

**THE VALUE OF EFFICIENCY MEASURES: LESSONS FROM
WORKFORCE DEVELOPMENT PROGRAMS**

Burt S. Barnow
George Washington University

John Trutko
Capital Research Corporation

Revised October 12, 2013

ABSTRACT

In 2007, the Office of Management and Budget (OMB) required the development of efficiency measures for all federal government programs as part of the effort to improve federal government program performance. As a result of Program Assessment Rating Tool reviews, OMB asked the Employment and Training Administration of the U.S. Department of Labor to develop and implementing an outcome-based measure or measures of efficiency for employment and training programs administered by the agency. In response to this OMB directive, in May 2008, ETA initiated a study to identify outcome-based efficiency measures for implementation by 11 ETA-administered programs: Workforce Investment Act (WIA) Adult Program; WIA Dislocated Worker Program; WIA Youth Activities Program; WIA National Emergency Grants Program; Trade Adjustment Assistance Program; Wagner-Peyser/Employment Service (ES) Program; Senior Community Service Employment Program National Farmworker Jobs Program Indian and Native American Program Work Incentive Grant Program; and Apprenticeship Program.

A key lesson that emerges from this study is that it is critical in selecting measures, standards, rewards, and sanctions to anticipate the behavioral changes that are likely to be induced by the performance management policies adopted and to structure the system so that the presence of efficiency measures does not result in undesirable behavior by programs, states, and grantees. To be implemented within three years, the study recommends that efficiency measures should be closely tied to the current outcome performance measures in effect under ETA's Common Measures framework. Though the report highlights some of the challenges of comparing efficiency measure results across programs, the Common Measures provide common definitions for outcome measures and thus increase the potential for making meaningful comparisons of efficiency measure results within individual programs (e.g., across states/subgrantees) and across at least some of the ETA programs of interest. This report also recommends use of program expenditures (rather than appropriations or obligations) as the measure of program costs in efficiency measures. Among the efficiency measures recommended for consideration in this report are cost per entered employment, cost per retained in employment, cost divided by post-program (average) earnings, and cost divided by change in earnings. The report concludes with a series of recommendations concerning the specific efficiency measures that should (and should not) be considered for implementation by each of the 11 ETA programs that are the focus of this study and, if adopted, how these measures should be used to monitor and enhance program performance.

I. INTRODUCTION

Performance measurement has long been an important tool in the United States. Poister and Streib (1999) observe that measuring workload and worker efficiency can be traced back to the beginning of the twentieth century, and that interest increased with the interest in program budgeting and program evaluation in the 1960s and 1970s, respectively. Poister and Streib (1999) note a decline in interest in performance measurement during the 1980s, and a subsequent resurgence in the 1990s. As Poister (2010, p. 100) points out, “performance measurement has become orthodox practice in the public sector, and increasingly in the sector as well.” Hatry (2010) notes that programs have long tracked expenditures and physical outputs, but he refers to outcome tracking as “the new kid on the block.” With such a long history of interest in tracking program costs as a management tool, it is somewhat surprising that measuring expenditures in a performance context is still somewhat controversial. This paper analyzes the use of efficiency measures for workforce development programs to illustrate why developing and using performance measures involving expenditures is more complex than might be expected and can sometimes lead to perverse incentives. The paper is based on a report prepared for the U.S. Department of Labor’s Employment and Training Administration (ETA) in response to a request from the Office of Management and Budget (OMB) for ETA to develop and implement outcome-based efficiency measures for workforce development programs administered by ETA.

Poister (2010, p. 102) defines two types of performance measure that involve program costs. First, *cost effectiveness* measures relate costs to outcomes; he illustrates the concept in the vocational rehabilitation field as cost per client placed in suitable employment and the cost per client successfully employed for six months or more. Poister defines *efficiency* measures as relating “...outputs to the resources used in producing them, most often focusing on the ratio of

the outputs to the dollar cost of the collective resources consumed.” Note that neither type of measure relates expenditures to the value of program impacts, as is done in cost-benefit analysis, and we will return to this issue later in the paper.

The next section of the paper reviews the literature on performance measurement in the workforce development area, particularly literature analyzing experience relating to the use of efficiency measures. The third section provides context and methods used for the study. The fourth section provides findings from the study, and the final section presents conclusions.

II. LITERATURE REVIEW

The Employment and Training Administration has had substantial experience with performance standards, and a number of studies have been conducted on the impacts of performance management on participants served, activities, costs, and program impacts. While most analysts note the strong rationale for developing performance measures for government programs, there has been considerable controversy in the literature regarding the benefits of performance management systems, particularly as they have been applied since enactment of the Government Performance and Results Act (GPRA) in 1993. This section reviews the literature on performance standards for workforce programs; most of the research was conducted on the performance standards system used under the Job Training Partnership Act (JTPA), WIA’s predecessor.¹ Although much of the literature on performance management points to its salutary

¹For a more in-depth review of the literature on performance standards in workforce programs, see Burt S. Barnow and Jeffrey A. Smith (2004). “Performance Management of U.S. Job Training Programs: Lessons from the Job Training Partnership Act.” 4(3): 247-287. Most of this section is based on Barnow and Smith (2004). For a critical review of the performance management movement, see Beryl A. Radin (2006). *Challenging the Performance Movement: Accountability, Complexity, and Democratic Values*. Washington, DC: Georgetown University Press.

effects, there is little doubt from the literature that instituting performance standards can have a strong impact on program behavior, and not always in the desired direction.

A. THE IMPACT OF PERFORMANCE STANDARDS ON WHO IS SERVED

The majority of the employment and training literature on performance incentives addresses the question of their effect on who gets served. Under JTPA, local service delivery areas (SDAs) had strong incentives to serve persons likely to have good labor market outcomes, regardless of whether those outcomes were due to JTPA because the performance measures were based on labor market outcomes. Similar incentives guide the WIA program. In fact, the absence of a regression model to adjust standards for serving individuals with labor market barriers should make these incentives stronger under WIA than they were under JTPA.

The literature divides this issue into two parts. First, do service delivery area (SDAs under JTPA and Workforce Investment Boards [WIBs] under WIA) respond to these incentives by differentially serving persons likely to have good outcomes, whether or not those good outcomes result from the effects of the program? This is the literature on “cream skimming.” Second, if there is cream skimming, what are its impact effects? Taking those with the best labor market prospects among those eligible could be economically efficient if the types of services offered by these programs have their largest net impacts for this group. In what follows, we review the literature on each of these two questions in turn. We focus here on the first issue. For a review of the relationship between cream skimming and impact, see Barnow and Smith (2004).

A few papers, all about the JTPA program, examine whether or not program staff cream skim in response to the incentives provided by the JTPA performance system. The key issue in this literature is the counterfactual: to what group of non-participants should the participants be compared in order to determine whether or not cream skimming has occurred? In all cases, the

studies proceed by comparing observable characteristics correlated with outcomes, such as education levels or participation in transfer programs such as Aid to Families with Dependent Children (AFDC) or Temporary Aid to Needy Families (TANF).

Anderson et al. (1992, 1993) compare the characteristics of JTPA enrollees in Tennessee in 1987 with the characteristics of a sample of JTPA eligibles in the same state constructed from the Current Population Survey. The literature suggests that less than five percent of the eligible population participated in JTPA in each year (see the discussion in Heckman and Smith, 1999), which allows wide scope for cream skimming. Both papers find modest evidence of cream skimming. In particular, the Anderson et al. (1993) analysis of program participation and post-program job placement suggests that if eligible persons participated at random, the placement rate would have been 61.6 percent rather than 70.7 percent, a fall of 9.1 percentage points.

Heckman and Smith (2004) address the issue of self-selection versus selection by program staff using data from the Survey of Income and Program Participation (SIPP) on JTPA eligibles combined with data from the National JTPA Study. They break the participation process for JTPA into a series of stages – eligibility, awareness, application and acceptance, and participation – and look at the observed determinants of going from each stage to the next. They find that some differences between program eligibles and participants result primarily from self-selection at stages of the participation process, such as awareness, over which program staff have little or no control. The evidence in Heckman and Smith (2004) suggests that while cream skimming may be empirically relevant, comparing the eligible population as a whole to participants likely overstates its extent, and misses a lot of substantive and policy-relevant detail.

The paper by Heckman, Smith, and Taber (1996) presents a contrasting view. They use data from the Corpus Christi, Texas SDA, the only SDA in the National JTPA Study for which

reliable data on all program applicants are available for the period during the experiment. In their empirical work, they examine whether those applicants who reach random assignment (i.e., were selected to participate in the program) differ from those who do not in terms of both predicted outcome levels (earnings in the 18 months after random assignment) and predicted program impacts (projected into the future and discounted). Heckman, Smith, and Taber (1996) argue that it is this stage over which program staff have the greatest control, although even here applicants may wander off if they find employment elsewhere, get in trouble with the law, and so on. The authors find strong evidence of negative selection on levels combined with weak evidence for positive selection on impacts. They attribute the former to a strong “social worker mentality” toward helping the hard-to-serve among the eligible that was evident in interactions with program staff at the Corpus Christi site. The Workforce Investment Act (WIA) program offers an interesting contrast to JTPA because the WIA performance standards are not adjusted by a regression model, and they therefore do not hold programs harmless for the characteristics of their participants. Because programs now have stronger incentives to enroll individuals with few barriers to employment, we would expect to observe enrollment shift toward this group. An internal (U.S. Department of Labor, 2002) study finds that this is precisely what appeared to be occurring, at least in the area scrutinized:

A brief survey of States by our Chicago Regional Office indicated that WIA registrations were occurring at only half the level of enrollment achieved by JTPA. While some of this may be due to start up issues, there are indications that the reduced registration levels are due to a reluctance in local areas to officially register people in WIA because of concerns about their ability to meet performance goals, especially the “earnings gain” measure. It appears that local areas in these States are selective in whom they will be accountable for. Some local areas are basing their decisions to register a person on the likelihood of success, rather than on an individual’s need for services.

A study by the U.S. General Accounting Office (GAO, 2002) confirmed these problems. The GAO report, based on a survey of 50 states, indicated “many states reported that the need to meet

performance levels may be the driving factor in deciding who receives WIA-funded services at the local level.”

Overall, the literature provides modest evidence that program staff responded to the incentives provided by the JTPA performance standards system to choose participants likely to improve their measured performance whether or not they benefited from program services, and studies of the implementation of WIA indicate that, if anything, the situation has been exacerbated by the new program. At the same time, the evidence from the Corpus Christi SDA indicates that staff concerns about serving the hard-to-serve could trump the performance incentives in some contexts.

B. RELATIONSHIP BETWEEN PERFORMANCE MEASURES AND PROGRAM IMPACT

Performance measures for a program may be of intrinsic interest, or they may be a proxy for some underlying factor of interest that is not easy to measure on a relatively quick and inexpensive manner. For example, Blalock and Barnow (2001) note that programs may wish to use program impact as a performance measure, but accurately measuring impact requires many years and the presence of a randomly assigned control group. Because this is not generally compatible with obtaining quick, inexpensive measures, programs often rely on proxy measures such as post-program earnings or the pre-post change in earnings. If the goal is to have performance measures serve as a proxy for impact, then it is necessary to assess how well the types of measures that are practical and have been used for the JTPA and WIA programs correspond with program impact.

Two studies have explored this issue for JTPA in recent years, and another study looked at the Job Corps. The studies by Barnow (2000) and by Heckman, Heinrich, and Smith (2002)

both made use of the fact that the National JTPA Study provided experimental impact findings in 16 local areas and included the data needed to construct performance measures similar to those used by ETA. However, the approach used to measure performance does not include a control group, so it is not surprising that the performance measures used are at best weakly correlated with program impact.²

The recent evaluation of the Job Corps that was based on a classical experimental design provided Schochet and Burghardt (2008) with an opportunity to analyze how closely Job Corps' performance standards track the program's impacts. Job Corps is a primarily residential program for highly disadvantaged out-of-school youth. Schochet and Burghardt indicate that during the evaluation period, program years 1994 through 1996, the performance measures included eight measures in three broad areas: (1) program achievement (reading and math gains, GED attainment rate, and vocational completion rate), (2) placement measures (placement rate, average wage at placement, and the percentage of placements related to training), and (3) quality/compliance measures (ratings of federal monitors). Because of the random assignment used to assign treatment status, impact can be estimated as the difference between treatment and control group values on the outcome measures. Schochet and Burghardt (2008) compared program impacts for Job Corps centers ranked in each third of the performance distribution. They concluded: "Our results indicate that at the time of the National Job Corps Study, measured center performance was not associated with impacts on key education, crime, and earnings outcomes."

²A related problem is that performance measures must use short-term post-program earnings to measure performance, but the impact of a program is best measured over a longer period. Barnow and Smith (2004) review the literature on the relationship between short-term earnings impacts and long-term impacts, and they find that most studies find a very weak relationship between the two.

C. STRATEGIC RESPONSES TO PERFORMANCE INCENTIVES

In addition to the substantive responses to performance incentives considered above, in which training centers changed what they actually did, local training programs can also attempt to change their measured performance without changing their actual performance. This behavior is referred to as a strategic response, or as “gaming” the performance system. Regardless of their differing goals, all types of organizations have an incentive to respond strategically to performance incentives, provided the cost is low, as doing so yields additional resources to further their own goals. The literature provides clear evidence of such gaming behavior under JTPA.

One important form of strategic behavior under JTPA was the manipulation of whether or not participants were formally enrolled. Under the JTPA incentive system, only persons formally enrolled counted towards site performance. In addition, for the first decade of JTPA’s existence, training centers had substantial flexibility in regard to when someone became formally enrolled. Clever SDAs improved their performance by basing enrollments on job placements rather than the initiation of services. For example, some SDAs boosted performance by providing job search assistance without formally enrolling those receiving it in the program. Then, if an individual found a job, the person would be enrolled, counted as a placement, and terminated, all in quick succession. Similarly, SDAs would send potential trainees to employers to see if the employer would approve them for an on-the-job training slot; enrollment would not take place until a willing employer was found.

There are several pieces of evidence regarding the empirical importance of this phenomenon. The first is indirect, and consists of the fact that DOL found it enough of a problem to change the regulations. Specifically, in 1992, the Department of Labor required that

individuals become enrolled once they received objective assessment and that they count as a participant for performance standards purposes once they received any substantive service, including job search assistance.

The flexibility of JTPA also allowed strategic manipulation of the termination decision. Because performance standards in JTPA were based on terminees, SDAs had no incentive to terminate individuals from the program who were not successfully placed in a job. By keeping them on the rolls, the person's lack of success would never be recognized and used against the SDA in measuring its performance. As the Department of Labor explains in one of its guidance letters, "Without some policy on termination, performance standards create strong incentives for local programs to avoid terminating failures even when individuals no longer have any contact with the program."³

Problems with local programs retaining participants on the rolls long after they stopped receiving services go back to the days of JTPA's predecessor, the Comprehensive Employment and Training Act (CETA). In one of their guidance letters, the Department of Labor observed that "monitors and auditors found that some participants continued to be carried in an 'active' or 'inactive' status for two or three years after last contact with these programs." For Title II-A of JTPA, DOL limited the period of inactivity to 90 days, although some commentators suggested periods of 180 days or more.

Courty and Marschke (1996, 1997, 2004) provide additional evidence on the strategic manipulation of termination dates using data from the National JTPA Study. The first type of evidence consists of the timing of termination relative to the end of services as a function of the employment status of the trainee as of the end of services. Assuming that the timing of termination responds mainly to the employment at termination standard in place during the time

³See [TEIN 5-93](#) available at: http://wdr.doleta.gov/directives/corr_doc.cfm?DOCN=770

their data were collected (rather than the wage rate or cost standards, which would be more difficult to game), they argue that sites should immediately terminate participants who are employed when their services end. In contrast, they should not terminate participants who are not employed at the end of their services; instead, they should wait and see if they later become employed, at which point they should then terminate them from the program. Not surprisingly, Courty and Marschke (1996, 1997, 2004) find that the sites in the National JTPA Study did exactly this.

D. USING RETURN ON INVESTMENT (ROI) AS A PERFORMANCE MEASURE

In making investment decisions, businesses must determine the best or most profitable use of scarce funds. In general terms, the metric that is used to compare potential investments is called *return on investment*, often referred to by its initials as ROI. Economists who try to determine the best use of public or private funds have developed a very similar concept that they refer to as *cost-benefit analysis*. Although the two concepts are closely related, an examination of several books on cost-benefit analysis and return on investment analysis yielded no cross references. Our interpretation is that return on investment analysis tends to focus more on the rate of return to investments, while cost-benefit analysis generally focuses more on net present values. (These terms are defined and described and below.) As public training programs have tried to learn from business practices, many analysts have advocated developing ROI or cost-benefit measures that can be used to judge how good investments in workforce programs are. The concept can be used to determine if a single program or project is worth carrying out, or it can be used to compare alternative projects or programs to rank them or to see if they are performing adequately.

The concept of ROI is quite straightforward—an ROI measure should relate the benefits of a project to its costs. In a book describing how to conduct ROI analyses for training programs, Jack Phillips (2003, p. 21) offers two definitions. He first suggests using a benefit-cost ratio, BCR, defined as $BCR = \text{Program Benefits} \div \text{Program Costs}$. Alternatively, he suggests that ROI can be measured as a percentage: $ROI = (\text{Net Program Benefits} \div \text{Program Costs}) \times 100$. In practice, however, measuring ROI involves many additional decisions and assumptions. The decisions that must be made to compute the ROI of a project include both technical issues (e.g., How do we account for nonmonetary benefits such as improved health? Should we compute an annual rate of return or the overall rate?) and philosophical and judgmental issues (e.g., From whose perspective are the returns calculated? What time period following the investment should be included in the calculations?). Moreover, there is disagreement among experts on the most appropriate way to measure ROI. While Phillips (2003, p. 21) states that computing BCR is acceptable, economists generally shun the use of benefit-cost ratios; one leading text (Boardman et al. 2006) concludes that “Thus, benefit cost ratios are subject to manipulation. For these reasons, *we recommend that analysts avoid using benefit cost ratios...*”

A second important issue in estimating program benefits is that we must generally have a control group or comparison group of people similar to the customers who did not participate in the program. In a workforce program, it is generally not enough to observe what customers earn before and after participation—some and often most of their post-program earnings would have accrued without the program. Thus, the best impact evaluations rely on a randomly selected control group of individuals who applied to the program or a comparison group of individuals who are similar to the participants but who did not receive the training. It is widely accepted that

having a randomly assigned control group is the best method of creating a group that is as similar as possible to the treatment group. Unfortunately, incorporating random assignment into a workforce program is expensive, time consuming, and sometimes considered unethical. This is particularly true if the program is to be evaluated each year.

The alternative of creating a comparison group has generated considerable controversy in recent years. Although sophisticated matching strategies, often based on propensity scores (the estimated probability of enrolling in the program given the person's characteristics), have been advocated by some researchers as being a reasonable alternative to random assignment, studies that have compared results from classical experiments with comparison groups generated by propensity score matching have frequently questioned how well the matching strategies achieve their goal of generating comparable groups and unbiased impact estimates.⁴ Thus, the choices in creating an appropriate group for comparing the participants may boil down to a strategy that is infeasible or one that may lead to biased estimates of the impact.

A further complication of using ROI as a performance measure is that to make comparisons across states or local areas, the services received by the control or comparison group must be the same across all areas. Even random assignment is insufficient to solve this problem if control group members are permitted to enroll in other available training opportunities (such as courses available at the community college). If some states have more alternative training programs available and the control or comparison group is not barred from

⁴ See Howard S. Bloom, Charles Michalopoulos, Carolyn J. Hill, and Y. Lei (2002). *Can Nonexperimental Comparison Group Methods Match the Findings from a Random Assignment Evaluation of Mandatory Welfare-to-Work Programs?* New York, NY: Manpower Demonstration Research Corporation; Jeffrey Smith and Petra Todd (2005). "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics* 125: 305-353; and Elizabeth T. Wilde and Robinson Hollister (2007). "How Close Is Close Enough? Testing Nonexperimental Estimates of Impact Against Experimental Estimates of Impact with Education Test Scores as Outcomes." *Journal of Policy Analysis and Management*. 26(3): 455-477. For a more optimistic appraisal, see Thomas D. Cook, William R. Shadish, and Vivian C. Wong (2008). "Three Conditions under Which Experiments and Observational Studies Produce Comparable Causal Estimates." *Journal of Policy Analysis and Management*. 27(4): 724-750.

enrolling in these other programs, the counterfactual will not be the same across states, and the ROI estimated in different states are not comparable.

E. EMPLOYMENT AND TRAINING ADMINISTRATION EXPERIENCE WITH EFFICIENCY MEASURES

The Employment and Training Administration also has previous experience with efficiency standards under WIA's predecessor, the Job Training Partnership Act (JTPA). Under JTPA, Section 106(b)(4) required that efficiency measures be prescribed for the JTPA Adult Program and that the efficiency measures be related to the outcome measures used. The National Commission for Employment Policy (NCEP) sponsored an evaluation of the effects of JTPA performance standards on participants, services, and costs.⁵ The study included quantitative statistical analysis of JTPA Annual Status Report (JASR) data linked to data on the characteristics of local program areas, as well as qualitative analysis based on interviews with 30 local programs and 87 service providers in eight states.

For the most part, the study found that the JTPA performance standards had the desired effects of holding programs harmless for differences in participant characteristics and local economic conditions. However, the study found that the cost standards had intrinsic problems and created some undesirable effects on participants served:

This evaluation found that the federal standards for the entered employment rate and wage rate for adults generally did not have unintended effects on clients or services. ... The federal cost standards, however, had the most unintended effects and were the least comparably measured of all the federal performance measures. The evaluation found that SDAs in states that placed more weight on the federal cost standard tended to serve fewer hard-to-serve clients and that [local areas] concerned about exceeding the cost standards tended to design less intensive services. At the same time, this evaluation found serious measurement problems with the cost standards. We found large differences in the extent to which [local programs] were leveraging JTPA funds, either by

⁵See Katherine Dickinson et al. (1988). *Evaluation of the Effects of JTPA Performance Standards on Clients, Services, and Costs*. Washington, DC: National Commission for Employment Policy Research Report No. 88-15.

using funds from other programs to help fund JTPA Title II-A programs or by using service providers that had alternative funding sources. As a result, it is difficult to compare the cost of services received by JTPA participants across [local programs]. (Dickinson, et al, p. 5).

Based on their findings from both the quantitative and qualitative components of the study, the authors recommended that alternatives to the cost measures be explored. The authors noted that as a result of concern about the unintended impacts of the cost standards, ETA set more lenient cost standards in PY 1988, but they concluded that this policy change would not eliminate the disincentive problems in states that emphasize exceeding rather than meeting standards.

In response to the research findings, the NCEP made a number of recommendations for changing the statutory provisions of JTPA dealing with performance standards. Taking note of the study's findings regarding the undesirable incentives and comparability of cost issues, the Commission's first recommendation was: "The Commission recommends that Section 106(b)(4), which requires the Secretary [of Labor] to prescribe performance standards relating gross program expenditures to various performance measures, be amended to direct that cost-efficiency be monitored by states." In August 1992, the JTPA statute was amended, and the amendments repealed the federal requirement for efficiency standards and prohibited Governors from using efficiency standards in making awards to local areas.

III. CONTEXT AND METHODS

In 2007, the Office of Management and Budget (OMB) required development of efficiency measures for all federal government programs as part of the effort to improve federal government program performance. As a result of ongoing Program Assessment Rating Tool (PART) reviews, OMB charged ETA with the responsibility of developing and implementing an outcome-based measure or measures of efficiency for employment and training programs

administered by the agency.⁶ In response to this OMB directive, in May 2008, ETA initiated a study aimed at identifying outcome-based efficiency measures for implementation by 11 ETA-administered programs:

- Workforce Investment Act (WIA) Adult Program;
- WIA Dislocated Worker Program;
- WIA Youth Activities Program;
- WIA National Emergency Grants (NEG) Program;
- Trade Adjustment Assistance (TAA) Program;
- Wagner-Peyser/Employment Service (ES) Program;
- Senior Community Service Employment Program (SCSEP);
- National Farmworker Jobs Program (NFJP);
- Indian and Native American Program (INAP);
- Work Incentive Grant (WIG) Program; and
- Apprenticeship Program.

The Work Incentive Grant program was in the process of being phased out during the period of study and is not addressed in detail here.

To develop and assess potential efficiency measures for the ETA programs, the following activities were undertaken:

- A literature review on efficiency measures;
- An initial round of interviews with officials in six states to obtain their views and experience with efficiency measures;
- Telephone interviews with officials from the Organization for Economic Cooperation and Development (OECD), the United Kingdom, and Canada to learn of their experience with efficiency measures;
- Analysis of efficiency measures used by the following federal agencies for workforce programs: the Department of Health and Human Services, the Department of Agriculture, the Department of Education, the Department of Housing and Urban Development, and the Department of Veterans' Affairs;
- A second round of interviews with officials in five states to follow up on issues related to implementing efficiency measures;
- Collection and analysis of cost and outcome data for the 11 ETA programs being studied to investigate the potential developing regression models to adjust expected

⁶Executive Office of the President, Office of Management and Budget, *Program Assessment Rating Tool Guidance No. 2007-7: Improving the Quality of PART Performance and Efficiency Goals*, issued December 12, 2007 (available at: http://georgewbush-whitehouse.archives.gov/omb/performance/guidance/part_guid_2007-07.pdf). The Program Assessment Rating Tool Guidance No. 2007-7 called for a government-wide initiative “to increase the government and public’s ability to assess and improve a program’s effectiveness and efficiency.”

- performance for states and to explore the feasibility of developing state-level measures; and
- Ongoing consultation with ETA national and regional staff as well as an expert panel.

This paper focuses on the literature review and the overall findings and recommendations from the study. Detailed findings are available in Trutko and Barnow (2010).

B. DEFINITION AND CRITERIA FOR SELECTION OF EFFICIENCY MEASURES

In its 2007 guidance letter to federal agencies, OMB noted that the “Government’s ability to determine a program’s effectiveness, and to direct attention to genuinely desired outcomes, is largely dependent upon the quality of the program’s performance and efficiency goals, i.e., their measures and targets.” OMB also provided a basic definition of and purpose for generating such efficiency measures – terming such measures as “efforts to provide the most benefits (outcomes and outputs) for the taxpayer dollars spent.” Finally, this OMB guidance letter underscored the urgency for federal agencies to refocus their attention on improving the types of efficiency measures used to assess ongoing program performance: “...a substantial portion of PART [Program Assessment and Rating Tool] efficiency measures need to be revised in order to meet current PART guidance.”⁷

Frequently, in employment and training programs (and more broadly for all government human service programs), efficiency measures use unit costs tied to one or more of the following: (1) participation (such as “cost per participant served or exited”); (2) delivery of services (such as “cost per participant trained”); or (3) outcomes (such as “cost per participant

⁷Executive Office of the President, Office of Management and Budget, *Program Assessment Rating Tool Guidance No. 2007-7: Improving the Quality of PART Performance and Efficiency Goals*, issued December 12, 2007 (available at: http://georgewbush-whitehouse.archives.gov/omb/performance/guidance/part_guid_2007-07.pdf).

entering or retaining employment”).⁸ An underlying concept for many efficiency measures is that a unit cost for a particular time period (usually a program or fiscal year) is produced by dividing program costs by the number of participants served/exited, the number of participants/exiters receiving a particular service, or the number of participants/exiters achieving a certain outcome.

While the definition of efficiency measures is relatively straightforward, there are a number of challenges to identifying appropriate and feasible measures for a single program – and these challenges multiply when consideration is given to applying an efficiency measure or measures across more than one program. The efficiency measure or measures applied to one or more programs should at a minimum meet the following four criteria:

- be fair to the programs being judged;
- encourage desired service delivery and program outcomes;
- discourage undesired strategies and behaviors; and
- maintain program quality, integrity, and fiscal responsibility.

In addition, the efficiency measures recommended for a specific program should be feasible and cost-effective to implement using participant, outcome, and cost data that are currently collected or that could potentially be collected in the future at a reasonable cost (such as data collected and reported as part of the Common Measures).⁹

⁸For example, GAO identifies several potential types of performance measures, making a distinction between three types of potential measures: “Performance measures may address the type or level of program activities conducted (process), the direct products and services delivered by a program (outputs), and /or the results of those products and services (outcomes).” United States General Accounting Office, *Glossary: Performance Measurement and Evaluation -- Definitions and Relationships*, GAO/GGD-98-26, April 1998 (available at: <http://www.gao.gov/special.pubs/gg98026.pdf>).

⁹For additional background on the Common Measures, see Training and Employment Guidance Letter No. 17-05, “[Common Measures Policy for the Employment and Training Administration’s \(ETA\) Performance Accountability System and Related Performance Issues](#)” (available at: http://wdr.doleta.gov/directives/corr_doc.cfm?DOCN=2195).

C. CANDIDATE EFFICIENCY MEASURES

In the early stages of this study, the research team identified and considered a range of candidate process and outcome measures that could potentially be combined with cost measures to produce one or more efficiency measures to be tracked across the 11 ETA programs.

Efficiency measures for short-term consideration – that is, measures that ETA was already collecting and potentially could be implemented within one to three years across some or all of the 11 ETA programs -- were the following:

- Cost per participant;
- Cost per exiter;
- Cost per entered employment;
- Cost per retained employment;
- Cost divided by increase in earnings (or cost per \$1 increase in earnings); and
- Cost divided post-program earnings (or cost per \$1 post-program earnings).

Efficiency measures identified in the early stages of the research study for longer-term consideration were:

- Cost per exiter or participant receiving a particular service (such as an intensive or training service);
- Cost per positive outcome (e.g., placed in a job, the military, a registered apprenticeship program, education program, or advanced training program);
- Cost per recognized credential received; and
- Return on Investment (ROI).

The four potential measures listed above have some desirable features, but they are not tied to currently used outcome measures. ROI has additional problems, as ROI can only be measured with a substantial follow-up period, perhaps many years, to be meaningful for workforce programs.

IV. FINDINGS

As discussed above, to generate an efficiency or unit cost measure, it is necessary to account for costs in the numerator of the measure and a unit of participation, service delivery, or outcome in the denominator of the measure. This section focuses on the cost portion of the efficiency measure calculation, examining the three options available for measuring costs, providing a recommendation on which cost type should be used, and detailing challenges to appropriately and consistently capturing costs when measuring program efficiency.

A. OPTIONS AVAILABLE FOR MEASURING PROGRAM COSTS

Based on discussions with ETA officials and the review of the literature, three potential types of “cost” data could be considered for the numerator of each efficiency measure:

- **Appropriations/Allotments** – “Appropriations” are defined by the Government Accountability Office as “*Authority given to federal agencies to incur obligations and to make payments from Treasury for specified purposes.*”¹⁰ Appropriations are generally the amount of funding made available by Congress for spending on a given program (such as the TAA program) during a fiscal year. Appropriations have been used by ETA in calculating “cost per participant” for programs such as WIA and Wagner-Peyser. Such measures have been used primarily as part of the planning process prior to the start of a program year to estimate the numbers of individuals who can likely be served for a given funding level. Allotments are the amount of appropriated funds distributed to a state or grantee based upon a legislative or regulatory formula.¹¹ Allotments exclude amounts retained by the federal government to administer programs. In programs that do not have a formula distribution, this term may be used to denote the discretionary amount planned to be distributed to a grantee/contractor.
- **Obligations** – According to the Government Accountability Office (GAO), “*...obligations reflect orders placed, contracts awarded, and other similar transactions during a fiscal year. As an expression of an agency’s total financial commitments for a given period, gross obligations portray the relative size of an organization, without regard to the type of underlying budgetary resource or when resulting outlays may*

¹⁰ See U.S. Government Accountability Office (2004). *Principles of Federal Appropriations Law*, Third Edition, Volume 1. Washington, DC: U.S. Government Accountability Office Report GAO-04-261, SP, p. 2-5.

¹¹ Although the term “allocation” is used quite often in an interchangeable manner with allotment, in the context of formula programs the appropriate term is allotment.

occur.”¹² Hence, obligations are funds that have been committed through contracts, grants, and other vehicles.

- **Expenditures** – Expenditures are funds paid or the amount of funds due (depending on whether a cash or accrual basis is used) for provision of goods or services pursuant to a grant or contract agreement. With regard to accounting for various program expenditures, ETA’s Office of Financial and Administrative Management (OFAM) noted (in an electronic response to a question posed by the research team) that “as far as the accounting process for the various levels, the process varies depending on what type of program you look at and how it operates. The general process is that DOL obligates federal funds to direct grantees, and requires those grantees to provide financial reports which include cash transactions, obligations for lower level grants or contracts, and cost reporting. There could be several levels of grants/contracts below DOL, but the direct grantees are responsible for summarizing all financial data for those sub-levels when reporting back to DOL.”¹³

Telephone interviews with six states and discussions with ETA program officials, as well as analyses of data on individual programs, confirmed our conclusion that expenditure data should be used when calculating efficiency measures for ETA programs. Expenditures are preferred to allotments or obligations because they are based on what is actually spent providing services for program participants, taking into consideration rescissions, transfers between programs, and expenditure of program funds across more than one program year. Exhibit 1 shows how the use of expenditures versus allotments can make a difference when calculating cost per entered employment for a program such as the WIA Adult program. For the United States overall, in PY 2005, expenditures in the WIA Adult program were slightly higher (about \$17 million or 1.8 percent) than allotments. Despite the relatively small difference at the national level, as shown in Exhibit 1, there can be substantial differences between allotments and expenditures at the state

¹²United States General Accounting Office (GAO), *Federal Budget: Agency Obligations by Budget Function and Object Classification for Fiscal Year 2003*, GAO-04-834, June 2004 (available at <http://www.gao.gov/new.items/d04834.pdf>).

¹³Also according to OFAM, the term “drawdown or payment” is sometimes used to reflect the “transfer of cash to a grantee/contractor based on grantee requests/contractor invoices to reimburse the grantee/contractor for expenditures on a valid grant/contract. Source of quotations in footnote and text is OFAM electronic response to a question.

**EXHIBIT 1: COMPARISON OF ALLOTMENTS AND EXPENDITURES IN
PROGRAM YEAR 2005**

State	PY 2005 Allotments	PY 2005 Expenditures	Difference (Expenditures - Allotments)	Percentage Difference
<i>Total</i>	\$972,406,996	\$989,481,877	\$17,074,881	1.8%
New Mexico	\$6,460,982	\$9,547,783	\$3,086,801	47.8%
Delaware	\$1,978,186	\$2,505,938	\$527,752	26.7%
Nevada	\$4,643,187	\$5,729,486	\$1,086,299	23.4%
Mississippi	\$12,201,673	\$15,006,538	\$2,804,865	23.0%
Hawaii	\$3,026,432	\$3,686,765	\$660,333	21.8%
New Jersey	\$21,742,322	\$25,210,314	\$3,467,992	16.0%
Kentucky	\$17,150,395	\$18,785,452	\$1,635,057	9.5%
North Carolina	\$30,650,748	\$33,559,283	\$2,908,535	9.5%
Louisiana	\$20,206,337	\$22,087,656	\$1,881,319	9.3%
Washington	\$22,810,203	\$24,588,371	\$1,778,168	7.8%
Florida	\$42,534,930	\$45,228,422	\$2,693,492	6.3%
Puerto Rico	\$35,811,897	\$38,032,686	\$2,220,789	6.2%
Utah	\$5,186,709	\$5,452,807	\$266,098	5.1%
Oregon	\$16,023,659	\$16,644,227	\$620,568	3.9%
Pennsylvania	\$33,565,397	\$34,786,979	\$1,221,582	3.6%
Michigan	\$41,989,813	\$43,393,133	\$1,403,320	3.3%
Texas	\$88,060,741	\$90,920,855	\$2,860,114	3.2%
Arizona	\$16,629,687	\$17,133,496	\$503,809	3.0%
Ohio	\$40,994,031	\$41,977,559	\$983,528	2.4%
New York	\$77,930,704	\$79,595,396	\$1,664,692	2.1%
Wisconsin	\$11,542,384	\$11,772,235	\$229,851	2.0%
Wyoming	\$2,640,294	\$2,676,683	\$36,389	1.4%
Alabama	\$17,442,093	\$17,665,302	\$223,209	1.3%
Illinois	\$44,436,912	\$44,847,898	\$410,986	0.9%
Idaho	\$2,801,747	\$2,815,094	\$13,347	0.5%
Minnesota	\$9,435,871	\$9,412,227	-\$23,644	-0.3%
California	\$137,225,360	\$136,577,973	-\$647,387	-0.5%
Montana	\$2,561,631	\$2,540,090	-\$21,541	-0.8%
Maryland	\$9,919,836	\$9,778,360	-\$141,476	-1.4%
Nebraska	\$2,499,885	\$2,460,076	-\$39,809	-1.6%
Missouri	\$15,743,342	\$15,480,533	-\$262,809	-1.7%
Virginia	\$12,535,527	\$12,322,012	-\$213,515	-1.7%
New Hampshire	\$1,900,876	\$1,857,842	-\$43,034	-2.3%
Kansas	\$6,673,042	\$6,497,283	-\$175,759	-2.6%
Rhode Island	\$2,824,329	\$2,714,604	-\$109,725	-3.9%
Tennessee	\$19,176,233	\$18,420,971	-\$755,262	-3.9%
Massachusetts	\$14,684,986	\$14,037,922	-\$647,064	-4.4%
Georgia	\$16,958,731	\$16,091,106	-\$867,625	-5.1%
Alaska	\$3,029,756	\$2,858,082	-\$171,674	-5.7%
Maine	\$3,069,783	\$2,869,752	-\$200,031	-6.5%
North Dakota	\$2,218,186	\$2,064,034	-\$154,152	-6.9%
Connecticut	\$7,592,184	\$6,955,344	-\$636,840	-8.4%
West Virginia	\$6,286,348	\$5,747,333	-\$539,015	-8.6%
Arkansas	\$9,574,708	\$8,710,671	-\$864,037	-9.0%
Iowa	\$4,265,476	\$3,854,824	-\$410,652	-9.6%
Oklahoma	\$9,753,985	\$8,582,068	-\$1,171,917	-12.0%
District of Columbia	\$3,876,655	\$3,388,419	-\$488,236	-12.6%
Colorado	\$13,187,525	\$11,452,728	-\$1,734,797	-13.2%
South Dakota	\$2,385,238	\$2,020,786	-\$364,452	-15.3%
Vermont	\$2,249,082	\$1,859,946	-\$389,136	-17.3%
Indiana	\$15,987,777	\$12,916,277	-\$3,071,500	-19.2%
South Carolina	\$18,329,181	\$14,360,256	-\$3,968,925	-21.7%

Note: Total excludes funds not distributed to states; data provided by OFAM.

level; for example, in 13 states, the percentage difference between expenditures and allocations is plus or minus 10 percent or more, and in one state, the difference was nearly 50 percent.

Actual expenditure patterns can be quite different from allotments and obligations because of transfers of funds, rescissions, and allotted funds not expended. Within a state (or at the local workforce level), patterns of expenditures can change based on priorities and needs of states and local workforce areas (based on changing economic conditions and needs of the individuals served by each program). In WIA, for example, funds can be transferred to some extent among the Adult, Dislocated Worker, and Youth programs; and Governors can move funds from one local workforce investment area to another if the funds are not being spent. Original allotments can also be affected by rescissions and supplemental appropriations. As another example, in PY 2008, the Department of Labor was required to apply three different rescissions, which resulted in an across-the-board cut of allotted funds to states.¹⁴ Such rescissions result in cutbacks of training and other services (thus, affecting overall expenditure levels).

¹⁴ The rescissions are to be applied to various WIA funds during FY 2008 as follows: (a) to the FY 2007 Advance fund year source, a 1 percent across-the-board rescission; (b) to the FY 2008 Advance fund year source (appropriated in advance in the FY 2007 appropriation act), a 1.747 percent across-the-board rescission; and (c) to unexpended balances of formula funds appropriated for PY 2005 and PY 2006 which includes fund year sources PY 2005, FY 2006, PY 2006, and FY 2007, a rescission of \$250 million, required by P.L. 110-161, applicable to the WIA Adult, Dislocated Worker and Youth formula programs only. For additional details, see U.S. Department of Labor, [Training and Employment Guidance Letter No. 24-07, "Rescissions During Fiscal Year 2008 for Workforce Investment Act \(WIA\) Programs, issued March 26, 2008](http://wdr.doleta.gov/directives/corr_doc.cfm?docn=2620), available at: http://wdr.doleta.gov/directives/corr_doc.cfm?docn=2620.

B. CHALLENGES INVOLVED IN PRODUCING ACCURATE AND TIMELY EXPENDITURE DATA TO PRODUCE VALID AND RELIABLE EFFICIENCY MEASURE RESULTS ACROSS STATES

There are a number of challenges and issues that arise in using expenditure data. Among the main challenges, most of which were identified in discussions with state administrators, are the following:

Challenge #1 - Varying Cost Allocation Methods Used Across States and Local Areas. The ways in which states and local workforce areas collect, allocate, and report on expenditures vary substantially. Two particular problems with respect to allocating and reporting expenditures were cited in discussions with states and ETA program officials, which could affect efficiency measure results (and undercut comparability of results across states and local areas). These two issues are discussed below.

Sharing of Costs across Programs and in the One-Stop Career Center System Clouds Calculating Efficiency Measures by Funding Stream

Over the past decade, there has been a mandate to build a comprehensive One-Stop Career Center system that brings together various programs within a single physical location and shares (to the extent possible and appropriate) the costs of serving individuals. States have created One-Stop delivery systems supported by Memoranda of Understanding (MOUs) and cost allocation plans in local areas with the intent of leveraging resources from Wagner-Peyser, WIA, TAA, and the Veterans Employment and Training Service (VETS), as well as other state and local partners. In such a system, individuals may receive services by one or more programs (either at the same time or sequentially), and various funding streams may be brought together to pay for services. For example, an individual enrolled in the WIA Dislocated Worker program may also be co-enrolled and simultaneously (or sequentially) receive services paid for by the WIA Adult, TAA, Wagner-Peyser, and/or non-DOL programs. Core unassisted services—when

a participant first comes into the One-Stop Career Center—may have been paid for with Wagner-Peyser, WIA, and/or other funding sources (e.g., county funding), while training services may be reimbursed by a combination of WIA Dislocated Worker and TAA funding (perhaps combined with and further subsidized with funds from a Pell Grant, state-funded training assistance program, or subsidized training at a community college.) Hence, the sharing of funds and co-enrollment of participants, which can greatly affect numbers of individuals served across programs, makes it complicated (at times nearly impossible) to isolate costs of serving a particular individual with a particular funding stream.

One possibility would be to develop a “composite efficiency measure” for the One-Stop system as a whole. One of the challenges to developing a composite measure for the One-Stop system is that some of the 11 ETA programs are not located at the One-Stop (e.g., in some states, stand-alone Wagner-Peyser/ES offices that operate in tandem with One-Stops still exist). A One-Stop is a place – not a program – that offers a varying constellation of voluntary and mandatory partners from one state to another (and even within states, sometimes there is variation within and across localities as well). As a result, comparing cost per placement or other outcome-based efficiency measures is complex. Further, the cost of supporting One-Stop operations comes from a variety of sources outside of ETA programs that may be difficult or impossible to track (e.g., county government or TANF programs may be contributing to paying the operating costs of the One-Stop, but are not included in the 11 ETA programs). An alternative to a composite measure for the One-Stop system would be a composite “system” measure that would assess efficiency across the 11 ETA programs, or possibly a broader array of programs including TANF, Food Stamps Employment and Training, and similar programs. An example would be to calculate cost per placement using all expenditures across the 11 ETA

programs in the denominator and an unduplicated count of job placements across the same ETA programs (taking into consideration co-enrollments across WIA, TAA, and other programs by counting co-enrolled individuals only once).

Complexities of Calculating Staff Time by Program Activity

Allocating staff time by activity is a particular challenge should DOL institute a cost per activity-type efficiency measure. Staff time is a major cost item for all of the programs. Some states use time sheets or random moment time studies to allocate staff time across funding streams. Among the states interviewed, one state uses random moment time sampling to allocate staff time across programs – WIA, Wagner-Peyser, etc. – at the state and local levels; however, within a program such as WIA, the state does not track staff time devoted to core, intensive, or training services. A second state uses a case study approach to break down costs by activity: under a special study focusing on a small number of local workforce areas in the state, a team of university researchers conducted site visits to One-Stop Career Centers to develop estimates of staff and other expenditures by program activity within the One-Stop setting (Moore et al. 2007). This proved to be a very time intensive, and the researchers uncovered idiosyncrasies in co-enrollment, sharing of costs, and accounting for costs that made it difficult to make valid comparisons across local areas.¹⁵

Challenge #2 - Need to Account for Expenditures of Funds Received Over Multiple Years. After programs receive an initial allotment of funding in a program, states and local workforce areas have several years to spend funds (for example, under the WIA program, states have up to three years to expend program funds after they have been allocated for a given program year). As a result, when collecting and analyzing expenditure data to support efficiency

¹⁵An alternative to collecting cost data by activity would be to take into account the activity mix of participants (e.g., the percentage of individuals trained) when setting performance standards for efficiency measures.

measure calculations, it is necessary to gather expenditure amounts from allotments for the current program year and past program years expended during the program year of focus. For example, for WIA, this means that to calculate cost per entered employment for PY 2006, expenditures that occurred during the program year (July 1, 2006 through June 30, 2007) of program funds received for the current program year (PY 2006), as well as the two previous program years (PY 2004 and PY 2005) would have to be gathered.

Challenge #3 – Expenditures for Individuals Served by the 11 Programs May Occur over Several Years. Customers are enrolled in programs for varying periods, depending upon the types of services delivered. For a labor exchange program, such as the Wagner-Peyser program, expenditures typically occur over a relatively short period of time – several weeks to several months. Such expenditures typically would be confined to one or two program years, depending upon when an individual enters the program (e.g., if an individual enrolled near the end of the program year, then participation could easily stretch across more than one program year). At the other extreme are programs that provide substantive training, such as the Apprenticeship, WIA, and TAA programs. Individuals enrolled in WIA and TAA may participate in training that extends over two or three program years (using funding stretching across multiple program years). Expenditures on apprentices may go on for up to five years, depending on the length of the apprenticeship program. To further complicate matters, the funding for training costs for apprentices comes from states, unions, and/or employers (rather than the federal government). The duration of program participation and expenditures for other ETA programs also varies considerably – for example, an older worker may be engaged in subsidized work under the SCSEP program for one year or longer. The overlapping of expenditures across two or more program years is not a big problem, but because an individual is

counted as exiting or entering employment only for the last year of participation in the program, it requires a simplifying “steady state” assumption (i.e., that although costs of serving an individual may lap across more than one year, unless there are extraordinary changes in patterns of expenditures or exiting, costs per outcome will even out over the long run and from year to year).

Challenge #4 – Lags Occur in Reporting Expenditures at Various Levels of Government, and Final Expenditure Data May Not Become Available for Several Years After the Fact. For programs with several administrative levels, there may be a significant time lag between when money is spent at the lowest program level (such as a service provider or contractor) and when it is reported to the local workforce area or grantee, the state, and finally, the federal government). States submit final annual expenditure data for the ETA programs generally within 90 days of the end of each program year. However, the expenditure data submitted by states can be revised after it is submitted to ETA – for example, financial audits may result in disallowance or modification of program expenditures for a year or even longer after the close out of a program year.

Challenge #5 – Inclusion or Exclusion of Subsidized Wages, Needs-Related Payments, and Stipends Can Substantially Affect Performance on Efficiency Measures. With regard to programs such as SCSEP (where about three-quarters of program expenditures are devoted to wage subsidies for program participants), TAA (where substantial amounts of program funds are expended on Trade Readjustment Allowances [TRA] or income supports), and NEG (where in the case of NEGs in response to disasters a large portion of fund are for wages for temporary subsidized work rather than training), the inclusion or exclusion of such payments can very significantly alter outcomes for efficiency measures such as cost per entered

employment.¹⁶ Moreover, whether a customer receives TRA or unemployment compensation (which would not be included as a TAA cost) depends in part on factors such as the state's unemployment rate and the reason for job loss. ETA officials overseeing SCSEP and TAA observed that definitions of the cost items included in the numerator of efficiency measures need to be carefully defined – and in the case of these two programs, should exclude wage subsidies and income supports.

Challenge #6 – Burden/Costs Are Likely to Be Imposed on States and Localities to Ensure Standardized Cost Data Are Collected; Technical Assistance and Training Likely to Be Necessary. Given the variation in practices across states and localities, collection of expenditure data in a form that is comparable across states/grantees and local programs would likely impose substantial burden and costs on reporting units (i.e., states, grantees, and localities), particularly if such cost data was furnished by program and by discreet activity. It will likely be necessary to provide guidance, training, technical assistance, and ongoing monitoring to ensure that expenditure data are provided in an appropriate format. Also, to avoid gaming by grantees, ETA may find it necessary to standardize some aspects of cost reporting so that measures are consistent across grantees and so that grantees do not make changes to their allocation of costs simply to improve measured performance.

Challenge #7 – Cost-of-Living Differences, Along with Differing Practices with Regard to Co-Enrollment, Sharing of Costs Across Programs, and Accounting Procedures, Make States Anxious About Comparisons of Efficiency Measure Results Across States and

¹⁶The Federal Trade Act provides special benefits under the Trade Adjustment Assistance (TAA) program to those who were laid off or had hours reduced because their employer was adversely affected by increased imports from other countries. These benefits include paid training for a new job, financial help in making a job search in other areas, or relocation to an area where jobs are more plentiful. Those who qualify may be entitled to weekly Trade Adjustment Allowances (TRA) after their unemployment compensation is exhausted. Additional background on TAA and TRA is available on the U.S. Department of Labor website, at <http://workforcesecurity.doleta.gov/unemploy/tra.asp>.

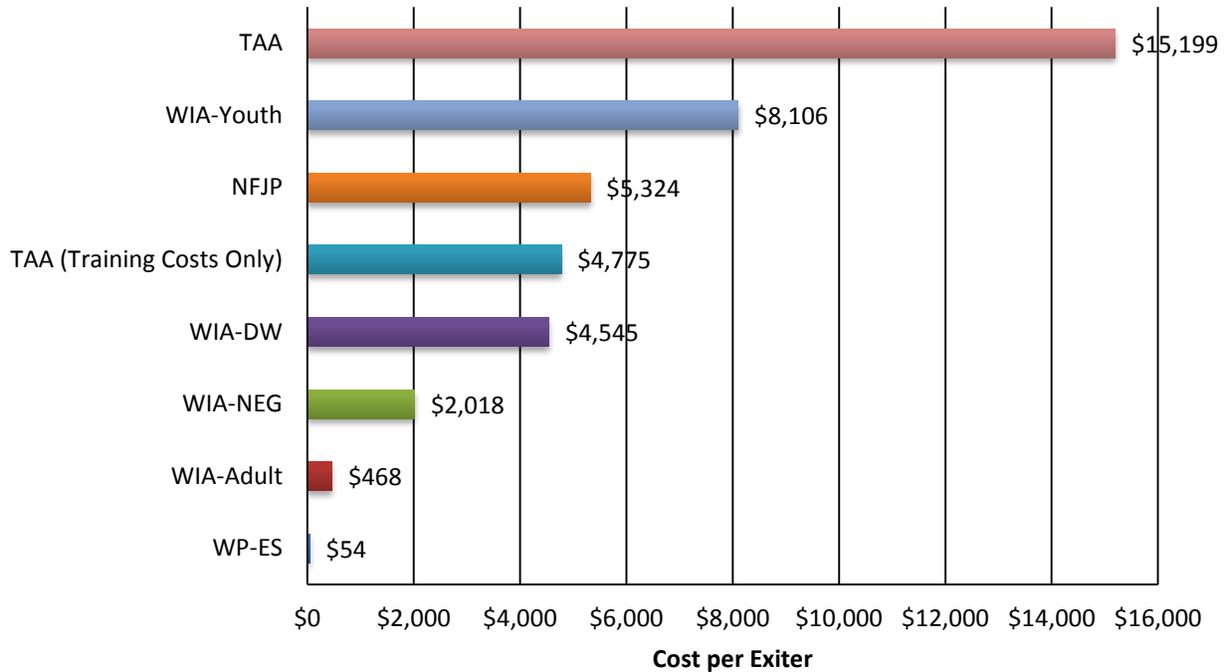
Localities. Cost-of-living differences across states and localities (especially staff and facility cost differences, which are major cost elements) could affect performance on efficiency measures. Different policies with regard to co-enrollment and cost sharing across states and local WIBs can have substantial effects on efficiency measure results. Most of the states interviewed for this study expressed concerns over the potential for inappropriate and unfair comparisons being made once efficiency measure results are available.

ETA program offices provided data on Common Measure performance outcomes, and the Office of Financial and Administrative Management (OFAM) provided data on program expenditures to support development of estimates of efficiency measure results for PY/FY 2007. The estimates provided in this section and the next section (on state-level results) are preliminary and great care should be taken in making cross program comparisons. Exhibit 2 provides preliminary analysis of results for efficiency measure results for candidate efficiency measures (based on the Common Measures) at the national level for PY/FY 2007.

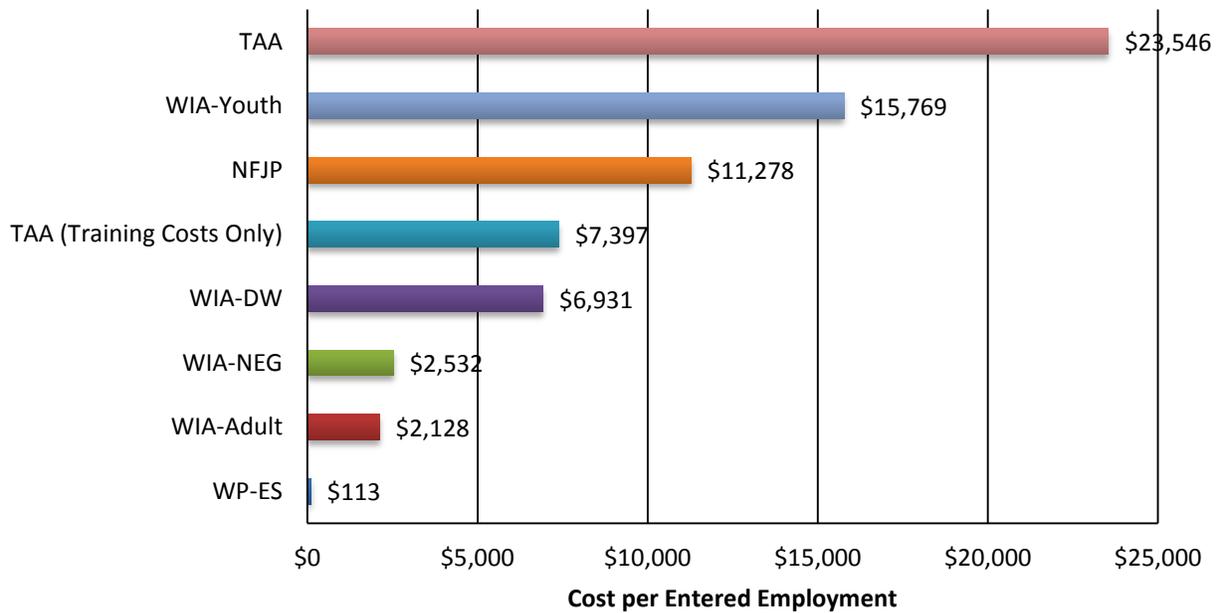
At the national level, not only is there considerable variation across programs, but within programs there is substantial year-to-year variation in efficiency measure results. The program displaying the greatest change over the three-year period is the WIA Adult Program: While expenditures declined slightly (6.2 percent) over the three-year period (PY 2005-07) for the WIA Adult Program, there was more than a tripling of the number of exiters (217 percent increase, from 626,182 to nearly 2 million exiters). As a result of the surge in the number of exiters, there was also very substantial percentage increases in the number of exiters entering employment (226 percent increase) and exiters retained in employment (144 percent increase). The increases in numbers of individuals flowing through the WIA Adult Program finding and retaining jobs are reflected in steep decreases in efficiency measure results between PY 2005 and PY 2007 – a 70

EXHIBIT 2: ESTIMATED RESULTS FOR FOUR OUTCOME-BASED EFFICIENCY MEASURES FOR ETA PROGRAMS, NATIONAL RESULTS, PY/FY 2007

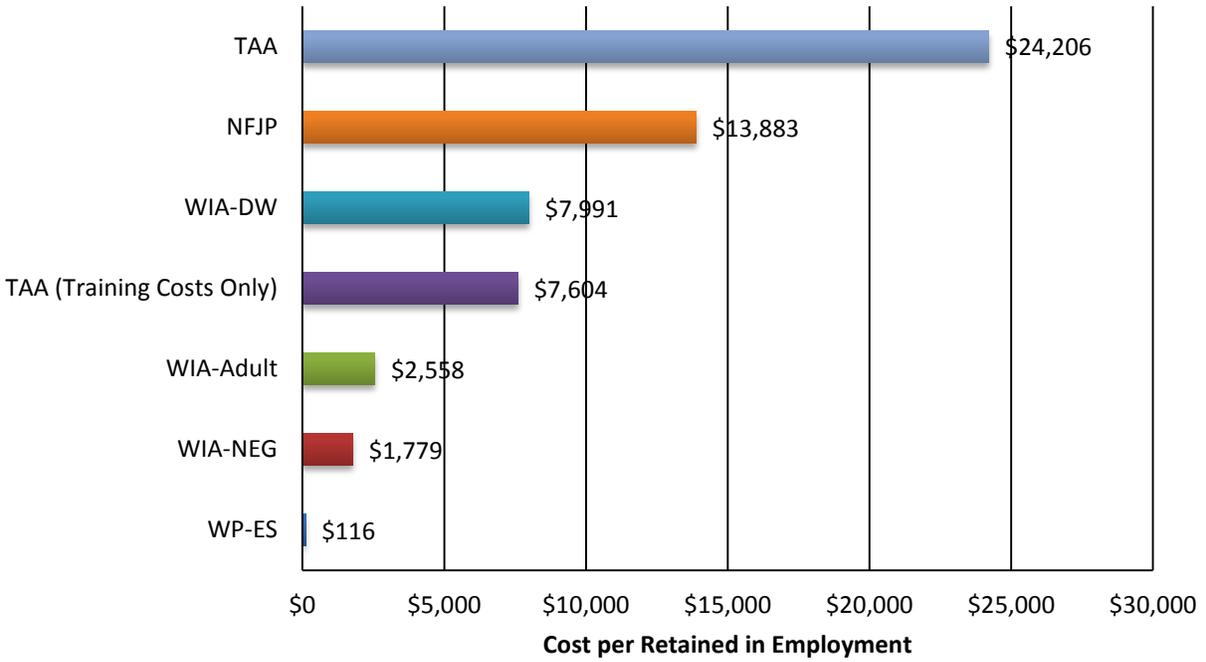
Cost per Exiter, by Program, National Total, PY/FY 2007



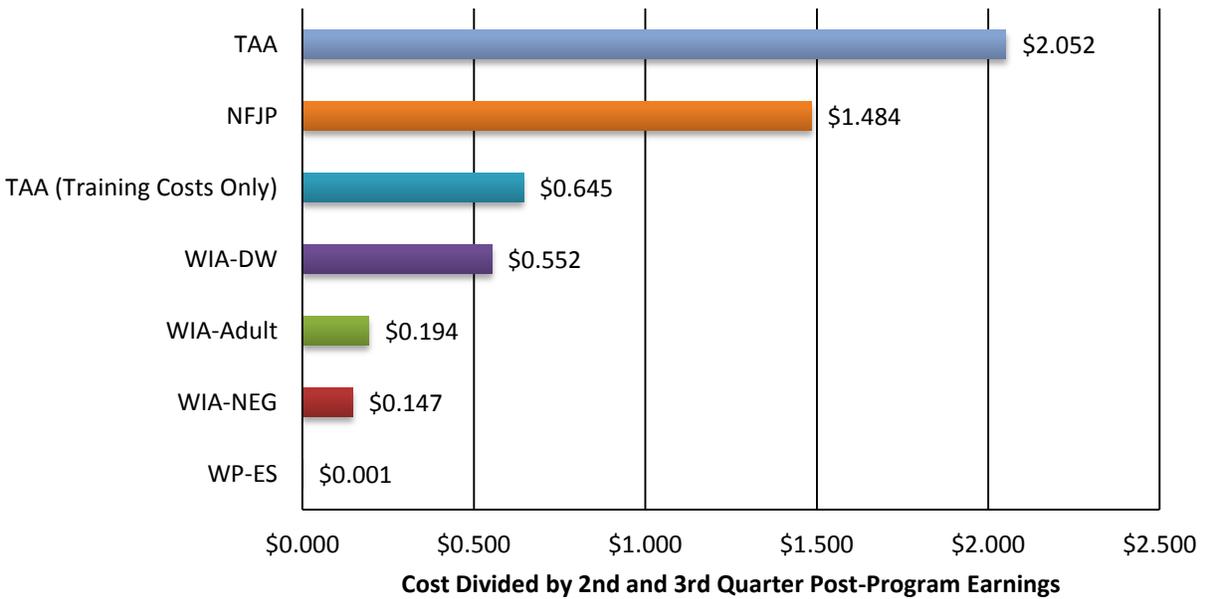
Cost per Entered Employment, by Program, National Total, PY/FY 2007



Cost per Retained in Employment, National Total, PY/FY 2007



Cost Divided by Post-Program Earnings, National Total, PY/FY 2007



Notes: Figures are based on data provided by OFAM on expenditures and ETA program offices for outcomes.

*For WIA-Youth, cost per placement in education or employment was used in place of cost per entered employment.

percent decrease in the cost per exiter, a 71 percent decrease in cost per exiter entering employment, and a 62 percent decrease in cost per exiter retained in employment. As discussed earlier, some states have aggressively moved in the direction of co-enrolling large numbers of Wagner-Peyser customers into the WIA program, which accounts for some (and perhaps much) of the surge in enrollment in the WIA program in recent years.

V. CONCLUSIONS

In assessing how well a program is performing, program costs or efficiency is an important consideration. In this paper we have analyzed the potential of establishing efficiency measures for employment and training programs. We found that previous experience with the use of efficiency measures led state and local governments to avoid training and focus instead on low-cost activities such as job search assistance. In addition, we have identified a number of challenges that will make it difficult to establish efficiency measures that provide a fair comparison across programs. If meaningful efficiency standards are to be established, it is necessary for the federal government to require state and local governments to standardize their cost reporting systems, which could involve significant costs. Decisions on how to count cash payments to participants for wages and transfer payments will significantly affect efficiency measures, and there is no clear correct way to select a strategy. Finally, in programs where the mix of services includes high-cost (e.g., training) and low-cost (e.g., job search assistance) activities, a comparison of costs in a performance measurement framework is likely to lead to less training than Congress and the Department of Labor would like; developing separate efficiency measures is one potential approach to deal with the problem, but that would require a great amount of new data collection.

In the long run, periodic program evaluations with well-designed impact studies and cost-benefit analyses can address most of the issues discussed here, but the type of data collection called for in a periodic national evaluation is not feasible for continuous performance measurement. Thus, we recommend that efficiency measures be explored on a low-stakes basis for now, and that representatives of the federal government, states, and local governments work together to improve data collection and develop feasible efficiency measures for workforce programs.

References

- Anderson, Kathryn, Richard Burkhauser, Jennie Raymond, and Clifford Russell. 1992. "Mixed Signals in the Job Training Partnership Act." *Growth and Change*. 22(3): pp. 32-48.
- Anderson, Kathryn, Richard Burkhauser, and Jennie Raymond. 1993. "The Effect of Creaming on Placement Rates under the Job Training Partnership Act." *Industrial and Labor Relations Review*. 46(4): pp. 613-624.
- Barnow, Burt S. 1992. "The Effects of Performance Standards on State and Local Programs: Lessons for the Job Opportunities and Basic Skills Programs." In C. Manski & I. Garfinkel (Eds.), *Evaluating Welfare and Training Programs* (pp. 277-309). Cambridge, MA: Harvard University Press.
- Barnow, Burt S. 1996. "Policies for People with Disabilities in U.S. Employment and Training Programs," in *Disabilities, Cash Benefits, and Work*, Jerry L. Mashaw, Virginia Reno, Richard Burkhauser, and Monroe Berkowitz Eds. Kalamazoo, Michigan: Upjohn Institute for Employment Research.
- Barnow, Burt. 2000. "Exploring the Relationship between Performance Management and Program Impact: A Case Study of the Job Training Partnership Act." *Journal of Policy Analysis and Management*. 19(1): pp. 118-141.
- Barnow, Burt S. and Jeffrey A. Smith. 2004. "Performance Management of U.S. Job Training Programs: Lessons from the Job Training Partnership Act." *Public Finance and Management*. 4(3): 247-287.
- Barnow, Burt S. and Christopher T. King. 2005. *The Workforce Investment Act in Eight States*. Albany, NY: The Nelson A. Rockefeller Institute of Government.
- Barnow, Burt S. and Carolyn Heinrich. 2010. "One Size Fits All? The Pros and Cons of Performance Standard Adjustments." *Public Administration Review*. 70(1): 60-71.
- Blalock, Ann and Burt Barnow. 2001. "Is the New Obsession with 'Performance Management' Masking the Truth About Social Programs?" In *Quicker, Better, Cheaper: Managing Performance in American Government*. Ed. Dall Forsythe. Albany, NY: Rockefeller Institute Press, pp. 485-517.
- Bloom, Howard S., Charles Michalopoulos, Carolyn J. Hill, and Y. Lei. 2002. "Can Nonexperimental Comparison Group Methods Match the Findings from a Random Assignment Evaluation of Mandatory Welfare-to-Work Programs?" New York, NY: Manpower Demonstration Research Corporation
- Cook, Thomas D., William R. Shadish, and Vivian C, Wong. 2008. "Three Conditions under Which Experiments and Observational Studies Produce Comparable Causal Estimates." *Journal of Policy Analysis and Management*. 27(4): 724-750.

Courty, Pascal, and Gerald Marschke. 1996. "Moral Hazard under Incentive Systems: The Case of a Federal Bureaucracy." In *Advances in the Study of Entrepreneurship, Innovation and Economic Growth*, Volume 7. Ed. Gary Libecap. Greenwich, CT: JAI Press, pp. 157-190.

Courty, Pascal, and Gerald Marschke. 1997. "Measuring Government Performance: Lessons from a Federal Job-Training Program." *American Economic Review*. 87(2): pp. 383-388.

Courty, Pascal, and Gerald Marschke. 2004. "An Empirical Investigation of Gaming Responses to Explicit Performance Incentives." *Journal of Labor Economics*. 22(1): pp. forthcoming22-56.

Courty, Pascal, and Gerald Marschke. 2007. "Making Government Accountable: Lessons from a Federal Job Training Program." *Public Administration Review* 67(5): pp. 904-916.

Dickinson, Katherin, et al. 1988. *Evaluation of the Effects of JTPA Performance Standards on Clients, Services, and Costs*. Washington, DC: National Commission for Employment Policy Research Report No. 88-15.

Hatry, Harry P. (2007). *Performance Measurement: Getting Results*. Washington, DC: The Urban Institute Press.

Heckman, James, Carolyn Heinrich and Jeffrey Smith. 2002. "The Performance of Performance Standards." *Journal of Human Resources*. 37(4): pp. 778-811.

Heckman, James, and Jeffrey Smith. 2004. "The Determinants of Participation in a Social Program: Evidence from JTPA." *Journal of Labor Economics*. 22(2): pp. 243-298.

Heckman, James, Jeffrey Smith and Christopher Taber. 1996. "What Do Bureaucrats Do? The Effects of Performance Standards and Bureaucratic Preferences on Acceptance into the JTPA Program." In *Advances in the Study of Entrepreneurship, Innovation and Economic Growth*, Volume 7. Ed. Gary Libecap. Greenwich, CT: JAI Press, pp. 191-218.

Idaho. 2007. "Workforce Investment Act – Annual Report for PY 2006." October 2007.

Marschke, Gerald. 2002. "Performance Incentives and Bureaucratic Behavior: Evidence from a Federal Bureaucracy." Working paper, Department of Economics, SUNY Albany, Albany, NY.

Moore, Richard, Phillip Gorman, and Andrew Wilson. 2007. "California One-Stop System Cost Study Report." Prepared for the California Workforce Investment Board by The California State University – Northridge.

Radin, Beryl 2006. *Challenging the Performance Movement*. Washington, D.C.: Georgetown University Press.

Phillips, Jack J. 2003. *Return on Investment in Training and Performance Improvement Programs*, Second Edition. New York: Butterworth Heinemann.

Poister, Theodore H. 2010. "Performance Measurement: Monitoring Program Outcomes" in Wholey, Joseph S., Harry P. Hatry, and Kathryn E. Newcomer, Editors. *Handbook of Practical Program Evaluation*, Third Edition. San Francisco: Jossey-Bass.

Poister, Theodore H. and Gregory Streib. 1999. "Performance Measurement in Municipal Government: Assessing the State of the Practice." *Public Administration Review*. 59(4): 325-335.

Rubenstein, Ross, Leanna Stiefel and Amy Ellen Schwartz. 2003. "Better than Raw: A Guide to Measuring Organizational Performance with Adjusted Performance Measures." *Public Administration Review* 63(5): 607-615.

Schochet, Peter Z., John Burghardt, and Sheena McConnell. 2006. *National Job Corps Study and Longer-Term Follow-Up Study: Impact and Benefit-Cost Findings Using Survey and Summary Earnings Records Data*. Princeton, NJ: Mathematica Policy Research.

Schochet, Peter Z. and John Burghardt. 2008. "Do Job Corps Performance Measures Track Program Impacts?" *Journal of Policy Analysis and Management*. 27(3): pp. 556-576.

Schultz, T. W. 1961. "Investment in Human Capital." *American Economic Review*. 51(1): 1-17.

Smith, Jeffrey and Petra Todd. 2005. "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics* 125: 305-353.

Social Policy Research Associates. 2004. *The Workforce Investment Act after Five Years: Results from the National Evaluation of the Implementation of WIA*. Oakland, CA: Social Policy Research Associates.

Social Policy Research Associates. 1999. *Guide to JTPA Performance Standards for Program Years 1998 and 1999*. Menlo Park, CA: Social Policy Research Associates.

John Trutko and Burt S. Barnow. 2010. *Implementing Efficiency Measures for Employment and Training Programs Final Report*. Washington, DC: U.S. Department of Labor, Employment and Training Administration, Occasional Paper ETAOP 2010-05.

U.S. Department of Labor, Office of Job Corps. 2001. *Policy and Requirements Handbook: Job Corps*. Washington, DC: U.S. Department of Labor, Office of Job Corps. Revised July 2007.

U.S. Department of Labor, Employment and Training Administration. 2001. *2002 Annual Performance Plan for Committee on Appropriations*. Washington, DC: U.S. Department of Labor, Employment and Training Administration.

U.S. Department of Labor. 2002. Employment and Training Administration. Training and Employment Guidance Letter No. 11-01. "Guidance on Revising Workforce Investment Act (WIA) State Negotiated Levels of Performance." Washington, DC: U.S. Department of Labor, Employment and Training Administration.

U.S. Department of Labor. 2002. "Summary Report on WIA Implementation." Washington, DC: Employment and Training Administration.

U.S. Department of Labor. 2007. Employment and Training Administration. Training and Employment Guidance Letter No. 19-06. "Negotiating Performance Goals for the Workforce Investment Act Title 1B Programs and Wagner-Peyser Act Program for Program Years 2007 and 2008." Washington, DC: U.S. Department of Labor, Employment and Training Administration.

U.S. General Accounting Office. 2002. Workforce Investment Act: Improvements Needed in Performance Measures to Provide a More Accurate Picture of WIA's Effectiveness. Washington, DC: U.S. General Accounting Office.

Wholey, Joseph S. and Harry P. Hatry. 1992. "The Case for Performance Monitoring." Public Administration Review. 52(6): 604-610.

Wilde, Elizabeth T., and Robinson Hollister. 2007. "How Close Is Close Enough? Testing Nonexperimental Estimates of Impact Against Experimental Estimates of Impact with Education Test Scores as Outcomes." *Journal of Policy Analysis and Management*. 26(3): 455-477.

Workforce Enterprise Services, Inc. 2007. "Chicago Workforce Board Cost per Participant Study Final Report; Cost Per Participant Comparison Study - Phase Two; Final Report."