

APPROACHES TO ADJUSTING WORKFORCE DEVELOPMENT PERFORMANCE MEASURES

Ray Marshall Center for the Study of Human Resources • Lyndon B. Johnson School of Public Affairs
The University of Texas at Austin • 512-471-7891 • <http://www.utexas.edu/research/cshr/>

This policy brief by Joe Siedlecki and Christopher T. King presents alternative approaches for adjusting performance measures for publicly funded workforce development programs and the rationale for making such adjustments. It also discusses the advantages and disadvantages of these approaches.

Introduction

While performance measurement as a management tool has had a long history, it is “primarily in the last two decades that public-sector performance management has shifted to an explicit focus on measuring outcomes and rewarding results.”ⁱ Workforce development, from the Comprehensive Employment and Training Act Amendments (CETA) of 1978 and the Job Training Partnership Act (JTPA) of 1982 to the Workforce Investment Act (WIA) of 1998, is a policy area that has long required states and localities to track outcomes and measure performance. Performance measurement in workforce policy has evolved over time. But the evolution has not always moved in exactly the same direction. For example, under JTPA, the U.S. Department of Labor’s Employment and Training Administration (USDOL/ETA) provided specific departure points for each of the key performance measures that could be adjusted locally based on national regression models. Under WIA, USDOL/ETA initially prescribed expected performance levels for states with very little negotiation, while doing little to encourage adjustments for varying local conditions.ⁱⁱ

Today, workforce performance measures establish the basis for accountability, offer evidence of program effectiveness, and provide customer feedback to foster continuous improvement. Accordingly, workforce development policy makers and program administrators have embraced performance measurement as a tool for improving their services and proving their worth. This embrace of performance measurement is not without significant obstacles. A major concern in the measurement of performance is ensuring accurate context for these measures. Measuring performance is difficult, but ensuring fairness in measurement adds another level of complexity. Workforce development service provision does not take place in a vacuum. The demographics of the population in a workforce development area, the geographic layout of the area, as well as its’ economic and social conditions may vary considerably. Adjusting performance outcomes to reflect these differences, while maintaining some level of consistency across areas, is a challenge for policy makers and administrators.

This issue brief provides an outline of the major alternatives for adjusting workforce development performance and explores a number of key issues facing policymakers and practitioners. The focus of the issue brief is on WIA though these same performance measurement issues apply to related workforce development and education programs as well.

Background

WIA requires that specific measures be used to track outcomes for three discrete populations: adults, youth, and dislocated workers. The U.S. Office of Management and Budget (OMB) has developed a set of four core measures the following populations as shown in Table 1.ⁱⁱⁱ

TABLE 1. OMB COMMON PERFORMANCE MEASURES FOR WORKFORCE PROGRAMS

<i>Adults & Dislocated Workers</i>	<i>Youth</i>
Entered Employment	Placement in Employment or Education
Retention	Attainment of Degree or Certificate
Earnings Increase	Literacy or Numeracy Gains
Efficiency (Cost Per Participant)	Efficiency (Cost Per Participant)

Source: USOMB (2004).

Additionally data must be collected on the customer satisfaction of jobseekers and employers.

States are held accountable for their performance in these specific areas. Failing to meet expected performance levels can result in a financial penalty in the form of a 5% reduction in the size of the WIA grant provided to the state. Conversely, meeting or exceeding expected performance levels may result in the awarding of additional funds or incentive grants.^{iv} Notably, many states also use these or very similar measures to monitor the performance of their local workforce areas.

In order to make objective decisions about which states receive incentives and penalties, WIA seeks to ensure that states collect and report comparable data on WIA service users. WIA requires that these data and performance measures be calculated using Unemployment Insurance (UI) wage records. This is a significant departure from data collection under the early years of JTPA, when program administrators were required to collect post-program information via telephone interviews.

State and local workforce development administrators, as well as GAO, have expressed concerns with performance measurement under WIA.^v More specifically, some administrators have decried the lack of direction from the U.S. Department of Labor regarding how and when to track WIA participants. Additional complaints include the high levels of expected performance that were, initially, based on a small body of largely non-comparable data and did not sufficiently take into account variations in local demographic or economic conditions. Finally, state officials worry that UI wage records have inherent time lags in their reporting that cannot be avoided. As a result, states have difficulty identifying performance deficiencies in time to correct them.

In an attempt to mitigate perceived problems with the OMB common measures, a handful of leading-edge states and local areas, working with the National Governors Association and university partners with USDOL/ETA support, have developed a performance measurement framework that is more comprehensive, as well as more consensus based.^{vi} The Integrated Performance Information (IPI) Project has provided USDOL/ETA with input from the states and local areas regarding potential workforce system performance measures, include employer outcomes, labor market outcomes for jobseekers, social welfare outcomes, customer satisfaction, skill gains, and return on investment, as well as subgroup or comparative information. In

addition, the IPI report explicitly calls for adjusting performance measures to account for variations in local demographic and economic conditions.

Rationale for Adjusting Performance Measures

Workforce performance measurement, especially when coupled with incentives and sanctions, tends to be a high-stakes proposition. Exceeding established performance levels could result in significant funding increases. Falling short of performance levels could lead to decreased funding or even sanctions. States report that the need to meet these performance levels may lead local staff to focus on providing services to those participants who are most likely to succeed in their job search or make wage gains.^{vii} Serving better situated or more advantaged groups—thus, limiting services to disadvantaged or more difficult-to-serve participants—in order to improve performance outcomes is known as “creaming.” One major argument for adjusting performance measures is to prevent or, at least, diminish the motivation for creaming.

State and local demographic and economic conditions can vary substantially. Using national or even state-level data to develop expected performance outcomes can penalize localities that operate in particularly harsh economic environments. Holding a rust-belt factory town to the same performance standard as a small city with many financial services type employers, even within the same state, is not only unfair, but also unproductive. If the rust-belt area failed to attain the performance levels then it could lose funding. This sanction would not likely lead to the improvement of services. Adjusting performance outcomes helps to “hold harmless” states and localities for unfortunate local economic conditions that are beyond their control. An analogous demographic example would be the concentration of a disadvantaged population (such as high school drop-outs) in a particular area.

Local economic conditions in a particular area can also change abruptly. In a particular workforce development area that was once economically stable, the downsizing by several large employers could potentially alter the demand for WIA-type services, while the supply of available local jobs also decreases. Adjusting performance, if it can be done in a timely manner, provides some amount of cushion for such unexpected and generally uncontrollable events.

The recent interest in performance measurement might be “masking the truth about social programs.”^{viii} The use of performance measures by a system or program does not necessarily ensure greater efficiency. Indeed, sometimes the implementation of performance measures is seen as an end, in and of itself. Avoiding a philosophical pitfall such as this requires policy makers and administrators to dedicate time and resources to determining the proper ways of adjusting performance.

In summary, the rationale for adjusting performance outcome levels includes encouraging the provision of services to traditionally hard-to-serve populations, holding local programs harmless for difficult local demographic or economic conditions that lie outside the control of program administrators, buttressing against sudden changes in local economic conditions, and ensuring a more accurate system of measurement and comparison.

Alternative Approaches to Adjusting Performance Outcomes

There are two major alternatives available for adjusting workforce development performance outcomes: regression-based models, and what have come to be known as negotiations-based models. Each is reviewed briefly here.

The idea of using regression-based models for use in workforce system performance measurement was conceived in the late 1970s during the development of the Work Incentive

Program approach. But, regression-based models and related approaches were not formally incorporated into workforce development performance management systems until their use in the early 1980s as part of JTPA.^{ix} As part of the CETA and early JTPA program, analysts developed regression models to measure and predict employment and earnings of program participants.^x

Using historical employment data, DOL developed a set of national standards that local JTPA service delivery areas were expected to meet or exceed. But a desire to hold local programs harmless for variations in local economic conditions led to the use of USDOL-designed regression models to adjust the level of expected performance in light of local economic conditions and the characteristics of the program participants. States had considerable discretion regarding performance expectations in JTPA. State officials could use either the national standards that had not been adjusted for local conditions, employ the standard DOL-designed regression model, or make adjustments to the USDOL model for further variations in local conditions. In summary, JTPA performance outcomes were developed using a USDOL regression model that used data from programs across the nation, but took into account local economic conditions and program participant characteristics.

The implementation of WIA changed the method in which performance standards are set. A regression model is no longer applied to the national “departure points.” Rather, expected performance outcomes for state and local areas are the result of negotiations between state and local officials and USDOL. Importantly, no automatic adjustments are made to take into account the local variations of economic conditions or the characteristics of program participants (Barnow and Smith 2004). Adjustments for such factors can be negotiated, but rarely were in the early stages of WIA implementation. State and local officials have expressed concern with their lack of bargaining power *vis a vis* USDOL. Many indicated that, for the most part, USDOL unilaterally established the standards. (Barnow and King 2005). States then tended to pass the WIA performance measures down to the local workforce areas, often without any further negotiation (D’Amico et al. 2005).

Variants of the negotiations-based and older regression-based expected models are currently being explored. John Baj at Northern Illinois University has developed the Federal Records and Evaluation Database (FRED) tool to assist in supplying comparative data for negotiations. The State of Michigan, with funding from USDOL’s Employment and Training Administration, has developed and tested a modified regression model that could be used as a tool for adjusting WIA performance outcomes. Researchers at the W. E. Upjohn Institute for Employment Research are conducting the research for this effort.

FRED is described as a “performance management and negotiation support tool.” The USDOL-funded online database (www.fred-info.org) emerged from the desire of USDOL, as manifested in the DOL Administrative Data and Research Evaluation (ADARE) project, to make the data it collects from state and local partners more accessible to the training and employment community. FRED enables analysis of the WIA Standardized Record Data (WIASRD), which is data submitted annually by states on WIA exiters' demographic characteristics, the services they received, and the outcomes they achieved after program exit. The data can be extracted for trend analysis, specific goal analysis, and/or even ad hoc reporting. Rather than setting expected standards or monitoring performance outcomes in a vacuum, FRED allows state and local officials to compare data and results for similar workforce areas throughout the country. Proponents of FRED as a negotiations tool note that using annual WIASRD data incorporates most of what a regression model attempts to control for. Unless local economic or participant demographic conditions suddenly and significantly shift, the picture painted with data for the most recent year tends to be a relatively good guide for the conditions in the current year or, as Baj puts it, “the best guide for predicting this year’s performance is last year’s performance.”

The goal of the Michigan project is to develop performance management adjustment tools that develop fair measures of local workforce program performance, add a new level of “value-added” measures and develop timely predictors of performance so that administrators may identify and correct problems before the final performance report is submitted to USDOL.^{xi} The major management tool is a regression model that attempts to control for local economic, demographic and other factors. Using individual-level data from state UI wage records, the regression model quantifies the impact that participant and economic conditions have on outcomes in local areas. Performance expectations are then set accordingly. This approach differs from the earlier JTPA approach in that the Michigan project *allows for the calculation of an adjustment factor for each discrete individual who receives services*, whereas the JTPA approach was based more on aggregate, area-based calculations. It is based on the Upjohn Institute’s pioneering work on worker profiling systems for UI claimants and on the front-line decision-support model in Georgia. The Michigan project approach provides timely data, seeks to create an accurate and fair expected performance standard, and allows for better targeting of resources at the micro level. The adjustment of expectations based on individual participant characteristics encourages programs to serve the traditionally hard-to-service populations.

Advantages of Adjusting Performance Outcomes

There are numerous advantages to adjusting performance outcomes. First, adjusting outcomes would help *align WIA operations and accountability with the letter and spirit of the law*. WIA legislation states that state level standards are supposed to be set “taking into account factors including differences in economic conditions, the characteristics of participants when the participants entered the program, and the services to be provided” (Section ##).

Second, and more importantly, adjusting performance outcomes might better hold state and local areas harmless for difficult economic conditions or participant characteristics as discussed above. Holding areas harmless for varying conditions could mitigate “creaming,” minimize program administrators’ incentives for “gaming” their reported performance levels, and encourage administrators to provide greater levels of service to harder-to-serve participants. In effect, *adjusting expected performance encourages greater access to services for those who need them most*.

Finally, more recent variants of standards setting—both FRED and the Michigan project—allow administrators *access to more accurate and consistent data* in a timely fashion. Using predictive models or comparisons across workforce development areas with similar economic and demographic characteristics should allow state and local program operators to more accurately compare outcomes, as well as to identify and address program and service deficiencies before final performance levels are submitted. In this manner, adjusted performance standards and the tool used to develop them also serve as better tools for continuous program improvement.

Disadvantages of Adjusting Performance Outcomes

The major drawback to adjusting performance outcomes is that setting lower performance thresholds *might not encourage state and local officials to reach beyond the goals*. This is not to say that these officials do not strive to improve their systems and services, but rather they juggle numerous tasks and have to meet many different performance standards. That said, adjusting expected performance outcomes does not necessarily mean each expected performance outcome might be lower. It may be the case that data support raising the expected performance outcome for a particular group of participants.

Adjusting performance outcomes may also tend to be overly fatalistic, *always settling for lower outcomes for harder-to-serve participant populations*. A “stretch goal” might be exactly what is necessary to help push members of a hard-to-serve population into a better situation from whence they came. Overly adjusted expected performance levels might inhibit such progress.

Finally, it must be noted that the regression models used in JTPA and studied in Michigan do not have strong predictive power (that is, they have relatively low R^2 scores): stated another way, the typical models developed for JTPA could only account a very small share of the variations in performance among areas. Adjusting performance based on relatively weak regression models might not even sufficiently control for the desired variables. Or it may leave out undiscovered or unmeasured influences on outcomes. This would result *in managing to the regression model, rather than on behalf of jobseekers or the local community*.

Concluding Observations

There appears to be widespread agreement among policy makers, administrators, and government officials that workforce development system performance measurement can be improved. Several options have been developed and piloted. Projects, such as the Integrated Performance Information Project, provide a meeting place for interested parties to discuss potential changes and enhancements to the existing system. The ongoing WIA reauthorization debate provides a forum for many such changes, as do the related ones for the Perkins and TANF legislation.

There are two major options for adjusting workforce development system performance outcomes: regression-based and negotiation-based models. Regression-based models, like those espoused by John Baj, use historical data to predict or model future/present outcomes. Regression-based models are useful in identifying economic and demographic conditions associated with particular outcomes but, in addition to being relatively weak, are essentially backward looking. Such models might serve better for managing rather than setting standards. For example, knowing that a particular economic or demographic event is highly associated with particular outcomes could prompt an administrator could shift workforce development strategies, rather than lower performance expectations.

Negotiation-based models, when supported with relevant data, have the ability to take into account historical information and trends while providing the latitude to determine goals that serve to encourage longer-term benefits to the community. Simply, negotiation-based models use data to set baseline, but allow room for developing “stretch goals.”

Aside from adjusting performance measures, policy makers should also consider providing more guidance to administrators in areas that have caused confusion or that have been prone to creaming. Currently legislation and direction from USDOL is unclear regarding when a potential program participant should be registered to be included in the data on program participants. Simply providing a clear definition might mitigate creaming. Additionally, policy makers might set aside incentive monies for those workforce development areas whom take on relatively large numbers of hard-to-serve participants.

For additional information, please contact Chris King at ctking@uts.cc.utexas.edu.

This brief was developed in conjunction with the Workforce Development System Mapping and Performance Improvement Project, an initiative of the Rockefeller Foundation.

Endnotes

ⁱ Heinrich, Carolyn J (2003), "Measuring Public Sector Performance and Effectiveness" in *The Handbook of Public Administration*, Guy Peters and Jon Pierre (eds.), London: Sage Publications, pp. 25-27.

ⁱⁱ See Barnow, Burt S. and Christopher T. King (2005), *The Workforce Investment Act in Eight States: Final Report*, Washington, D.C.: US Department of Labor, ETA Occasional Paper 2005-01, February; Barnow, Burt S. and Jeffrey A. Smith (2004), "Performance Management of U.S. Job Training Programs," In Christopher J. O'Leary, Robert A. Straits, and Stephen A. Wandner, Eds., *Job Training Policy in the United States*, Kalamazoo, Michigan: W. E. Upjohn Institute for Employment Research; and D'Amico, Ronald et al. (2004), *The Workforce Investment Act after Five Years: Results from the National Evaluation of the Implementation of WIA*, Oakland, CA: Social Policy Research Associates, June. Prepared for the U.S. Department of Labor, Employment and Training Administration, Office of Policy Development, Evaluation, and Research.

ⁱⁱⁱ U.S. Office of Management and Budget (2004). *Memorandum for Heads of Executive Departments and Agencies & Addendum*, Washington, DC, April 24. Online. Available: http://www.whitehouse.gov/omb/memoranda/m02-06_addendum.html#h2

^{iv} U.S. Government Accountability Office (2004). "Workforce Investment Act: States and Local Areas Have Developed Strategies to Assess Performance, But Labor Could Do More To Help", GAO-04-657, Washington, DC, May.

^v See Barnow and King (2005), Barnow and Smith (2004), and USGAO (2004).

^{vi} Wilson, Bryan (2005). *Integrated Performance Information for Workforce Development: A Blueprint for States*, Olympia, WA: Washington State Workforce Training and Education Coordinating Board, February.

^{vii} USGAO (2004).

^{viii} Blalock, Ann B. and Burt S. Barnow (2001). "Is the New Obsession with 'Performance Management' Masking the Truth About Social Programs?" In Dall W. Forsythe, Ed., *Quicker, Better, Cheaper? Managing Performance in American Government*, Albany, New York: Rockefeller Institute Press.

^{ix} See Barnow and Smith (2004).

^x See Dickinson et al. (1988).

^{xi} Eberts, Randall W. and Christopher J. O'Leary (2004), *A New WPRS Profiling Model for Michigan*, Kalamazoo, MI: The W. E. Upjohn Institute for Employment Research, Working Paper 04-102.