W.E. UPJOHN INSTITUTE FOR EMPLOYMENT RESEARCH

Employment Research Newsletter

Volume 18 | Number 2 Article 2

4-1-2011

The Performance of Performance Standards

Carolyn J. Heinrich
University of Wisconsin

Citation
Heinrich, Carolyn, J. 2011. "The Performance of Performance Standards." Employment Research 18(2): 4-5. https://doi.org/10.17848/1075-8445.18(2)-2

Carolyn J. Heinrich

The Performance of Performance Standards

Incentives and accountability for government performance have become so central to contemporary government reform agendas across the globe that public management scholars have proclaimed a new era of "government by performance management" (Moynihan and Pandey 2005, p. 422). Elements of these recent reforms include establishing performance measures and standards to facilitate increased accountability to the public, pay for performance, organization-wide performance bonuses, and competitive performance-based contracting; reducing "red tape" and promoting more transparent management; and devolving government functions and incentivizing innovation. A core objective of incorporating performance measures and standards into public sector incentive systems is to create clear expectations for government performance (while loosening the reins of bureaucratic control) and allow for overt assessment of results. If they are to be more than data collection exercises, however, performance management systems also need to incorporate a means for incentivizing or rewarding individuals, teams, or entire organizations for their achievements relative to performance goals.

In the United States, performance standards systems and bonuses are (or have been) used in Food Stamps and welfare-to-work programs, employment and training programs, child welfare agencies and child support enforcement programs, Medicaid and SCHIP programs, and other social programs. Performance incentive systems in public bureaucracies are also advancing in Europe and other parts of the world, with some governments such as Australia and the Netherlands now implementing incentive systems with fully (100 percent) performance-contingent pay and contracting arrangements (Finn 2008; Struyven and Steurs 2005) As the use of performance measures and incentive systems has expanded in the public sector, so has the number of studies calling attention to their challenges and unintended effects, although there is relatively little rigorous empirical evidence of their implications for government outcomes.

Among the earliest introductions of incentive systems in government agencies was the Job Training Partnership Act (JTPA) performance standards system in 1982, described by Klerman (2005, p. 347) as one of the "most mature implementations of performance-based management." It is also one of the most studied systems, in part because of the randomized experimental evaluation of the JTPA program that produced important information for assessing the performance of this performance standards system. Policymakers have looked to the results of these studies to inform and guide changes in the design and operation of performance standards systems in other government programs, as well as to improve these systems in the Workforce Investment Act (WIA) programs that replaced JTPA.

At the same time, one of the motivations for assembling the research presented in our new book, *The Performance of Performance Standards* (published by the Upjohn Institute), is that despite decades of study and practice, some of the important lessons that have been learned do not appear to be reflected in the current design and implementation of performance standards systems. Bevan and Hood (2006, p. 7), for example, describe the development and use of performance targets in the English public health care system, along with the perverse incentives they generated, as "hitting the target and missing the point."

It may be that some of the empirical evidence from past studies has not been effectively communicated or penetrated policymaking and public management circles deeply enough. Or it may be that some of the fundamental lessons have been ignored or deferred in pursuit of other objectives (political or otherwise). Or, as James Heckman and Jeffrey Smith comment, it may be that policymakers who have mandated such systems (and administrators involved in their implementation) have not fully appreciated the challenges of designing a performance management system that generates incentives for improving performance and impacts. For example, even if a government designed a performance management system that initially suggested a strong correlation between performance measures and desired outcomes, over time, its effectiveness may decline as program managers learn how to game the measures and other limitations of the measures and system design become known.

The essays we present use U.S. employment and training programs as a "laboratory" for investigation. Drawing on a variety of data sources on these incentives systems, we explore how performance standards and incentives affect the behavior of public managers and agency employees, their approaches to service delivery, and ultimately, the outcomes for participants. Both the JTPA and WIA programs have allowed state and local administrators and their governing boards substantial discretion, within broad limits, to determine performance goals, standards, and bonus systems. This administrative flexibility is reflected in the range of incentive systems that states have implemented over time. It is this variation in incentive systems among states that serves as the grist for our empirical mill and is used to extract general lessons that can be applied on a wider scale to both existing and newly developing performance incentive systems.

In our investigation of formal incentive structures and organizational behavior within U.S. employment and training programs, we focus on the following questions that are broadly applicable to any public sector performance management and incentive system:

- How do performance standards and measures operate to include or exclude individuals with different characteristics in public programs?
- How do performance standards and measures affect the types of public services offered and received?
- How do the processes for setting standards and weights for performance goals and for recognizing and rewarding performance affect system incentives and bureaucratic responses?
- Are the performance standards, measures, and incentives effective in motivating bureaucratic behavior toward the achievement of program goals?
- Do short-term outcome measures used in performance standards systems predict long-term impacts of programs on participants?
- What problems or unintended effects are associated with the design and implementation of performance standards systems in the public sector?
- What other lessons do we learn from the implementation of performance standards systems and the variation in rules and guidelines governing their administration over time?

In undertaking research to address the questions above, we were fortunate to have access to data superior in scope and detail to much of the data used in the existing literature or available on a regular basis for assessing program performance. Our research benefitted from detailed longitudinal, microlevel data that were collected in the National JTPA Study (NJS) and through other administrative data sources. In addition, we collected and analyzed complete information about state-level variation in the JTPA and WIA performance standards systems. We show that state incentive systems are highly complex and differ widely across states and over time within states, and are not easily characterized by small dimensional summary measures as used in previous studies. This wealth of data, to which we applied a variety of analytical/empirical strategies, is essential in assessing the implications of features and changes in performance standards and incentive systems in different contexts and across time.

The broad findings and lessons that we draw from this research are as follows. First, individuals and organizations respond to incentives, but sometimes the responses are perverse. In the first iteration of a performance standards system's design, well-meaning designers of the system are unlikely to fully anticipate the responses of program administrators and frontline workers to system incentives, or the many possible ways they might influence measured performance without necessarily adding to (or possibly even detracting from) program value or impact. Incentive system designers will have to expect to regularly review and revise the rules and incentives they create if they want to avoid inefficient and unintended responses.

Second, the short-term outcome measures that are commonly used in performance standards systems are only weakly related to the true long-run impacts of the programs. Researchers and policymakers have yet to identify performance measures that will promote key, long-term program objectives while simultaneously generating more readily available performance information for ongoing program management. This will continue to be one of the most vexing challenges for performance standards system designers for some time to come.

Third, the "cream-skimming" issue, or concern about the trade-off between efficiency and equity in access to programs, has been overstated. While there is some evidence of a trade-off between serving the most disadvantaged and allocating program resources most efficiently, it appears to be modest at best. Personal choices and informational constraints play a more important role in accounting for demographic differences in program participation than administrative discretion, and thus, program administrators should consider investing more in increasing awareness among the eligible population.

Clearly, demand on the part of policymakers or the public for greater accountability and a results-oriented government is not diminishing. The design and implementation of performance standards and incentive systems in the public sector will continue to be a dynamic pursuit, and it is our hope that the lessons distilled in this volume will have a role in shaping and speeding their evolution, as well as in ultimately improving government performance.

References

Bevan, Gwyn, and Christopher Hood. 2006. "What's Measured Is What Matters: Targets and Gaming in the English Public Health Care System." *Public Administration* 84(3): 517–538.

Finn, Dan. 2008. *The British "Welfare Market": Lessons from Contracting Out Welfare to Work Programmes in Australia and the Netherlands.* York: U.K: Joseph Rowntree Foundation. http://www.jrf.org.uk/sites/files/jrf/2306-welfare-unemployment-services.pdf (accessed June 10, 2010).

Klerman, Jacob Alex. 2005. "Measuring Performance." In *High Performance Government: Structure, Leadership, Incentives*, Robert Klitgaard and Paul C. Light, eds. Santa Monica, CA: RAND Corporation, pp. 343–380.

Moynihan, Donald P., and Sanjay K. Pandey. 2005. "Testing How Management Matters in an Era of Government by Performance Management." *Journal of Public Administration Research and Theory* 15(3): 421–439.

Struyven, Ludo, and Geert Steurs. 2005. "Design and Redesign of a Quasi-Market for the Reintegration of Jobseekers: Empirical Evidence from Australia and the Netherlands." *Journal of European Social Policy* 15(3): 211–229.

Carolyn J. Heinrich is the director of the La Follette School of Public Affairs and a professor at the University of Wisconsin–Madison.