Assessing the Performance of Performance Standards in Public Bureaucracies

James Heckman; Carolyn Heinrich; Jeffrey Smith


Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28199705%2987%3A2%3C389%3AATPOPS%3E2.0.CO%3B2-7


Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/aea.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.
Assessing the Performance of Performance Standards in Public Bureaucracies

By James Heckman, Carolyn Heinrich, and Jeffrey Smith*

Performance-standards systems for public bureaucracies have been advocated to improve efficiency in government. They are motivated by the perception that government bureaucracies are inefficient, and that part of this inefficiency arises from poorly defined or conflicting agency goals and from the absence of incentives to motivate the activity of bureaucrats. By defining goals and providing incentives for achieving these goals, performance-standards systems attempt to bring to public-sector agencies the type of discipline that markets bring to firms. These systems are a cornerstone of Vice President Al Gore's "Reinventing Government" initiative. Performance-standards systems now govern many federal job-training programs and are also expanding into state welfare and employment agencies, yet little is known about the effectiveness of these systems.

The case for performance standards rests on the implicit premises that (a) agencies have, or should have, a specified goal or set of goals and (b) the goals can be quantified so that success or failure relative to the goals can be measured. When the public sector performs a task where there are good private-sector substitutes, premises (a) and (b) are satisfied. In this case, where the outputs and performance of a bureaucracy are easily measured, a market test provides the ideal performance standard, and no other formal standard is required.

In this paper, we consider the use of performance-standards systems in the more difficult setting where premises (a) and (b) fail to hold. Thus we consider the case where the goals of a public agency are not easily defined or may be in conflict, and performance relative to the goals is not easy to measure. Specifically, we consider the performance-standards system that governs the programs providing job training to the disadvantaged under the Job Training Partnership Act (JTPA).

The JTPA performance-standards system, established in 1983, is the oldest major federal performance-standards system. Recently, it has served as a prototype for performance-standards systems in other programs. The standards are a matter of public record, and with some effort, their incentive effects can be quantified in a precise way. There is substantial time-series and cross-section variation in the incentives built into this program because of the latitude given to individual states to modify and augment the federal standards. The data on incentives and outcomes for this program are richer than those available for most private organizations. They provide insight into the effectiveness of performance-standards incentives both inside and outside of the public sector. Our analysis summarizes a forthcoming monograph (Heckman, 1997) where the arguments presented here are developed at greater length.

Two legislated goals of the JTPA program, which are defined below, potentially conflict with each other. In addition, one major goal, the net gain to program participation, is difficult to measure. For this program, the implicit premises that justify performance standards are not satisfied. We consider the following questions. (i) Do bureaucrats respond to

* Heckman: Department of Economics, University of Chicago, 1126 E. 59th St., Chicago, IL 60637, Center for Social Program Evaluation, Harris School of Public Policy Studies, and American Bar Foundation; Heinrich: Center for Social Program Evaluation, Harris School of Public Policy Studies, 1155 E. 60th St., Chicago, IL 60637, and American Bar Foundation; Smith: Department of Economics, University of Western Ontario, Social Science Centre, London, Ontario, Canada N6A5C2, and Center for Social Program Evaluation. This research was supported by the Russell Sage Foundation, NSF-93-21-048, a grant from the W. E. Upjohn Foundation, and a grant from the Joyce Foundation to the American Bar Foundation. We thank George Baker for his helpful comments.

1 Avinash K. Dixit (1996) and James Q. Wilson (1989) discuss these sources of bureaucratic inefficiency.
incentives? Are they equally effective for managers and caseworkers? (ii) Do the incentives point the bureaucrats in the "right" direction and get them to perform at "appropriate" levels? Specifically, are the short-run outcome measures used in the performance-standards system positively related to the long-term goal of earnings gains? Which of the multiple and conflicting goals of the program are favored, if any? (iii) How much wasteful activity is induced by performance-standards systems as the affected bureaucrats attempt to "game" them? Before presenting our answers, we first describe the JTPA program and the potential for conflict in achieving its goals.

I. The JTPA Performance-Standards System

The Job Training Partnership Act (Public Law 97-300) mandates the provision of employment and training opportunities to "those who can benefit from, and are most in need of, such opportunities" [Section 141 (c)]. Since benefit and need are two different things, the potential for conflict between efficiency and equity is written into the law authorizing the program. This conflict in objectives is the outcome of a political process involving compromise among competing factions with different objectives. Wilson (1989) and Dixit (1996) describe this phenomenon in detail.

Although authority for JTPA originates in the U.S. Department of Labor (DOL), the program is decentralized. DOL defines eligibility for the program, distributes federal funds for state programs, and defines basic performance standards. The standards are currently stated in terms of trainee employment, wage rates, and earnings levels three months after trainees leave the program. At the time most of our data were collected, these outcomes were measured at the date the trainee was terminated from the program.

States modify the federal performance standards (within limits) and have implemented a variety of incentive schemes to reward good performance relative to the standards. Some states institute a winner-take-all tournament where the best center gets the entire award budget for the state. Other states provide marginal payments above a threshold, as well as payments if a specific threshold is exceeded. Most centers subcontract for service delivery and use subsidiary performance-standards systems in their contracts with vendors to motivate them to conform with the standards.

Performance-based incentive awards to training centers cannot be spent on bonuses or salary increases for program managers or caseworkers. They can only be used to augment the budget of the training center. Because incentive payments are to the center, and not to individual workers, incentives are muted and may affect managers differently than their employees. Such restrictions on the use of incentive pay are typical of public agencies.

Assessing the efficiency of the program requires assessing the gain to trainees and the social cost of providing training. Construction of the required counterfactual is a challenging and controversial task, as is the estimation of social cost. However, the problem of constructing this counterfactual is no more daunting than assessing what difference a corporate executive has made to the profits of a firm. It is easy to ascertain the outcomes of program participants in terms of measured levels, but it is much more difficult to determine what their outcomes would have been had they not participated in the program and thus to determine the value added of a program (see Heckman and Smith, 1996). Thus it is not surprising that performance-standards targets are written in terms of measured levels and not unobserved gains net of costs, the proper definition of efficiency. In addition, performance-standards systems typically use short-run rather than long-run measures in order to provide quick feedback to program managers, even though long-run measures more closely approximate the notion of social benefit.

The use of employment and earnings levels as outcome measures in place of unmeasured value added gives rise to potential gains from "cream-skimming." Training centers may select persons with high anticipated levels of the target outcomes at the date of their termination from the program. These persons may receive little value added from the program and also may be among the least disadvantaged of those eligible for the program.
It may happen, however, that the socially efficient enrollment strategy is to enroll the least disadvantaged, for whom the value added net of cost may also be the greatest. In this case, if the incentives induce managers to enroll the least disadvantaged, they may promote social efficiency. There is no necessary conflict between equity and efficiency. We discuss empirical evidence on this issue below.

II. Empirical Evidence on the Effects of the JTPA Performance-Standards System

Our research uses data collected at the 16 JTPA training centers participating in the National JTPA Study, a recent experimental evaluation of the JTPA program, (see Larry Orr et al., 1995), as well as more detailed data from an intensive study of a large training center in Illinois. We have extensive information on JTPA-eligible populations, applicants, and participants, and on program outcomes.

When performance measures and organizational goals diverge, “dysfunctional” behavior by agency employees may occur. George Baker (1992) and Susan Bernstein (1991) give examples of how managers “game” or “play” incentive systems to maximize their personal rewards when their objectives diverge from those of the organization. Dixit (1996) suggests that moral-hazard and shirking problems are more likely to arise in government agencies because multiple principals complicate agency relationships, and the multiplicity of objectives gives scope for employees to indulge their own preferences.

In their study of JTPA centers, Pascal Courty and Gerald Marschke (1997a, b) report that managers respond to incentives and strategically organize their “trainee inventories” to maximize measured program performance. Enrollees who secure unsubsidized employment are terminated as rapidly as possible to maximize program employment rates measured three months after they complete the program. The termination of unemployed enrollees is postponed for up to three months in an effort to gain credit for employment enrollees might find on their own. We note below that the employment status of participants at termination is weakly correlated with their long-term employment impacts. Thus the incentives induce training centers to pursue a target that is not correlated with long-run gains for its trainees.

Use of employment and earnings levels to define the standards, rather than value added net of social costs, illustrates the point made by Baker (1992), who argues that few organizations’ stated objectives are sufficiently verifiable to be used directly in incentive contracts and reward systems. Readily measured targets substitute for the ideal goals, and in most cases the performance standards are not highly correlated with organizational objectives.

This problem is demonstrated in Heckman and Smith (1995a) for the JTPA performance-standards system. Employment and earnings levels three months after trainees complete the program are used as targets for the performance system. They are at best weakly related to long-run employment and earnings gains measured using data from a social experiment. Many measures used as standards are actually negatively related to the program’s long-run impacts on trainees. Also the JTPA outcome measures are less strongly correlated with program impacts 30 months after enrollment in the program than at 18 months, and changing the date at which outcomes are measured (from the time of program discharge to three months after termination) does not improve the correlation of these measures with long-run gains. The performance-standards system motivates administrators to pursue outcomes unrelated, or even negatively related, to a major program goal.

Cream-skimming induced by performance standards is a frequently discussed concern. JTPA administrators have a strong incentive

---

2 The evidence presented in Heckman and Smith (1995a) indicating that performance standards motivate managers to pursue goals unrelated or negatively related to a major program goal is inconsistent with a version of Baker’s (1992) model with multiplicative errors but is not inconsistent with a version of his model with additive errors. In the second case, levels are not informative about optimality conditions; thus in this case their evidence is inconclusive about the proposition that the standards are optimally constructed, because marginal and average returns are not positively correlated.
to enroll only those applicants likely to improve center performance on the standards, irrespective of their contribution to net value added or to “fairness.” Bureaucrats have the ability to do this because, on average, less than 5 percent of the JTPA-eligible population can be served with the funding available in a given program year, and the training centers have flexibility in recruiting and selecting from among the eligible population. Kathryn Anderson et al. (1993) allege that cream-skimming is a serious problem in JTPA programs, contributing to inequities in service delivery and lowered program impacts.

These critics ignore the point that cream-skimming promotes efficiency if the more employable (the “cream”) among JTPA-eligibles achieve larger value-added gains through program participation. Targeting the most disadvantaged (i.e., the bottom 20 percent of the skill distribution) substantially decreases the social efficiency of the program (Heckman and Smith, 1995a). Serving the remainder of the skill distribution of current participants results in little, if any, loss of efficiency. Therefore, program bureaucrats confront a nonmonotonic trade-off between enrollment of the disadvantaged and efficiency, depending on how far they go down the skill distribution in recruiting trainees.

The question of how cream-skimming arises is also examined in Heckman and Smith (1995a). It is not enough to compare participant and eligible populations in assessing the extent of cream-skimming as previous analysts have done. Such comparisons confound the effects of applicant self-selection decisions with the effects of administrative choices. In Heckman and Smith (1995b), participation rates are decomposed into components due to eligibility, awareness, application, and acceptance into the program. Voluntary choices and lack of information about JTPA programs among eligibles, not bureaucratic preferences, play the major role in accounting for demographic disparities in program participation.

III. The Influence of Performance Incentives at Two Training Centers

When agency goals are vague and do not define clear objectives, the tasks performed by a bureaucracy are often defined by the agency employees and not the agency directors (Wilson, 1989). The behavior of the employees of bureaucracies is influenced not only by incentives controlled by the agency, but also by their attitudes, ideology, and the standards of external reference groups, such as the professional mores of social workers. The more vaguely a job is defined, the more likely it is that these other factors influence employee job performance. We examine whether performance standards tend to improve the agency director’s control and align the incentives of bureaucrats with the legislated goals of the agency, or whether they induce bureaucrats to pursue new goals unrelated to the stated objectives of the program.

From our analyses of local JTPA agencies in Corpus Christi (Texas) and Cook County (Illinois), we have gained substantial insights into how performance-standards systems can be shaped by bureaucratic decisions. In the Cook County training center, which is a technocratically elite JTPA center, performance incentives are strongly reinforced in performance-based contracts for providers that deliver program services. Both caseworkers and program managers are acutely aware of contractually defined performance expectations, and there is particular administrative emphasis on achieving high placement rates at low costs (Heinrich, 1995, 1997). An analysis of the center’s service-provider contract-award decisions (Heinrich, 1996) finds that service providers’ past performance relative to cost-per-placement standards is the most influential factor in these decisions. Training-center administrators also establish considerably higher performance requirements in the contracts they write with their vendors than training centers are actually held to in the state system (Heinrich, 1996). This buffer provides some surety that the center will meet state standards.

Performance standards also influence caseworkers’ participant selection and service-assignment decisions (Heinrich, 1995). Although caseworkers indicate in interviews, and in exercises simulating the participant-selection process, that they desire to help the most disadvantaged, analyses of their actual participant selections show that the probabil-
ity of meeting the standard is the most statistically significant and numerically influential factor in selecting applicants. Overall, strong performance incentives encourage the selection of more job-ready applicants and the provision of less-intensive training services, such as job-search assistance, which produce smaller earnings gains for participants. On the other hand, the federally mandated eligibility rules serve as a check to cream-skimming, because they limit the pool of program-eligible individuals to those with low incomes or to participants in means-tested government transfer programs.

Cook County is exceptional in its concern about performance standards up and down the chain of command. Bureaucrats there get monthly updates on their performance against targets. Based on his interviews with more than 20 federal, state, and local job-training officials, Burt Barnow (1992 p. 299) concluded that "it is extremely unlikely that more than a minority of training centers understand the incentives" produced by this system. In addition, he notes that the performance information gathered by states to determine training-center incentive awards is not a management tool for everyday use and is rarely used to provide feedback to local managers for improving program performance, confirming that Cook County is unique among these centers.

Corpus Christi is a much more typical site, which appears to fit the norm described by Barnow. For that training center, performance incentives play a very different role in caseworker decision-making than in Cook County (Heckman et al., 1996). Case workers manifest strong preferences for serving the most disadvantaged (and least employable). These preferences dominate any concern that caseworkers have about performance standards. Pre-program mean wage rates and earnings are substantially higher for rejected applicants than for those who are accepted into the program. Corpus Christi caseworkers prefer to serve the least employable among the applicants, in apparent defiance of the performance-standards system. Thus, while the managers may manipulate the books in the manner described by Courty and Marschke (1997a, b) their employees tend to seek out the applicants who are least likely to contribute to center success on the performance standards.

In this training center and, we conjecture, in most others, performance standards operate as a partial check on the preferences of caseworkers with a social-worker mentality. In Heckman et al. (1996), it is argued that hiring caseworkers with a social-worker mentality reduces the wage cost of operating the training center. Such persons receive satisfaction from providing services to the most disadvantaged and from helping their "clients" manage their lives. In return for these benefits, they accept lower wages and are loyal to the organization. Workers motivated in this way are unlikely to be interested in meeting some externally imposed standard largely unrelated to their own value of helping the poor. Performance standards tend to offset these motivations.

IV. Conclusions

Our answers to the questions posed in the introduction are as follows. (i) Bureaucrats respond to performance incentives, but fears of "cream-skimming" to meet the standards are exaggerated. Managers strive to meet the standards, but their employees are not always in sympathy with these objectives.

(ii) The incentives built into JTPA performance standards do not point in either of the two directions indicated by the legislation that authorized the program. Value added net of social cost is one objective that is too difficult to measure on a regular basis. The short-run performance measures that are used in its place are either uncorrelated with or negatively correlated with net value added, especially in the long run. With the exception of programs in a few states, current incentives do not reward enrollment of the least advantaged into the program, but the social-worker mentality of most "street-level" bureaucrats probably assures that this goal is achieved. However,

3 Awards to the center can be targeted toward the disadvantaged. JTPA participants trained using performance-standards money are often exempt from performance standards. Thus, caseworkers may select applicants who are more employable as a means of raising incentive funds that can be used to help the least employable.
when performance standards are taken very seriously and intensely pursued, as in Cook County, they tend to subvert the goal of enrolling the least advantaged into the program, although they do not induce a trade-off between equity and efficiency. Over a broad range of disadvantage, the net value added of the program is the same for everyone.

Finally, in response to question (iii), we find some evidence that bureaucrats "game" the performance-standards system in an attempt to maximize their center's measured performance. There is no direct empirical evidence that gaming lowers the quality of the training produced by the program, but it is likely that it does so.

The goal of bringing market-like incentives to government or private bureaucracies has much rhetorical appeal, especially in an era of tight budgets. In assessing the case for a performance-standards system, however, it is important not to confuse a focused effort with a productive one. When the output is difficult to measure, as is true in most government bureaucracies and in many private ones, installation of specific goals may focus effort but may send the bureaucrats marching in the wrong direction.

Rules that restrict the relationship between individual performance and individual pay of bureaucrats and their employees mute the role that can be played by performance-standards incentives. Agency goals are not clearly defined. This ambiguity is part of the legislative compromise that brings most agencies into existence. It is unreasonable to expect that externally imposed performance standards can solve the problems of governance and direct activity toward socially productive goals in bureaucracies that serve many masters with conflicting or ill-defined goals, unless the standards effectively repeal the legislation creating the program. To the extent that they do so, performance standards may serve to improve the effectiveness of the program relative to the goals they embody. However, this approach to policy reform raises basic questions of whether administrative decrees should supersede legislative intent. Inefficiency may be built into a program as part of a legislative compromise. A faction opposing a program may sabotage it by giving it vague or contradictory goals. In that case, making the program more effective at achieving one goal favors one faction over another. On the other hand, if the inefficiency built into the program is just a consequence of ignorance, performance standards can improve the performance of the bureaucracy and will be acceptable to all factions in the legislature.

REFERENCES


Heckman, James J. and Smith, Jeffrey A. "The Performance of Performance Standards:

------. "Determinants of Participation in a Social Program: Evidence from JTPA." Unpublished manuscript, University of Chicago, 1995b.


------. "Do Bureaucrats Make Effective Use of Performance Management Information?" Unpublished manuscript, University of Chicago, 1996.


Orr, Larry; Bloom, Howard; Bell, Stephen; Lin, Winston; Cave, George and Doolittle, Fred. The National JTPA Study: Impacts, benefits, and costs of Title IIA. Bethesda, MD: Abt Associates, 1995.

LINKED CITATIONS
- Page 1 of 2 -

You have printed the following article:

Assessing the Performance of Performance Standards in Public Bureaucracies
James Heckman; Carolyn Heinrich; Jeffrey Smith
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28199705%2987%3A2%3C389%3AATPOPS%3E2.0.CO%3B2-7

This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.

[Footnotes]

2 Incentive Contracts and Performance Measurement
George P. Baker
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28199206%29100%3A3%3C598%3AICAPM%3E2.0.CO%3B2-F

References

The Effect of Creaming on Placement Rates Under the Job Training Partnership Act
Kathryn H. Anderson; Richard V. Burkhauser; Jennie E. Raymond
Stable URL:
http://links.jstor.org/sici?sici=0019-7939%28199307%2946%3A4%3C613%3ATEOCOP%3E2.0.CO%3B2-T

Incentive Contracts and Performance Measurement
George P. Baker
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28199206%29100%3A3%3C598%3AICAPM%3E2.0.CO%3B2-F

NOTE: The reference numbering from the original has been maintained in this citation list.
Measuring Government Performance: Lessons from a Federal Job-Training Program

Pascal Courty; Gerald Marschke


Stable URL:

http://links.jstor.org/sici?sici=0002-8282%28199705%2987%3A2%3C383%3AMGPLFA%3E2.0.CO%3B2-2

NOTE: *The reference numbering from the original has been maintained in this citation list.*