers improve their performance in response to the incentive structure, a growing percentage might achieve the highest possible scores for certain measures. PBAS designers should plan for this eventuality, e.g., by replacing topped-out measures with more-challenging ones or by requiring service providers to maintain a high level of performance for topped-out measures in order to qualify for incentives.

**Provide Resources to Support Provider Improvement.** It can be valuable to devote program resources to support efforts at improvement. This might involve infrastructure investments or education for providers on becoming more effective.

## Evaluation

Ironically, given the spirit of accountability embodied in the PBAS approach, most of the cases reviewed in this study have not been subjected to rigorous evaluation. We believe that it is vitally important to rectify this lack of evaluation. Only through careful monitoring and evaluation can decisionmakers detect problems and take steps to improve the functioning of the PBAS over time.

**Consider Using a Third Party to Evaluate the PBAS.** Not all organizations that implement a PBAS possess the necessary methodological expertise to conduct a sound programmatic evaluation. Additionally,

many implementing organizations, for understandable reasons, will tend to be biased in favor of positive results. For these reasons, it is beneficial to rely on an independent and qualified third party to conduct an evaluation of the PBAS.

**Structure the Evaluation of a PBAS Based on Its Stage of Development.** When a system is first developed, it might be most helpful to evaluate implementation activities (e.g., whether appropriate mechanisms for capturing and reporting performance measures have been developed). As the system matures, the focus should shift to evaluating the effects, in terms of observed provider behavior and service outputs, of the performance measures and incentive structure. An evaluation should focus on outputs only after performance measures and incentives have been in place long enough to influence behavior.

**Examine the Effects of the PBAS on Both Procedures and Outputs.** One approach for doing so is to develop a logic model, a visual representation of the ways in which the PBAS is intended to influence provider behavior. This model can then become the basis for thoughtful monitoring and evaluation and make it easier to plan the evaluation of a PBAS based on its stage of development.

**Use the Strongest Possible Research Design Given the Context in Which the PBAS Exists.** Options, sorted in order of decreasing rigor, include randomized control trials, regression discontinuity designs, nonequivalent-group designs, lagged implementation designs, and case studies. If certain design aspects are flexible, it might be possible to implement variations in the PBAS coupled with common evaluation frameworks to provide rigorous comparison and help choose the most effective options. Such variations could include different performance measures, different types of incentives, or different incentive levels (e.g., significant versus modest financial rewards).

**Implement Additional, Nonincentivized Measures to Verify Improvement and Test for Unintended Consequences.** A PBAS might induce service-provider responses that lead to improved performance scores without corresponding improvement in the underlying objectives (e.g., a teacher might invest instructional effort on test-taking strategies that lead to improvement on standardized test scores that overstates actual student gains in mastery of the broader subject matter). To

detect when this might be occurring, it can be helpful to include non-incentivized measures intended to test similar concepts (e.g., additional math and reading exams in alternative test formats to check whether there has been a comparable level of improvement). Nonincentivized measures can also be used to examine whether a focus on the incentivized measures within the PBAS is causing other areas of performance to be neglected.

**Link the PBAS Evaluation to a Review and Redesign Process.** The true benefits of evaluation come not from simply understanding what is working and what is not, but rather from applying that understanding to improve the functioning of the PBAS. Evaluation should thus be embedded within a broader framework for monitoring and continuing to refine the PBAS over time.

## Areas for Further Research

Because so few of the PBASs that we examined have been subjected to rigorous testing and evaluation, there are a number of fundamental questions that our study cannot answer about the design, implementation, and performance of PBASs. Policymakers would benefit from research—both within individual sectors and across sectors—on the short- and long-term impact of PBASs, the elements of a PBAS that are most important in determining its effectiveness, and the cost and cost-effectiveness of PBASs, particularly in comparison to other policy approaches.

## Concluding Thoughts

This study suggests that PBASs represent a promising policy option for improving the quality of service-delivery activities in many contexts. The evidence supports continued experimentation with and adoption of this approach in appropriate circumstances. Even so, the appropriate design for a PBAS and, ultimately, its prospects for success are highly dependent on the context in which it will operate. Because PBASs are

typically complex, getting all of the details right with the initial implementation is rare.

Ongoing system evaluation and monitoring should be viewed, to a far greater extent than in prior efforts, as an integral component of the PBAS. Evaluation and monitoring provide the necessary information to refine and improve the functioning of the system over time. Additionally, more-thorough evaluation and monitoring of PBASs will lead, gradually, to a richer evidence base that should help future decisionmakers understand (1) the circumstances under which a PBAS would be an effective and cost-effective policy instrument and (2) the most appropriate design features to employ when developing a PBAS for a given set of circumstances.