

ORIGINAL ARTICLE

Open Access

Do Public Employment and Training Programs Work?

Carolyn J Heinrich¹, Peter R Mueser^{2*}, Kenneth R Troske³, Kyung-Seong Jeon⁴ and Daver C Kahvecioglu⁵

* Correspondence:

mueserp@missouri.edu

²Department of Economics and Harry S. Truman School of Public Affairs, University of Missouri, IMPAQ International, LLC, and IZA, 118 Professional Building, Columbia, MO 65211, USA

Full list of author information is available at the end of the article

Abstract: We estimate impacts on earnings and employment of the two primary adult workforce support and training programs under the U.S. Workforce Investment Act (WIA) using administrative data on 160,000 participants from 12 states for up to four years following program entry. We find that participants in the WIA Adult program, who typically enter with poor work histories, realize improved employment levels and increased average quarterly earnings of several hundred dollars. Earnings gains for Dislocated Worker program participants are appreciably smaller, although these participants do experience employment gains.

JEL codes: I38, J08, J24

Keywords: Job training; Program evaluation; Workforce investment act; Matching

1. Introduction

The U.S. labor market currently faces the fallout from one of the most severe economic downturns since the 1930s. Even with economic growth returning, unemployment remains well above normal, with most observers predicting that the jobless rate will not return to pre-recession levels for several years. The length of time workers are unemployed also rose to unprecedented levels (Bureau of Labor Statistics 2012). The American Recovery and Reinvestment Act of 2009 allocated nearly \$2 billion in additional funds to the Workforce Investment Act (WIA) in an effort to raise individual skill levels and improve job seekers' prospects and outcomes.

Enacted in 1998, a central goal of WIA was to create a new, comprehensive workforce investment system. WIA differed from its predecessor, the Job Training Partnership Act (JTPA), in the introduction of a system of centralized centers designed to improve coordination of employment services; the inclusion of job search with training in an integrated sequence; the use of training vouchers; and significant changes in governance structures at the state and local level. It also shifted responsibility for some types of activities believed to contribute little to performance outcomes (such as adult basic education) to other programs. Initially, WIA also reduced the share of low-income individuals served and appreciably reduced the number of adults receiving training relative to JTPA (Osterman 2007).

Despite the emphasis placed on WIA in addressing current economic problems, little is known about its impacts on labor market outcomes. In 2007, the U.S. Office of Management and Budget assigned the WIA program low marks for its evaluation efforts, suggesting that existing evaluations had not been of sufficient scope and rigor to

determine WIA's impact on participants' employment and earnings. The evaluation we present of the WIA programs serving adults was initiated by the U.S. Department of Labor (DOL) in response to the administration's concern to obtain rigorous information on the effectiveness of the U.S.'s largest, publicly-funded training programs for adults. The evaluation mandate set a relatively short timeframe for producing results, which compelled a nonexperimental design.

Our analysis covers program entrants between July 2003 and June 2005. The data are from 12 states that cover approximately 160,000 WIA participants and nearly 3 million comparison group members. Within each state, we employ matching methods to compare WIA program participants with individuals who have not participated in the WIA program but who are observationally equivalent in terms of demographic characteristics, prior participation in employment programs, and labor market experiences. Participants and comparison group members are compared within labor market areas to ensure that they are facing similar local labor markets, and measures of employment and earnings are fully comparable for program participants and the comparison group.

The study results show that, for the average participant in the WIA Adult program, participation is associated with a several-hundred-dollar increase in quarterly earnings. Adult program participants who obtain training have lower earnings in the months during training and the year after exit than participants who don't obtain training, but they catch up within 10 quarters, ultimately registering large total gains. The marginal benefits of training are over \$400 in earnings per quarter three years after program entry. The earnings of Dislocated Workers are depressed over several quarters following entry into WIA. As a group, their earnings ultimately match or overtake the comparison group, but the benefits they obtain are smaller than for those in the Adult program. Although the absence of cost data precludes a comprehensive benefit-cost analysis, we conclude that the Adult program very likely satisfies such a test, whereas the Dislocated Worker program does not.

In the next section, we provide a brief literature review focusing on studies examining WIA and related job training programs. Next, we describe the structure of WIA Adult and Dislocated Worker programs, followed by discussion of our data and methods. Subsequent sections examine outcomes for the Adult program and the Dislocated Worker program. We then consider selected subgroup analyses and conclude.

2. Literature

Perhaps the most influential study of job training was the random assignment evaluation of WIA's predecessor, the JTPA program, focusing on disadvantaged workers—those with unstable work histories and low earnings—who participated in the program in the late 1980s (Orr et al. 1996). The study found statistically significant but modest effects of job training on ultimate earnings for adult disadvantaged workers. There are no large-scale random assignment studies of job training for dislocated workers, i.e., individuals who became unemployed after losing jobs (often after extended periods of stable employment).

A growing number of nonexperimental studies in Europe, covering a variety of programs providing both job search assistance and job training, take advantage of exceedingly detailed administrative data. Card et al. (2009), in an international meta-analysis of training program evaluations, found that longer-term job training programs tended to have

small or negative impacts on employment or earnings in periods of less than a year, presumably reflecting “lock-in” effects due to withdrawal from the labor market during training, but that impacts often turned positive in the second or third years. They also concluded that “research designs used in recent nonexperimental evaluations are not significantly biased relative to the benchmark of an experimental design” (p. 26). A meta-analysis by Greenberg et al. (2006) reached a similar conclusion.

Hollenbeck et al. (2005) examined labor market outcomes in seven states for WIA participants who had completed the program in the period July 2000-June 2002, but data limitations and the focus on the early years of program implementation (WIA was adopted in 2000 in most states) raise questions about the validity of reported estimates. Three other studies covering two additional states (summarized in Hollenbeck 2011) examine labor market outcomes for WIA, but again data issues limit inference¹.

Andersson et al. (2013) report results of a preliminary analysis of WIA training in two states. Their work, which is limited to comparisons of WIA participants receiving training with WIA participants receiving other services, provides a replication of a subset of the results reported below. Andersson et al. linked state data to the Longitudinal Employer Household Dynamics (LEHD) data and found moderate positive impacts on employment and earnings for adults, but did not find methodological benefits from the adding conditioning variables from the LEHD (compared to data available in state UI records). Decker (2011) provides a useful review of other impact studies of WIA and related programs.

A substantial literature attempts to identify nonexperimental methods that can replicate results of random assignment experiments (Heckman et al. 1999; Bloom et al. 2005; Mueser et al. 2007; Cook et al. 2008; Pirog et al. 2009). It is widely recognized that controls for basic demographic characteristics and prior labor market activities of individuals are critical, that comparisons must be between program participants and the comparison group within labor markets, and that outcome measures must be measured in fully comparable ways. The methods employed in our study correspond closely to these recommendations. The approach presented here is replicated in an evaluation of the Trade Adjustment Assistance Act undertaken by Mathematica Policy Research (Schochet et al. 2012).

3. The structure of WIA adult and dislocated worker programs

WIA legislation specifies three levels of service in both the Adult and Dislocated Worker programs. The first level is core services, which include staff-assisted job search and placement, provision of labor market information, and basic counseling. The next level is intensive services, which involve comprehensive assessment, more extensive counseling and career planning, and possibly short courses. The third level is training services, which are usually provided by outside vendors (often community colleges, proprietary schools or nonprofits) through a voucher called the Individual Training Account (ITA). All participants who enter the program receive core services, whereas selected individuals progress to intensive and, possibly, training services. Access to WIA core services is not limited by federal legislation, but higher levels of service are formally restricted to those in targeted groups, which includes workers with poor work histories (Adult program) and those who have been laid off (Dislocated Worker program). In addition, both state and local authorities administering WIA (the state workforce agencies, and the Workforce Investment Boards or WIBs) have latitude in implementing WIA

programs, with WIA program staff retaining power to determine who is accepted into the program, and, of those accepted, who receives intensive and training services (Social Policy Research Associates 2004; Rockefeller Institute of Government 2004). Programs generally face budget constraints that limit the number of participants who can receive training, so training slots may be a scarce resource allocated by staff.

During program year 2004 (July 2004–June 2005), published tabulations show that nationwide about one in five WIA participants received only core services, and about two in five received training services. Of those engaging in training, up to 10 percent received on-the-job training and another 5 percent received basic skills training; the remainder received occupational and other training, including customized training designed to meet the needs of particular employers. Between a third and a half of participants exited WIA in less than 26 weeks, and a similar proportion remained in the program for at least a year (Social Policy Research Associates 2006). Those who obtained training spent approximately 25–40 percent longer in the program than those who did not².

Although many individuals would be eligible for either the Adult or Dislocated Worker program, the average participants in these programs differ quite dramatically in terms of gender, age, race and prior work experience. Given that the programs also tend to serve different functions, we analyze each separately.

4. Data and method of analysis

Study sample

Beginning in December 2007, workforce development agencies in all 50 states were contacted to request their participation in the study. Agreements through which necessary data were released to the researchers were reached with 12 states: Connecticut, Indiana, Kentucky, Maryland, Missouri, Minnesota, Mississippi, Montana, New Mexico, Tennessee, Utah, and Wisconsin³. These states account for about a fifth of the nearly 600 U.S. Workforce Investment Areas and a similar proportion of the participants in WIA's two main job training programs serving adults. Although we cannot claim to estimate a "national" average impact of WIA—in fact, no experimental or nonexperimental evaluation has generated impact estimates based on a representative sample for WIA or any of its predecessor programs—the sample of WIA participants we consider suitably reflects the diversity of local Workforce Investment Board areas, in terms of geography, population characteristics, and organizational configuration.

All analyses are based on state administrative data, with files identifying program participants and comparison group members drawn from each state. Estimates are obtained by methods that match a program participant (a treated case) with a comparison case that has the same characteristics and the same prior labor market and program experience. Such matching is undertaken separately for the Adult and Dislocated Worker programs by gender within each state. Three sets of estimates are obtained in each case. The first estimates the impact of the program taken as a whole, without regard to services received. Here the treated group is WIA participants and the comparison group is individuals who filed for Unemployment Insurance (UI) benefits (nine states) or participated in Employment Service (ES) activities (three states), and did not enter WIA. The UI and ES comparison groups are very similar in important respects, and both receive minimal services that are universally available.

The second set of estimates focuses on WIA participants who were receiving UI benefits at the time they entered WIA. The comparison group is UI recipients who did not enter WIA, and this analysis is limited to the nine states where UI is the comparison group. The third set of estimates identifies the extent to which training, per se, is associated with employment and earnings outcomes. Here the treated sample is WIA participants who received training, and the comparison group is WIA participants who did not receive training.

To what degree can these results be generalized to WIA in the country as a whole? The clearest threat to generalization would be if states with high-performing programs had been selected (or had selected themselves) into the study. Although this possibility cannot be ruled out with certainty, previous work suggests that there are no easily-observable factors that predict program impact (Orr et al. 1996; Heckman et al. 2011), particularly when considering program impacts occurring years after participation. It would appear unlikely that more effective programs would be able to identify themselves, let alone pursue a strategy of arranging to provide data. We suspect that state administrative and data handling idiosyncrasies played a dominant role in determining state-agency willingness to provide data for the study.

Data sources

Data on WIA participants were obtained from annual Workforce Investment Act Standardized Record Data (WIASRD) or closely related files maintained by states for administrative purposes. In most states, the data files include those who exited the program from July 2003 through June 2007, along with an individual identifier allowing a match with other state data. We focus on WIA participants who entered the program between July 2003 and June 2005 (inclusive), a pre-recession sample. No information is available on individuals who did not exit the program by June 2007, but our tests suggest that few participants were omitted for this reason.

Administrative files for the comparison group, with individual identifiers, were obtained from each state. These data were also used to control program participation prior to the quarter of program entry for both participants and comparison group individuals. In all but three states, at least six quarters of data are available prior to the first quarter of program participation for all participants and comparison group cases.

State Wage Record files maintained as part of the UI system provide quarterly earnings for all employees in UI-covered firms within a state. Although these data omit earnings from informal employment, employment outside the state, and employment in firms not subject to UI reporting requirements, studies show that estimates of program impacts from these data are not seriously biased as compared with those based on employee reports of earnings (Kornfeld and Bloom 1999; Wallace and Haveman 2007). Our earnings data extend through calendar year 2007, providing information on earnings and employment for participants and comparison group members up to 16 quarters following participation. Wage record information prior to WIA entry allows the construction of employment histories of participants and comparison group members. Earnings are adjusted for inflation to the first quarter of 2006. Individuals are defined as employed in a quarter if they have nonzero earnings⁴.

In addition to work histories, gender, age, education and race are available as control variables. Local labor market is captured using dummies identifying groups of counties

for residence or location of services received. It is widely recognized that controls for individuals' detailed labor market experiences in the two years immediately prior to program participation are critical (Heckman and Smith 1999), and we obtain such information from wage record data. All analyses are performed separately by gender. Where possible, WIA participants are matched with comparison cases (ES or UI recipients) observed in the same quarter that participants enter the program.

Descriptive statistics

Table 1 reports statistics for WIA participants and the comparison group. Our sample of WIA Adult program participants comprises 95,580 unique individuals, yielding a total of 97,552 entries. Dislocated Worker program participants comprise 63,515 individuals, with 64,089 total program entries. The rightmost column identifies the number of individuals who participate in comparison programs (UI claimants or ES participants) and are available to be matched to program participants. The upper entry indicates that approximately 2.9 million unique individuals are available, contributing nearly 6.2 million quarters of program activity. Since the units of analysis for the comparison group are quarters of program activity, we provide statistics for these units.

We see in Table 1 that individuals who participated in the WIA Adult program are more likely to be female and black and are also appreciably younger than individuals in the comparison population. The data on past employment for these groups provide

Table 1 Summary statistics for WIA participants and comparison group

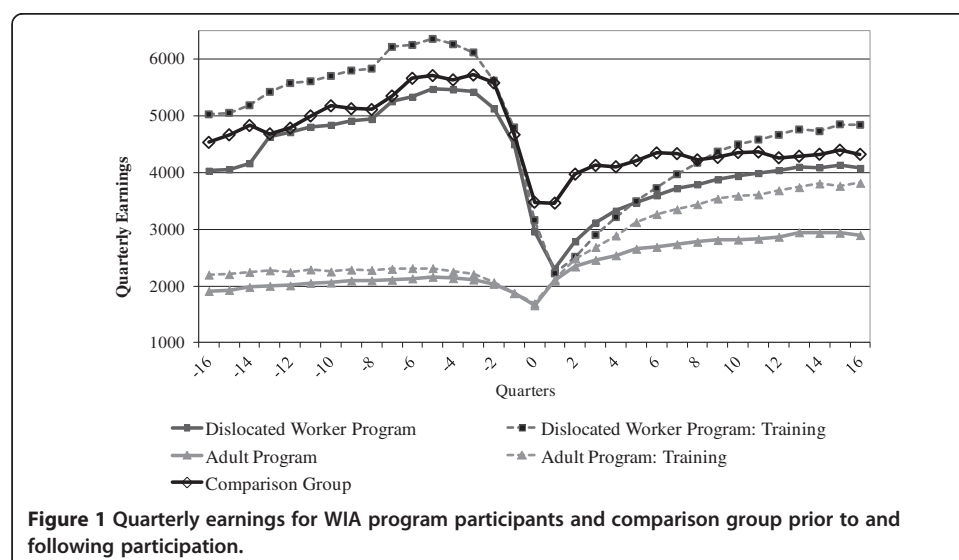
	WIA adult		WIA dislocated worker		Comparison group
	Overall	Training	Overall	Training	
Sample Size					
Unique individuals	95,580	27,325	63,515	20,002	2,929,496
WIA entries, or quarters of comparison program participation	97,552	27,840	64,089	20,195	6,161,510
Demographic					
	Mean	Mean	Mean	Mean	Mean
Male	0.420	0.356	0.482	0.456	0.585
Black	0.445	0.277	0.330	0.198	0.171
Hispanic	0.031	0.072	0.022	0.043	0.064
Age	32.70	32.16	40.24	40.46	39.59
Years of education	12.27	12.43	12.55	12.63	12.42
Employment transitions					
Employment-employment	0.297	0.307	0.462	0.456	0.476
Employment-not employed	0.208	0.241	0.281	0.335	0.279
Not employed-employed	0.325	0.297	0.183	0.149	0.225
Not employed-not employed	0.168	0.151	0.070	0.053	0.040
Program experience					
WIA in prior two years	0.052	0.035	0.041	0.034	0.020
Comparison program participation in prior two years	0.211	0.297	0.409	0.551	0.668

Note: Employment transitions are based on quarterly earnings data in prior five quarters and current (entry or participation) quarter, coded into four exclusive categories. Employed-employed identifies cases with all positive earnings and not employed-not employed identifies cases with no earnings. Other cases are coded by the most recent employment transition.

evidence that participants in the WIA Adult program have weaker labor market attachment than comparison program participants. In contrast, WIA Dislocated Worker program participants have similar labor market attachment to those in the comparison program. Employment differences are confirmed in Figure 1, which provides average earnings for participants in the two WIA programs and for the comparison group. The horizontal axis identifies quarters relative to the quarter of entry into WIA (for WIA participants), or quarter of participation in the comparison activity (for comparison group members). Earnings for the comparison group in prior quarters are much higher than for those who enter the WIA Adult program, while prior earnings of WIA Dislocated Worker are quite similar to those of the comparison group.

The bottom panel of Table 1 shows that 4 to 5 percent of WIA entrants had previously participated in WIA (in either program). About a fifth of Adult program participants had prior comparison program (UI or ES) experience, compared to over two-fifths of Dislocated Workers. By definition, a comparison case participates in the comparison program in the specified quarter; about two-thirds of such individuals had participated in that program in the prior two years. Table 1 also shows that, within each program, participants who receive training are more likely to be female and much *less* likely to be black than participants who do not receive training.

Despite the differences noted above, the patterns of earnings for WIA and comparison group participants show marked similarities. For both groups, the most notable pattern in Figure 1 is a decline in earnings that occurs in the several quarters prior to program entry, a pattern that has been called the “Ashenfelter dip” (Heckman and Smith 1999). This reflects the fact that individuals often enter such programs following a period of labor market setbacks. The fact that the comparison group has a similar basic pattern in earnings to the two WIA programs implies that there will be sufficient cases to match with WIA participants on the basis of prior employment. Equally important, the common patterns suggest that there may be similarities in the individual employment circumstances faced by the comparison and treatment groups, so that unmeasured factors may be similar as well.



Methods

We use propensity score matching to estimate program impacts for those who participate in the program. We control for standard demographic factors (gender, race/ethnicity, age, education), calendar quarter of program entry (or quarter of participation for the comparison group members), disability and veteran status, employment information based on wage record data over the two years prior to program entry, including earnings and industry, and program participation history (UI or Wagner-Peysor, and WIA) up to four years prior to WIA entry.

Like other matching and related methods, the approach assumes that the outcome that would occur in the absence of the treatment is conditionally independent of the treatment⁵. Although the conditional independence assumption cannot be tested directly, we apply a specification test that examines earnings in the tenth and sixteenth quarters prior to program entry, points in time before earnings and employment measures that are controlled. If controls are successful, then these prior earnings should not differ for treated and matched comparison cases. In contrast, if differences in stable factors that influence subsequent earnings exist between the treatment and matched comparison groups, we would expect there to be differences in the conditional means of earnings in these earlier periods.

If the specification test suggests unmeasured stable factors are important, difference-in-difference fixed effects estimators can provide unbiased impact estimates under some assumptions. For example, this would be the case if program participants were selected on stable personal characteristics that had similar impacts on earnings or employment prior and subsequent to treatment. Mueser et al. (2007), for example, argue that difference-in-difference estimates are valid as measures of a training program's impact in an environment in which prior earnings of participants differ from matched comparison group members. More generally, however, depending on the processes underlying earnings dynamics and program participation, these estimates may have biases that are not present in cross-sectional matching estimates. The difference-in-difference estimator needs to be understood as one of several estimators that are valid under alternative assumptions (Smith and Todd 2005; Jung and Pirog 2011).

Matching strategy

We use many-to-one caliper matching with replacement, based on the propensity score. The estimate of program impact, identifying the average effect of the treatment on the treated, can be written:

$$E(\Delta Y) = \frac{1}{N} \sum_{i=1}^N [Y_{1i} - \bar{Y}_{0j(i)}],$$

where $\bar{Y}_{0j(i)}$ is the average outcome for all comparison cases that are matched with treated case i , Y_{1i} is the outcome for case i , and N is the number of treated cases. Sometimes referred to as "radius matching", this approach does not limit the number of cases that are matched with a given participant, as long as those cases are within a specified radius, measured in terms of the propensity score. Mueser et al. (2007) found that methods like this one, which use all the available data, produced more precise impact estimates than one-to-one matching or other methods that discard potentially similar cases in the comparison group. The method is closely related to

propensity score stratification or interval matching (Rosenbaum and Rubin 1983). For difference-in-difference estimates, Y_{1i} is replaced by the difference between earnings following participation and earnings prior to program participation, and $\bar{Y}_{0j(i)}$ is replaced by the average difference for the matched comparison cases over the same period. In this application, matching is based on a constant radius expressed as the difference in the log-odds of the propensity score between treated and comparison cases.

The particulars of the matching parameters varied across states, reflecting differences in WIA participants and comparison group sample size, and the distribution and coding of variables. Matching is only possible when there are comparison group cases whose values on the control variables correspond with those for each participant case. We implemented matching criteria for each state that assured our matched comparison sample corresponded closely with the treated cases on all control variables.

One advantage of the very large comparison group of UI and ES participants available in this study is that it was possible to find cases matching almost all participants. In the analysis that compares WIA participants with comparison group members, the number of participants that could not be matched was very small, generally in the range of 2–7 percent. In the analysis focusing on training, where WIA training participants were matched with a comparison group of WIA participants who did not receive training, however, the proportion excluded was much higher, almost 50 percent for males participating in the Adult program. This is because it was necessary to omit analyses in several states with high proportions of individuals receiving training, as there were too few WIA participants without training to allow a meaningful matching analysis. Although such partial coverage calls into question the generalizability of these results, omitted states do not appear to be selective in any clear way, except that they represent small states with relatively large proportions trained.

We employed a “bias adjustment” procedure, which fits a linear model and uses this to adjust final estimates for any differences in the means for treated and comparison cases (Abadie and Imbens 2006), but this had little effect on reported estimates. This is what we expect, given the high quality of our matches (Cook et al. 2008). The Additional file 1 presents additional details on the matching methods and our results.

Conventionally, standard errors of propensity score matching estimates are obtained using bootstrap methods, but with large samples such as those available to this study, it is not feasible to calculate bootstrap standard errors for all estimates. We have chosen to report conditional standard errors using methods recommended by Imbens and Wooldridge (2008). We undertook limited comparisons with bootstrap standard errors and found similar results. Additional details are included in the Additional file 1.

We calculate estimates of impacts of the WIA programs on average inflation-adjusted earnings and employment for each state in the 16 quarters following program start. The mean impact across states is obtained by weighting the estimate for a given state by the number of participants who were matched in that state, providing the average impact for matched WIA participants. Associated with each state impact estimate is a conditional standard error, which is combined across states in the conventional way to form the standard error for the weighted average.

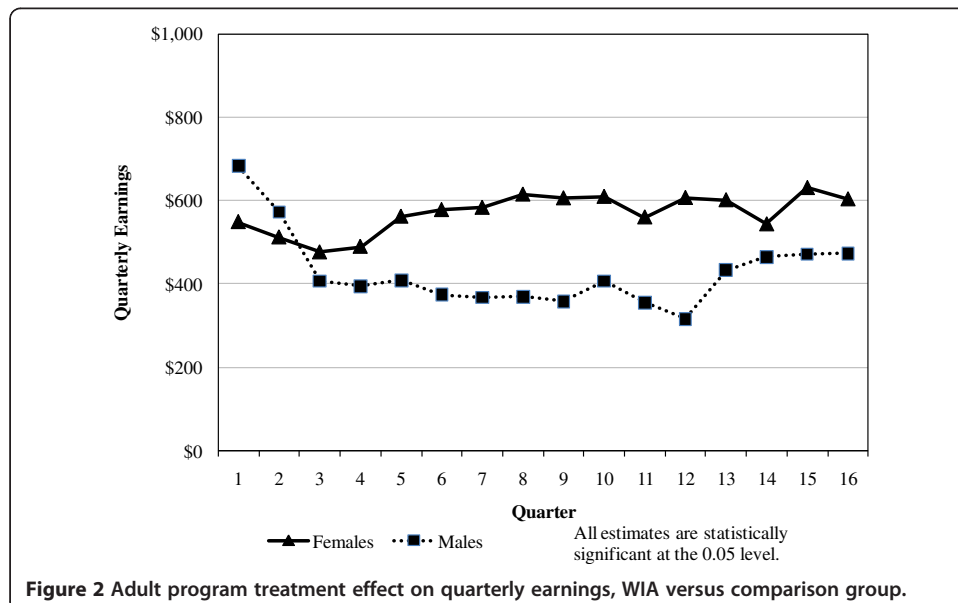
5. Impact estimates for adult program

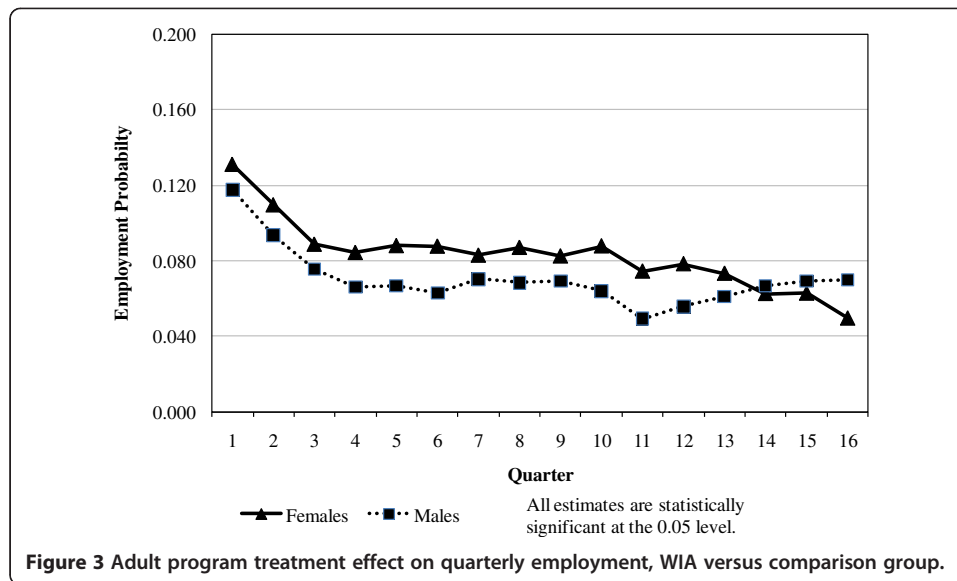
Figure 2 provides estimates of average Adult program impacts for women and men⁶. The horizontal axis extends from 1 to 16, identifying the quarter following program entry. The vertical axis is in dollars, indicating the difference between average earnings in a quarter for the WIA Adult program participants and matched comparison program participants.

Estimates imply that, for both genders, participants generally earn between \$400 and \$600 more per quarter as a result of program participation. For women, the estimate over most of the 16 quarters is between \$500 and \$600 per quarter, or about 30 percent of earnings, whereas for men there is a decline in the first three quarters, with the level settling in the range of \$400, or about 15 percent of earnings⁷. In their random assignment evaluation of the JTPA program, Orr et al. (1996) estimated inflation-adjusted impacts on quarterly earnings two years after program entry both for male and for female participants in the JTPA program in the \$300-350 range.

Figure 3 provides analogous estimates for employment—identified by nonzero income in the quarter—using the same method as that for estimating earnings impacts. Each value can be interpreted as the difference between the employment rate for Adult program participants and the matched comparison cases. For example, the estimate 0.13 for females in the first quarter after participation implies that the employment rate for participants is 13 percentage points higher than that for matched comparison cases. The basic pattern of results is quite similar to that for earnings. In particular, female participants' levels of employment—relative to the comparison group—decline from 13 percentage points to about 8 points within a year and ultimately to about 6 percentage points. Male impacts are one or two percentage points lower until the last three quarters. We estimate employment proportions for both men and women of about 0.5-0.6 in the absence of the program, so employment after the first year is increased by up to 15 percent.

There are substantial differences in the proportion of individuals receiving training across the state programs. The total resources per participant may be higher in states





with more active training programs, so long-run program impact could be higher if more intensive services produce greater impacts. In addition, a large share of any benefit may occur with a greater lag, since training returns presumably accrue over a more extended period than less time-intensive services. We tabulated impact estimates for the seven states that provided training to more than half of their participants; in these states taken together, 68 percent of Adult program participants received training. Effect estimates for the first several quarters after program entry in these seven states were very similar to the aggregate for all states, but earnings were higher in subsequent quarters by about \$200 per quarter. This provides some evidence that high-training states produce benefits that endure longer. The basic pattern for employment impacts is similar to that for earnings.

Our findings imply strong and substantial program impacts with little or no lag. In contrast, previous studies suggest that participants obtain little benefit initially—possibly experiencing earnings reductions due to “lock-in” effects—as they engage in training activities that supplant employment. In our data, the mean duration for participation in the program is between two and three quarters, so we expect that program participation would hinder participants’ early employment and earnings, consistent with previous studies. One possibility is that differential selection induces these observed estimates, in which case WIA participants have unmeasured attributes that make them more likely than those in the comparison program to obtain employment or higher earnings. This explanation implies that staff admission criteria or participant choice selected program entrants who would have obtained higher earnings and employment levels in the absence of participation than comparison group members with similar attributes and similar prior employment and program experiences.

As noted above, estimates based on the tenth and sixteenth quarters *prior* to entry provide a specification test for this kind of selection. These estimates, presented in Table 2, indicate whether, conditional on controls, there are differences in these prior earnings between WIA participants and the comparison cases. Note that controls are

included for earnings and employment patterns in the eight quarters prior to entry, so that our matching methods ensure that there are no differences in the two years prior to program entry. However, if there are stable factors that improve the employment prospects for treated cases relative to matched comparison cases, earlier earnings would be higher for the WIA cases. Positive estimates in this specification test would then suggest that estimates of the impact on subsequent earnings could be spurious.

Reported estimates produced by this specification test are not significantly positive, implying that levels of pre-program earnings and employment are *not* higher for WIA participants; in most cases, the differences are small (Table 2, line 1). The largest estimates in absolute value are for male WIA participants 16 quarters earlier, and in this case it appears that WIA participants had earnings about \$100 *below* those of the comparison group. Analogous estimates for quarterly employment (line 2) similarly suggest there is no stable factor causing employment levels to be higher for WIA participants. We can therefore infer that selection on stable individual characteristics is not causing spurious positive impact estimates.

Still, transient factors could also induce bias. Among comparison group members, those who receive UI benefits may have reduced incentives to obtain employment, since benefits are contingent on remaining unemployed. UI recipients classified as awaiting recall are not required to search for employment, and other UI recipients may have little interest in getting a job—despite formal requirements—until benefits are about to expire. WIA participants, in contrast, have chosen to enter a program with the purpose of improving their employment prospects.

Table 2 Specification tests for impact estimates using prior earnings and employment

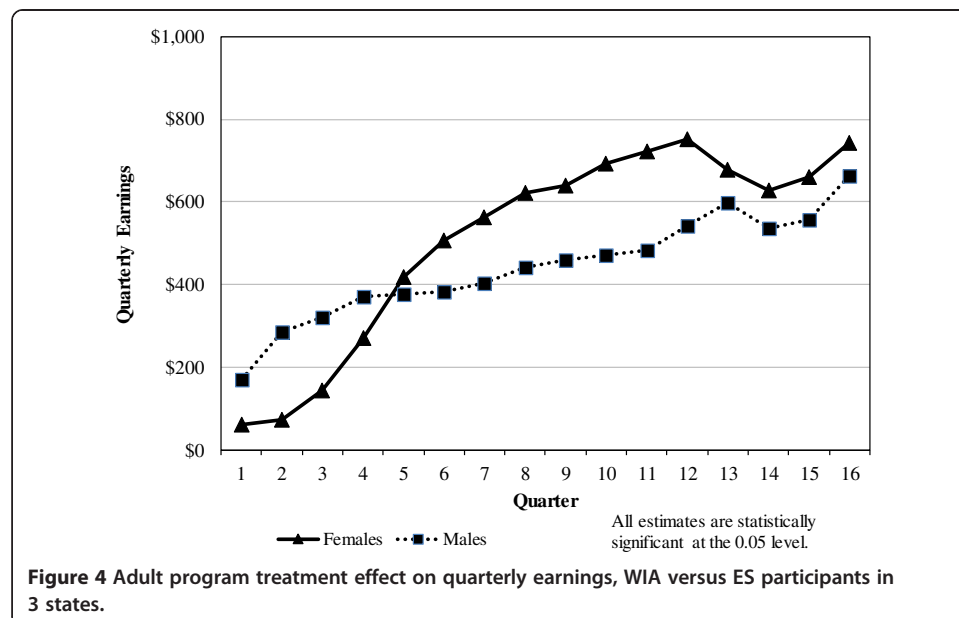
	Females		Males	
	Prior Quarter 10	Prior Quarter 16	Prior Quarter 10	Prior Quarter 16
	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
1. Adult program, predicting quarterly earnings (see Figure 2)	-48 (23)	6 (39)	11 (31)	-102 (58)
2. Adult program, predicting quarterly employment (see Figure 3)	-0.0143 (0.0062)	-0.0005 (0.0071)	-0.0078 (0.0060)	-0.0007 (0.0085)
3. Adult program, predicting quarterly earnings, ES States (see Figure 4)	-7 (44)	2 (40)	60 (69)	-71 (61)
4. Adult program receiving UI, predicting quarterly earnings	-26 (65)	-29 (89)	71 (103)	445 (204)
5. Adult program training, predicting quarterly earnings (see Figure 5)	13 (51)	104 (90)	-38 (92)	174 (187)
6. Dislocated Worker program, predicting quarterly earnings	76 (31)	233 (69)	67 (41)	267 (106)
7. Dislocated Worker program, predicting quarterly employment	0.0073 (0.0034)	0.0219 (0.0059)	0.0085 (0.0039)	0.0141 (0.0066)
8. Dislocated Worker program receiving UI, predicting quarterly earnings	56 (40)	305 (113)	40 (58)	382 (183)
9. Dislocated Worker program training, predicting quarterly earnings (see Figure 8)	0 (94)	-212 (205)	43 (137)	28 (305)

If the benefits provided by UI were responsible for inflating impact estimates, this bias would be less important for the other comparison group, those obtaining ES services. Although there is overlap in the UI and ES populations because most UI recipients are required to register for ES, UI recipients awaiting recall are exempt from this requirement. In addition, the ES sample includes self-motivated job seekers who are not receiving UI benefits. Figure 4 provides earnings impact estimates for the three states where ES recipients form the comparison group. The impact estimates in the first few quarters after entry for these states are smaller than for the full sample, and we see they increase over time. Estimates are generally in the same range after the fourth or fifth quarter. Given that UI benefits are normally limited to six months, this is consistent with expectations. Results with employment as the dependent variable are similar.

We conclude that the impacts on earnings and employment in the quarters immediately after WIA entry reported in Figures 2 and 3 could be at least partly due to differences in the incentives faced by WIA participants and the UI claimant comparison group rather than to the effects of program participation. However, although estimates in the first two or three quarters after program entry may be biased, our tests for selection and incentive explanations do not suggest that impact estimates for later quarters are spurious.

Impacts for UI recipients

In the discussion above, we suggested that UI claimants may face different incentives than many WIA Adult program participants, especially in the initial quarters after program entry when most UI claimants are eligible for benefits. To examine whether program effects may differ for this group, we undertook analyses limiting the treated group to those receiving UI benefits when they entered WIA. Adult program participants receiving UI benefits at the point of entry account for fewer than 10 percent of



entries during the period of our study⁸. These analyses showed that in the first two quarters following entry into the Adult program, women receiving UI benefits earned less as a result of their WIA participation, suggesting a lock-in effect during program participation. During quarters 7–13, earnings effects averaged about \$100 for participants, implying a less than a 5 percent earnings increment. These results differ from those for all female Adult participants (Figure 2), which show immediate impacts in the range of \$500, with small increases in later quarters.

The results for males also showed substantial negative initial effects extending for the first five quarters. By quarter 10, impact estimates were nearly \$200, implying about a 5 percent earnings increment. Even this modest impact estimate may be largely spurious. Table 2 (line 4) implies that male earnings 16 quarters prior to program entry were appreciably higher for participants than for the comparison group, so observed differences may reflect the impacts of stable factors rather than the program. The comparable estimates predicting employment for females and males receiving UI benefits were broadly consistent with those reported for earnings.

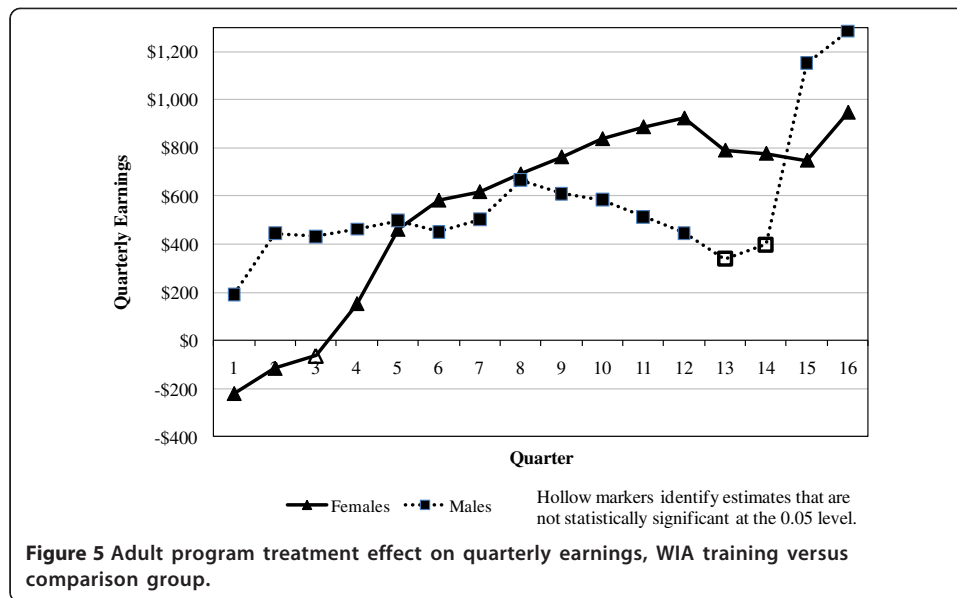
These results suggest that impacts for Adult WIA participants receiving UI benefits are substantially smaller than for the full population of Adult program participants, consistent with the view that the benefits of WIA for those who lose “good” jobs may be smaller than for workers with poorer work histories.

Impacts of training

Vocational skills training are the heart of the WIA Adult program. Although a variety of training opportunities are widely available outside of WIA, for many WIA Adult participants, the alternatives are more costly. Acceptance into WIA alters the type and extent of training these individuals ultimately obtain. As noted above, our estimates of the impact of training are based on comparing WIA Adult program participants who obtain training with Adult program participants who do not receive training.

Our estimates identify the incremental impact of WIA training relative to services received by the comparison group, who are also WIA participants but who did not receive WIA training. There is the possibility that some in this comparison group received job training outside the program. We have no direct measure of the extent of such training. In their study of the JTPA program, Orr et al. (1996, p. 97) reported that nearly a quarter of men and a third of women in the control group, who were prohibited from receiving JTPA services, received roughly comparable employment and training services elsewhere. If similar rates of training participation hold during the period of our study, our estimates identify the impact of providing training relative to available alternatives that are likely to be accessed by as much as one-third of our comparison group members.

Figure 5 presents estimates of training effects. For females, they imply a \$200 earnings reduction in the first quarter after program entry, as would be expected if time in training limited initial employment options. Earnings catch up three or four quarters later, with a positive effect of over \$800 by 10 quarters, implying an earnings increment of about 30 percent. In contrast, males who receive training appear to experience positive initial impacts—in the range of \$200 immediately after entry—with the increment



remaining in the \$500-600 range, or about 10–20 percent of earning, for the next 10 quarters⁹.

Analyses for employment (not presented) show that initial employment for women is reduced by about 5 percentage points as a result of training, and the employment rate then recovers four quarters after entry. By the tenth quarter, the impact of training on employment is about 5 percentage points. The pattern is similar for men, except that the increment is close to zero for six or seven quarters before it increases. The patterns of results do not vary substantially by whether states train a large share of their participants, or for ES states.

The differences in impact estimates between men and women may be due partly to the kinds of training they receive. Among males receiving training and exiting the Adult program in program year 2005, 37 percent received on-the-job training, in contrast to 15 percent for females (Social Policy Research Associates 2007). On-the-job training is less likely to depress initial employment and earnings than classroom training but may have smaller impacts on ultimate earnings. In our sample, we also observe that, of Adult program participants who obtain training, women spend, on average, over three months longer in the program than men.

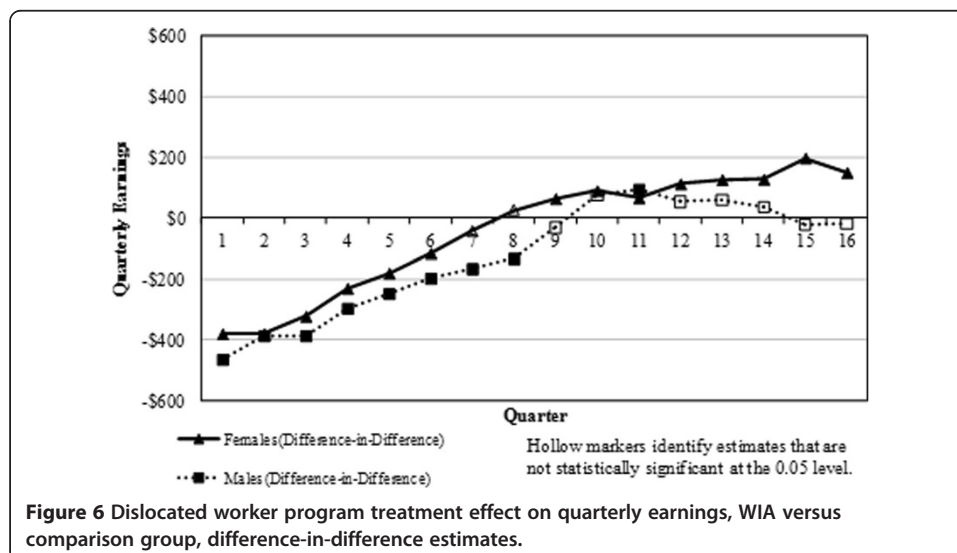
There is some indication that selection into training on the basis of stable differences may affect results, since earnings and levels of employment 16 quarters before program entry are higher for participants than the comparison group for both genders (see Table 2, line 5), but none of these differences is statistically significant. In addition, our estimates of the impact of training need to be treated with caution because they apply to a somewhat limited sample. About a third of women and nearly half of men receiving training were omitted from the analysis because they could not be matched with Adult program participants who did not receive training. It is unclear whether estimates of impact reported here are valid for omitted individuals.

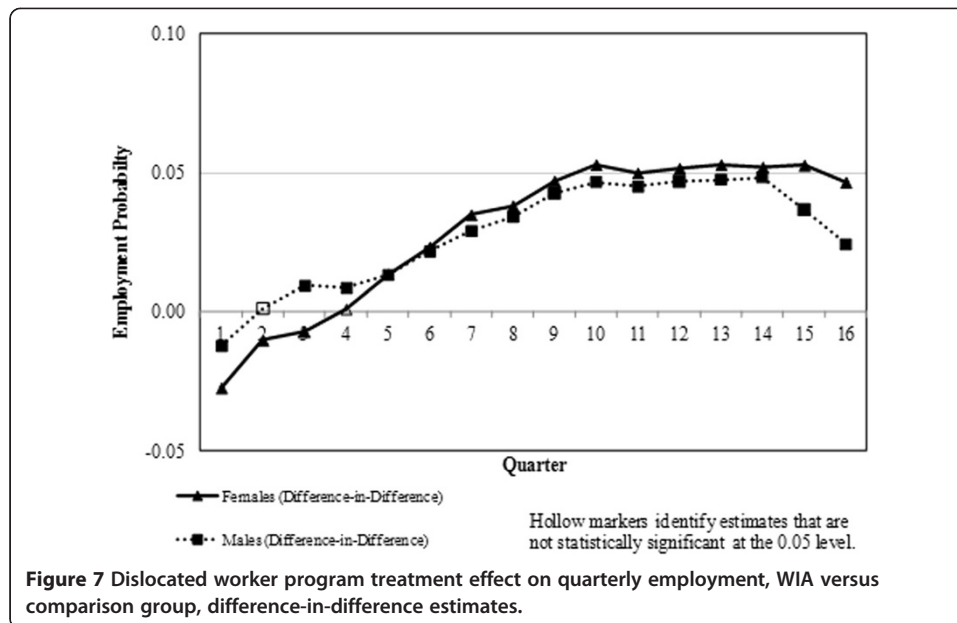
6. Impact estimates for dislocated worker program

Specification tests indicate that there are substantial differences between Dislocated Worker program participants and those in the comparison group 16 quarters earlier, with participant earnings over \$200 more and standard errors implying these differences are statistically significant (Table 2, line 6). Prior employment levels are also several percentage points higher for program participants (line 7). These results suggest that even if program participants' earnings are higher than those of the comparison group after participation, this would not necessarily reflect program impact but could instead identify unmeasured factors that are reflected in subsequent earnings and employment.

Given that differential selection for the participant and comparison cases may cause bias in matching estimates, difference-in-difference methods provide useful alternative estimates. As noted above, the difference-in-difference estimator provides a valid estimate of program impact if selection into the program is on the basis of stable characteristics affecting prior and subsequent earnings and employment that are not captured by variables that have been controlled¹⁰. Here we calculate the difference-in-difference estimate by subtracting the earnings difference 16 quarters prior to entry (for the matched samples) from the simple matching estimate. Intuitively, this approach simply adjusts for prior existing earnings differences on the assumption that they would re-appear after training even if the training had no real impact on earnings.

The difference-in-difference estimates, shown in Figure 6, imply that participants catch up to nonparticipants with a long delay and that ultimate impacts on earnings are modest, with the positive effect never over \$200 or 5 percent of earnings for women, and less than \$100 or 2 percent of earnings for men. Difference-in-difference estimates for employment (Figure 7) are more supportive of the program, with estimated program impacts on employment only about 25 percent smaller than those in the simple model, implying that participants ultimately increase their chance of employment by as much as 5 percentage points. It is worth pointing out that in order for program participation to increase employment without affecting average earnings, it





must cause the earnings of employed individuals to decline, essentially reducing average earnings for those who would be employed in the absence of the program.

Impacts for UI recipients in dislocated worker program

Nearly a third of WIA Dislocated Worker participants in our sample were receiving UI benefits at the point when they entered the program. Focusing on this subgroup allows us to control for possible incentive effects of UI receipt. The basic analysis corresponds to that presented above but with both program participants and the comparison group limited to individuals receiving UI benefits in the nine states with the UI comparison group.

As with the estimates in the prior section, specification tests presented in Table 2 (line 8) imply that program participants have substantially higher earnings 16 quarters prior to program participation than matched comparison group members. The simple estimates indicate that earnings even in the fourth year after participation do not exceed those of the comparison group by much more than \$200 for either men or women. Since the specification tests imply that even such modest earnings benefits could be due to selection, our results with this selected group confirm our findings that the WIA Dislocated Worker program has little impact on participants' earnings. Impacts on employment are more supportive of the program, although impact estimates are somewhat smaller than those reported in the prior section.

Impacts of training for dislocated workers

Estimates of the impact of training are based on a comparison of WIA Dislocated Worker participants who obtain training with those who do not. Figure 8 shows that initial earnings are reduced for participants in quarters 2 through 4 by \$1,100 for females (a decline in earnings of about one third), and \$800 for males (a decline of about 20 percent). After quarter 10, earnings impact estimates for both males and females

approach zero, although the standard errors are large. A very similar pattern exists for employment, and thus we do not present those results.

These estimates suggest that WIA Dislocated Worker program participants who enter training suffer large earnings losses in their first two years after program entry. Such negative effects are consistent with large training lock-in effects. Estimates of effects on earnings and employment three to four years after program entry—more than 18 months after program exit for most participants—show no evidence that training produces benefits. These conclusions must be tempered, however, by a recognition that sampling error alone could obscure impacts. In addition, 28 percent of women receiving training were omitted from the analysis because no matching comparison case could be found; the analogous figure for men is 38 percent. Hence, the results may not be representative of the full population of those receiving training.

7. Subgroup analyses

In recognition of the role gender plays in the labor market, for all the analyses above we have reported impact estimates separately for men and women. In auxiliary analyses, we also estimated impacts separately for nonwhites, Hispanics, individuals under 26 years of age, those 50 or older, and male veterans, in each case by gender. There are several reasons to focus on these subgroups. First, members of some of these subgroups, such as nonwhite minorities, tend to make up a larger share of participants in WIA than their share in the overall labor force. Second, these groups may face special challenges or barriers in the labor market that could affect the impact of any training they receive. Previous research has shown that some of these groups have lower returns to education and training.

Our analysis of nonwhites showed patterns in both WIA programs that closely paralleled those for the full sample for both men and women. Hispanics generally displayed larger long-term effects than the population as a whole, although these differences were not statistically significant. Estimates of impacts in the Adult program for those under age 26 were very similar to those for the full population for both women and men. In contrast, in the Dislocated Worker program, estimates were somewhat

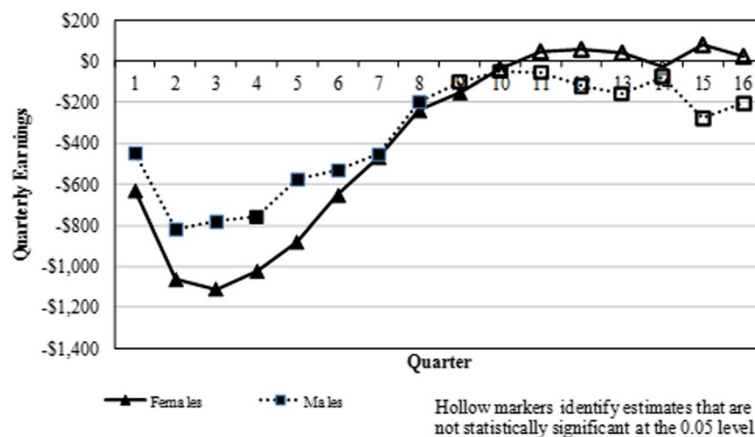


Figure 8 Dislocated worker program treatment effect on quarterly earnings, WIA training versus comparison group.

higher for this younger group. Older WIA participants (those age 50 and over) displayed patterns that largely matched the full population, although here sampling error made the comparison difficult.

Overall, we found little evidence of important differences in program impacts for any of these subgroups, albeit recognizing that there may well be differences that are not statistically discernible.

8. Conclusions

Although our results are complex, we believe it is possible to draw implications regarding the efficacy of the two WIA programs. The benefits of the programs depend critically on outcomes after the first two years. As a way to summarize impacts, we calculate the average quarterly earnings increment in quarters 11–16 following program entry. Although there is no way to be sure whether the earnings and employment benefits we observe would persist over an extended period (Smith 2011), these measures give a sense of the likely long-term benefits. These averages are based on estimates underlying Figures 2 and 3 (Adult program), and Figures 6 and 7 (Dislocated Worker program).

In the WIA Adult program, we estimate that earnings were increased by \$591 per quarter for women, amounting to about 25 percent of earnings over that period. For men, the average quarterly earnings increment was \$419, or about 15 percent of earnings. The employment increment was between 6 and 7 percentage points both for men and women, an increase of up to 12 percent in employment levels. Estimates for both men and women were statistically significant. Since our estimates provide little evidence of earnings losses during the period of program participation, these returns suggest that the program benefits participants.

Our estimates for the WIA Dislocated Worker program are not as encouraging. As noted in our discussion of the detailed results, we believe that the difference-in-difference estimates are most likely to provide meaningful measures of program impact. Although positive, these estimates of impacts on earnings are not statistically significant, amounting to only \$131 per quarter for women (3 percent of earnings) and \$36 for men (1 percent). Our estimates do suggest increases in employment for both men and women, in the range from 4 to 5 percentage points, an increment of as much as 8 percent, but this suggests that the program is associated a decline in earnings for those who would be employed in the absence of the program.

It is important to ask whether the net benefits we find satisfy a benefit-cost test. Costs incurred in serving WIA program participants in this study are not available—nor are there accurate average costs for those entering the programs over a particular period, either for states or for the nation as a whole. However, approximate average program costs over the long run can be inferred from published sources. Average per capita direct expenditures of the Adult program (including the costs of ITAs) aggregated for our 12 states are in the range of \$2400–\$2700 and Dislocated Worker expenditures are in the range \$2800–\$3200¹¹.

There are two important reasons that these figures may differ from true social costs. Because the program provides some services that would be obtained elsewhere, it reduces expenses—either by participants or others—that would otherwise be incurred; this causes social costs to be smaller than actual direct costs. In their benefit-cost analysis of the JTPA program, Orr et al. (1996) find that such substitution is important;

they estimate that it reduces social costs by nearly half of the total of training and other program costs (see Orr et al., pp. 97, 189, 269). On the other hand, some social costs are omitted from our direct cost measures. Individuals receiving certain WIA services may draw on other subsidies, such as when they receive training at publicly-subsidized community colleges. Orr et al. include such subsidies in their analysis, whereas the expenses we present above do not. Hence, the measures we cite are subject to biases in both directions, and it would not be surprising if our estimates of costs differed from true social costs by 30 percent in either direction.

In addition, it is necessary to identify any forgone earnings associated with participation in the program. Our estimates do not suggest that participants in the Adult program earned substantially lower earnings than comparison group members during the period of participation. Although, as noted above, we suspect there may be bias in estimates for the first two quarters after beginning the program, even moderate forgone earnings would not influence our conclusions. In the case of the Dislocated Worker program, our estimates suggest forgone earnings are substantial.

Notwithstanding the limitations in our cost estimates, the Adult program clearly satisfies a benefit-cost standard for both men and women if the earnings impacts, estimated at over \$400 per quarter, continue for a period of just three or four years, which seems plausible. In contrast, our best estimates of the impact on earnings for the Dislocated Worker program imply that impacts on earnings for women (\$131 per quarter) would have to be very long-lived to exceed direct costs, even without considering forgone earnings. Estimated benefits for men (\$36 per quarter) could never cumulate to exceed costs at any reasonable rate of interest. As noted above, however, estimates of impacts on employment are more supportive of the Dislocated Worker program. If the program succeeds in increasing the number of individuals with jobs, and the most disadvantaged workers are the gainers, such program impacts could justify the program even if it fails a benefit-cost standard¹².

As we did not find notable differences in program effects for demographic subgroups, we do not see any basis for targeting the programs to any of these subpopulations. On the other hand, given that the Adult program, which focuses on disadvantaged individuals, is estimated to have greater impacts than the Dislocated Worker program, and both programs appeared to have smaller impacts on those receiving UI benefits, focusing attention on the least advantaged may be the best investment of program resources.

One important limitation of our analyses is that they are unable to distinguish among effects of the highly varied services that are offered under the two WIA programs, but rather evaluate each program as kind of “black box”. As such, our results do not provide guidance as to which kinds of activities are most likely to be worthwhile and how programs can improve their efficacy. On the other hand, our results may provide an evaluation at the level which is relevant for federal budget decisions. Control over the details of program implementation is limited, so a primary lever for policymakers at the federal level will be the decision of whether the program—taken as a whole—should be continued or expanded.

Our results also underscore the importance of understanding the long-run impacts of these programs. Program administrators frequently have available point-in-time information from a performance management system rather than data analysis tools to examine individual employment and earnings histories and trajectories identified this

study. It will likely take a few years to fully observe the impacts of the substantially increased expenditures of the Obama administration on employment and training programs, or any subsequent budget cuts, given that program impacts “mature” over time and sometimes increase in magnitude and sometimes diminish. In other words, from a public policy perspective, the outcomes of greatest interest from our recent increased investment in workforce development programs may not be apparent for years. Yet our results also suggest that if we are to make sound policy decisions about whether to continue to invest in these programs (and human capital) in the future, it will be important to evaluate the longer-term outcomes of these programs. Our analysis, which is consistent with findings of prior studies, including random assignment experiments (Orr et al. 1996), suggests that investing in the training of disadvantaged adults will generate returns that exceed costs to the public.

Endnotes

¹The Hollenbeck studies used individuals who exited the program. Insofar as administrators can control exits, they may choose which participants will exit or they will choose the date of exit strategically, so that outcomes following date of exit may not be representative of participant outcomes in general. Using data from one state, Hollenbeck (2009) compared impact estimates based on exit date with those using a sample of entry dates. Estimates did differ from these two approaches, but the differences did not imply a consistent bias either in favor or against the program.

²Based on tabulations for our 12 states.

³These agreements were made with the condition that individual state results would not be released.

⁴Some studies define employment as earnings above some threshold, for example \$100. Since the number of individuals with such low earnings is very small, such an approach does not alter results in any substantial way.

⁵See Rosenbaum and Rubin (1983). For general discussions of matching methods see Rubin (2006).

⁶We also obtained estimates of impacts of the Adult program participants separately in each of the 12 states. Readers interested in these detailed analyses are referred to Heinrich et al. (2008). Here we focus on averages across participants in the 12 states, which reduces the substantial sampling error and averages across idiosyncratic state differences.

⁷Mean earnings and employment for the matched comparison group are provided in Additional file 1: Table A3.

⁸These estimates are provided in Additional file 1: Figure A1.

⁹Although estimates reported in Figure 5 for males imply an increase in the final two quarters in which we have data, standard errors are large and we suspect this observed spike reflects sampling error.

¹⁰As noted above, if these assumptions are not met, difference-in-difference estimates may be biased. For comparison, estimates based on a simple difference structure are provided in Additional file 1: Figures A2 and A3.

¹¹These figures are based on total expenditures in the indicated programs for July 2003-June 2005 as detailed in U.S. Department of Labor (2009). We have formed two

measures, one dividing total expenditures by the number of reported exits during this period, and the other dividing by the number of entries identified in our data. The ranges reported above reflect the differences in these measures.

¹²In common with most benefit-cost analyses, ours ignores general equilibrium issues, including the possibility that workers who obtained employment because of the program would displace others. Such effects would have to be quite large to affect our conclusions.

Additional file

Additional file 1: Technical Details of Estimation.

Competing interests

The IZA Journal of Labor Economics is committed to the IZA Guiding Principles of Research Integrity. The authors declare that they have observed these principles.

Acknowledgements

We wish to thank participants in seminars at the Australian National University, the Institute for the Study of Labor, Bonn (IZA), the Melbourne Institute for Applied Economic and Social Research, and participants in the Association for Public Policy Analysis and Management annual meetings, the Centre for European Economic Research (ZEW) workshop on social exclusion, the European Association of Labor Economists annual meetings, the Institute for Poverty Summer Research Workshop (Wisconsin), and the Missouri Economic Conference, and in particular for comments by Burt Barnow, Marco Caliendo, Paul Decker, Cory Koedel, Andrew Leigh, Sheena McConnell, Jeffrey Smith, and Arne Uhlenhorff. The analyses presented here include and extend work supported by the U.S. Department of Labor (DOL) and presented in "Workforce Investment Act Non- Experimental Net Impact Evaluation" (IMPAQ International, Final Report, December 2008, Department of Labor ETAOP 2009–10). The authors wish to acknowledge the central role in this project played by the staff at IMPAQ, including Nicholas Bill, Shirisha Busan, Goska Grodsky, Eileen Poe-Yamagata, and Ted Shen. Jacob Benus served as project director. Thanks are due to the many state agency staff who worked to provide data, to David Stevens who facilitated provision of data for Maryland, and to Suzanne Troske, who supported data processing in Kentucky. Jonathan Simonetta oversaw the project for DOL. The paper has not been reviewed by DOL; conclusions presented are the sole responsibility of the authors.
Responsible editor: V. Joseph Hotz.

Author details

¹Lyndon B. Johnson School of Public Affairs and Department of Economics (affiliate), University of Texas, P.O. Box Y, Austin, TX 78713-8925, USA. ²Department of Economics and Harry S. Truman School of Public Affairs, University of Missouri, IMPAQ International, LLC, and IZA, 118 Professional Building, Columbia, MO 65211, USA. ³University of Kentucky and IZA, 255G Gatton Business and Economics BLDG, Gatton School of Business, Lexington, KY 40506-0034, USA. ⁴Department of Economics, University of Missouri, 118 Professional Building, Columbia, MO 65211, USA. ⁵American Institutes for Research, 1000 Thomas Jefferson Street, Northwest, Washington, D.C. 20007, USA.

Received: 29 July 2013 Accepted: 9 October 2013

Published: 25 October 2013

References

- Abadie A, Imbens GW (2006) Large sample properties of matching estimators for average treatment effects. *Econometrica* 74:235–267. doi:10.1111/j.1468-0262.2006.00655.x
- Andersson F, Holzer HJ, Lane JL, Rosenblum DJ, Smith JA (2013) Does federally-funded job training work? Nonexperimental estimates of WIA training impacts using longitudinal data on workers and firms. Cambridge, MA: NBER Working Paper No. w19446, <http://ssrn.com/abstract=2330064>. Accessed 15 Jul 2012
- Bloom HS, Michalopoulos C, Hill C (2005) Using experiments to assess nonexperimental comparison-group methods for measuring program effects. In: Bloom HS (ed) *Learning more from social experiments: Evolving analytic approaches*. Russell Sage, New York
- Bureau of Labor Statistics (2012) Labor force statistics (CPS), Duration of unemployment <http://www.bls.gov/cps/cpsaat29.pdf>. Accessed 15 Jul 2012
- Card D, Klueve J, Weber A (2009) Active labor market policy evaluations: A meta-analysis. IZA Discussion Paper # 4002, <http://ftp.iza.org/dp4002.pdf>. Accessed 15 July 2012
- Cook TD, Shadish WR, Wong VC (2008) Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. *J Policy Anal Manage* 27(4):724–750
- Decker PT (2011) Ten years of WIA research. In: Besharov DJ, Cottingham PH (eds) *The Workforce Investment Act: Implementation experiences and evaluation findings*. Upjohn, Kalamazoo
- Greenberg D, Michaelopoulos C, Robins P (2006) Do experimental and nonexperimental evaluations give different answers about the effectiveness of government-funded training programs? *J Policy Anal Manage* 25(3):523–552
- Heckman JJ, Heinrich CJ, Smith JA (2011) Do short-run performance measures predict long-run impacts? In: Heckman JJ, Heinrich CJ, Courty P, Marshke G, Smith J (eds) *The performance of performance standards*. Upjohn, Kalamazoo

- Heckman JJ, LaLonde RJ, Smith JA (1999) The economics and econometrics of active labor market programs. In: Ashenfelter O, Card D (eds) *Handbook of Labor Economics*, vol 3. North Holland, New York, pp 1865–2095
- Heckman JJ, Smith JA (1999) The pre-programme earnings dip and the determinants of participation in a social programme: Implications for simple programme evaluation strategies. *Econ J* 109:313–348. doi:10.1111/1468-0297.00451
- Heinrich CJ, Mueser PR, Troske KR (2008) Workforce Investment Act Non-Experimental Net Impact Evaluation: Final report. <http://www.nawdp.org/Content/NavigationMenu/ResearchReports/2009-10-WIANon-ExperimentalNetImpact.pdf>. Accessed 18 July 2013
- Hollenbeck K (2009) Does the Workforce Investment Act Work? Upjohn Institute discussion paper. Association for Public Policy Analysis and Management (APPAM), Washington, DC, <http://research.upjohn.org/confpapers/59/> Accessed 30 September 2013
- Hollenbeck K (2011) Short-term net impact estimates and rates of return. In: Besharov DJ, Cottingham PH (eds) *The Workforce Investment Act: Implementation experiences and evaluation findings*. Upjohn, Kalamazoo
- Hollenbeck K, Schroeder D, King CT, Huang W (2005) Net impact estimates for services provided through the Workforce Investment Act. In: *On the use of administrative data for workforce development program evaluation*, ETA Occasional Paper 2005–06. U.S. Department of Labor, Employment and Training Administration, <http://research.upjohn.org/externalpapers/8/> Accessed 15 November 2005
- Imbens GW, Wooldridge JM (2008) Recent developments in the econometrics of program evaluation. IRP Discussion Paper # 1340–08. University of Wisconsin Madison, <http://www.irp.wisc.edu/publications/dps/pdfs/dp134008.pdf>. Accessed 16 Sept 2008
- Jung H, Pirog MA (2011) Nonexperimental impact evaluations. In: Besharov DJ, Cottingham PH (eds) *The Workforce Investment Act: Implementation experiences and evaluation findings*. Upjohn, Kalamazoo
- Kornfeld R, Bloom HS (1999) Measuring program impacts on earnings and employment: Do unemployment insurance wage reports from employers agree with surveys of individuals? *J Labor Econ* 17(1):168–197
- Mueser PR, Troske KR, Gorislavsky A (2007) Using state administrative data to measure program performance. *Review Econ and Stat* 89(4):761–783
- Orr LL, Bloom HS, Bell SH, Doolittle F, Lin W, Cave G (1996) *Does training for the disadvantaged work? Evidence from the national JTPA study*. Urban Institute Press, Washington DC
- Osterman P (2007) Employment and training policies: New directions. In: Holzer HJ, Nightingale DS (eds) *Reshaping the American workforce in a changing economy*. Urban Institute Press, Washington, DC, pp 119–154
- Pirog MA, Buffardi AL, Chrisinger CK, Singh P, Briney J (2009) Are the alternatives to randomized assignment nearly as good? Statistical corrections to non-randomized evaluations. *J Policy Anal Manage* 28(1):169–172
- Rockefeller Institute of Government (2004) *The Workforce Investment Act in eight states: State case studies from a network evaluation* (2 volumes). ETA Occasional Papers 2004–02 and 2004–03. U.S. Department of Labor, Employment and Training Administration, http://www.doleta.gov/reports/searcheta/occ/eta_occasional_papers.cfm. Accessed 2 Jan 2005
- Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70:41–55
- Rubin DB (2006) *Matched sampling for causal effects*. University Press, Cambridge
- Schochet PZ, D'Amico R, Berk J, Dolfin S, Wozny N (2012) Estimated impacts for participants in the Trade Adjustment Assistance (TAA) Program Under the 2002 Amendments. Final Report. ETA Occasional Papers 2012–10. U.S. Department of Labor, Employment and Training Administration, http://wdr.doleta.gov/research/FullText_Documents/ETAOP_2013_10_Participant_Impact_Report.pdf. Accessed 12 Sept 2012
- Social Policy Research Associates (2004) *The Workforce Investment Act after five years: Results from the national evaluation of the implementation of WIA*. Report prepared for the U.S. Department of Labor, http://www.doleta.gov/reports/searcheta/occ/papers/SPR-WIA_Final_Report.pdf. Accessed 12 Sept 2005
- Social Policy Research Associates (2006) *2004 WIASRD Data Book*. Report prepared for the U.S. Department of Labor, http://www.doleta.gov/performance/results/pdf/PY_2004_WIASRD_Databook.pdf. Accessed 15 Sept 2007
- Social Policy Research Associates (2007) *PY 2005 WIASRD Data Book: Final*. Prepared for the U.S. Department of Labor, http://www.doleta.gov/performance/results/pdf/PY_2005_WIASRD_DataBook_Rev%208-14-2007.pdf Accessed 16 Sept 2008
- Smith JA, Todd PE (2005) Does matching overcome LaLonde's critique of nonexperimental estimators? *J Econometrics* 125(1–2):305–353
- Smith JA (2011) Improving impact evaluation in Europe. In: Besharov DJ, Cottingham PH (eds) *The Workforce Investment Act: Implementation experiences and evaluation findings*. Upjohn, Kalamazoo
- U.S. Department of Labor, Employment and Training Administration (2009) *WIA state annual reports & summaries*. PY2003 and PY2004 (downloadable files), <http://www.doleta.gov/performance/results/Reports.cfm?#wiasstann>. Accessed 30 August 2009
- Wallace GL, Haveman R (2007) The implications of differences between employers and worker employment/earnings reports for policy evaluation. *J Policy Anal Manage* 26(4):737–753

doi:10.1186/2193-8997-2-6

Cite this article as: Heinrich et al.: Do Public Employment and Training Programs Work? *IZA Journal of Labor Economics* 2013 2:6.