DISCUSSION PAPER SERIES

IZA

IZA DP No. 4569

New Estimates of Public Employment and Training Program Net Impacts: A Nonexperimental Evaluation of the Workforce Investment Act Program

Carolyn J. Heinrich
Peter R. Mueser
Kenneth R. Troske
Kyung-Seong Jeon
Daver C. Kahvecioglu

November 2009

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

New Estimates of Public Employment and Training Program Net Impacts: A Nonexperimental Evaluation of the Workforce Investment Act Program

Carolyn J. Heinrich
University of Wisconsin

Peter R. Mueser
University of Missouri, IMPAQ International, LLC, and IZA

Kenneth R. Troske
University of Kentucky and IZA

Kyung-Seong Jeon
University of Missouri

Daver C. Kahvecioglu
IMPAQ International, LLC

Discussion Paper No. 4569
November 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 4569
November 2009

ABSTRACT

New Estimates of Public Employment and Training Program Net Impacts: A Nonexperimental Evaluation of the Workforce Investment Act Program*

This paper presents nonexperimental net impact estimates for the Adult and Dislocated Worker programs under the Workforce Investment Act (WIA), the primary federal job training program in the U.S, based on administrative data from 12 states, covering approximately 160,000 WIA participants and nearly 3 million comparison group members. The key measure of interest is the difference in average quarterly earnings or employment attributable to WIA program participation for those who participate, estimated for up to four years following entry into the program using propensity score matching methods. The results for the average participant in the WIA Adult program show that participating 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 those who don't receive training, but they catch up within 10 quarters, ultimately registering large total gains. The marginal benefits of training exceed, on average, $400 in earnings each quarter three years after program entry. Dislocated Workers experience several quarters for which earnings are depressed relative to comparison group workers after entering WIA, and although their earnings ultimately match or overtake the comparison group, the benefits they obtain are smaller than for those in the Adult program.

JEL Classification: J24, J48

Keywords: nonexperimental program evaluation, workforce investment act

Corresponding author:

Peter Mueser
Department of Economics
University of Missouri
Columbia, MO 65211
USA
E-mail: mueserp@missouri.edu

* 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 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 Jeffrey Smith, Burt Barnow, Marco Caliendo, Andrew Leigh, and Arne Uhlendorff. The analyses presented here include and follow from work undertaken for the U.S. Department of Labor 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 the Department of Labor. Notwithstanding the support provided by these individuals and the U.S. Department of Labor, the analysis and interpretation presented in the current paper are the sole responsibility of the authors.

I. Introduction

In the midst of one of the most severe economic downturns in the last 70 years, it is not difficult to grasp the reality and implications of increasing labor market volatility that has affected both low-wage workers and more experienced and skilled workers in recent decades. In fact, these trends of stagnating economic mobility, dislocation and longer-term joblessness have been well-documented (Appelbaum, Bernhardt and Murnane, 2003; Bradbury and Katz, 2002; Holzer, 2004; Osterman, 2007). At the same time that U.S. workers have been facing these growing labor market challenges, however, public expenditures on employment and training services have been declining. For example, in fiscal year 2007, the total U.S. federal government appropriations for Workforce Investment Act (WIA) programs—youth employment, adult job training, dislocated worker assistance, Job Corps and other national activities—was $4.4 billion, down 18 percent from fiscal year 2005. Furthermore, within the WIA program, the number of adults receiving training has declined appreciably relative to WIA's predecessor, the Job Training Partnership Act (JTPA) program (Frank and Minoff, 2005).

Enacted in August 1998, the central goal of WIA was to create a new, comprehensive workforce investment system. WIA is distinguished from the JPTA program primarily by the introduction of a One-Stop service delivery system designed to improve coordination and integration of services, its use of Individual Training Accounts in training services, and significant changes in governance structures at the state and local level. In implementation, WIA has reduced the share of low-income individuals served by one-third, decreased the length of time spent in training and the expenditures per trainee (in addition to the proportion receiving training), and shifted responsibility for some types of activities believed to contribute little (or negatively) to performance outcomes, such as adult basic education, to other programs (Osterman, 2007). Thus, important changes in both investments in and the implementation of public employment training programs have taken place in the last decade, and yet surprisingly little is known about the impact of WIA and its components on labor market outcomes.

To date, evaluations of WIA have been limited.[1] The U.S. Office of Management and Budget assigned the WIA program relatively low marks for its evaluation efforts (using its Program Assessment Rating Tool), suggesting that independent evaluations had not been of sufficient scope and rigor to determine WIA's impact on participants' employment and earnings. Although the U.S. Department of Labor (DOL) recently initiated a project to experimentally evaluate the WIA program, results will not be available for at least seven years. Given the current policy context, in which more than 2.7 million workers were added to the unemployment rolls in the last year and the Obama administration and other policymakers are calling for expanded public investments in employment and training to increase individual skill levels and their success in

[1] Social Policy Research Associates (2004) and Rockefeller Institute of Government (2004, Barnow and King, 2004) undertook studies based on interviews with a small sample of state and local agency staff and a review of administrative data during the first three years that the program was implemented but undertook no systematic study of participant outcomes. Hollenbeck et al. (2005) examined outcomes in seven states for WIA participants who had completed the program during the period July 2000-June 2002, the first two years of implementation in most states. Given that over a third of participants require more than a year to complete the program, this sample would have been severely censored.

the labor market,[2] we argue that rigorous evidence on WIA's impact and effectiveness is needed now.

This study employs nonexperimental matching methods to evaluate the WIA Adult and Dislocated Worker programs using data from 12 states that cover approximately 160,000 WIA participants and nearly 3 million comparison group members. Within each state, we compare WIA program participants with a matched comparison population of individuals who have not participated in the WIA program but who are observationally equivalent across a range of demographic characteristics, social welfare benefit receipt and labor market experiences. Comparison group members are drawn from those who have participated in the U.S. Employment Service (ES) under Wagner-Peyser legislation or who receive Unemployment Insurance benefits. Participants and comparison group members, both of whom are at junctures in their careers when they are either facing employment crises or are considering alternative vocational options, are compared within state and state-established workforce investment areas to insure that they are facing similar local labor markets, and measures of employment and earnings are fully comparable for program participants and the comparison group, which research suggests should increase the success of matching methods in identifying program impacts (Heckman, LaLonde and Smith, 1999; Glazerman, Levy and Meyers, 2003; Bloom, Michalopolos and Hill, 2005; Mueser, Troske and Gorislavsky, 2007). Indeed, in a recent meta-analysis of training programs, Card et al. (2009) concluded that results for experimental studies were not significantly different from those of nonexperimental studies, and they suggest that "research designs used in recent non-experimental evaluations are not significantly biased relative to the benchmark of an experimental design."[3]

This study adds to a fast-growing literature that evaluates active labor market programs, most of it based on nonexperimental methods and the evaluations of European programs, largely because of the availability of extensive administrative data with detailed information about program participation and employment over many years. Card et al. (2009) based their meta-analysis on 199 estimates obtained from 97 recent studies, with over four-fifths from continental Western Europe. In general, this literature is moderately supportive of the benefits of job training and related active labor market programs. Card et al. observe that job training programs, especially longer-duration programs, tend to have very small or negative impacts on employment measures in periods of less than a year, presumably reflecting "lock-in" effects, but have positive effects in the second or third years (see also Dyke et al., 2006; Hotz et al., 2006). One useful benchmark is the random assignment evaluation of the JTPA program for participants in the late 1980s. Program enrollees experienced minimal incremental effects in the two quarters after random assignment, but the increment in quarterly earnings increased to $300-350 (2006 dollars) by the tenth quarter (Orr et al., 1996, p. 107). Studies investigating whether program effectiveness

[2] Source: http://origin.barackobama.com/issues/urban_policy/, accessed January 2009.

[3] Two recent studies also consider the success of nonexperimental studies in reproducing experimental results. Cook, Shadish and Wong (2008) compare nonexperimental and experimental studies evaluating a wide range of interventions, and they are skeptical about the success of nonexperimental methods in identifying program impacts. However, the studies of job training programs they cite do not consider nonexperimental methods that satisfy the above requirements. Piekes, Moreno and Orzol (2008) are also pessimistic about nonexperimental methods in evaluating job training programs, but their comparisons involve very small sample sizes, and we do not believe that meaningful inferences can be based on them.

varies with the business cycle have generated mixed evidence. Kluve's (2007) review of studies found little or no difference in estimated impact based on the business cycle, whereas Lechner and Wunsch's (2006) ten-year study of a German job training program found substantially greater program impacts during economic downturns.

The results of this study 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 those who don't receive training, but they catch up within 10 quarters, ultimately registering large total gains. The marginal benefits of training exceed, on average, $400 in earnings each quarter three years after program entry. For Dislocated Workers, their earnings are depressed (relative to comparison group workers with the same characteristics and work histories) 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.

Still, in the absence of data drawn from a representative sample of the population of WIA participants, this study cannot claim to estimate a "national" average impact of WIA. In fact, no experimental or nonexperimental evaluation study has done this for WIA or any of its predecessor programs. Nonetheless, the sample of WIA participants considered here suitably reflects the diversity of local Workforce Investment Board areas, in terms of both geography and environment, including states from each major region in the U.S. and coverage of urban and rural areas; and in terms of operations, with programs that train varying proportions of their participants and manage delivery of services through a variety of organizational configurations in One-Stop centers. The states in this study account for about a fifth of the nearly 600 Workforce Investment Areas in the U.S. and a comparable proportion of the participants in WIA's two main job training programs serving adults.

II. Overview of WIA Adult and Dislocated Worker Programs

We evaluate two WIA programs: the Adult program, serving largely disadvantaged individuals, and the Dislocated Worker program, serving those who have lost jobs. Although the Adult program is designed largely for individuals who are unemployed, employed individuals are eligible to participate if participation allows them achieve economic self-sufficiency. The target population for the Dislocated Worker program is workers facing layoffs and those eligible for unemployment insurance, although other individuals who have lost their jobs are eligible if staff decide they fall in several broad categories.[4] Participation in the WIA programs is voluntary, but access is restricted, as program staff must admit all participants as well as authorize services that are provided. The study analyses focus on individuals entering WIA in the period July 2003-June 2005 (program years 2003 and 2004), which allows sufficient time after the program's initial startup (July 2000 in most states), while providing an extended follow-up period.

[4] Eligibility criteria can be found at http://www.doleta.gov/programs/general_info.cfm (accessed August 2009).

Although legislative requirements establish a general programmatic structure, states and local areas have a great deal of latitude in implementing the WIA programs.[5] States have further specified rules, and in keeping with the spirit of local control in WIA legislation, they have also left many decisions to the local agency, the Workforce Investment Board (WIB). Legislation does not define economic self-sufficiency, so whether an employed individual requires services is left largely to local discretion. In the first few years of WIA implementation, many local WIBs faced pressure to improve client employment outcomes under the program's performance standards, which measured employment and earnings of participants in the year after exit. Incentives to "cream skim" were documented by the U.S. Government Accountability Office (2002), and the point at which individuals were formally registered in WIA differed substantially across sites, with some areas (in the extreme) registering participants after brief investments of staff time and others waiting until service details were determined.

For both the Adult and Dislocated Worker programs, WIA legislation specifies three levels of service. All participants who enter WIA receive core services, which include staff-assisted job search and placement, provision of labor market information, and basic counseling, corresponding closely to the staff-assisted services offered by state offices as part of the U.S. Employment Service (ES) under Wagner-Peyser legislation. Once individuals receive core services, staff may recommend that they receive intensive services, which involve comprehensive assessment, more extensive counseling and career planning, and possibly short courses. Participants in intensive services may then be recommended to receive training services. Under WIA, most training is provided by separate organizations—including community colleges, proprietary schools, nonprofits servicing the disadvantaged, and others—through a voucher called the Individual Training Account (ITA).

Initially, many localities (WIBs) understood that WIA required a "work first" structure, in which participants would only be permitted to progress to a higher level of service if substantial time at a lower service level failed to yield employment. DOL subsequently made clear that this was not the case, noting that the length of time in a given service was not specified. In implementing WIA, it appears that only a small share of WIBs followed a strict work first policy, yet during its first several years, the importance of sequencing was highly variable across WIBs, with some requiring participants to undergo meaningful lower-level service activities and others referring participants to intensive or training services with little delay. The exact division between core and intensive training is somewhat ambiguous in the legislation, resulting in a fair bit of variation across local areas.

Federal regulations state that access to WIA core services is "universal," but local staff almost always limit admission to both the Adult and Dislocated Worker programs.[6] However, given that ES services are very similar to WIA core services, at least in terms of their basic structure,

[5] For a discussion of actual implementation, see the Social Policy Research Associates study of WIA implementation (2004, sections VI and VII), which examined practices in 21 states and 38 local WIBs, and the Rockefeller Institute of Government (2004; Barnow and King, 2004) study of eight states and 16 local WIBs. As WIA practices have shifted somewhat in recent years, we focus on the period of the study, program years 2003 and 2004, and our discussion draws primarily from these sources.

[6] In this context, "universal access" allows states to provide WIA core services to anyone; it does not require it. Since the period of our study, several states have adopted policies of open enrollment into WIA.

individuals needing such services who are not accepted into the WIA program are normally referred to ES—which is usually available at the same site. In some sites, policies are in place to refer all individuals who obtain exclusively core services to ES, and WIA enrolls only individuals who are authorized to receive intensive or training services.

Legislation also specifies that, in allocating intensive and training services, if demand exceeds available funding, staff must give preference to low-income individuals. However, it does not specify how low income is to be defined, and studies have found that in only a minority of areas is a particular cutoff applied. Despite the structure of the ITA as a voucher, WIA program staff retain power to determine who will receive the voucher and, in consequence, how it is used. Staff are generally required to assure that training prepares participants for jobs in high demand, although how this is implemented is highly variable. While the extent of counselor involvement in the training decision is clearly variable across sites, the legislation and practice place emphasis on the participant's involvement in the training choice, or, as one staff member put it, "case managers guide, but customers decide" (Social Policy Research Associates, 2004, p. VI-20).

Those locations that follow the spirit of the sequential service mandate might be expected to provide training primarily to individuals who had been unsuccessful in obtaining employment using less intensive services, causing negative selection into training. On the other hand, in most sites, as many as a third of those who enter WIA have a particular training goal prior to program entry (they are often referred to WIA by the training provider), and in general, WIB staff make an effort to accommodate them. It is unclear whether such individuals will be positively or negatively selected. Finally, staff are under pressure to provide training to individuals who they believe will benefit from the training and whose employment outcomes will aid the performance measures. Insofar as counselors can identify those who will ultimately succeed in the labor market, we would expect positive selection.

In the period of our study, nationwide about one in five WIA participants received only core services, and about two in five were coded as receiving training services. Of those who received training, up to 10 percent received on-the-job training and another 5 percent received basic skills training. Nearly 90 percent were coded as receiving occupational and other training, which includes an unknown amount of customized training for employers. About half of all training was funded by ITAs. Very little is known about the character of the training offered, hours per week, etc., but approximately two-thirds of training recipients received some kind of credential during the period of our study. Between a half and a third of participants exited WIA in less than 26 weeks, whereas a similar proportion remained in the program for at least a year.[7] Most WIBs limited the funding for an ITA, but limits varied dramatically across states and across WIBs within a state. Maximum time limits for training activities usually applied as well, most often one or two years, although generally individuals could remain in the program for longer periods, waiting for services or receiving multiple services.

Expenditure per participant in WIA differs dramatically across states. For the period of our study, the average state spent about $5,000 for each participant exiting the program, but the

[7] These figures are based on participants exiting the program April 2004-March 2005 (Social Policy Research Associates, 2006).

lowest average state expenditure was about $1,000 and the largest about $15,000 (U.S. Department of Labor, 2009).[8]

Although there is potential overlap between Adult and Dislocated Worker program participants, in practice, they differ quite dramatically in terms average age, gender, race and prior work experience. Given that the two programs serve very different functions, each are analyzed separately. The analysis here does not distinguish core and intensive levels of service.

III. Study Sample, Data, Measures and Method of Analysis

1. Study Sample

In December 2007, the U.S. Department of Labor issued a notice requesting that state workforce agencies provide access to administrative data for use in an evaluation of WIA activities funded under federal legislation. Agencies in all 50 states were contacted and efforts were made to negotiate agreements by which necessary data would be released to the researchers. Funds were made available to cover state expenses, and states were promised that individually-identifiable state results would not be released. Ultimately, agreements were reached and necessary data were provided by 12 states: Connecticut, Indiana, Kentucky, Maryland, Missouri, Minnesota, Mississippi, Montana, New Mexico, Tennessee, Utah, and Wisconsin.[9]

As noted in the introduction, we employ an econometric matching method, in which program participants are matched with individuals in a comparison group based on observed variables. All analyses are based on state administrative data, with files identifying program participants and comparison group members, as well as employment data, drawn from each state.

Treatment and comparison samples. In estimating the overall WIA program impact for participants receiving core/intensive services, we use a comparison group drawn from either Unemployment Insurance (UI) claimants or from U.S. Employment Service (ES) participants (i.e., individuals who register with the state's job exchange service and receive some services under Wagner-Peyser legislation). Of the 12 states in our analysis, nine have UI claimant comparison data while three have comparison data from ES participants. Appendix A describes some differences in these two comparison groups and the advantages and disadvantages of using one versus the other in the analysis. Estimates of the incremental impact of training use a comparison group consisting of WIA participants who did not receive training services, i.e., of those receiving only core or intensive services.

Table 1 shows treatment samples and the groups used in each comparison, which are identified by rows. Columns (a) and (b) indicate for which programs the comparison is undertaken, whereas (c) and (d) identify the treatment and comparison groups.

---

[8] For the 12 states in our study, the average expenditure per exit was about $4000 (U.S. Department of Labor, 2009).

[9] The primary contractor on the project was IMPAQ International, LLC, whose staff contacted all states and entered into agreements with nine of them. Three of the states provided data through the Administrative Data Research and Evaluation Project under separate contracts with the Department of Labor.

Row 1 lists comparisons of WIA participants—regardless of services received—with comparison group individuals who have filed for UI benefits or received ES services. These comparisons provide measures of the impact of the WIA program taken as whole. Row 2 lists comparisons that consider all WIA participants who were receiving UI benefits at the point 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. Row 3 identifies the comparison between those individuals who participate in WIA training services and other WIA participants. This comparison identifies the extent to which training, per se, is associated with employment and earnings outcomes. All comparisons are undertaken within a given state.

Table 1
Treatment and Comparison Samples

	WIA Program Group		Sample Group	
	Adult	DW	Treatment	Comparison
	(a)	(b)	(c)	(d)
1.	X	X	WIA	UI Claim or ES
2.	X	X	WIA Receiving UI	UI Recipients
3.	X	X	WIA Training	WIA Core/Intensive

Generalizing results. The analyses in this study provide estimates of average impact for participants in WIA Adult and Dislocated Worker programs in 12 states that provided data. To what degree can these results be generalized to the remainder of the states? Sampled states come from all the main geographic regions in the country, and they include five of the 30 largest U.S. cities. Given the decentralized structure of the WIA program, differences between Workforce Investment Boards within a state are often very large, and differences within a state between areas due to demographic and economic environments may dwarf between-state differences, so the sample in this study is less restricted than might initially be assumed.

The clearest threat to generalization would be if states were selected (or had selected themselves) on the basis of actual program performance. In this case, the 12 states might display impacts that were wholly unrepresentative of the remaining states. Although this possibility cannot be rejected on statistical grounds, previous work suggests that there are no easily-observable factors that predict program impact (Orr et al., 1996; Smith, Whalley and Wilcox, 2007), particularly when considering lagged program impacts. State administrative and data handling idiosyncrasies may have played a dominant role in determining willingness to provide data for the study.

2. Data Sources and Measures

The base data for the 12 states include annual Workforce Investment Act Standardized Record Data (WIASRD) or closely related data files obtained from each state that provide information on all participants exiting the WIA program within a program year (July-June). For most states, the data files extend through June 2007 (Program Year 2006). These data also include an individual identifier to allow a match with other state data. The focus of the current analysis is on WIA participants who entered the WIA program in the period July 2003-June 2005. In most cases no information is available on individuals who did not exit the program by June 2007.[10]

As indicated above, the comparison group is constructed from Unemployment Insurance (UI) claim data for nine states, and from U.S. Employment Service (ES) data in the other three states. 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 such information are available prior to the first quarter of program participation.[11]

Unemployment Insurance (UI) Wage Record data provide quarterly earnings for all employees in UI-covered firms within a state. Data extend through calendar year 2007, which, when matched with WIASRD information and information for individuals in the comparison groups, generate the study's primary outcomes measures. These include earnings and employment for participants for up to 16 quarters following participation and for comparison group members in the same periods. These data also include earnings prior to WIA participation, facilitating the construction of employment histories of participants and comparison group members.[12] Early quarters after program entry may show negative effects of training on earnings and employment, "lock-in" effects reflecting participants' involvement in program activities rather than employment. Later earnings effects are expected to be positive, as skills obtained during the program combine with job experience. All earnings have been adjusted for inflation to correspond with the first quarter of 2006.

It has long been recognized that controls for the standard demographic characteristics such as gender, age, education and race are important. Such information is available in the current study. Local labor market is captured using aggregates of county of residence or service, or where county is not available, the local Workforce Investment Area. It is also widely recognized that the details of the labor market experiences of individuals in the period immediately prior to program participation are critical.[13] Wage record data provide information on employment status at the time of initial program involvement and for prior years. Additional relevant variables include controls for veteran status and prior earnings.

---

[10] Two of the twelve states provided WIA exit data extending through only June 2006. Since WIA participants who did not exit the program by this date are omitted, a larger share of individuals are omitted in these states. Because of data problems in two other states, the study examined program entries for periods other than July 2003-June 2005, one for calendar year 2003 and the other for January 2004-June 2005.

[11] The ES sample has the advantage that it includes any individual who chooses to obtain services without regard to prior employment history, whereas UI provides benefits only to those who have sufficient prior work experience. As a practical matter, negotiating use of the ES data was more complex, and it was not possible to arrange for use of ES data in most states.

[12] In one state, wage record data extend only through June 2007. Wage record data are available for at least four quarters prior to the first quarter of analysis in every state, and in all but three states a full two years of wage record data are available for all WIA entry dates considered.

[13] In particular, movements into and out of the labor force and between employment and unemployment in the 18 months prior to program participation are strongly associated with both program participation and expected labor market outcomes (Heckman, LaLonde and Smith, 1999; Heckman and Smith, 1999).

Analyses are performed separately by gender. Where possible, WIA participants who enter in a given quarter are also matched with individuals in the comparison sample who have contact with their respective programs in the same quarter, providing an exact match on quarter of entry.[14]

3. Descriptive Statistics

Table 2 provides sample sizes and means for WIA participants and the comparison group in the 12 states. A total of 95,580 unique individuals entered the WIA Adult program during the observation window. Since about 2 percent entered the program more than once, the total number of entries was 97,552. Similarly, 63,515 individuals entered the Dislocated Worker program, producing a total of 64,089 total program entries.[15] The rightmost column identifies the number of individuals who participate in comparison programs 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.[16]

Turning to the next panel, we can see that individuals who participated in the WIA Adult program are more likely to be female and minority than participants in the comparison program; they are also appreciably younger and have slightly less education. These differences reflect the fact that participants in the WIA Adult program tend to be economically disadvantaged, whereas participants in the comparison program (UI claimants or ES participants) are individuals who have recently lost jobs. Therefore, individuals in the comparison program have the characteristics of individuals with relatively strong labor market attachments—white, male, older workers with more education. Comparing participants in the WIA Dislocated Worker program with the comparison group, it is clear there are fewer differences—participants in the WIA Dislocated Worker program are more likely to be female and are slightly older, but differences are smaller.

The data on past employment and earnings for these groups provide further evidence that participants in the WIA Adult program have weaker labor market attachments and are more economically disadvantaged than comparison program participants. Participants in the WIA Adult program are less likely to have worked continuously in the six prior quarters (30 percent versus 48 percent) and are much more likely to have not worked in any of the six quarters (17 percent versus 4 percent) prior to entering the program; they also have much lower annual earnings in the two years prior to entering the program and in the two subsequent years. In contrast, participants in the WIA Dislocated Worker program have similar labor market attachment and only slightly lower earnings than those in the comparison program.

[14] Comparison group individuals may contribute more than one unit as potential matches if they had contact with the program in multiple quarters. In such cases, when a later quarter for a comparison case is chosen to match with a WIA participant, prior quarters of participation in the comparison program must correspond for these cases. Further detail is provided in Section IV.

[15] Where an individual entered the program more than once during a quarter, this was coded as a single entry. Data cleaning also eliminated multiple entries when these appeared to be due to data entry errors or when they pertained to the same set of services.

[16] The matching methods employed here consider all quarters of comparison program participation, allowing a given individual to be matched to WIA participants in more than one quarter.

Table 2
Summary Statistics for WIA Participants and Comparison Group in 12 States

	WIA Adult			WIA Dislocated Worker			
	Overall	No Training	Training	Overall	No Training	Training	Comparison Group
Sample Size							
Unique individuals	95,580	68,255	27,325	63,515	43,513	20,002	2,929,496
WIA entries, or quarters of comparison program participation	97,552	69,712	27,840	64,089	43,894	20,195	6,161,510
Demographic	Mean	Mean	Mean	Mean	Mean	Mean	Mean
Male	0.420	0.445	0.356	0.482	0.494	0.456	0.585
Black	0.445	0.512	0.277	0.330	0.391	0.198	0.171
Hispanic	0.031	0.014	0.072	0.022	0.013	0.043	0.064
Age	32.70	32.91	32.16	40.24	40.14	40.46	39.59
Years of education	12.27	12.21	12.43	12.55	12.52	12.63	12.42
Employment							
Employment-employment	0.297	0.294	0.307	0.462	0.465	0.456	0.476
Employment-not employed	0.208	0.195	0.241	0.281	0.256	0.335	0.279
Not employed-employed	0.325	0.336	0.297	0.183	0.199	0.149	0.225
Not employed-not employed	0.168	0.175	0.151	0.070	0.078	0.053	0.040
Earnings second year prior	8,507	8,203	9,306	19,402	17,782	23,487	20,156
Earnings in prior year	8,149	8,050	8,398	20,499	19,450	22,779	21,584
Earnings following year	9,426	9,128	10,171	11,527	11,840	10,845	15,649
Earnings second year after	10,846	9,916	13,175	14,572	14,213	15,352	17,102
Program Experience							
WIA in prior two years	0.052	0.058	0.035	0.041	0.044	0.034	0.020
Comparison program participation in prior two years	0.211	0.178	0.297	0.409	0.353	0.551	0.668

The bottom panel of the table shows that 4 to 5 percent of WIA entrants had previously participated in WIA (either the Adult or Dislocated Worker program), and that the number participating in the comparison program was substantial. About a fifth of Adult program participants had prior comparison program experience, compared to over two-fifths of Dislocated Workers. By definition, a comparison case participates in the comparison program in the specified quarter; the table shows that about two-thirds of such individuals had participated in that program in the prior two years.

Comparing columns 2 and 3, and columns 5 and 6, we see that participants who receive training services are more likely to be female and much less likely to be black than participants who do not receive training services. Differences in education are very small. Based on prior earnings, those receiving training services appear to have had greater labor market success, but measures of employment imply only small differences in employment activity.

Notwithstanding these differences, there are important similarities in the patterns of earnings for treated and comparison cases. Figure 1 graphs quarterly earnings for WIA Adult program participants and the sample of individuals in the comparison group. Figure 2 provides comparable plots for the Dislocated Worker program. In these figures the negative numbers on the horizontal axis indicate quarters prior to program entry; quarter 0 is the quarter an individual begins participating in a program; and the positive numbers indicate quarters after entry into the program. In each plot, separate lines are provided for all WIA Adult participants, participants who receive training services, and WIA participants who do not receive training services.

For comparison cases, the horizontal axis identifies quarters relative to the quarter of participation. The most notable pattern in both figures is the decline in earnings that occurs in the several quarters prior to program entry, a pattern that has been called the “Ashenfelter dip” (Ashenfelter, 1978; Heckman and Smith, 1999). This reflects the fact that individuals often enter such programs following a period of setbacks in employment. In attempting to find a comparison group, this pattern may set program participants apart from potential comparison individuals. It is therefore significant that there is a decline preceding program participation for the comparison group.

That the comparison group displays a similar basic pattern in earnings to the two WIA programs confirms that there will be sufficient numbers of individuals to match with WIA participants on the basis of prior employment. Equally important, it also suggests that there may be similarities in the individual employment environments faced by the comparison and treatment groups, implying that unmeasured factors may be similar as well. Nonetheless, it bears repeating that even with great care in matching, there is no guarantee that all unmeasured differences in factors affecting outcomes between the treated and matched comparison groups will be eliminated. Specification tests are an essential part of the analyses that follow.

The aggregate numbers presented in Table 2 hide differences across states in programs. The total number of participants entering the Adult and Dislocated Worker programs during the period of the study varies across states from as little as 1,500 to over 50,000. For this reason, we also present an analysis in Appendix B that examines whether the patterns are similar in various subsets of the states. Where patterns are similar, this suggests that results are not driven by a small number of large states. For example, one important difference in the character of the programs is reflected in the proportion of individuals who receive training. Seven of the state programs provide training to more than 60 percent of participants, one state provides training to about half of its participants, and the remaining states provide training to less than 40 percent.[17]

[17] In our sample, those programs that are more likely to provide training to participants in the Adult program are also likely to provide training to participants in the Dislocated Worker program.

Figure 1
Quarterly Earnings for WIA Adult Program and Comparison Program Participants in 12 States, Prior to and Following Participation

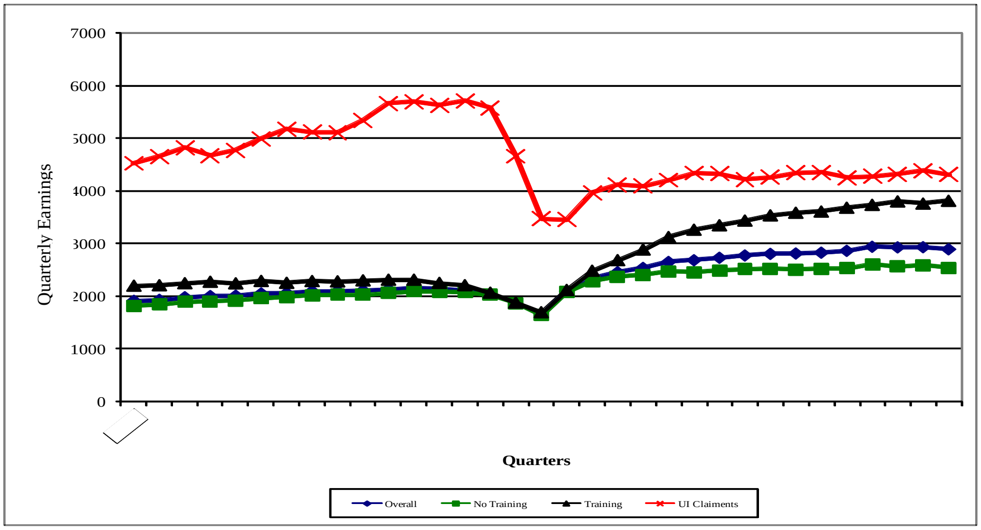
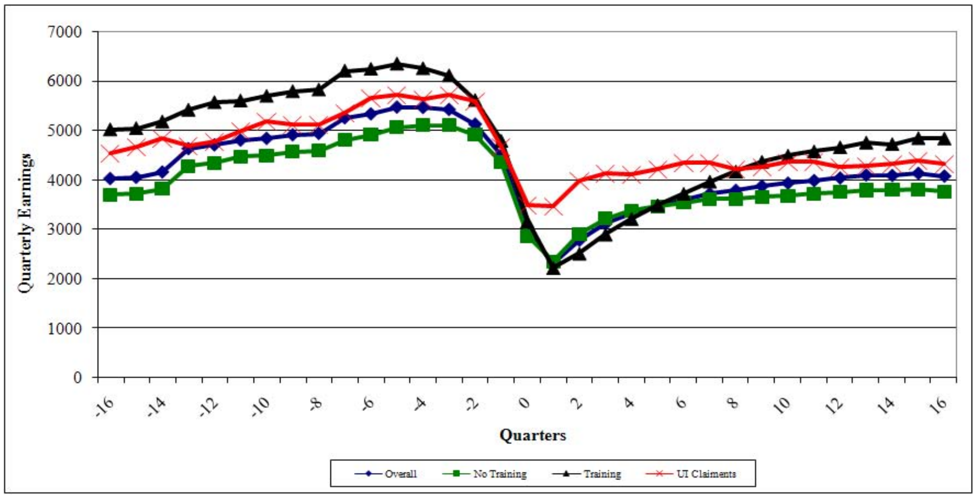


Figure 2
Quarterly Earnings for WIA Dislocated Workers and Comparison Program Participants in 12 States, Prior to and Following Participation



4. Method of Analysis

We estimate the impact of participation in the WIA Adult or Dislocated Worker programs on outcomes for those who participate, that is, the effect of the treatment on the treated. We use propensity score matching,[18] which, like other matching and related methods, assumes that the outcome that would occur in the absence of the treatment is conditionally independent of the treatment (Rosenbaum and Rubin, 1983). Although the conditional independence assumption cannot be tested directly, we apply a reasonable specification test that examines prior earnings. If subsequent earnings in the absence of the treatment would have been the same for treated and comparison groups conditional on measured characteristics, we would expect prior earnings to be the same as well. Conversely, if differences in stable factors that influence earnings exist between the treatment and comparison group, we expect there to be differences in the conditional means. In practice, the test based on this comparison amounts to estimating the "effect" of program participation on prior earnings. If there is no significant effect, this suggests that there are no stable factors influencing income that differ for the treated and control group.

Where the specification test fails, individual fixed effects estimators provide an alternative approach to controlling for differences across individuals who participate in WIA.[19] So long as such differences have stable effects on earnings, this specification can eliminate bias. Despite the benefits of the difference-in-difference estimators, depending on the processes underlying earnings dynamics and program participation, estimates may have biases that are not present in cross-sectional matching. The difference-in-difference estimator needs to be understood as one of several estimates that make different assumptions.

*Matching Strategy***.** The estimator of program impact that we use here is many-to-one caliper matching with replacement based on the propensity score. The impact estimate can be written

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

where $\overline{Y}_{0\,j(i)}$ is the average outcome for all comparison cases that are matched with 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 "close" enough, measured in terms of propensity score. The method is recommended in Dehejia and Wahba (2002), and is closely related to propensity score stratification or interval matching (Rosenbaum and Rubin, 1983).[20] In order to obtain difference-in-difference estimates, Y_{1i} is replaced by the difference between earnings following participation and earnings prior to program participation, and $\overline{Y}_{0\,j(i)}$ is replaced by the average

---

[18] See Rosenbaum (2002), Imbens (2004), or Rubin (2006) for general discussions of matching methods. Caliendo and Kopeinig (2008) provide practical guidance for propensity score matching.

[19] Smith and Todd (2005a) spell out the basic approach, which they describe as "difference-in-difference" matching.

[20] 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.

difference for the matched comparison cases over the same period.

The matching is performed on the basis of propensity score, so matched cases may have different values on independent variables. If the propensity score is correctly parameterized, in the limit as the sample size growth, the distribution of values for all included variables should be the same for the treated and matched comparison cases, meaning that for a given treated cases there is a matched case with the same values on all variables. Matching is based on a constant radius expressed as the difference in the log-odds of the propensity score between treated and comparison cases. Additional details on how the matching is performed, including the matching variables and diagnostics, are discussed in Appendix C.

Standard errors. Conventionally, standard errors of propensity score matching estimates are obtained using bootstrap methods. With large samples such as those available to this study, it is not feasible to calculate bootstrap standard errors for all estimates (see, e.g., Lechner, 2001). In pilot analyses undertaken for two states, we compared several approaches that have been suggested for estimating standard errors. The first, recommended by Imbens and Woodridge (2008) and Imbens (2008), produces a conditional standard error, which provides an estimate of the variation in an impact estimate conditional on the independent variables.[21] Abadie and Imbens (2006a) suggest an approach for estimating the unconditional standard error, which provides an estimate of the variation in the impact estimate that would result if the sample were chosen repeatedly from the full universe, with values on independent variables varying from sample to sample. When implemented in two states, these standard error estimates were very similar (with mean differences of less than 1 percent, and absolute differences of less than 1 percent in one state and 2.5 percent in the other state). Clearly. no substantive conclusion depends on the choice between these measures, and thus, we report the conditional standard in our analysis. Additional details on the standard error estimates are included in Appendix C.

*Matching Diagnostics***.** The matching model specification was determined separately for each of the comparisons by gender within each of the 12 states. Three dimensions were of concern: the appropriate radius value, the specification for independent variables, and whether exact matches would be required for quarter of entry.

The analysis began with a default specification. First, the sample was divided into subgroups by quarter of entry, producing a total of eight groups over the two program years, the period used in most states. In fitting the logit for the propensity score, all available independent variables were included in linear form along with selected interaction measures identifying particular patterns of employment and program participation over time. The radius was set so that any WIA entrant was matched to all comparison cases for which the log odds of the propensity score was within 0.1.

All entry subgroups were then combined and tests were performed to determine if the means for the independent variables for the treated cases differed from the matched comparison cases. In addition, tests of statistical significance were performed on differences between means for the squares of the continuous variables and selected interactions between the variables. In most

[21] In the implementation here, following Imbens (2008), it is conditional on the propensity score estimate.

states, there were between 70 and 100 initial variables that were tested, approximately 25 square terms, and up to 200 interaction terms. Approximately 5 percent of differences are expected to be statistically significant at the 0.05 level if individuals from a common universe were assigned randomly to the groups.. A matching procedure was viewed as successful if fewer than 8 percent of the differences were statistically significant.[22]

Appendix C presents additional details on the matching diagnostics and the results of these tests. The appendix table reports the proportion of cases that were omitted from the analysis for each of the treatment and comparison samples (1-3) shown in Table 1. For comparisons 1 and 2 (WIA participants vs. comparison group members), the loss of cases is small, generally in the range of 2-7 percent. For comparison 3, however, the proportion excluded is comparatively high, almost 50 percent for Adult males. 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 selected in any clear way, except that they represent small states with relatively large proportions trained.

IV. Results of Impact Estimation for Adult Program

We obtain estimates of WIA program impacts on average inflation-adjusted earnings and employment in the 16 quarters following program start. Once impact estimates specific to a state are obtained, the mean across states is estimated by weighting the estimate for a given state by the number of participants who were matched in that state. The resulting weighted mean provides an estimate of the average impact for matched WIA participants who entered the program during the periods specified in the states of this analysis. Bias adjustment was applied to the reported impact estimates, although given that the matches are very close, it makes very little difference.[23]

For each state, estimates were also obtained for the difference between the treated sample and the matched comparison sample on outcome variables in *prior* quarters. We focus on earnings and employment 10 quarters prior to program entry, and earnings and employment 16 quarters prior to program entry. These estimates provide a specification test for our model; if the program "effect" on prior measures is significantly different from zero, this implies that participants are different in ways not captured by the matching criteria. Estimates of the impact on subsequent outcomes would therefore be suspect.

[22] An additional concern was the proportion of WIA entries that were successfully matched. It may not be possible to find matching comparison cases for treated cases whose characteristics place them in the sparsely populated portion of the comparison case space. In general, if at least 90 percent of the WIA entries were matched, the matched proportion was taken as acceptable.

[23] Following Abadie and Imbens (2006a), we fit a linear model in the comparison sample, and then use coefficients estimated in this sample to adjust for any differences in independent variable means that exist between the treatment and matched comparison cases. Where sample sizes were very small, the bias adjustment was not possible, and estimates reported do not include the bias adjustment.

In the case where the model fails such a specification test, one approach is to calculate a difference-in-difference estimate, which can be obtained by subtracting the prior earnings effect from the conventional estimates. As noted above, such estimates can be unbiased if differences between treated and control cases are due to fixed differences. Such estimates rely, however, on assumptions that are difficult to test.

Associated with each state impact estimate is an estimated conditional standard error, which is combined in the conventional way to form the standard error for the weighted average across states.

1. Adult Program Impacts

The first set of analyses focuses on how earnings and employment for all Adult program participants are affected by the program for each of the 12 states. Table 3 provides summary statistics for the estimates. For each state, the simple average of impact estimates for quarters 1-5 is presented, as well as the average for quarters 11-16.[24]

There are substantial differences in average impacts in the first five quarters, but the finding of a substantial positive statistically significant effect is widespread. Of the 24 estimates (12 states x 2 genders), 18 are positive and statistically significant, and only one is negative and statistically significant. The same basic pattern applies for quarters 11-16, although effects are larger for most states.

Although the results in Table 3 show that there is substantial variation across states that cannot be explained by sampling error, sampling error is still large in many cases. Figures 3 and 4 provide estimates for women and men, respectively, combining the estimates from all 12 states. 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.[25] Also on the graph are dashed lines that show the confidence interval for each estimate. The lower dashed line subtracts twice the conditional standard error from the estimate, and the upper dashed line adds twice the standard error. [26] Also presented in this figure are the estimates of "impact" on earnings 10 quarters prior and 16 quarters prior to program entry, providing a specification test of the model.

[24] As noted above, data use agreements preclude revealing state identities. States are ordered by the size of the average effect in quarters 11-16 for females. For some states, impact estimates are only available for a subset of quarters 11-16, and, in these cases, averages are based on available quarters. In three states, no estimates after quarter 10 are available and the quarter 10 estimate is presented.

[25] Data are available for 10 quarters in all states, but estimates for quarters 11-16 are based on the subset of these states where that information is available, varying with the particular comparison. In order to prevent differences in impact by state from directly altering the trends, we have adjusted values for quarters greater than 10. The adjustment involves augmenting or reducing the impact estimates for the subset of states by a constant so that estimates for the subset correspond to the latest quarter where the fuller set of state estimates is available. In order to capture the uncertainties of this adjustment procedure, the confidence intervals presented in the graphs represent the widest combination of confidence intervals of both the adjusted and unadjusted effect estimates.

[26] These correspond to the 95.5 percent confidence interval.

The estimates reported in the figures imply that, for both genders, participants earn between $400 and $700 more per quarter than matched individuals in the comparison program. For women, the impact estimate over most of the 16 quarters is between $500 and $600 per quarter, whereas for men there is a decline in the first three quarters, with the level settling in the range of $400.

Table 3
Adult Program Treatment Effect on Quarterly Earnings by State: WIA versus Comparison Program

	Females		Males	
State	Quarters 1-5	Quarters 11-16	Quarters 1-5	Quarters 11-16
1	-140	-165	-187	235
2	208*	168*	69	-35
3	528*	396*	452*	290
4	409*	418*	354*	-12
5	302*	588*	475*	835*
6	302*	624*	359*	483*
7	476*	909*	120	197
8	-129*	949*	38	371*
9	241*	1094*	360*	964*
10	721*	1198*	892*	840*
11	1187*	1283*	1233*	892*
12	908*	1426*	1203*	1211*

*Statistically significant at the 0.05 level.
Note: Average effects for specified quarters. Where estimates are not available for a given state, the average is calculated on available quarters. For three states, estimates are not available for quarters 11-16, and the reported estimate applies to quarter 10.

Figure 3
Adult Program Treatment Effect on Quarterly Earnings for Females, WIA versus Comparison Group

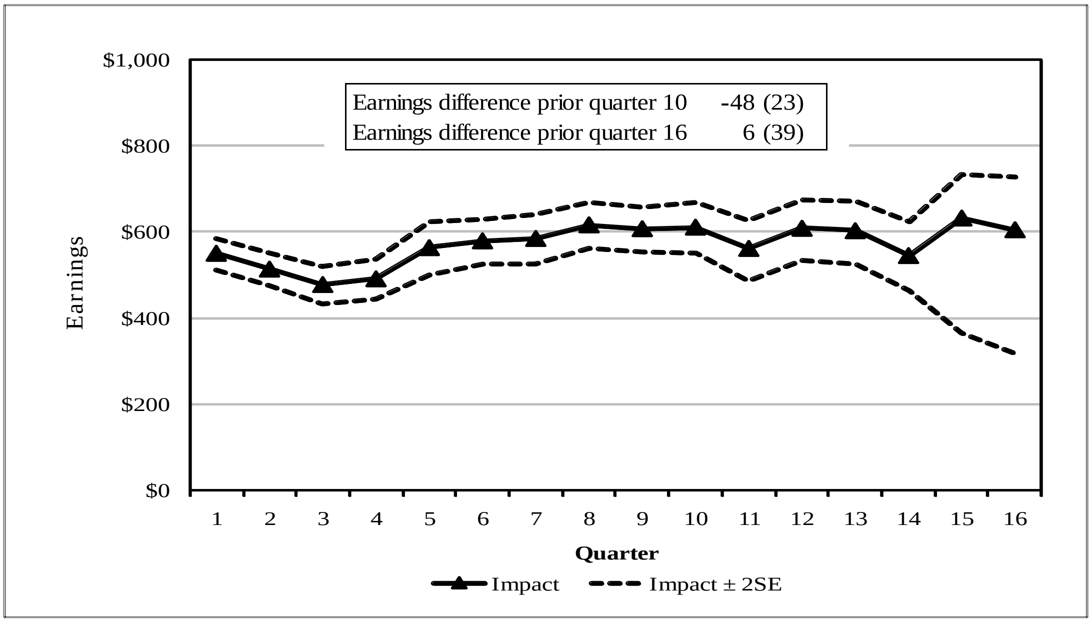
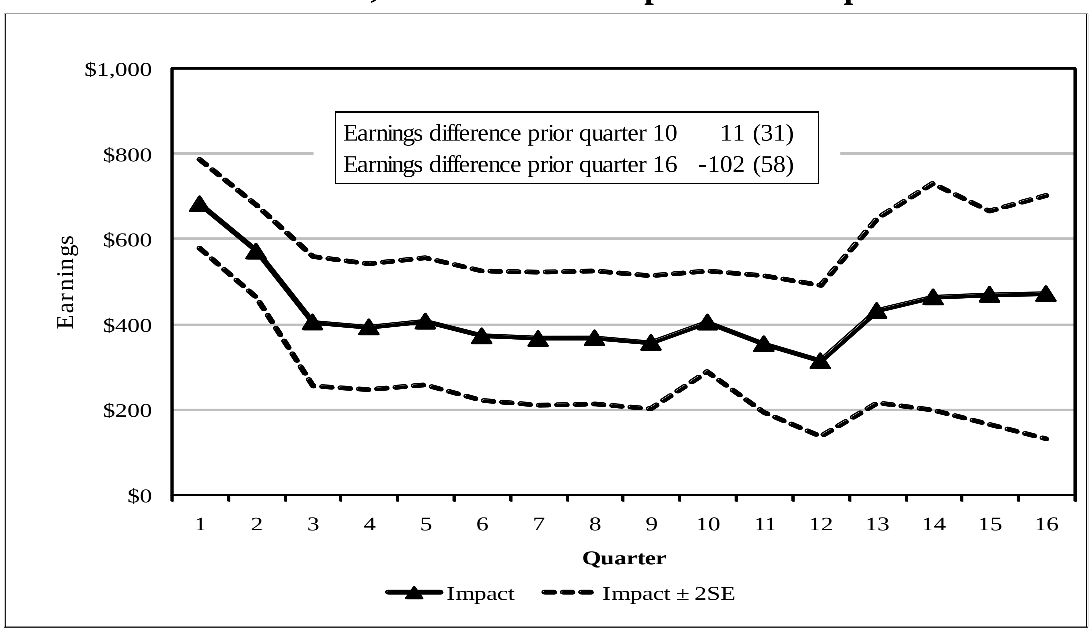


Figure 4
Adult Program Treatment Effect on Quarterly Earnings for Males, WIA versus Comparison Group



Given that average quarterly earnings for WIA Adult participants are between $2000 and $3000 in the first year after program entry, ultimately approaching $4000 after four years, the effects are in the range of 10 to 30 percent.[27]

Figures 5 and 6 provide analogous estimates for employment. The method used to obtain estimates reported in these figures is identical to that for earnings, except that here the proportion employed (identified as having received positive earnings in the quarter) is taken as the dependent variable. Each value can be read as the difference between the employment rate for Adult program participants and the matched comparison cases. For example, the value is 0.13 for females in the first quarter after participation, implying 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 points. Male impacts are one or two percentage points lower.

As noted earlier, there are substantial differences in the proportion of individuals receiving training across the state programs. It might be expected that patterns of effects would differ for programs with different levels of training. First, the total resources per participant are appreciably higher in such states, so long-run program impact could be higher if more intensive services produce greater impacts. Second, a large share of the value may well occur with a greater lag, since training benefits presumably accrue over a more extended period. Figures 7 and 8 provide impact estimates for the seven states that provide training to more than half of their participants. In these states taken together, 68 percent of Adult program participants receive training.

The initial effects—during the first several quarters after program entry—in these seven states are very similar to the aggregate for all states. In contrast, however, among females, growth in earnings up through the first 10 quarters is notable, with the ultimate quarterly earnings increment reaching $1,100. Although there is no growth in impacts for men, neither is there a precipitous initial decline as occurs for the whole sample. In short, there is at least weak evidence suggesting that high-training states produce benefits that endure longer. The basic pattern for employment impacts is similar to that for earnings, and so that graph is not presented.

Taken at face value, these results imply that the program has strong and substantial impacts with little or no lag. These could reflect aggressive actions by program staff to help workers obtain employment initially, with training assuring benefits that accrue over an extended period. Skeptics will argue, however, that the findings of such large initial impacts call into question the appropriateness of the comparison group and ultimately the validity of the results. With most training programs, initial participants experience reductions in earnings as they engage in training activities that supplant employment that would otherwise occur. In these data, the

[27] Recall that Orr et al. (1976) estimated impacts on quarterly earnings two years after program entry for participants in the JTPA program in the $300-350 range.

Figure 5
Adult Program Treatment Effect on Quarterly Employment for Females, WIA versus Comparison Group

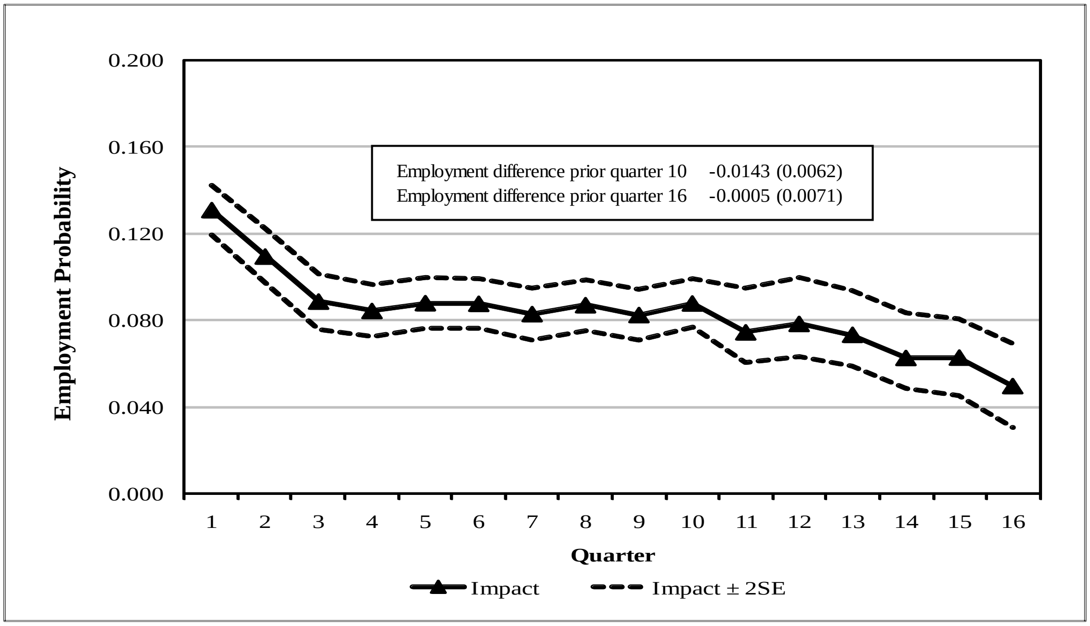


Figure 6
Adult Program Treatment Effect on Quarterly Employment for Males, WIA versus Comparison Group

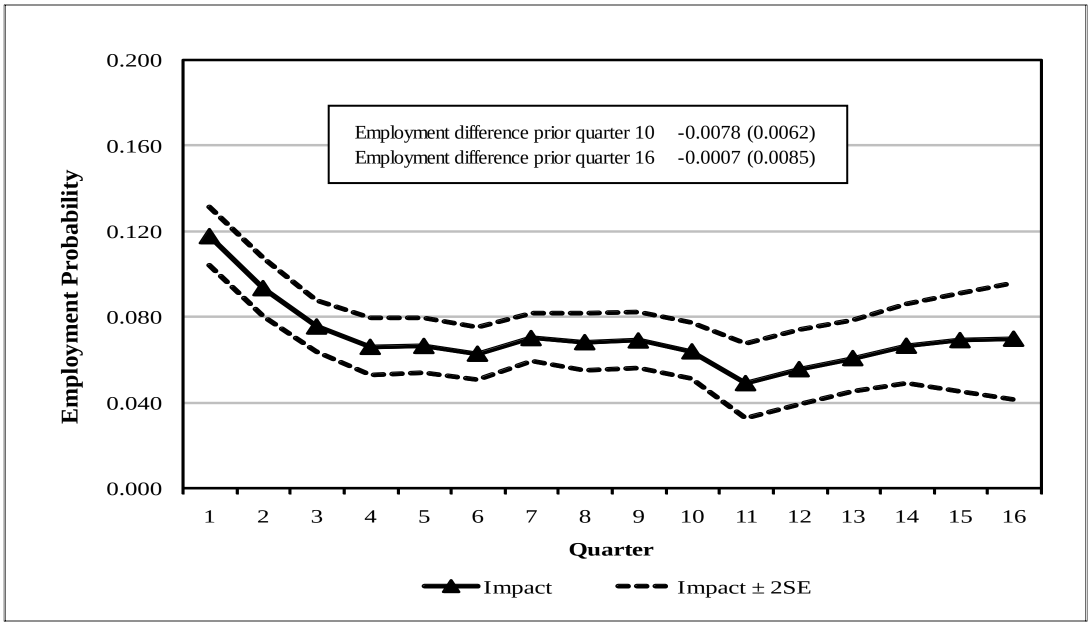


Figure 7
Adult Program Treatment Effect on Quarterly Earnings for Females, WIA versus Comparison Group in 7 High-Training States

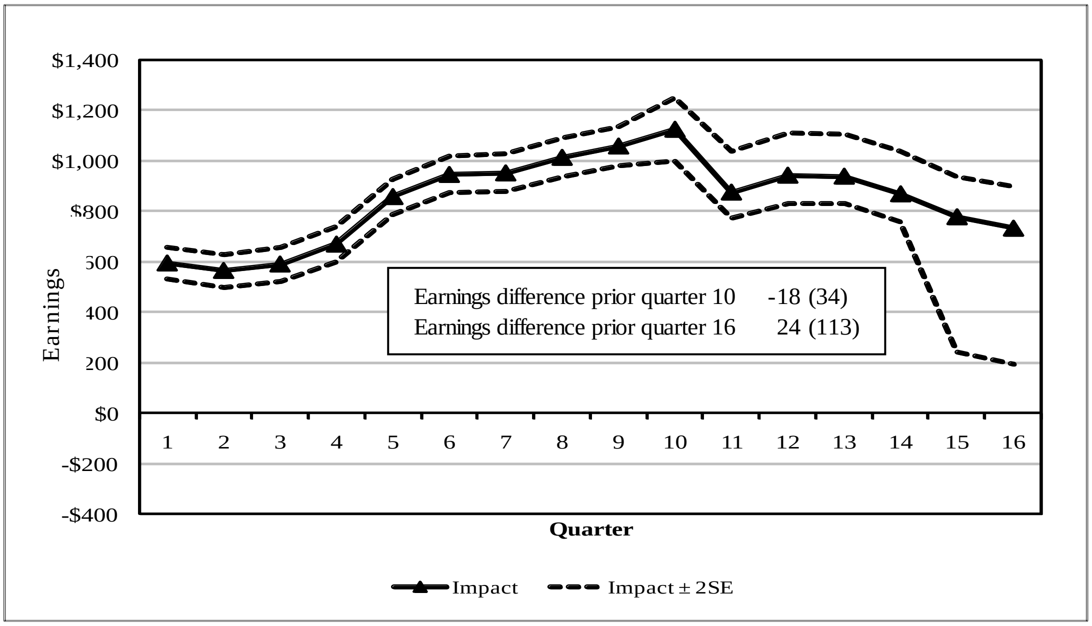
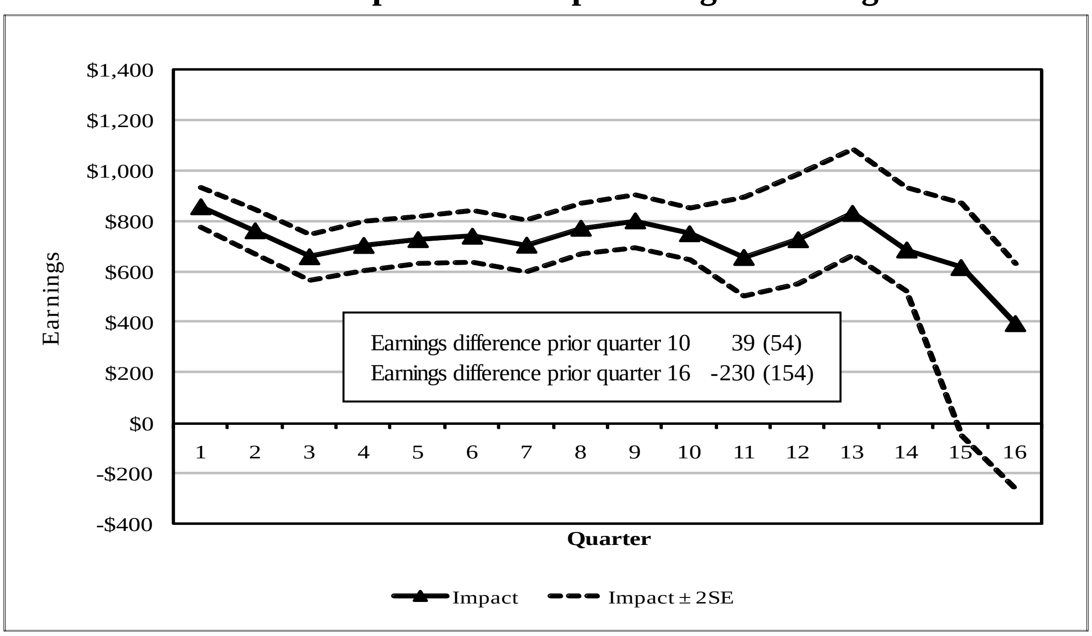


Figure 8
Adult Program Treatment Effect on Quarterly Earnings for Males, WIA versus Comparison Group in 7 High-Training States



median participant exits the program around three quarters after entering. Hence, it is clear that estimates of program impact in the two quarters after entry identify a point in time when most people are still participating in the program.

In order for selection to cause these results, it must be the case that WIA participants have unmeasured attributes that make them more likely than those in the comparison program to obtain employment or higher earnings. Staff admission criteria or participant choice would need to select entrants who were appreciably more likely to obtain employment than other individuals with similar characteristics, employment and program participation histories.

A first test for selection is provided by analyses that predict prior earnings. Although controls are included for earnings in the eight quarters prior to entry, 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. We calculate the difference in earnings or employment between treated and comparison cases for measures applying to the tenth and sixteenth quarters prior to entry, presenting these estimates as inserts in the figures. These difference estimates show that earnings and employment are *not* higher for WIA participants; in most cases, the differences are small (see Figures 3 and 4). The largest difference are for male WIA participants 16 quarters earlier, for which it appears that WIA participants had earnings about $100 *below* those of the comparison group. Such a difference would tend to downwardly bias estimates; estimates from a difference-in-difference model would produce program impact estimates that were $100 greater. For males in the seven high-training states (Figure 8), earnings are $230 lower. It is therefore clear that if selection is causing spurious positive impact estimates, selection is unlikely to be based on stable individual characteristics.

One alternative explanation would be that there are transient differences between WIA participants and others. The comparison group members receiving unemployment compensation may include a substantial portion of individuals who are not seeking employment. UI recipients classified as awaiting recall are not required to search for employment, and many others may have little interest in getting a job—despite formal requirements—until benefits are about to expire. According to this view, those obtaining UI benefits are in a phase where their short-term employment levels are expected to be depressed, reflecting the incentives due to UI benefits, which would discontinue if a job was obtained. WIA participants, in contrast, have chosen to select into a program with the purpose of improving their employment prospects.

If the problem stems from the differences between WIA participants and UI claimants, it might be expected that such differences would be less important for the other comparison group, those seeking ES services. Although most UI claimants are required to register for ES services, those awaiting recall are exempt from this requirement, so the ES sample removes one group whose interest in employment may be modest. Since any individual seeking support for employment search can obtain ES services, this sample is expected to include self-motivated job searchers.

Figure 9
Adult Program Treatment Effect on Quarterly Earnings for Females, WIA versus ES Participants in 3 States

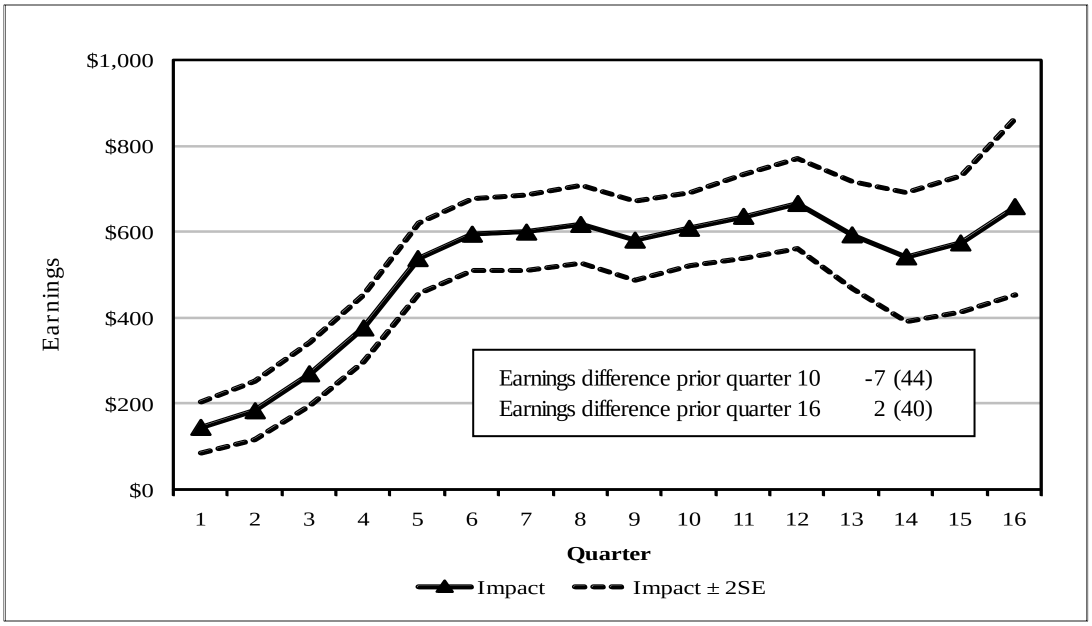
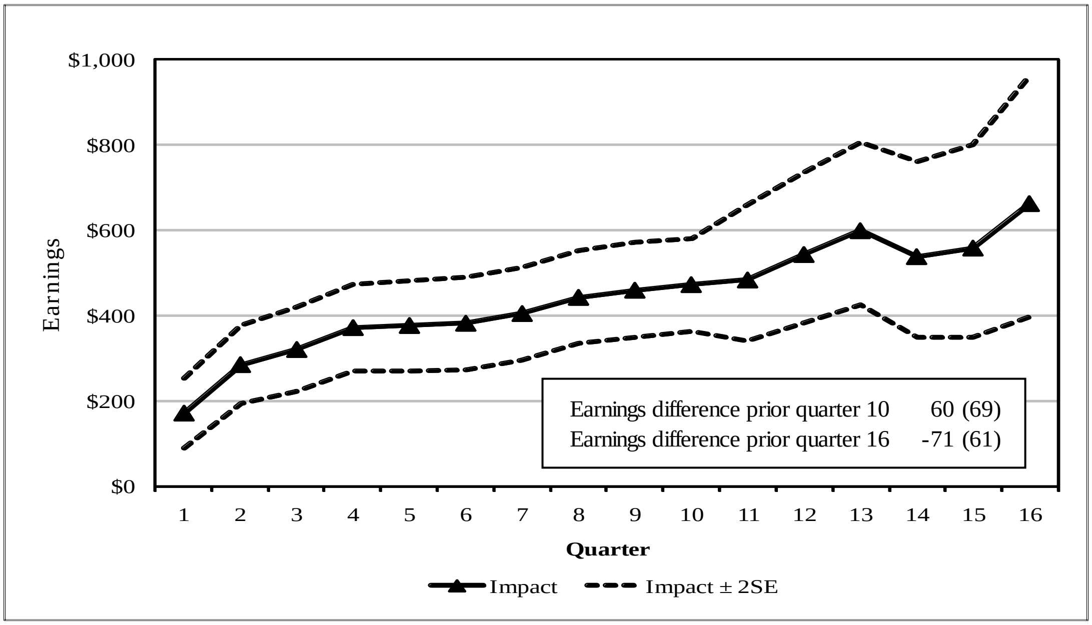


Figure 10
Adult Program Treatment Effect on Quarterly Earnings for Males, WIA versus ES Participants in 3 States



Figures 9 and 10 provide earnings impact estimates for the three states where ES recipients form the comparison group. There are important differences in these patterns as compared with the full sample of states. Perhaps most notable, impacts in the first few quarters after entry are somewhat smaller, in the range of $200 for both men and women. There is a fairly steady increase in program impact up through the last quarters.

These results therefore support the view that the large impacts on earnings and employment in the quarters immediately after WIA entry could be at least partly due to differences between WIA participants and the UI claimant comparison group rather than to the effects of program participation. Of the nine states for which UI claimants are the comparison group, initial program impact is similarly small in only two of them.

In conclusion, estimates of overall program impact are positive in almost all states, although variation is substantial. While the patterns suggest the possibility that estimates in the first two or three quarters after program entry may be biased, selection and incentive explanations do not suggest that impact estimates for later quarters are spurious.

2. Impacts for UI Recipients

In the discussion above, we considered the possibility that UI claimants, the comparison group used in nine states, 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. In the analyses here, we limit the treated group to those who are receiving UI benefits when they enter the WIA program.

The structure of these analyses is essentially the same as that for all Adult participants, with the exception that only WIA participants who were receiving UI benefits at the time of entry into WIA are included. In addition, the comparison group has been modified to include only UI claimants who actually received UI benefits, dropping the small number of individuals who applied for benefits but did not receive them. In common with the above analysis, prior UI participation and employment patterns are controlled, so UI eligibility in future periods should correspond closely for participants and matched comparison group members.

Adult program participants who receive UI benefits at the point of entry account for less than 10 percent of entries during the period of our study. Although this is an important group, impacts in this group need not be representative of others in the program. We would expect UI recipients to be more successful in the labor market than the average Adult participant, and they would tend to be similar to those in the Dislocated Worker program.

Figure 11 shows that in the first two quarters following entry into the Adult program, women earned less than those in the matched comparison group, suggesting a lock-in effect during program participation. During subsequent quarters, earnings effects and are in the range of $100 for participants. (Given that the sample for quarters 14-16 is limited and standard errors are large, we place little emphasis on these estimates.) Looking at estimates of earnings “effects” for prior quarters, we find little evidence that selection on stable factors is of importance.

Figure 11
Adult Program Treatment Effect on Quarterly Earnings for Female UI Recipients: WIA versus UI in 9 States

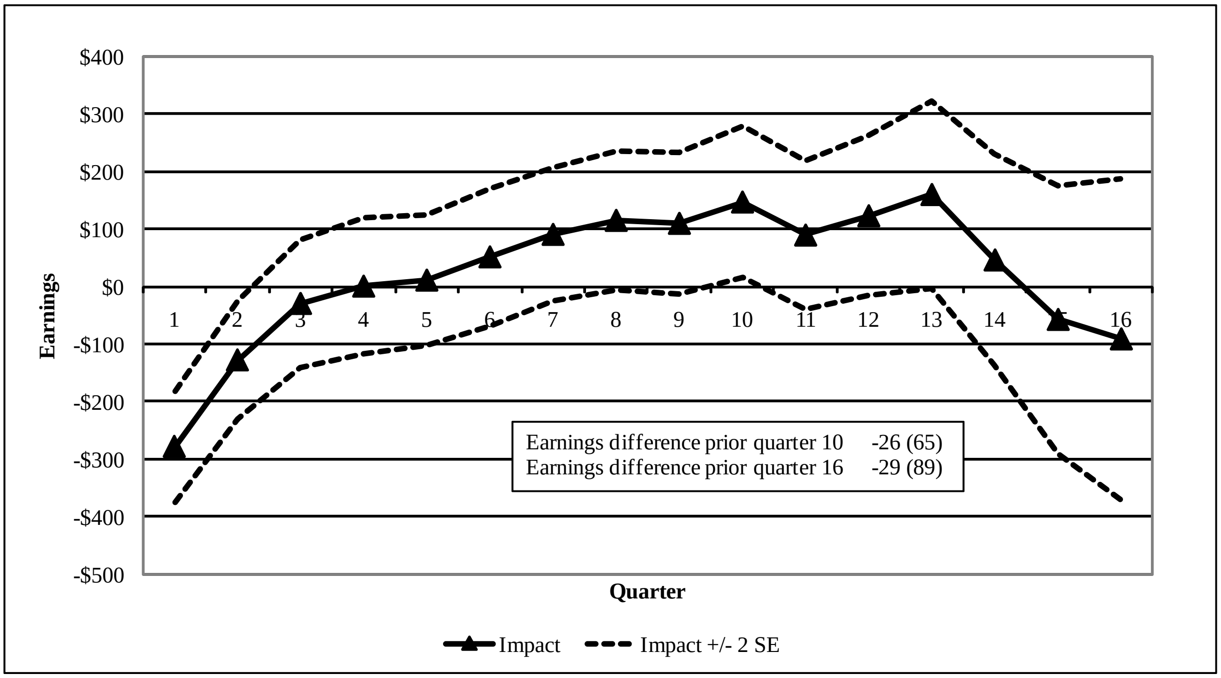
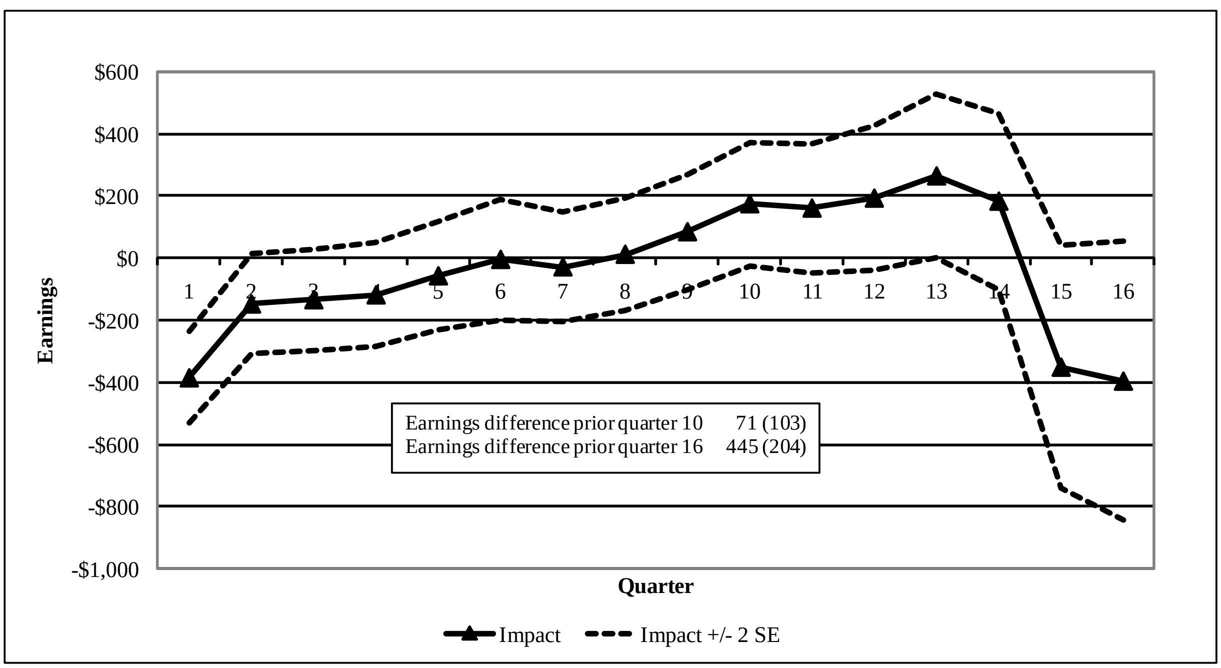


Figure 12
Adult Program Treatment Effect on Quarterly Earnings for Male UI Recipients WIA versus UI in 9 States



These results are markedly different from those for female Adult participants (Figure 3), which show immediate impacts in the range of $500, with little trend in later quarters. Impact estimates in the three states with the ES comparison group (Figure 9) also show higher impacts, although they are similar to the current result in that they show an initial upward trend. The difference between Figures 3 and 11 may reflect bias in the former estimates. However, anticipating the results in the analysis of the Dislocated Worker program below, we suspect that the lower quarterly impacts for quarters 4-16 indicate that the benefits to UI recipients are smaller than for the most disadvantaged individuals, who would not be eligible for UI benefits.

Results for males (Figure 12) also show substantial negative initial effects, in this case extending for the first five quarters. By quarter 10, impact estimates are nearly $200. However, prior earnings differences between participants and the control group appear to be very large. In prior quarter 16, earnings are appreciably higher for participants than the comparison group, with a difference of over $400. Unfortunately, in this analysis, this estimate is based on only two of the 12 states, so less emphasis should be placed on these estimated differences.

The comparable estimates for employment for females and males receiving UI benefits (not presented) are broadly consistent with those reported above. The initial lock-in effect is observed for females, although it only lasts one quarter. Employment effects in quarters 2-5 imply a 2 percentage point increment for participants, and a somewhat larger increment—perhaps as much as 4 percentage points—in later quarters. There is no evidence that female participants are positively selected, but male program participants may be positively selected, based on prior employment differences.

Taken together, these results suggest that impacts for Adult WIA participants receiving UI benefits are substantially smaller than for the full population of participants. In the discussion below of the Dislocated Worker program, we present evidence suggesting that the average impact in that program may be smaller than for the Adult program. This supports the view that the benefits of WIA for those who lose a "good" job may be smaller than for workers with generally poor work histories.

3. Impacts of Training

The heart of WIA services is the basic and vocational skills training provided to individuals. Although a variety of training opportunities are widely available outside of WIA, for many WIA Adult participants, the alternatives available are more costly. It is clear that acceptance into WIA alters the type and extent of training these individuals ultimately obtain.[28]

[28] There is no way to determine the extent to which comparison group members receive training or related services outside the WIA program. Reported impact estimates are therefore incremental relative to services received by the comparison group, some of whom undoubtedly receive job 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 received roughly comparable employment and training services outside the JTPA program. For those in the treatment group, who were offered JTPA services, the number actually receiving services was 50-60 percent.

Figures 13 and 14 present impact estimates of training based on comparison 3, where the comparison group is individuals in the same WIA program who did not receive training. Earnings impact estimates for females imply a $200 decrement in the first quarter after program entry, as would be expected if time in training limited employment options initially. Earnings, however, catch up three or four quarters later, with a positive increment over $800 by the end of 10 quarters. In contrast, males who receive training appear to experience positive impacts—in the range of $200 immediately after entry—with the increment remaining in the $500-600 range for the next 10 quarters.[29]

Figures 15 and 16 show estimates for employment impacts. For women, initial employment is about 5 percentage points lower for those receiving training, and the employment rate for participants only catches up four quarters after entry. By the tenth quarter, the increment is in favor of training recipients by about 5 percentage points. For men, the pattern is quite similar, although the increment is close to zero for six or seven quarters after program entry. Ultimately, the increment is slightly smaller than for women, in the range of 3-4 percentage points. Interestingly, the pattern of results does not vary substantially by whether states train a large share of their participants, nor are results substantially different for ES states. (These results are not presented.)

Differences in patterns for men and women may partly reflect the types of training they receive. A study of exits for program year 2005 found that of males exiting from the WIA Adult program, 37 percent received on-the-job training, in contrast to 15 percent for females (Social Policy Research Associates, 2007). Classroom training would be expected to reduce initial earnings and employment by more than on-the-job training and possibly provide greater earnings with a delay. There is some indication that selection into training on the basis of stable differences may affect results, since prior earnings and levels of employment are higher for both genders, but none of these differences is statistically significant.

Summary of WIA Adult Program Impacts. Taken at face value, the results reported above imply large and immediate impacts on earnings and employment for individuals who participate in the WIA Adult program. Those who obtained training services have lower initial earnings, but they catch up to others within ten quarters, ultimately registering larger total gains. Although there is evidence that estimates of effects in initial quarters following program entry may be biased, we do not believe a selection story can be constructed to explain away estimated effects for later quarters. In particular, growth in earnings for those receiving training would appear to reflect growth that has been widely observed in related programs.

We estimated impacts separately for various subgroups (and males and females within them), focusing on those that are overrepresented among WIA participants or who face special challenges or barriers to working in the labor market, to wit, nonwhites, Hispanics, those under 26 years of age, those 50 or older, and veterans (males only). For the most part, estimated effects for these subgroups were similar to those for all WIA participants: There is essentially no evidence of substantial differences in impact between these subgroups. On the other hand,

[29] The very high estimates in quarters 15 and 16 should be discounted given the large standard errors.

sampling error for many of these groups is quite large, so the statistical power of these tests is modest.

Figure 13
Adult Program Treatment Effect on Quarterly Earnings for Females, WIA Training versus Comparison Group

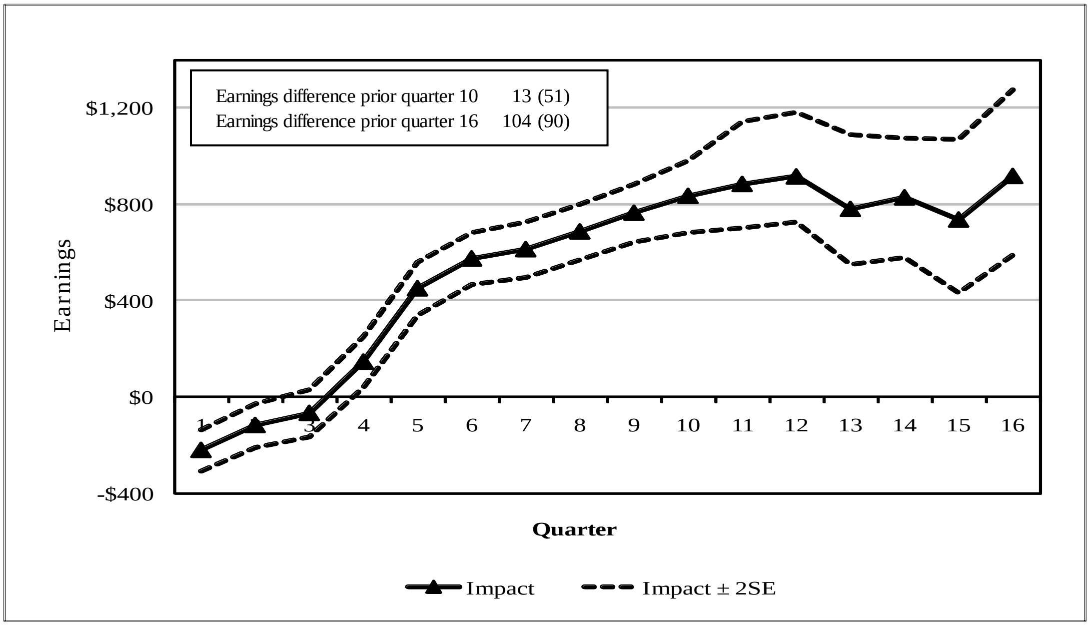


Figure 14
Adult Program Treatment Effect on Quarterly Earnings for Males, WIA Training versus Comparison Group

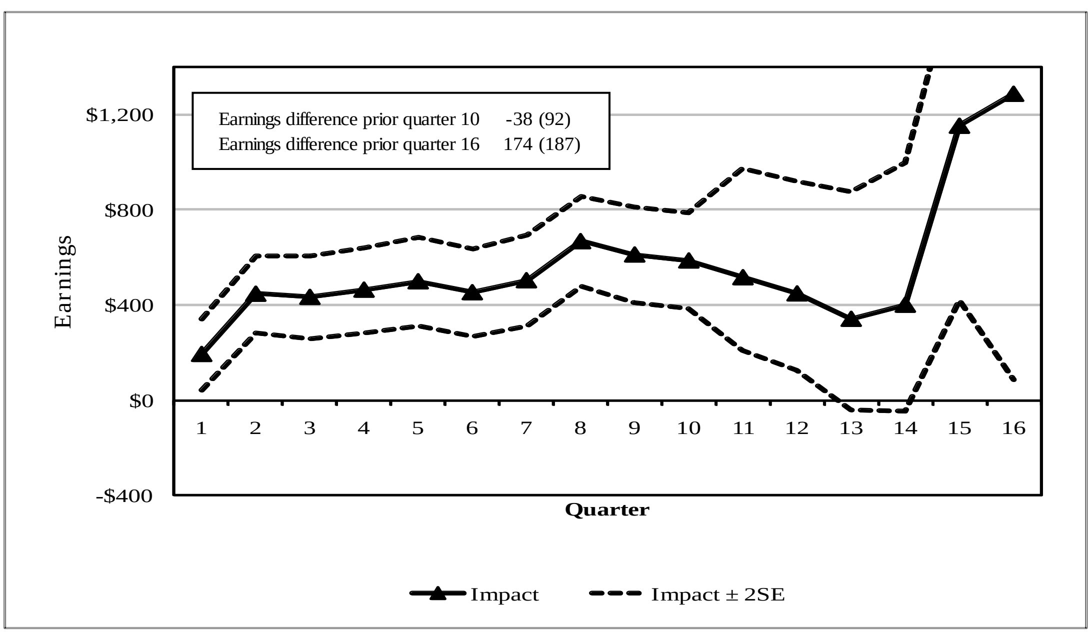


Figure 15
Adult Program Treatment Effect on Quarterly Employment for Females, WIA Training versus Comparison Group

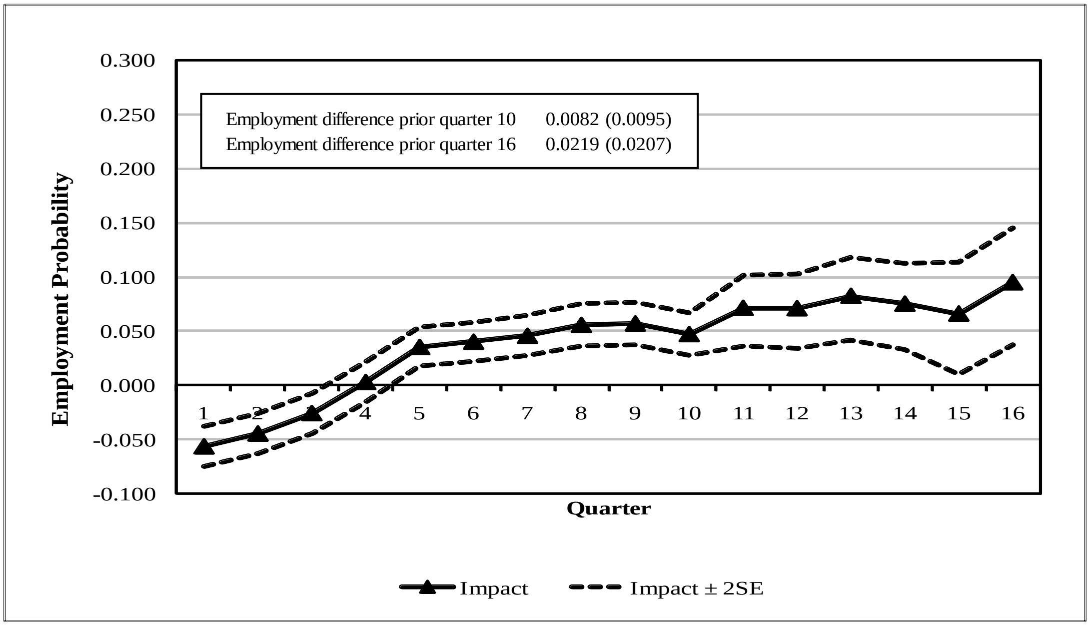
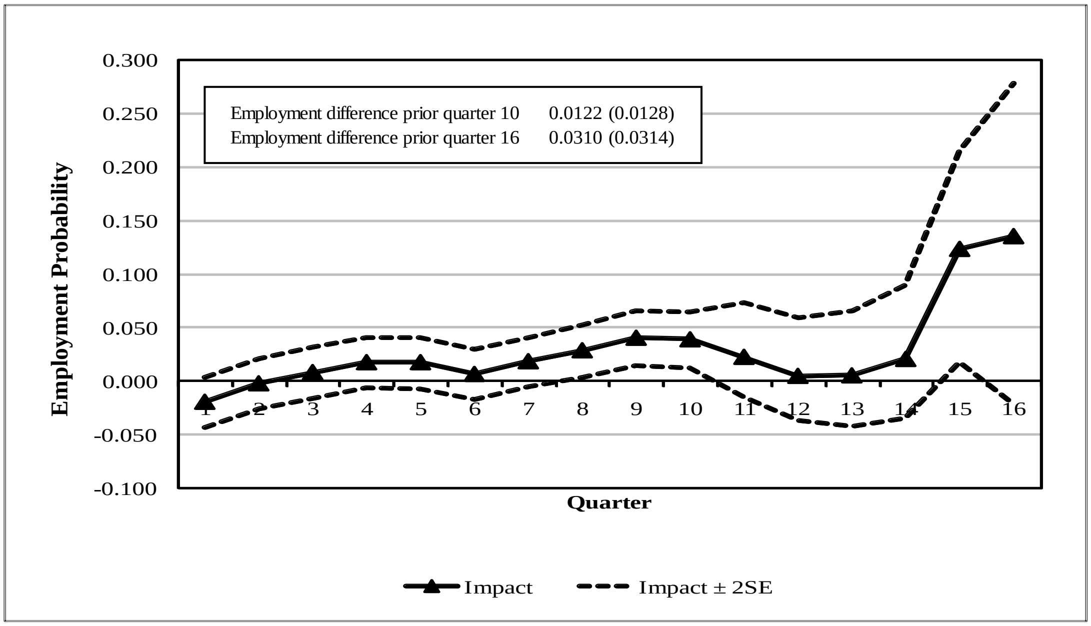


Figure 16
Adult Program Treatment Effect on Quarterly Employment for Males, WIA Training versus Core/Intensive



V. Results of Impact Estimation for Dislocated Worker Program

1. Dislocated Worker Program Impacts

Estimates of state-specific effects for participants in the Dislocated Worker program are provided in Table 4. Impact estimates in the initial quarters are much smaller than comparable estimates for the Adult program. Five states display impact estimates for at least one gender that are negative and statistically significant, implying that those who participate in the program experience lower earnings during the first five quarters after program entry as a result of their program participation. In only three states is the estimate for these quarters positive and statistically significant for at least one gender.

For quarters 11-16, estimates of impacts in most states are positive and statistically significant for at least one gender. Only two states have values that are negative and statistically significant for either men or women. Despite these apparently positive impact estimates, there are some indicators that these estimates should not be taken at face value. For the ten states where a specification test is possible, we generally find a positive impact on prior earnings, implying that individual participants in the Dislocated Worker program may be advantaged relative to the comparison group in ways not captured by control variables.[30] However, sampling error is too large to allow this possibility to be investigated in any detail for individual states.

Table 4
Dislocated Worker Program Treatment Effect on Quarterly Earnings by State: WIA versus Comparison Program

	Females		Males	
State	Quarters 1-5	Quarters 11-16	Quarters 1-5	Quarters 11-16
1	-348	143	-561*	130
2	-1813*	-345*	-1666*	-283*
3	142*	368*	211*	254*
4	150	693*	90	624
5	577*	914*	481	941*
6	59	670*	21	568*
7	-619*	286*	-832*	327
8	-1251*	38	-1657*	-517*
9	-191	780*	72	897*
10	-541*	215	783	776
11	888*	1292*	674*	1270*
12	-53	992*	-115	899*

*Statistically significant at 0.05 level. Average effects for specified quarters, calculated on available quarters. For 4 states, estimates are not available for quarters 11-16; reported estimate based on quarter 10.

[30] Due to data limitations, in five states we use earnings 10 quarters prior to program entry, and in five states we use earnings 16 quarters prior to entry. Among the 20 state-by-gender estimates, 17 are positive.

Figures 17 and 18 graph estimated program impacts for participants in all 12 states in the Dislocated Worker program. Participant earnings in the quarter following entry are $200-$300 below the comparison group, but relative earnings show an increasing trend over the 16 quarters of follow-up analysis. In the fifth or sixth quarter after program entry, participant earnings are equal to those of the comparison group. Ultimately earnings grow to exceed those of comparison group workers by up to $400 per quarter. Despite the similarity in basic pattern, male earnings peak at around 10 quarters, whereas female earnings appear to grow until the end of the four-year window.

Figures 19 and 20 show that, for women, initial employment is approximately 2 percentage points below the comparison group, and catches up within about three quarters, and ultimately, employment is nearly 8 percentage points above the comparison group. In contrast, for men, there is no initial employment difference, although the growth over time is smaller, with the positive increment after three years peaking at about 6 percentage points.

Results for various subsets of states (not presented) support these basic conclusions. In the three states for which ES participants are the comparison group, the pattern is almost exactly the same. The seven states offering high levels of training display a rather more extreme pattern. In the first three to four quarters following program entry, quarterly earnings are more than $800 below the comparison group. Relative earnings do increase, but they only equal the comparison group earnings after eight or nine quarters. Earnings exceed the comparison group earnings by about $200 three years (12 quarters) after program entry. Interestingly, participants in these seven states also experience a substantially reduced employment rate in the initial quarters after program entry, but, after 10 quarters, employment rates for both genders are about 5 percentage points above the comparison group.

It is important to recognize that Dislocated Worker program participants are usually relatively high-wage individuals who are faced with permanent job loss. Their initial negative impact estimates imply that their earnings are below unemployed workers with similar prior incomes and work histories. This is what would be expected if involvement in training activities precludes or reduces employment, inducing lock-in effects. Earnings growth observed over the three following years is consistent with the attainment of skills with training.

Such an interpretation is based on the assumption that Dislocated Workers are similar in unmeasured ways to the comparison group. Our specification test, based on predicting prior earnings, suggests this is not the case. There are substantial differences between the participant and comparison groups 16 quarters earlier, with participant earnings more than $200 higher, and standard errors imply that these estimates are statistically significant. Prior employment is also several percentage points higher for program participants. That matched participants have higher prior earnings suggests the possibility that their earnings in later periods may not reflect program impact but rather unmeasured factors that become apparent in the three years after program entry.

Figure 17
Dislocated Worker Program Treatment Effect on Quarterly Earnings for Females, WIA versus Comparison Group

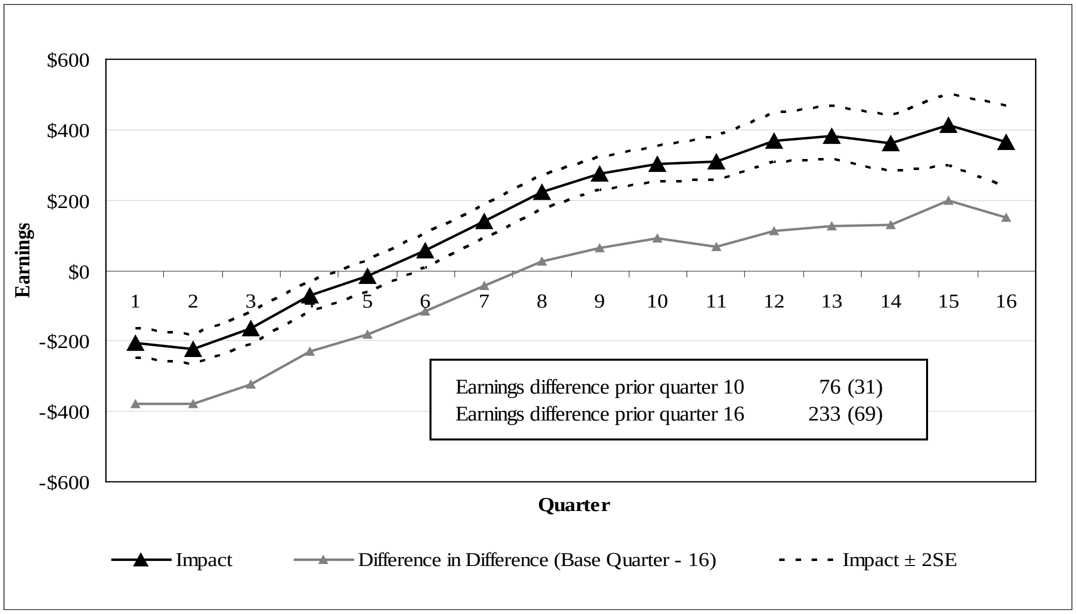


Figure 18
Dislocated Worker Program Treatment Effect on Quarterly Earnings for Males, WIA versus Comparison Group

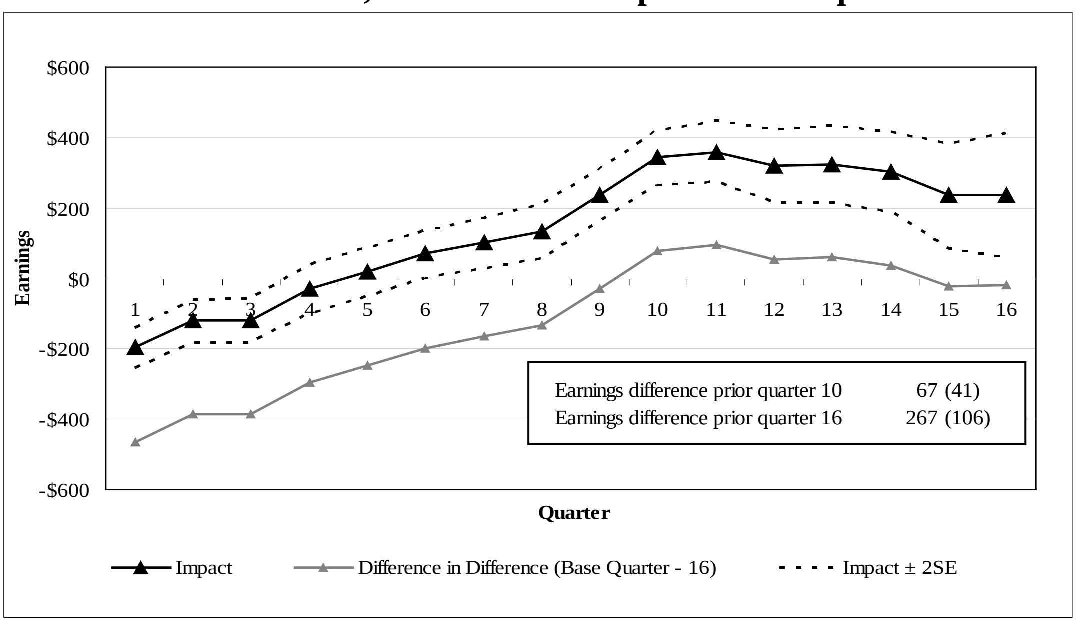


Figure 19
Dislocated Worker Program Treatment Effect on Quarterly Employment for Females, WIA versus Comparison Group

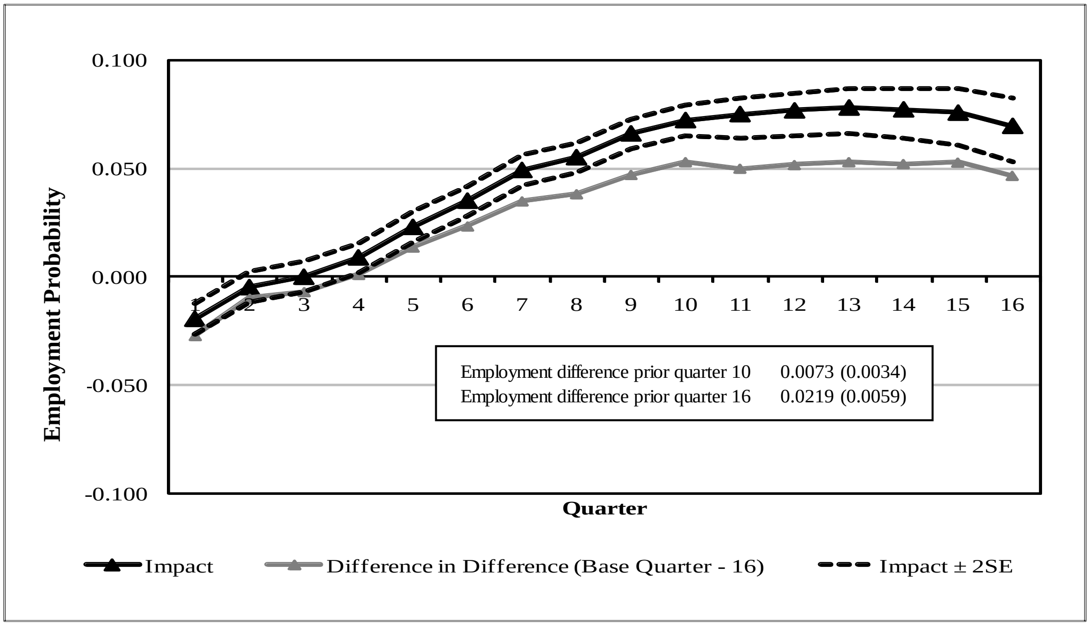
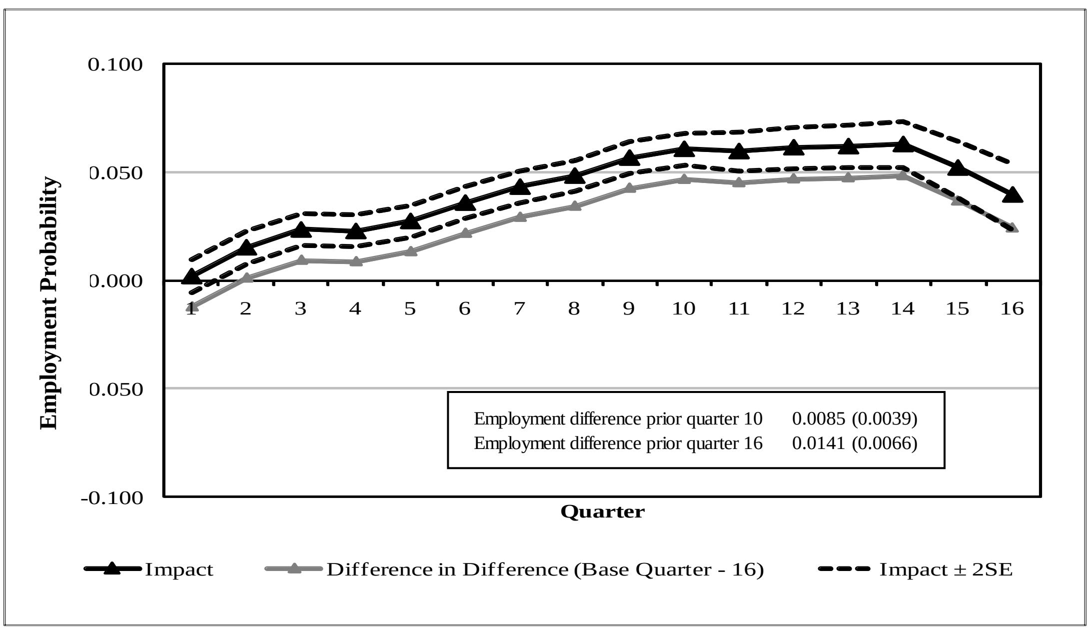


Figure 20
Dislocated Worker Program Treatment Effect on Quarterly Employment for Males, WIA versus Comparison Group



Although these findings contribute to greater uncertainty as to the true impact, some indication of the possible extent of the bias is provided by the difference-in-difference estimates, which subtract the prior quarter 16 increment. As discussed above, this estimator provides a valid estimate of program impact if selection into the program is on the basis of stable characteristics that are not captured by variables that have been controlled. The difference-in-difference estimates imply that participants' earnings catch up to those of nonparticipants with a longer delay and that the ultimate impact on earnings is more modest. For women, earnings exceed those of nonparticipants only after eight quarters, and the positive increment is never over $200. For men, the crossover point is after nine quarters, and the increment is generally less than $100.

Taken together, these results imply those who participate in the Dislocated Workers program have earnings below nonparticipants for an extended period. Among women, participants are also less likely to be employed. Participants' earnings and employment levels overtake those of nonparticipants in two to three year's time. Prior earnings differences between participants and the comparison group suggest that there may be selection on unmeasured stable factors, so that ultimate impact estimates may overstate program effects. On the other hand, it is worth noting that initial earnings differentials may well reflect selection on short-run differences. If individuals whose immediate employment opportunities are particularly poor are likely to enter the Dislocated Worker program, their lower initial earnings could at least partly reflect selection.

2. Dislocated Worker Impacts for UI Recipients

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 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. Given that the program is largely targeted at individuals who have lost jobs, this subsample is quite similar to others in the program.

The results of this analysis, presented in Appendix D, show that WIA participant earnings do not catch up until seven or eight quarters after program entry. The initial negative effect is in the range of $700 for both men and women, and the maximum positive impact is also lower, at about only $200 for each group. As in the estimates reported in the prior section, the specification tests implies that program participants have higher prior earnings than matched comparison group members, so even these modest positive impacts may be spurious.

3. Impacts of Training for Dislocated Workers

The incremental impact of training is based on a comparison of WIA Dislocated Worker participants who obtain training with those who do not. Figures 21 and 22 show that initial earnings for those obtaining training are below those of other program participants for eight quarters for women and for more than ten quarters for men. Differences are $1,100 for females in quarters 2 through 4, and $800 for males. A very similar pattern exists for employment, and thus, we do not present these results here.

Although the initial negative impact estimate is easily statistically significant, the confidence interval is large relative to estimated impacts after quarter 10. Estimates for both males and females could easily differ from the point estimate by $200 due to sampling error. Also of possible concern is the difference in earnings prior to entry into the program. For females, the individuals who select into training have lower earnings relative to other WIA participants in the sixteenth quarter prior to participation, suggesting the estimates of effects could be downwardly biased. This difference is not, however, statistically significant, so evidence of selection is inconclusive. Estimates for states offering high proportions of training are not substantively different.

Figure 21
Dislocated Worker Program Treatment Effect on Quarterly Earnings for Females, WIA Training versus Core/Intensive

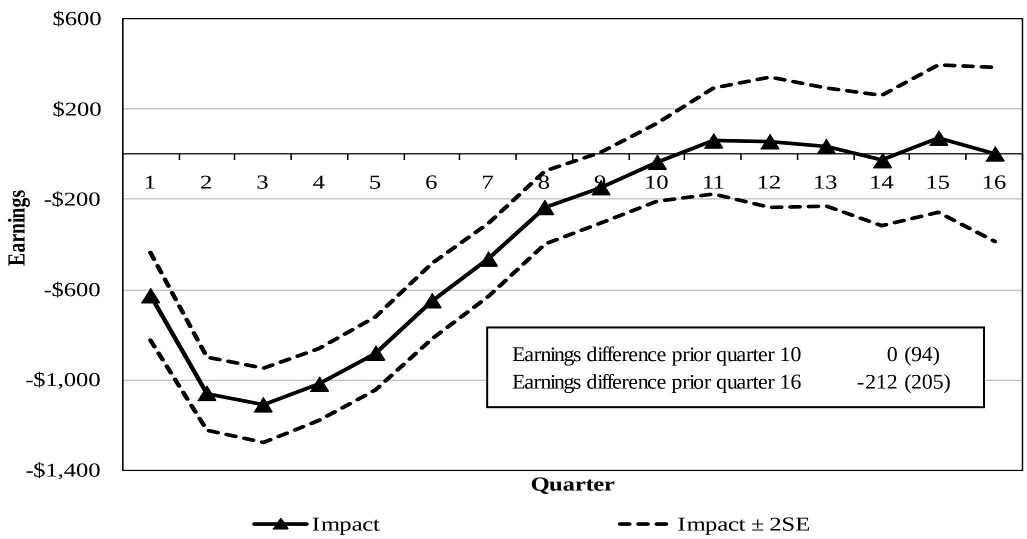
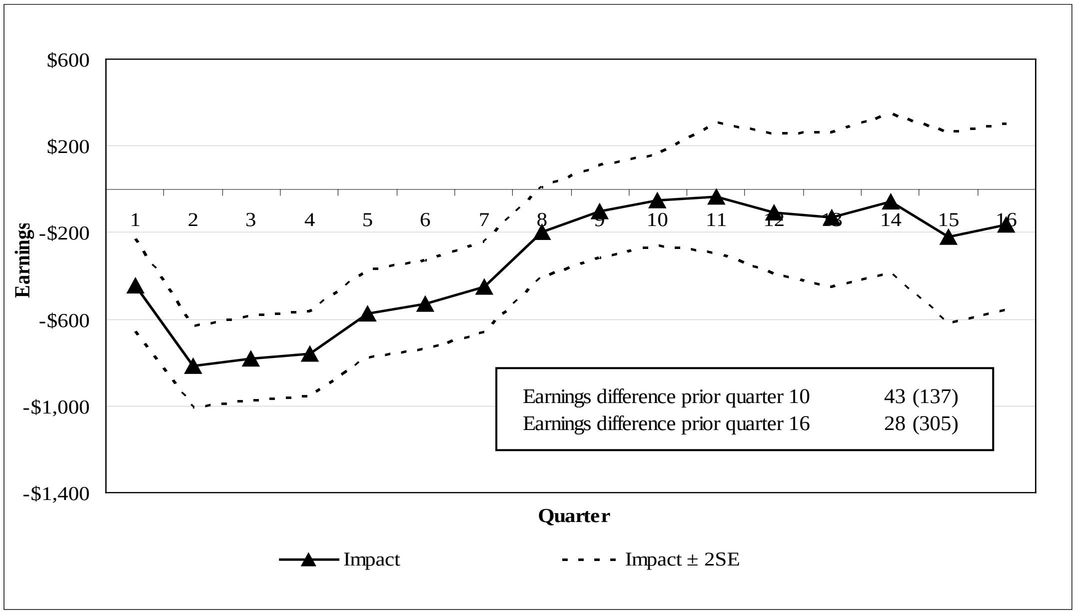


Figure 22
Dislocated Worker Program Treatment Effect on Quarterly Earnings for Males, WIA Training versus Core/Intensive



Taken at face value, point estimates suggest that WIA Dislocated Worker program participants who enter training experience large earnings losses relative to others in their first two years after program entry. Although consistent with a large training lock-in effect, these effects could be at least partly due to selection on short-term employment prospects, with those who land jobs leaving the program without obtaining training. 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 little evidence that training produces substantial benefits. These negative conclusions must be tempered, however, by a recognition that sampling error alone could obscure substantial impacts.

Summary of WIA Dislocated Worker Program Impacts. Dislocated Workers are likely to face serious difficulties in obtaining reemployment, and the kinds of services WIA offers may require time to produce impacts. The pattern of results is consistent with these expectations. However, the extent of any benefits that accrue from participation is particularly hard to judge. Some specification tests suggest that results may be biased toward finding positive program impacts. Difference-in-difference estimates are therefore appreciably smaller than the primary reported estimates. These estimates imply that program participants' earnings do not reach the level of earnings of comparable nonparticipants until more than two years after participation. Perhaps more important, the growth in earnings, relative to nonparticipants, slows at that point. As a result, these estimates imply that the gains from participation are, at best, very modest, even three to four years after entry. Overall, it appears possible that ultimate gains from participation are small or nonexistent. Insofar as there are impacts, females are more likely to benefit than are males.

Where employment is taken as the outcome of interest, estimates of program impact are more supportive of the program. Although the specification tests again suggest that there are unmeasured differences between the treated and matched comparison group, the difference-in-difference estimates of program suggest at least a moderate positive impact.

We estimated impacts separately for nonwhites, Hispanics, individuals under 26 years of age, those 50 or older, and male veterans, finding no evidence that there are any important differences in program impacts for any of these subgroups. As in the case of subgroup analysis for the Adult program, sampling error is substantial, and there may be differences that are not statistically discernable.

VI. Conclusions and Implications

The estimates of WIA program impact presented here are based on administrative data from 12 states, covering approximately 160,000 WIA participant and nearly 3 million comparison group members. Given variation in the way that states have implemented WIA, it was anticipated that estimates of WIA program impact would differ across states. The findings of the analysis confirmed notable differences across states in estimated program impacts and patterns of impacts. Such differences likely reflect local economic environment, program structure, labor force composition, and possibly complex interactions of these factors. Some of the observed differences may be due to various statistical artifacts, reflecting selection effects occurring both in the WIA programs and the comparison programs. Yet many differences are undoubtedly real and identify differences in program efficacy.

There are also similarities in the patterns of estimated impacts across states. Women appear to obtain greater benefits for participation in both the Adult and Dislocated Worker programs, with the quarterly earnings increment exceeding that for males. The value of training appears to be greater as well, especially over the long run. Adult program participants who obtain training services have lower initial earnings, but they catch up to others within ten quarters, ultimately registering total gains of over $800 for females and $500-$600 for males. These results are consistent with findings of prior studies, including random assignment experiments (Orr et al., 1996), and generally suggest that it is highly likely that WIA participants gain from their participation.

The Adult program is more likely to produce tangible benefits for participants than the Dislocated Worker program. Dislocated Workers experience several quarters following entry into WIA with earnings that are below those for the matched comparison group. Although our estimates suggest their earnings ultimately equal or even overtake the comparison group, our statistical tests suggest that the actual earnings increment from participation may be much smaller or even nonexistent. We also find that the benefit of obtaining training for this group is probably small or nonexistent.

In conclusion, overall WIA program net impacts were estimated to be positive in almost all states, although important variation across programs and specific services clearly exists. While the possibility that these estimates may be partly spurious cannot be ruled out, none of the selection explanations considered would fully explain away these patterns of positive estimated program effects.

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

There are important policy implications of these results that go beyond a simple judgment of whether the program is effective. Program administrators typically look at the cross-sectional or "point-in-time" information that is available to them from performance management systems on a regular basis. They do not have at hand the data analysis tools to examine individual employment and earnings histories and trajectories for more than eight years (33 quarters that include up to 16 quarters of follow-up data) for both program participants and a comparison group, as in this study. The results of this evaluation show that program impacts typically "mature" over time, sometimes increasing in magnitude and sometimes diminishing. Insofar as this work underscores the fact that long-term impacts are of significance and that outcomes of interest may not be apparent for years, it will help to refocus training activities.

Although the WIA participants in our analysis are not a representative sample, there are several reasons why these results are likely to apply to the WIA program more generally. The sample includes about one in five WIA participants during the period of study and the data contain a diverse sample of states. It is reasonable to conclude that the processes by which WIA operates in the 12 states included in the study are similar to those in other states. The results of this study therefore provide information on the long-run impact of the WIA program of potential value in allocating WIA program resources both at the local and national level.

References

Abadie, Alberto and Guido W. Imbens. 2006a. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* 74:1 (January), 235-267.

Abadie, Alberto, and Guido W. Imbens. 2006b. "On the Failure of the Bootstrap for Matching Estimators." Unpublished paper, John F. Kennedy School of Government, Harvard University ().

Appelbaum, Eileen, Bernhardt, Annette and Richard J. Murnane. 2003. *Low-Wage America: How Employers Are Reshaping Opportunity in the Workplace*. New York: Russell Sage.

Ashenfelter, Orley C. 1978. "Estimating the Effect of Training Programs on Earnings." *Review of Economics and Statistics* 60 (February), 47-57.

Barnow, Burt and Chris King. 2004. "The Changing Workforce Development Landscape: Report on the Operation of the Workforce Investment Act." U.S. Department of Labor, December.

Black, Dan A., Jeffrey A. Smith, Mark C. Berger, and Brett J. Noel. 2003. "Is the Threat of Reemployment Services More Effective Than the Services Themselves? Evidence from Random Assignment in the UI System." *American Economic Review* 93(4), 1313–27.

Bloom, Howard S., Charles Michaelopoulos and Carolyn J. Hill. 2005. "Using Experiments to Assess Nonexperimental Comparison-Groups Methods for Measuring Program Effects," in Howard S. Bloom (Ed.), *Learning More from Social Experiments: Evolving Analytic Approaches*, New York: Russell Sage, 173-235.

Bradbury, Katharine and Jane Katz. 2002. "Are Lifetime Incomes Growing More Unequal? Looking at New Evidence on Family Income Mobility." *Regional Review* (Federal Reserve Bank of Boston) 12(4), 3-5.

Caliendo, Marco, and S. Kopeinig. 2008. "Some Practical Guidance for the Implementation of Propensity Score Matching." *Journal of Economic Surveys* 22:1, 21-72.

Card, David, Jochen Kluve, and Andrea Weber. 2009. "Active Labor Market Policy Evaluaitons: A Meta-Analysis." IZA Discussion Paper No. 4002, February.

Cook Thomas D., William R. Shadish and Vivian C. Wong. 2008. "Three Conditions Under Which Experiments and Observational Studies Often Produce Comparable Causal Estimates: New Finidngs from Within-Study Comparisons." *Journal of Policy Analysis and Management* 27 (Autumn), 724-750.

Dehejia, Rajeev H., and Sadek Wahba. 2002. "Propensity Score-Matching Methods for Nonexperimental Causal Studies." *Review of Economics and Statistics* 84:1 (February), 151-161.

Dyke, Andrew, Carolyn Heinrich, Peter R. Mueser, Kenneth R. Troske, and Kyung-Seong Jeon. 2006. “The Effects of Welfare-to-Work Program Activities on Labor Market Outcomes.” *Journal of Labor Economics* 24:3 (July), 567-608.

Frank, Abby and Elisa Minoff. 2005. “Declining Share of Adults Receiving Training Under WIA Are Low-Income or Disadvantaged.” Washington, DC: Center for Law and Social Policy.

Fredriksson, Peter and Per Johansson. 2008. “Dynamic Treatment Assignment: The Consequences for Evaluations Using Observational Data.” *Journal of Business and Economic Statistics* 26:4, 435-445.

Galdo, Jose, Jeffrey Smith, and Dan Black. 2008. “Bandwidth Selection and the Estimation of Treatment Effects with Unbalanced Data.” Unpublished, May.

Glazerman, Steven, Dan M. Levy and David Myers. 2003. “Nonexperimental versus Experimental Estimates of Earnings Impacts.” *Annals of the American Academy of Political and Social Science* 589 (September), 63-93.

Heckman, James J., Robert J. LaLonde, and Jeffrey A. Smith. 1999. “The Economics and Econometrics of Active Labor Market Programs,” in Orley Ashenfelter and David Card (Eds.) *Handbook of Labor Economics*, Vol. 3. Amsterdam: North Holland.

Heckman, James J., and Jeffrey A. Smith. 1999. “The Pre-programme Earnings Dip and the Determinants of Participation in a Social Programme: Implications for Simple Programme Evaluation Strategies.” *Economic Journal* 109 (July), 313-348.

Heckman, James J. and Petra E. Todd. 2009. “A Note on Adapting Propensity Score Matching and Selection Models to Choice Based Samples.” IZA Discussion Paper No. 4304, July.

Heinrich, Carolyn J. 2007. “Demand and Supply-Side Determinants of Conditional Cash Transfer Program Effectiveness.” *World Development* 35 (January), 121-143.

Heinrich, Carolyn J., Peter R. Mueser and Kenneth R. Troske. 2008. “Workforce Investment Act Non-Experimental Net Impact Evaluation.” Final Report, Employment and Training Administration, Department of Labor, ETAOP 2009-10, December. Available at: wdr.doleta.gov/research/keyword.cfm?fuseaction=dsp_resultDetails&pub_id=2419&mp=y

Holzer, Harry. 2004. “Encouraging Job Advancement among Low-Wage Workers: A New Approach.” Policy Brief #30. Washington, DC: The Brookings Institution.

Hollenbeck, Kevin Daniel Schroeder, Christopher King and Wei-Jan Huang. 2005. “Net Impact Estimates for Services Provided through the Workforce Investment Act.” Employment and Training Administration Occasional Paper (ETAOP 2005-06), October.

Hotz, V. Joseph, Guido W. Imbens and Jacob A. Klerman. 2006. “Evaluating the Differential

Effects of Alternative Welfare-to-Work Training Components: A Reanalysis of the California GAIN Program." *Journal of Labor Economics* 24 (July), 521-566.

Imbens, Guido W. 2004. "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review." *Review of Economics and Statistics* 86:1 (February), 4-29.

Imbens, Guido W. 2008. "Estimating Variances for Estimators of Average Treatment Effects." Unpublished, Harvard University (September).

Imbens, Guido W., and Jeffrey M. Wooldridge. 2008. "Recent Developments in the Econometrics of Program Evaluation." Institute for Research on Poverty Discussion Paper no. 1340-08, University of Wisconsin.

Kluve, Jochen. 2007. "The Effectiveness of European ALMP's." In Jochen Kluve et al., *Active Labor Market Policies in Europe: Performance and Perspectives*. Berlin and Heidelberg: Springer, 153-203.

Lechner, Michel. 2001. "Identification and Estimation of Causal Effects of Multiple Treatments Under the Conditional Independence Assumption." In *Econometric Evaluation of Active Labour Market Policies*, ed. M. Lechner and F. Pfeiffer (Heidelberg: Physica,), 43–58.

Lechner, Michael and Conny Wunsch. 2006. "Are Training Programs More Effective When Unemployment Is High?" IZA Discussion Paper No. 2355, October.

Mueser, Peter R., Kenneth R Troske, and Alexey Gorislavsky. 2007. "Using State Administrative Data to Measure Program Performance." *Review of Economic and Statistics* 89: 4 (November), 761-783.

O'Leary, Christopher J. 2004. "Evaluating the Effectiveness of Labor Exchange Services." Chapter 5 in David E. Balducchi, Randall W. Eberts, and Christopher J. O'Leary (Eds.), *Labor Exchange Policy in the United States* (Kalamazoo, Michigan: Upjohn Institute).

Orr, Larry L., Howard S. Bloom, Stephen H. Bell, Fred Doolittle, Winston Lin and George Cave. 1996. *Does Training for the Disadvantaged Work? Evidence from the National JTPA Study*. Washington D.C.: The Urban Institute Press.

Osterman, Paul. 2007. "Employment and Training Policies: New Directions." In *Reshaping the American Workforce in a Changing Economy*, Harry J. Holzer and Demetra Smith Nightingale (eds.), Washington, DC: Urban Institute Press, 119-154.

Peikes, Deborah N., Lorenzo Moreno and Sean Michael Orzol. 2008. "Propensity Score Matching: A Note of Caution for Evaluators of Social Programs." *The American Statistician* 62 (August), 222-231.

Rockefeller Institute of Government. 2004. *The Workforce Investment Act in Eight States: State*

Case Studies from a Network Evaluation (2 volumes), Employment and Training Administration, Department of Labor (ETAOP 2004-02 and ETAOP 2004-03), February.

Rosenbaum, Paul R., and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika* 70 (), 41-55.

Rosenbaum, Paul R. 2002. *Observational Studies.* New York: Springer-Verlag.

Rubin, Donald B. 2006. *Matched Sampling for Causal Effects*. Cambridge: Cambridge University Press.

Sianesi, Barbara. 2004. "An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s." *Review of Economics and Statistics*, 86:1, 133-155.

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, June.

Social Policy Research Associates. 2006. *2004 WIASRD Data Book.* Report prepared for the U.S. Department of Labor, February.

Social Policy Research Associates. 2007. *PY 2005 WIASRD Data Book: Final*. Prepared for the U.S. Department of Labor, August.

Smith, Jeffrey A. and Petra E. Todd. 2005a. "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics* 125 (March-April), 305-53.

Smith, Jeffrey A. and Petra E. Todd. 2005b. "Rejoinder." *Journal of Econometrics* 125 (March-April), 365-75.

Smith, Jeffrey, Alex Whalley and Nathanial Wilcox. 2007. "Are Program Participants Good Evaluators." Unpublished, University of Michigan.

U.S. Government Accounting Office. (2002) Improvements Needed in Performance Measures to Provide a More Accurate Picture of WIA's Effectiveness, GAO Report #02-275.

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?#wiastann (accessed August 2009).

Zhao, Zhong. 2003. "Data Issue of Using Matching Methods to Estimate Treatment Effects: An Illustration with NSW Data Set," Peking University Working Paper No. #203004.

Appendix A: Comparison Groups

As indicated in section II.2, we use comparison groups drawn from either Unemployment Insurance (UI) claimants or from U.S. Employment Service (ES) participants in the impact evaluation. There is substantial but not complete overlap between the UI claimant population and those receiving ES services. In most states, the majority of UI recipients are required to register for ES services, but some claimants do not face this requirement. Conversely, although a large share of ES recipients are or have recently received UI benefits, anyone seeking services to aid in job search is eligible to receive ES services. Nationwide, about a third of individuals seeking ES services are UI recipients who are required to register with ES (O'Leary, 2004).

Generally, the level of services participants receive in both programs is relatively low, and one may view such individuals as representing a "no treatment" control. Alternatively, given that ES and related services are widely available, they may be viewed as representing a highly relevant "counterfactual" that reflects program options faced by individuals in the absence of WIA. Voluntary participants in ES may well receive benefits that are large compared to the costs, and benefit-cost studies have found that these services are cost effective. Studies show that requiring UI recipients to participate in job search or related services increases the likelihood of employment, although the activities themselves are probably less important than the incentives to obtain employment created by these requirements (Black et al., 2003; O'Leary, 2004).

One important shortcoming of UI recipients as a comparison group is that recipients must have prior earnings above a minimum in order to receive UI benefits. As a result, it may be difficult to find appropriate UI recipient matches for some WIA participants. For those states where it is available, those who apply for UI benefits but who do not receive them are included in the comparison sample, allowing for the possibility that some rejected applicants can serve as matches for WIA participants with weak employment histories.

Although a WIA entry is identified as a treated case even when the individual subsequently reenters WIA, any potential comparison case where the individual subsequently enters WIA is omitted. This approach reduces contamination bias, and the comparison cases can be viewed as analogous to controls in a random assignment study that are precluded from participating in the program. The danger of this approach, however, is that, if future participation is an indication of employment difficulties, omitting participants from the comparison group will improve average employment outcomes for the group, inducing a negative bias in impact estimates. In European systems where program participation is very likely for those who fail to obtain employment, this bias may be severe. Sianesi (2004) argues for including those who receive treatment at a later point in the comparison group but reinterpreting the effect estimate as an indicator of participating at a given point in time (not a net impact estimate). In an analysis that examines program impact on entry into employment, Fredriksson and Johansson (2008) show how to correct for subsequent program participation in the comparison group when return to work and program entry can be treated as competing risks. Unfortunately, their approach is not feasible in the current analysis. However, the proportion of comparison group individuals who are eliminated due to omitting subsequent WIA participants is only in the order of 2-4 percent, so the likely impact on estimates is expected to be small.

Appendix B: Summary Statistics for State Programs

Below, Table B.1 presents summary statistics for the seven state programs providing training services to at least 60 percent of participants and Table B.2 provides statistics for the balance of the state programs. The two sets of programs differ in terms of average size. The seven high-training programs average approximately 3,200 Adult participants and 2,400 Dislocated Workers, for a total of approximately 38,000 participants in the seven programs. Approximately two-thirds of all recipients in this group receive training. The comparison group comprises 1.6 million individuals who contribute 3.5 million records available for matching.

Despite differences in WIA program structure, overall patterns for the WIA and the comparison program are similar in Table B.1 and Table B.2. In the set of low-training states, nearly 120,000 WIA participants are identified, and, in aggregate, just 17 percent of WIA participants receive training services. Over 1.3 million comparison program participants contribute 2.7 million matching units. Even when there are differences in levels reflecting variation across state populations, patterns in Tables B.1 and B.2 correspond closely. For example, the proportion that is black is approximately half as great in the set of high-training states, but this difference exists for both WIA and comparison group participants.

Table B.1
Summary Statistics for WIA Participants and Comparison Group in 7 States with High Training Rates

	WIA Adult			WIA Dislocated Worker			
	Overall	No Training	Training	Overall	No Training	Training	Comparison Group
Sample size							
Unique individuals	22,646	7,114	15,532	16,520	5,027	11,493	1,601,399
WIA entries, or quarters of comparison program participation	22,694	7,133	15,561	16,536	5,036	11,500	3,479,550
Demographic	Mean	Mean	Mean	Mean	Mean	Mean	Mean
Male	0.381	0.421	0.363	0.471	0.510	0.453	0.580
Black	0.190	0.256	0.160	0.095	0.124	0.083	0.132
Hispanic	0.107	0.089	0.115	0.068	0.073	0.067	0.090
Age	33.46	35.77	32.40	41.73	42.99	41.17	39.67
Years of education	12.32	12.23	12.36	12.60	12.69	12.57	12.45
Employment							
Employment-employment	0.280	0.269	0.286	0.475	0.494	0.467	0.497
Employment-not employed	0.240	0.214	0.252	0.347	0.300	0.367	0.253
Not employed-employed	0.309	0.361	0.285	0.121	0.150	0.108	0.215
Not employed-not employed	0.161	0.140	0.170	0.042	0.037	0.044	0.031
Earnings second year prior	9,526	10,303	9,218	26,156	27,732	25,462	20,701
Earnings in prior year	8,352	8,626	8,226	24,618	26,062	23,985	21,435
Earnings following year	10,579	11,603	10,109	11,906	16,321	9,973	15,693
Earnings second year after	12,903	12,452	13,109	16,428	19,742	14,976	17,092
Program Experience							
WIA in prior two years	0.016	0.017	0.016	0.007	0.010	0.006	0.014
Comparison program participation in prior two years	0.325	0.307	0.332	0.667	0.643	0.676	0.706

Table B.2
Summary Statistics for WIA and Comparison Program Participants in 5 States with Low Training Rates

	WIA Adult			WIA Dislocated Worker			
	Overall	No Training	Training	Overall	No Training	Training	Comparison Group
Sample size							
Unique individuals	72,934	61,141	11,793	46,995	38,486	8,509	1,328,097
WIA entries, or quarters of comparison program participation	74,858	62,579	12,279	47,553	38,858	8,695	2,681,960
Demographic	Mean	Mean	Mean	Mean	Mean	Mean	Mean
Male	0.431	0.448	0.346	0.486	0.492	0.460	0.591
Black	0.522	0.541	0.426	0.412	0.426	0.350	0.221
Hispanic	0.008	0.006	0.018	0.006	0.005	0.013	0.030
Age	32.46	32.58	31.87	39.73	39.77	39.52	39.48
Years of education	12.26	12.21	12.50	12.53	12.50	12.71	12.38
Employment							
Employment-employment	0.302	0.296	0.334	0.458	0.462	0.442	0.449
Employment-not employed	0.198	0.192	0.226	0.258	0.250	0.291	0.313
Not employed-employed	0.329	0.333	0.313	0.205	0.205	0.202	0.238
Not employed-not employed	0.170	0.179	0.127	0.080	0.083	0.064	0.051
Earnings second year prior	8,232	8,011	9,414	17,783	16,945	21,609	19,269
Earnings in prior year	8,088	7,985	8,614	19,066	18,593	21,183	21,777
Earnings following year	9,076	8,846	10,248	11,395	11,260	11,998	15,592
Earnings second year after	10,223	9,627	13,258	13,926	13,496	15,848	17,114
Program Experience							
WIA in prior two years	0.061	0.062	0.057	0.049	0.047	0.058	0.027
Comparison program participation in prior two years	0.181	0.166	0.258	0.351	0.330	0.441	0.628

As noted above, in three of the states the comparison program is U.S. Employment Service (ES) participants rather than Unemployment Insurance (UI) claimants. Table B.3 presents tabulations for states where the ES program is the comparison group. Whereas in Table 2 (in the text) prior earnings for the comparison group (three-quarters of them UI claimants) were two to three times the level of earnings for Adult program participants, ES earnings are only 50 percent higher. Furthermore, whereas ES earnings are much lower than Dislocated Workers earnings, UI claimants' earnings are equal or greater than Dislocated Worker earnings. Overall, it is clear that the ES serves a population that is closer to that of Adult workers than does the UI program. In the three ES states, 17 percent of WIA Adult program participants had no employment during the six quarters up to participation, whereas the number is 14 percent for the comparison group of ES participants. In the full sample, the Adult figure is similar to that reported in Table B.3, but for the comparison group, the number is only 4 percent.

It is unclear whether differences between the UI and ES comparison samples are important for this analysis. Both samples are very large, and, given that detailed earnings and employment information is available in both, good matches for most WIA participants are available in either sample.

Table B.3
Summary Statistics for WIA and ES: Program Participants in 3 States

	WIA Adult			WIA Dislocated Worker			
	Overall	No Training	Training	Overall	No Training	Training	Comparison Group
Sample size							
Unique individuals	14,715	7,394	7,321	11,288	5,741	5,547	884,894
WIA entries, or quarters of comparison program participation	15,582	7,984	7,598	11,432	5,834	5,598	1,561,121
Demographic	Mean	Mean	Mean	Mean	Mean	Mean	Mean
Male	0.366	0.389	0.343	0.432	0.423	0.442	0.543
Black	0.405	0.551	0.251	0.232	0.285	0.177	0.244
Hispanic	0.033	0.028	0.038	0.033	0.029	0.036	0.040
Age	34.02	36.00	31.93	42.59	44.16	40.95	36.90
Years of education	12.31	12.13	12.50	12.79	12.91	12.67	12.20
Employment							
Employment-employment	0.242	0.221	0.265	0.421	0.464	0.376	0.333
Employment-not employed	0.272	0.255	0.291	0.400	0.354	0.448	0.321
Not employed-employed	0.320	0.374	0.263	0.134	0.144	0.124	0.298
Not employed-not employed	0.166	0.150	0.182	0.044	0.038	0.051	0.136
Earnings second year prior	8,919	9,260	8,564	28,632	30,544	26,561	14,406
Earnings in prior year	7,332	7,294	7,371	26,178	28,068	24,207	13,408
Earnings following year	8,713	9,763	7,609	12,966	16,111	9,689	10,343
Earnings second year after	11,201	11,185	11,218	17,118	19,607	14,525	11,766
Program Experience							
WIA in prior two years	0.030	0.035	0.024	0.017	0.022	0.011	0.008
Comparison program participation in prior two years	0.486	0.480	0.493	0.559	0.531	0.587	0.604

Appendix C: Matching Method Details and Diagnostics

Matching basics. A very small number of individuals enter WIA more than once, and each WIA entry is treated as a separate case in our implementation of matching. Measures of overall program impact therefore identify the incremental impact of program entry. In practice, the number of multiple entry individuals is so small that results are insensitive to the way they are treated.

The propensity score $P(X)$ is estimated using a logit specification with a highly flexible functional form allowing for nonlinear effects and interactions. It is necessary to test to insure that the estimated propensity score is successful in balancing values of matched treatment and comparison cases. Following the matching, tests for statistically significant differences between variable means for the treated cases and the weighted comparison sample are performed to insure that the score in fact balances the independent variables (see Smith and Todd, 2005b).

As noted in the text, the radius matching method is applied to the log odds of the propensity score. The vast majority of studies using propensity score matching measure the proximity of cases as the absolute difference in the propensity score. Matching on the log odds of the propensity score has the advantage that results are invariant to choice-based sampling, where the treated and comparison cases are sampled at different rates, as is the case here. In addition, since the logit is used to predict propensity score, the log odds are a linear combination of the independent variables, and a constant radius in the log odds translates into the same metric at different propensity score levels (see Smith and Todd, 2005b, and Heckman and Todd, 2009). Our comparison samples are very large, often more than 50 times greater than the sample of treated cases, and so propensity scores for most cases—both treatment and comparison—are small, generally less than 0.05. Matching on the log odds of the propensity score has the advantage that it "spreads out" the density of very low propensity scores. Preliminary experiments showed that matching was much more successful with the log odds.

Although propensity score matching is designed to reduce differences in characteristics between treated and matched comparison cases, for certain variables that are particularly important in determining the outcome, it may be prudent to use exact matching. As noted above, analyses are undertaken separately by gender, so that a male is never matched with a female. Labor market opportunities and other experiences may also be influenced in a direct way by seasonal and other time factors, and so most analyses employ exact matching by calendar quarter, assuring that impact estimates are based on a comparison of individuals during the same time period. In some cases, the sample is too small for this approach, and different quarters are combined, but in each case, dummies for quarters are included as matching variables.

Matched samples. A comparison case is defined as a quarter in which an individual had contact with the comparison program, either filing a claim or receiving UI benefits (where UI claimants make up the comparison pool) or receiving some job search service (where ES participants make up the comparison pool). A particular individual in the comparison sample therefore contributes a case for every quarter of participation. Because details of program participation in prior quarters are controlled, each quarter contributed by a given comparison individual differs in terms of the attributes used for matching. For example, if an individual receives UI benefits in two consecutive quarters but has no prior UI experience, the case corresponding to the first quarter may match with WIA entries occurring in that quarter that do not have prior UI experience. The comparison case corresponding to the second quarter of UI experience may match WIA entries occurring in that second quarter where that WIA participant also received UI benefits in the prior quarter. Hence, a given comparison individual will offer multiple potential matches, reflecting differences in the flow of experience over time.

In comparison 1 (see Table 1 in text), by definition treated and comparison cases have different experiences during the quarter of participation. Both are likely to be experiencing similar employment difficulties, but it is not possible to match them according to the specifics of their program participation in that quarter. However, it is clear that matching by the details of prior experience may be of substantial value. Comparison 2 assures that both treated and comparison cases receive UI benefits, but it requires that we limit the analysis to a subset of participants. For comparison 3, since the estimate focuses on the impact of training, both treatment and control cases are WIA participants. The approach therefore controls not only for the prior experience in the alternative program but also for any experience in the current quarter.

Matching variables. Control variables include calendar quarter of program entry (exact match in most cases), gender (exact match), age, years of educational attainment, race/ethnicity (separate categories for nonwhites and Hispanics), disability status, veteran status (for males), local labor market (local WIA area or other county-based measure), employment information based on wage record data over the two years prior to program entry, including employment transitions and earnings, industry of most recent employment, program participation history up to four years prior to WIA entry (WIA; UI or ES; and TANF in one state), and time since layoff. Veteran status is not available in five of our sample states, and industry is not available in three states. Prior program participation is available for up to four years in five states, for five to twelve quarters in eight states, and is available for at least two quarters for all but a handful of cases in all states. Disability (following the definition stated in the Americans with Disabilities Act of 1990) and time since layoff are only used in comparison 3, since comparable variables were not available for WIA and comparison groups. Further details of variable coding can be found in Heinrich, Mueser and Troske (2008).

We fit a logit model using all the control variables to produce the propensity score, which is then used to match WIA participants to the comparison group. If the propensity score fully and properly identifies the probability that an individual is a treated case, then matching on this score will yield exactly matching distributions—at least in terms of expectation—on all independent variables. In order to ensure that WIA entrants match the comparison group members as closely as possible in all relevant ways, a highly flexible parameterization that includes nearly 100 independent variables (including dummy variables) is used.

Critical to these methods is that controls for the WIA and comparison samples be coded in a fully comparable fashion. In the case of demographic and geographic measures, policy differences in program data gathering and coding procedures were of concern. Every effort was made to ensure comparability between treatment and comparison group variable coding and to restructure variables when necessary. Since individual employment information was based on UI wage records that were matched with both WIA and comparison cases, it is clear that they are measured comparably. Similarly, information on prior experience in the comparison program was available in a symmetrical fashion for WIA and comparison group cases.

Radius Choice. The choice of radius involves a trade-off between potential bias and statistical stability. When the radius is too small, although any comparison case matched to a given treated case may be almost identical in terms of propensity score, other comparison cases that may be quite similar to a given treated case are lost. Conversely, where the radius is too large, comparison cases will not be sufficiently similar to treated cases. The weighted cross-validation method outlined in Galdo, Smith and Black (2008), which is designed to minimize the mean integrated standard error (MISE) for comparison case means, was used as an aid to choosing the optimal radius. In contrast to approaches that choose bandwidth based on the distribution of the comparison cases, Galdo et al. recognize that, in estimating the effect of the treatment on the treated, the MISE must be based on weighting by the distribution of the treated cases.

Galdo et al. provide several methods for implementing their approach. We employed the MISE based on weighting the comparison sample by the inverse odds of the propensity score. Experiments on two states were undertaken, using earnings in the fifth and tenth quarters after program entry as the dependent variables. MISE was estimated for radius values of 0.001, 0.003, 0.01, 0.03, and 0.1. For large samples, radius values larger than 0.1 were extremely time-consuming to estimate, so, in most cases, estimating MISE for larger radius values was not feasible. In general, the larger radius values produced smaller MISE values and also matched a larger portion of the treated cases. Based on these results, in the analysis that follows, the radius value was set to 0.1 as an initial default.

However, in choosing the radius for the analyses reported below, these considerations do not override the need that the approach successfully match on independent variables. Where a specification fails the balancing tests, the danger of bias requires that impact estimates be rejected, since there is no assurance that estimates obtained from an unbalanced match are appropriate for all dependent variables. Nonetheless, in deference to the above findings, larger radius values are chosen where tests suggest similar levels of balance.

Standard errors. As indicated in section III.4, we compared several approaches that have been suggested for estimating standard errors, one recommended by Imbens and Woodridge (2008) and Imbens (2008) that produces a conditional standard error, and another suggested by Abadie and Imbens (2006a) for estimating the unconditional standard error. Although both of these standard error estimates are asymptotically correct given the assumptions on which they are based, the design of the analysis is expected to increase sampling error in ways they do not capture. Methods used in pilot analyses for estimating the bootstrap standard errors attempted to incorporate all sources of error that could influence the estimates. (Although Abadie and Imbens, 2006b, show that bootstrap standard errors are asymptotically biased for certain matching estimators, there is no work indicating whether bootstrap standard errors for radius matching methods are consistent. See Imbens and Wooldridge, 2008). In particular, each replication was selected from the original samples of unique individuals, forming units of analysis from these individuals and undertaking the matching procedure on the resulting sample. Bootstrap standard errors were estimated both using the initial matching estimates and after undertaking bias adjustment. In each case, estimates were based on 25 replications, reflecting the computer-intensive nature of these replications. Across the two states chosen for these analyses, bootstrap standard errors were estimated for earnings and employment outcomes (up to 16 quarters after program entry), for each of the comparisons by gender (6 comparisons by 2 genders). Hence, over 300 estimates were obtained in each state.

On average, the bootstrap standard errors we calculated for the two states were quite close to the conditional standard errors. In one state, the bootstrap standard errors averaged 4 percent higher and in the other state they averaged 7 percent higher. There was some variation in the difference across estimates. The absolute percentage difference between the bootstrap and conditional standard errors was between 15 and 20 percent in these states. These experiments with bootstrap standard errors do not suggest that the results obtained using the conditional standard errors are misleading. These conclusions are similar to those of Heinrich (2007), who reported only modest differences between alternative standard error estimates.

Matching diagnostics details. As indicated above, tests were performed to determine if the means for the independent variables for the treated cases differed from the matched comparison cases, and a matching procedure was viewed as successful if fewer than 8 percent of these differences were statistically significant. If the initial specification successfully balanced the sample and matched a sufficient proportion of treated cases, the specification was accepted. In such a case, although interaction terms were not included in the logit specification, given the success of the balancing tests, their inclusion was not necessary. If the specification failed the balancing test, interaction terms were added to the logit specification. If balance continued to be a problem, the radius size was reduced, first to 0.03 and then to 0.01, forcing cases to match more closely. This approach often reduced the number of treated cases that matched.

If fewer than 90 percent of the WIA entries were matched, subgroups based on quarters of WIA entry were combined, in some cases combining quarters into two groups representing program years and in other cases into a single group. The benefit of matching separately by entry quarter is that a WIA participant is always matched with a comparison case in the same quarter, so that prior and subsequent employment measures all apply to the same period. However, by requiring matches to occur for cases within the same quarter, this approach limits the number of potential matches for each entry, and may reduce the likelihood of a match.

In most cases, this search produced a specification that passed these balancing criteria and matched over 90 percent of the treated cases. However, in a few comparisons, it was not possible to find a specification meeting these criteria, and a specification was ultimately accepted that matched as few as 80 percent of the WIA entries. In some cases, the balancing criteria were relaxed, so that specifications were accepted for which as many as one in ten mean differences were statistically significant. In a small number of instances where the number of treated cases was relatively large, even small differences in variable means were statistically significant. Specifications were then accepted for which more than one in ten differences was statistically significant so long as the mean absolute standardized difference was less than 0.02.

In almost all cases involving comparisons 1 and 2 (comparing WIA participants with the comparison program), it was possible to find a specification that met the above standards. In a few cases, an acceptable balance was obtained only by reducing the radius to the point where too few treated cases were matched. In a substantial number of the states, matching failed for one or more of the groups considered in comparison 3. Comparison 3 focused on training, with the treated group defined as WIA participants receiving training services and the comparison group as WIA participants receiving only core or intensive services. Several of the smaller programs provided training to a large share of their participants, so that few comparison cases were available. For example, several programs had fewer than 500 male participants in the Dislocated Worker program who had received training, and fewer than 200 others. In such cases, even where those receiving training were not very different from others, no close matches were available for a substantial number of treated cases.

For each comparison, Table C.1 lists the total number of treated cases by gender and program, as well as the proportion of cases that were omitted from the analysis because comparison matches were not available. The second column identifies cases that were omitted because the comparison was not performed in certain states, due to small sample sizes. The third column indicates the proportion of cases (in a state where the comparison was performed) that were omitted because they did not match. Both percentages are expressed relative to the total number of treated cases, so the sum of these percentages indicates the total loss of cases due to matching problems. For comparisons 1 and 2 (WIA participants vs. comparison group members), in most cases the total loss (omitted plus failed to match) is in the range of 2-7 percent, although it is close to 11 percent for female adult participants in comparison 1. The omitted cases are generally those with very minimal employment activity in the two years prior to participation, for which no similar comparison cases (usually UI claimants) were available. In the case of comparison 3, the total proportion excluded is much greater, nearly 50 percent for males in the Adult program and over 30 percent in others. As noted above, this reflects the fact that small states with high proportions of individuals receiving training provide very few comparison cases that are available for analysis.

Table C.1
Matched Treated Cases Available for Estimation

Total: 12 States		Cases Omitted (Percent)		
Comparison	Treated Cases	State Analysis Omitted	Failed to Match	Cases Analyzed
Adult Program				
Females				
1. WIA vs Comparison Group	56,612	4.9%	5.7%	50,657
2. WIA UI Recipient vs. UI	4,058	0.0%	2.1%	3,972
3. WIA Training vs. WIA Core/Intensive	17,941	24.4%	5.6%	12,564
Males				
1. WIA vs Comparison Group	40,940	0.0%	5.0%	38,894
2. WIA UI Recipient vs. UI	3,042	0.0%	2.1%	2,978
3. WIA Training vs. WIA Core/Intensive	9,899	43.2%	5.5%	5,087
Dislocated Worker Program				
Females				
1. WIA vs Comparison Group	33,174	0.0%	3.9%	31,876
2. WIA UI Recipient vs. UI	13,684	0.0%	3.1%	13,261
3. WIA Training vs. WIA Core/Intensive	10,984	25.0%	5.3%	7,655
Males				
1. WIA vs Comparison Group	30,915	0.0%	4.0%	29,690
2. WIA UI Recipient vs. UI	12,253	0.0%	2.8%	11,905
3. WIA Training vs. WIA Core/Intensive	9,211	32.4%	5.9%	5,676

Appendix D: Results of Dislocated Worker Program Impact Analysis for UI Recipients

Figures D.1 and D.2 present the estimates of the impact of the Dislocated Worker program on quarterly earnings for female and male UI recipients in nine states. The graphs show that WIA participant earnings are well below those of the comparison group in the first few quarters after entry, and do not catch up until seven or eight quarters after program entry. The initial negative effect is in the range of $700 for both men and women, and the maximum positive impact estimates—which occurs 10-14 quarters after WIA entry—are also lower, only about $200 for each group.

In one respect, results are very similar to those of the previous analysis. Earnings 16 quarters prior to program entry are appreciably higher for participants than for the comparison group. Although the sample of states on which this comparison is based is limited, it opens up the possibility that even these modest positive impacts could be spurious.

Impacts on employment (Figures D.3 and D.4) are slightly more supportive of the program, suggesting that the program ultimately increases the chance of employment for both men and women by as much as 4 percentage points. We cannot, however, reject the possibility that this largely reflects stable differences between participants and the comparison group.

Figure D.1
Dislocated Worker Program Treatment Effect on Quarterly Earnings for Female UI Recipients: WIA versus UI in 9 States

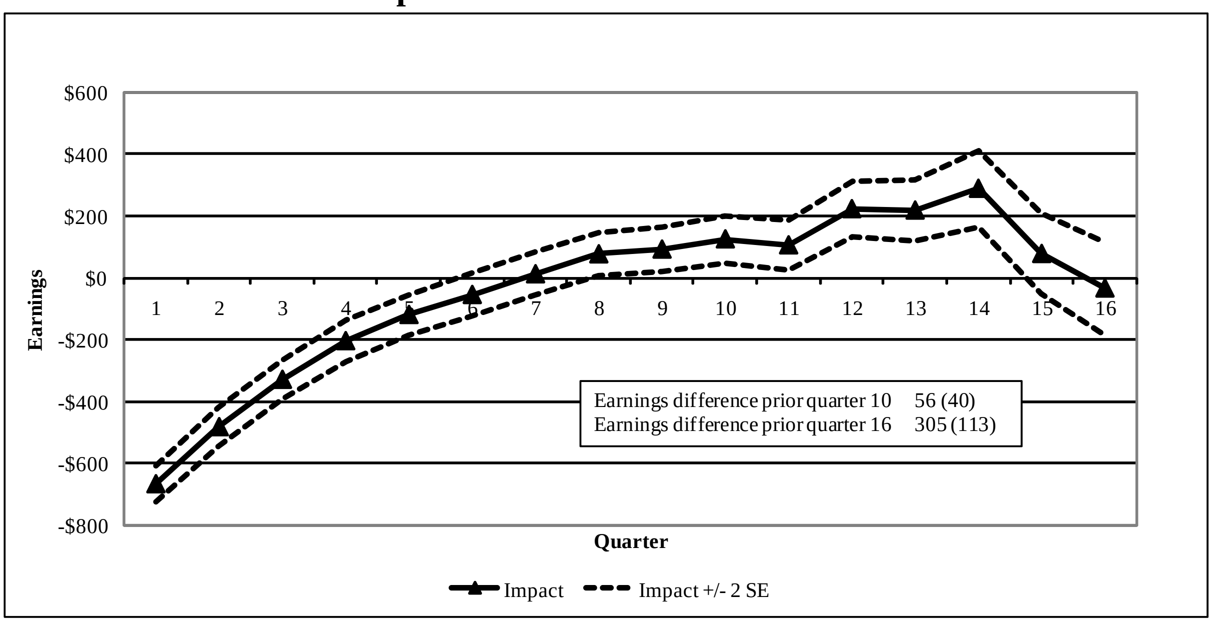


Figure D.2
Dislocated Worker Program Treatment Effect on Quarterly Earnings for Male UI Recipients: WIA versus UI in 9 States

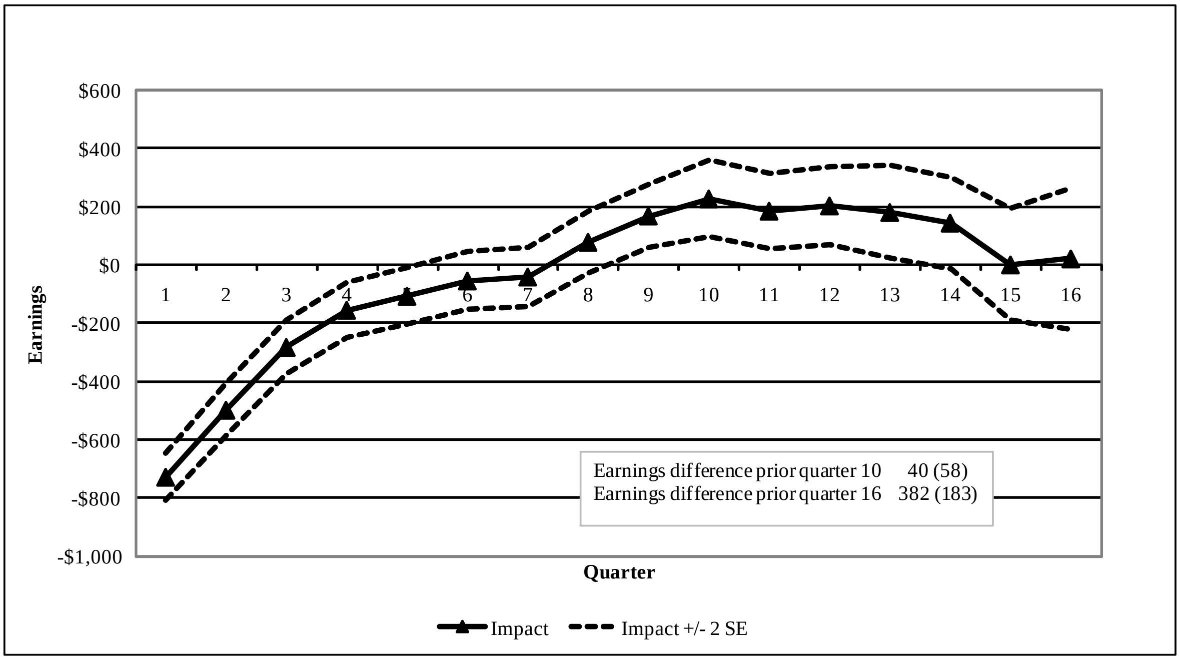


Figure D.3
Dislocated Worker Program Treatment Effect on Quarterly Employment for Female UI Recipients: WIA versus UI in 9 States

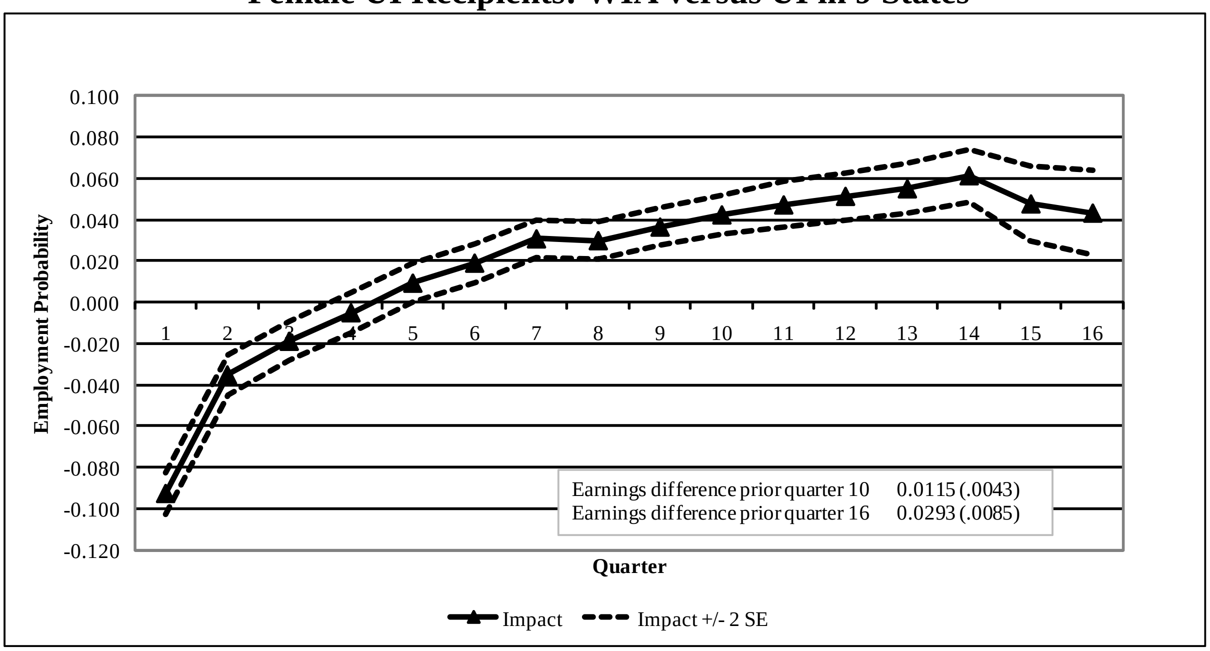


Figure D.4
Dislocated Worker Program Treatment Effect on Quarterly Employment for Male UI Recipients: WIA versus UI in 9 States