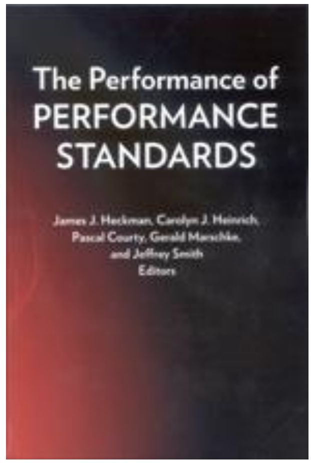
W.E. UPJOHN INSTITUTE FOR EMPLOYMENT RESEARCH

Upjohn Institute Press

Lessons for Advancing Future Performance Standards Systems



James J. Heckman
University of Chicago

Carolyn J. Heinrich
University of Wisconsin

Jeffrey Smith
University of Michigan

Chapter 10 (pp. 305-309) in:
The Performance of Performance Standards
James J. Heckman, Carolyn J. Heinrich, Pascal Courty, Gerald Marschke, and Jeffrey Smith, eds.
Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 2011
DOI: 10.17848/9780880993982.ch10

The Performance of Performance Standards

James J. Heckman
Carolyn J. Heinrich
Pascal Courty
Gerald Marschke
Jeffrey Smith
Editors

2011

W.E. Upjohn Institute for Employment Research
Kalamazoo, Michigan

Library of Congress Cataloging-in-Publication Data

The performance of performance standards / James J. Heckman . . . [et al.], editors.
p. cm.
Includes bibliographical references and index.
ISBN-13: 978-0-88099-292-3 (pbk. : alk. paper)
ISBN-10: 0-88099-292-1 (pbk. : alk. paper)
ISBN-13: 978-0-88099-294-7 (hardcover : alk. paper)
ISBN-10: 0-88099-294-8 (hardcover : alk. paper)
1. Government productivity. 2. Performance standards. 3. Civil service—Personnel management. I. Heckman, James J. (James Joseph)
JF1525.P67P476 2011
352.6'7—dc22
2011007877

300 S. Westnedge Avenue
Kalamazoo, Michigan 49007-4686

Cover design by Alcorn Publication Design.
Index prepared by Diane Worden.
Printed in the United States of America.
Printed on recycled paper.

10
Lessons for Advancing Future Performance Standards Systems

James J. Heckman
Carolyn J. Heinrich
Jeffrey Smith

The economic recession that started in 2007 led to renewed interest in public employment and training services. At the same time, the accompanying financial crises also elevated concern for how dwindling government budgets could be spent more efficiently and effectively to maximize returns to public investments in training. As discussed in the introductory chapter, revamping incentive systems in government is a critical step toward improving government performance and our future economic outlook.

The chapters in this volume marshal some of the most detailed evidence available on how performance standards and incentives influence the behavior of program administrators and staff and contribute to program outcomes or unintended consequences. Since each chapter presents a self-contained summary of its main findings, we do not review the details of each one. Instead, this conclusion presents three main lessons learned from these essays and discusses some of their policy implications.

LESSON 1: AGENCIES RESPOND TO INCENTIVES

Concerns that performance incentives are disregarded by government employees because award levels are low or because benefits are diffused are not justified. Low-powered cash incentives may, in fact, be high-powered because of the value of the budgetary awards in establishing the reputation of bureaucrats and the recognition that comes with them. Bonuses and award money create leverage in the social services community and are thus frequently highly prized (Heinrich 2007). The evidence reported by Courty and Marschke in Chapter 7 demonstrates that agencies made placement, enrollment, and termination decisions in ways that were consistent with maximizing their performance as measured by the JTPA performance standards system. The evidence presented by Heinrich in Chapter 8 reveals that although caseworkers claimed to discount performance standards in decision making, they nevertheless selectively enrolled into JTPA people who were likely to contribute to the placement goals rewarded under the performance standards system.

Courty and Marschke and Heinrich also present a dark side to the behavior elicited by the JTPA performance standards system. Training centers showed remarkable ingenuity in manipulating the JTPA accounting system and reporting requirements in their efforts to boost their measured performance. Practices included enrolling persons receiving job search assistance or on-the-job training only after they found a job, using short-term training arrangements in order to maximize the probability of counting a successful placement, holding persons who did not find jobs in dead-end job clubs, and releasing poorly performing trainees from the program at strategic accounting dates when it did least damage to training center performance records. These all represent behaviors that make perfect sense in terms of the performance standards, but they do nothing to raise participant earnings or increase the equity with which program services are distributed.

The problem of regulating job-training programs—or any government agency—is that enforceable regulations cannot be written too finely, and simple rules can and will, as shown in this volume, be subverted. Along these lines, performance standards systems designers also need to grasp the dynamic properties of performance incentive systems, as discussed in Chapter 5 and illustrated throughout this volume. An incentive system designer's understanding of how individuals will respond to performance standards will inevitably be imperfect prior to their introduction, and it is only as performance measures and targets are tried, evaluated, modified, and/or discarded that their responses become known (Courty and Marschke 2007). Of course, this type of monitoring to assess a measure's effectiveness and possible distortion requires a considerable investment on the part of incentive system designers, one that has probably been underestimated in the past. And as Heinrich and Marschke (2010) point out, learning on the part of bureaucrats will advance over time as well, as they come to better know the distinct weaknesses of performance measures and how they can be exploited through their day-to-day experiences. Incentive system designers will have to expect to regularly review and revise the rules and incentives they create if they want to avoid the inefficient behavior documented in the chapters in this volume.

LESSON 2: CURRENT PERFORMANCE STANDARDS DO NOT PREDICT LONG-RUN GAINS

While the chapters by Courty and Marschke and Heinrich indicate that the JTPA performance standards system effectively motivated agency staff and service providers to meet short-run performance standards, they do not indicate whether the measures themselves are appropriate to induce the achievement of the primary program goals, i.e., increasing the earnings of program participants (or the value added of the program). Chapter 9 demonstrates that the short-run performance measures featured by the JTPA performance standards system were weakly and often perversely related to the long-term effects of the program on the earnings and employment of participants. Yet it is these long-term effects that constituted the true goal of the program's services. The analysis also reveals that the measurement of wages and employment at a later point following termination represented no improvement in the performance monitoring system. Neither set of measures was strongly positively related to long-run program impacts on earnings and employment, suggesting that the performance standards system did not promote the program's key objectives of long-run gains in earnings and employment for participants. We see no reason to expect that the relationships between the performance measures and programs have changed under WIA and thus no reason to think that this problem has gone away.

From the perspective of policymakers and taxpayers who would like to maximize the value from government dollars spent on public programs, the goal of incentive system designers should be to choose performance measures so that the effects of bureaucratic actions on measured performance are aligned with the effects of those same actions on value added. However, we frequently lack the information required to realize this objective in practice. Empirical research, as Heckman and Smith suggest, has focused primarily on estimating measures of statistical association between performance and value added, where the value added of programs has been assessed through the use of randomized experiments or through sophisticated statistical modeling. The dynamics of performance standards systems, however, may limit the usefulness even of these estimates, as the alignment between a performance measure and value added may decrease after a performance measure is introduced and bureaucrats respond by exploring *all* strategies for raising it, not just those that also increase program value added. Clearly, this will continue to be one of the most vexing issues for performance standards system designers for some time to come.

LESSON 3: THE CREAM SKIMMING PROBLEM IS OVERSTATED

The charge of cream skimming has arisen frequently in public job training programs, including the Workforce Investment Act program. In Chapters 3, 6, and 9, we define various uses of this term and present evidence that fears about cream skimming are overstated. For most demographic groups, experimental estimates of the earnings impacts of participation in JTPA are uniform over broad skill levels. Only at the lowest skill levels is there any evidence that impacts were smaller than at higher levels of the skill distribution. While there is some evidence of a trade-off between serving the most disadvantaged within eligible populations and allocating program resources most efficiently, it appears to be a modest one given the modest benefits realized by program participants.

Cream skimming is usually defined as arising from purposive screening behavior by bureaucrats. Chapters 6 and 8 present information about disparities in program participation rates among different demographic groups. Even among eligible persons, there are substantial differences in program participation rates. It is not enough to compare participation rates among the eligible to determine if cream skimming is an important factor. At issue is how much of the difference in participation rates conditional on eligibility is due to the voluntary choices of individuals, how much is due to their lack of information about the program, and how much is due to the decisions of bureaucrats. The decompositions of participation rates into components due to eligibility, awareness, and acceptance that are presented in Chapter 6 reveal that personal choices and lack of information play the major role in accounting for demographic disparities in participation. Administrative discretion also has some role to play in determining the participation rates of different groups in public job training programs, but it is not the dominant factor.

The evidence in Chapters 6 and 8 indicates that those whose characteristics make them more employable and more easily placed were more likely to be screened into the JTPA system. However, the analysis of Chapter 6 reveals that the same characteristics that make a person more attractive in terms of achieving objectives within the job placement system also made that person more aware of the program and more willing to apply to it. Thus, if resources are going to be used to improve equity in access to programs with voluntary enrollment, incentive system designers and program administrators should consider investing more in (or providing incentives for) increasing awareness of the program among the eligible population.

We realize that there will be some limits to the generalizability of these findings and the lessons of this collection of studies to other public program contexts, and that there is inherent variability in the potential of performance standards systems for improving government performance in the many different contexts in which they have been introduced or considered. At the same time, it is clear that there is no diminishing of demand on the part of policymakers or the public for greater accountability and a results-oriented focus of government. The research in this volume suggests that there is considerable work to do in addressing the flaws of current public sector performance standards systems as well as investing in research that will guide improvements and advancements to these systems as their use expands in the public sector.

References

Courty, Pascal, and Gerald Marschke. 2007. "A General Test for Distortions in Performance Measures." *Review of Economics and Statistics* 90(3): 428–441.

Heinrich, Carolyn J. 2007. "False or Fitting Recognition? The Use of High Performance Bonuses in Motivating Organizational Achievements." *Journal of Policy Analysis and Management* 26(2): 281–304.

Heinrich, Carolyn J., and Gerald R. Marschke. 2010. "Incentives and Their Dynamics in Public Sector Performance Management Systems." *Journal of Policy Analysis and Management* 29(1): 183–208.