An Evaluation of the Workforce Investment Act Adult Program in Minnesota: Lessons from the Financial Crisis

Ellen Bendewald
Nicholas Maryns
Raymond Robertson

An Evaluation of the Workforce Investment Act Adult Program in Minnesota: Lessons from the Financial Crisis

Ellen Bendewald
Minnesota Department of Employment and Economic Development

Nicholas Maryns
InsightWorks

Raymond Robertson
Texas A&M University
and IZA

ABSTRACT

An Evaluation of the Workforce Investment Act Adult Program in Minnesota: Lessons from the Financial Crisis

Using matched administrative data, this paper estimates the program effects of Minnesota's WIA Title I-B Adult program to make two contributions to the literature. First, this paper applies a wide range of non-experimental approaches and generates similar estimates of program effects that are consistently larger than most other evaluations. Second, comparing workers job seeking in 2007-2008 with those job seeking in 2009-2010 suggests that the workers job seeking in a growing job market experience more persistent program effects.

JEL Classification: J28, J38, J68

Keywords: pseudo-experimental, job training, WIA-Adult

Corresponding author:

Raymond Robertson
Bush School of Government and Public Service
Texas A&M University
4220 TAMU College Station
Texas 77843-4220
USA
E-mail: robertson@tamu.edu

Introduction

The Workforce Investment Act (WIA) of 1998 has been one of the most prominent pieces of job training legislation in the United States.[1] Not surprisingly, the programs it authorizes – including the Adult and Dislocated Worker programs – have been widely evaluated. While evaluations of WIA programs generally find that participants' wages increase (Imbens and Wooldridge (2009) and Decker and Berk (2011), the range of estimated impact is quite large.[2] For example, Hollenbeck (2009) and Andersson et al. (2013) find positive impacts, but Moore and Gorman (2009), for example, find little evidence of the positive impact on WIA Adult participants. Estimates vary by estimation approach, program, demographic characteristics, and outcome variable.

While many factors could explain the varied results, four explanations emerge as primary contenders. First, studies vary in their empirical approach, including both the empirical methodology and access to data (time-series or cross-section, and comparison group). Selection bias implies that comparing wages (or other outcome measures) before and after program participation leads to inaccurate results. Comparing participants with non-participants after the program is problematic because program participants may be different (e.g. more motivated) than those who chose not to participate, and it is possible, and perhaps even likely, that such differences (and not the program per se) explain differences in post-participation outcomes. To address these challenges, randomized experiments, propensity score matching, and difference-in-difference (DiD) models are three of the most common approaches used in evaluation. Comparing results from a variety of empirical approaches would be a valuable contribution to the literature.

Second, although mandated and funded federally, WIA programs operate on a state-by-state basis.

Therefore, most evaluations focus on statewide impacts of WIA programs. States have power over aspects of WIA, such as the percentage of applicants that will be accepted into training programs. Heinrich, Mueser, and Troske (2008) use data from 12 states and find that selectivity can range from less than 40% to over 60% of applicants. Lusby (2010) shows that there is neither significant negative impact nor significant positive impact of WIA legislation on enrollment at technical and community colleges based on the data from Louisiana. Chrisinger (2013) finds that benefits associated with WIA cannot be distinguished from those from the people receiving only less-intensive Labor Exchange services in Washington State. As a result, additional state-level evaluations that describe state-specific program characteristics are valuable contributions to the literature.

Third, demographic characteristics matter. Hegewisch and Luyri (2010) find that even through both male and female employees who participated in WIA services benefit from the trainings, the average earnings for women are still significantly lower than average earnings for men. Moore and Gorman (2009) find that demographic factors, especially ethnicity, will influence the effect of WIA programs.

Finally, the time of displacement in the business cycle may also affect program outcomes. Jacobson et al. (1993), in a seminal paper in the study of displacement on wages, found that the negative effects of displacement for U.S. workers were large and long-lasting. Interestingly, subsequent papers found either no effects or even positive effects of displacement. For example, Abbring et al. (2002) found no change in wages in the United States and Bender et al. (2002) found positive wage changes following displacement in France and Germany. These differences suggest that that both time and location of displacement may contribute to post-displacement earnings. Kaplan et al. (2005) find that time and place significantly affect post-displacement earnings in Mexico. Couch et al. (2011) find that the results of displacement vary over the business cycle in Connecticut. Perhaps not surprisingly, workers who are displaced in areas or times with lower unemployment rates have higher post-displacement wages (including in some cases positive outcomes). Comparing the results of job training programs for workers displaced during the recession and during recovery would also be

a valuable contribution to the literature because it is possible that the success of job training programs depends on the time and place workers seek jobs after training.

In addition to generating results that help us understand the potential determinants of program success, this study also adds government benefits ("welfare") to earnings as an outcome variable. There is a sizable debate in policy circles about whether or not training programs actually reduce benefit use, especially in the short run, because participating in a training program may bring workers "into the system" and introduce them to benefits they may not have known about. Moore and Gorman (2009) point out the shortcoming of the universal "average earnings" measure. The inappropriate measure may lead the

One-Stops and local area WIBs to shift their resources from relatively disadvantaged job seekers to relatively advantaged job seekers, which are helpful to improve their outcome measures. As a result, studies that expand the set of outcome variables would be valuable contributions to the literature.

This paper contributes to this literature in four ways. First, it compares the results from a variety of relatively recent empirical approaches. The main results, generated with kernel-density propensity-score matching difference-in-difference estimation, are compared to results from regression adjustment, inverse probability weights, augmented inverse-probability weights, nearest neighbor matching, and combinations of these approaches. Second, it adds Minnesota to the list of states that have WIA evaluations.[3] Minnesota may be unique is several ways, including its relatively homogeneous population, relatively strong economy, and possibly a unique governance approach. The third is that it compares results of the program during the financial crisis and during the recovery. Finally, it adds benefits to the usual outcome variables of income and employment.

---

[3] We are not the first to study Minnesota's WIA program. Heinrich, Mueser, Troske, Jeon, and Kahvecioglu (2011) include Minnesota in the group of states they study and use annual earnings gains and the quarterly employment increment to captures the expected long-term benefits of training. It is difficult, however, to identify Minnesota's results indivudally in their study. Nevertheless, we extend their results by adding the comparison of empirical techniques, different business cycle phases, and adding benefits as an outcome variable.

The paper's structure follows the structure of the four contributions outlined above. In the next section, we describe the empirical approaches. We then describe the Minnesota context and data. The presentation of the main results adds benefits to wages and presents the results from the alternative estimation methods. The final section concludes.

Empirical Approach

This study takes a quasi-experimental approach to evaluating Minnesota's WIA Adult program. The program evaluation literature, comprehensively reviewed by Imbens and Wooldridge (2009), centers on attempts to accurately identify the most relevant comparison for program participants. Selection issues imply that comparing wages (or other outcome measures) before and after program participation leads to inaccurate results. Comparing participants with non-participants after the program is also problematic because program participants may be different (e.g. more motivated) than those who chose not to participate, and it is possible, and perhaps even likely, that such differences (and not the program per se) explain differences in post-participation outcomes. Randomized experiments, propensity score matching, and difference-in-difference (DiD) models are three of the most common approaches used in evaluation.

The U.S. Job Training Partnership Act (JTPA), which preceded WIA, included a randomized evaluation component, but such experiments are rare. While theoretically rigorous, they are not without their own concerns. For example, often experimental design assumes no treatment for the control group and can include participants that enroll but do not receive treatment in the treatment group, biasing estimates of program effect downward (King and Heinrich, 2011). Others have raised the possible ethical concern that if there is a reasonable belief that the program might be helpful (even to the point that the evaluation's goal is to estimate the magnitude of the positive effect) then putting people in the control group denies them of the opportunity to improve their lives. A third concern about such experiments is the cost: they can be expensive

to set up and conduct. Since a randomized evaluation requires an experimental set-up and implementation, the experimental approach is not feasible for existing programs.

Perhaps the most popular alternative falls under the umbrella of propensity score matching (PSM) methods.[4] Dehejia and Wahba (1999, 2002) suggest that PSM approaches can be used to evaluate programs that were not implemented with a random assignment evaluation component. Recent attempts to evaluate programs (e.g. Hollenbeck et al. 2005) rely almost exclusively on PSM methods.

Of course, PSM approaches have attracted their share of criticism. Heckman et al. (1998) argue that if the required PSM assumptions are inaccurate, matching methods may make the selection issue worse. Smith and Todd (2005), in particular, raise several concerns about the PSM approach. One of the concerns is that Dehejia and Wahba's (1999, 2002) results are sensitive to the specification of the equation used to estimate the propensity score and to achieve the balance between the treatment and comparison groups required for successful matching.

The third approach is known as difference-in-differences (DiD). Imbens and Wooldridge (2009) highlight Smith and Todd's (2005) argument that that the difference-in-difference (DiD) approach provides an alternative to the PSM approach that addresses the same concern driving the PSM approaches. The DiD approach involves comparing workers in and out of the program before and after the study. The main idea is that any unobserved differences between participants and non-participants would affect wages prior to participating in the program. As long as these differences are constant (that is, they are unique to the individual program participants), changes in the difference between participants and non-participants before and after the study (that is, the difference in the difference between the workers) can be reasonably attributed

---

[4] While most readers are probably familiar with PSM methods, it is worth offering a brief explanation for those who may not. Instead of randomly assigning participants to either a treatment or control group, PSM creates comparison groups to match the participant group on a number of observable variables (e.g. race, sex, education status, and previous wages). If these observable characteristics are correlated with unobserved worker traits (like productivity and effort), then PSM may generate an appropriate comparison group. In this approach, the probability of program participation is estimated as a function of observed individual characteristics. This probability is known as the "propensity score." Individuals who participated in the program are then matched with individuals who did not participate based on the similarity of the propensity scores.

to the program. Smith and Todd (2005) find that the DiD approach "did exhibit better performance than cross sectional estimators" in which the cross-sectional estimators include the PSM approaches described above.

Much of the debate in the program evaluation literature compares experimental design with PSM. Surprisingly, it is not clear that experimental design, while theoretically superior, generates results that are significantly different than other methods. Card et al. (2010) perform a meta-analysis of labor market evaluations and find no statistically significant difference in the results produced by experimental and observational studies. King and Heinrich (2011) build on this result and suggest that having an experimental design is not necessarily superior to observational approaches. Other studies, such as Heinrich et al. (2008) and Andersson et al. (2013) and argue that quasi-experimental methods, specifically using propensity score matching to create comparison groups, are not only a valid approach but also produce more appropriate estimates of change in wages. This assumption is supported by meta-analyses from Greenberg et al.'s (2006) analysis of 31 studies and Card et al.'s (2010) analysis of 199 studies. Both conclude that experimental and quasi-experimental studies produce no statistically different estimates of program effectiveness.[5] Greenberg et al. (2006) regress programs' outcomes on the type of evaluation and find that the evaluation type coefficient is small in magnitude with small standard errors. Card's studies also find a statistically insignificant effect of experimental design but with large standard errors. Given the increased cost in dollars and time of conducting an experimental evaluation, quasi-experimental design may be an efficient and arguably reasonably accurate choice for program evaluation.[6]

While the DiD approach performs better than PSM, PSM studies are still common because the DiD approach is much more data-intensive than PSM approaches. The DiD approach requires comprehensive

---

[5] Greenberg et al. only use non-experimental designs that include a correction for selection bias while Card et al. do not specify which programs they allow into the quasi-experimental category.

[6] King reflects an uncommon shift in opinion on this topic in two studies. While in 2004 he states "[d]espite enhancements in quasi-experimental methods for evaluating training programs in recent years…, the most reliable and credible evidence of the impacts of training comes from well-designed and structured experiments relying on randomly assigned treatment and control groups", his opinion shifts by 2011. Heinrich and King (2011) reads, "research designs used in quasi-experimental studies are generally unbiased" (p. 66) and furthermore "quasi-experimental designs are more likely to estimate the impact of the 'treatment on the treated' (p. 14).

information about workers before and after the program, as well as information from a comparison group of workers before and after the program. These data requirements are rarely met, making the DiD approach relatively less common in the academic literature. Our main approach relies on the DiD approach.

In addition to the DiD approach, however, we also include results from several other approaches. We describe them intuitively here. Catteneo et al. (2013) and Huber (2015) provide more detailed descriptions. Regression adjustment (RA) estimators model the outcome in such a way as to account for the non-random treatment assignment. The basic intuition behind this approach is that we can model the selection into the treatment group with observed covariates and then predict the unobserved potential outcomes. The average treatment effect (ATE) that is conditional on the observed covariates can then be estimated. When we limit the observations to those in the treatment group, we can estimate the average treatment effect on the treated (ATET).

A second alternative empirical approach is the inverse probability weighting (IPW) estimator. This approach is also designed to deal with non-random treatment assignment. The main difference between this approach and the RA approach described above is that the RA approach requires that we specify a model for the outcome and the IPW approach is a way to model the assignment into treatment (again based on observed covariates). The IPW approach incorporates weights that account for the non-random selection. The process uses the predicted probabilities of being in each group (treatment / comparison) as weights when estimating the treatment effects. This approach also can be used to generate the ATE and ATET estimates. When The IPW estimator with an additional bias-correction term is known as the Augmented IPW (AIPW) estimator. The bias correction term is important for those cases in which the treatment models are not correctly specified.

The RA and IPW (and AIPW) estimators are alternatives to the matching approach described earlier. As mentioned earlier, we primarily rely on propensity score matching. An alternative matching approach is known as nearest-neighbor matching. Like the RA and IPW estimators, the neighbor matching approach uses

information from observable covariates. Unlike these approaches, however, the nearest neighbor matching (NNM) estimator compares covariate values across individuals and matches them based on the specified covariates. Drukker (2015) describes this approach and uses the example of males being matched with males to illustrate the idea of matching on covariate values. NNM also includes a bias adjustment to remove the bias that may emerge from matching on more than one continuous variable.

The key difference between our DiD results and those from these alternative approaches is that the alternative approaches are designed for, and therefore here applied to, cross section data limited to post-treatment observations. The use of these estimators is often determined by the data availability. We now turn to a more detailed description of our data.

Data Description

Our treatment group consists of workers who participated in Minnesota's Workforce Investment Act (WIA) Title 1-B Adult program. The history and details of the national program are thoroughly described by Heinrich et al. (2011). The WIA Adult program provides employment and training assistance to adults who face significant barriers to employment. Minnesota's Adult program prioritizes individuals who receive public assistance, individuals living with low incomes, and veterans within these groups. For each customer, the overarching goal is employment or enhancement within his or her occupation. Generally, Adult program customers work to increase their earnings, retain employment, and diversify their occupational skills.[7] The Adult program provides services through a network of 48 WorkForce Centers. Adult program counselors meet with customers, provide active case management, and coordinate training. In contrast to some other states, Minnesota tracks "core" services in a separate data system from "intensive" and "training" services; the participant group we use in this study excludes those who only received "core" services. Compared to other

---

[7] See WIA Annual Report, Program Year 2013, available at: www.doleta.gov/Performance/Results/AnnualReports/PY2013/MN-PY13_WIA_AnnualReport.pdf

studies, then, we expect a greater net impact resulting from these more intensive program services.

The treatment population in this study includes any individual who exited the program between July 1, 2007 and June 30, 2008 (the "2007-2008 cohort") or between July 1, 2009 and June 30, 2010 (the "2009-2010 cohort"). Figure 1 shows the national, Minnesota, and St. Paul-Minneapolis unemployment rates for our relevant sample period. The 2007-2008 cohorts entered their programs during a period of low unemployment (the recession officially began December 2007). The 2009-2010 cohorts experienced much higher unemployment rates. The results of Kaplan et al. (2005) and Couch et al. (2011) suggest that the unemployment rate may affect post-displacement experiences and therefore it is possible that the effects of training programs are also affected. If so, we would expect that the wage gains would be larger for the first cohort.

The data used in the analysis come from several sources. The Minnesota Department of Employment and Economic Development (DEED) is our main data provider. DEED has worker-level administrative data on all employment covered by the Unemployment Insurance program. All employers in Minnesota are required to report individual-level wage data for every employee per quarter. For each quarter we have accurate total wage information (meaning that we sum earnings across all jobs) for every individual.[8] Wage data completeness is vital to our examination of the effect of job training because it ensures accurate matching between treatment and comparison groups, and it eliminates the need for programs to report (follow) the wages of participants after they exit the program.[9]

Demographic data for training participants come from the job training programs. Once again because these questions are program-specific, we have harmonized responses to basic questions in our analysis. When collapsing these categories, we attempt to retain the highest common specificity in the definitions to give the

---

[8] Several notable exceptions include sole proprietors, unincorporated firms, and Minnesota residents working outside the state. However, almost all (approximately 98%) of working Minnesotans are included in this resource.

[9] Follow up is a challenge because programs often do not maintain contact with participants after exit.

most descriptive benefit. Finally, treatment programs report social security numbers (SSNs) as well as basic program information[10] for each participant.

Comparison group individuals are drawn from three separate sources: those who apply for Unemployment Insurance (UI), those who register on MinnesotaWorks.net (MNW), and those who register with the WorkForce Center Customer Registration System (CRS). All three aid unemployed workers, albeit with a low intensity. None of these services involve more than cursory personalized job search assistance, and they therefore provide a logical pool from which to draw our comparison group (see Heinrich et al., 2008 and Andersson et al., 2009).

Unemployment Insurance is a temporary partial wage replacement for workers who are unemployed due to layoffs. Workers may be paid up to 50 percent of their average weekly wage, subject to a state maximum (currently $597) for up to 26 weeks. We include individuals who applied for Unemployment Insurance benefits in the same time period that our treatment groups were entering programs, whether or not they drew down any benefits. Certain Unemployment Insurance applicants were excluded from comparison groups: individuals receiving pensions, those who had re-filed for administrative or technical reasons (only first-time filers were included), those working out-of-state or for the military or federal government (since there may be gaps in wage detail data for these individuals), and those filers deemed ineligible for non-monetary reasons (e.g. individuals who may have quit or been discharged from employment). Applying and receiving Unemployment Insurance payments generally occurs online; however, a small portion of those receiving Unemployment Insurance are also required to attend job search workshops in person at a WorkForce Center, where they may access any of the other services offered there (discussed below).

---

[10] Including start dates, exit dates, and types of training provided.

In addition to Unemployment Insurance applicants, we also include individuals who registered an account at a WorkForce Center (via CRS) or with MinnesotaWorks.net in same timeframes that our treatment groups were entering programs. Any individual can access services at Minnesota's WorkForce Center locations, the state's WIA One Stop centers. Customers are provided with access to useful websites, software, and other job, career, or educational resources. They are informed of upcoming seminars, job fairs, and other events. Customers must be registered to use the Center's computers and to attend seminars and job fairs, and registration does not imply the receipt of staff-assisted services. MinnesotaWorks.net is Minnesota's labor exchange system, administered by the Minnesota Department of Employment and Economic Development (DEED). Registered job seekers can post up to five resumes to be searched by employers. They can also search for job openings and be contacted by e-mail when new job postings meeting their search criteria are found by the system. Employers can post job openings and search for job candidates, recruit job seekers online, and elect to receive emails when new resumés are found that match their requirements.

The Minnesota Department of Human Services (DHS) provided four kinds of welfare payment data: MFIP or MFMF is the Minnesota Family Investment Program (MFIP) cash grant. MFFOOD or MFFS is the MFIP food portion (Minnesota has a waiver to combine food assistance with cash assistance in MFIP. Without such a waiver, SNAP and TANF are separate programs). The third is Food Support (FS), which was recently renamed SNAP in MN to match the federal name. The final is the Diversionary Work Program (DWP), a four-month employment-focused program that some MFIP applicants get diverted to. They can transfer to MFIP after the four months if eligible at that time.

The DHS data also include information on case (family) size. For MFIP and DWP, the case is basically a nuclear family, caregiver(s) plus child(ren) minus those who are ineligible for some reason. For stand-alone FS (i.e., cases getting food assistance excluding MFIP cases getting the food portion – nearly all MFIP cases get the food portion), a household includes everyone in a household eligible for SNAP who prepares food and

eats together. All these programs are distributed via the EBT card, but the food amounts can only be spent on approved food items.[11]

Original data sources varied in the ways they presented every necessary variable. For example, gender was coded as a binary 1 or 0 in some datasets and Male or Female in others. Although variables like gender are relatively straightforward, more nuance is required to harmonize different definitions of education levels. Generalizing education data from five different sources requires reducing all data to the least common denominator. Unfortunately, some detail was lost when combining codes. For example, education is coded as groups of years representing familiar levels rather than in years. Finally, while some programs offer specific details on the types of services provided, not all do; definitions of services vary across programs and datasets.

Harmonized demographic data are then combined with wage detail data, reported quarterly to the state by employers, for all individuals in our sample. Pre-treatment wage data can be used as an instrument for factors that are challenging to quantify (e.g. initiative).

Our analysis uses program entrance date rather than exit for several reasons. The analysis began under the assumption that most individuals who exit the program at the same time would be clustered around the same enrollment time for that program. We discovered that variation in enrollment times were much larger than anticipated. Adjusting the unit of analysis to enrollment date rather than exit date allows for similar comparisons across all programs. When a participant enters the program, we 'start the clock' (so to speak) and measure effects from that point in time. The average program length of WIA Adult is under one year.

We control for the possibility that serial correlation biases our results (Bertrand et al. 2004) by reducing our comparison to "pre" and "post" treatment. For the "pre" period, we take the sum of all earnings for each individual i in quarters 5-8 prior to entering the program for workers in the treatment group. For the

---

[11] For more program info, go to http://mn.gov/dhs/people-we-serve/children-and-families/economic-supports.

comparison group, we take the sum of wages in quarters 5-8 prior to entering comparison group status (e.g. becoming unemployed if the individual is drawn from the sample of unemployed workers).   Our "post" period is the sum of wages in the 5th through 8th quarter after entering the status that qualifies the individual for treatment or comparison group status.

Table 1 contains the summary statistics for the WIA Adult cohorts. The sample sizes are not particularly large. The demographics are reasonably similar to those from other studies, but are not identical across cohorts. The education and gender composition are more similar than the racial and education composition. Note that the second cohort is larger and has lower earnings, which may reflect the lasting impact of the financial crisis.

Matching

We match participants and non-participants on three main factors: their demographic profile, their time of program enrollment, and their pre-enrollment wages. The actual variables used to estimate the propensity scores are gender, age, race (whether or not the individual listed their race as White[12]), veteran status, education, residence (whether or not the individual lived in the seven-county Minneapolis – St. Paul metropolitan area), and benefits.

To match comparison and treatment samples, we use a two-tier approach. The comparison group sample sizes are several orders of magnitude larger than the treatment group sample sizes. To get smaller comparison group samples, we first apply a randomized restriction algorithm that randomly drops comparison group individuals that are sufficiently dissimilar to the treatment group means. Specifically, we first specify a bandwidth around the treatment group means. Observations in the comparison group that are outside this

---

[12] The size of all racial groups in our samples are too small to provide accurate identification.

bandwidth are dropped with a probability that reflected the size of the comparison universe relative to the treatment sample size (the mode probability is 0.5). This process is repeated until the comparison group means are close to the treatment group means in the sense that the balancing condition is satisfied or nearly satisfied.

Table 2 contains the pre-match diagnostic statistics. Comparison and treatment pre-treatment means for earnings, gender, age, race, veteran status, and benefits are not statistically different. Education and metro area status are significantly different for the earlier cohort, but the differences for education means are very small. The percent metro is not statistically different in the second cohort. Where the covariates are not balanced, we rely on the Difference-in-Difference approach to provide unbiased estimates. This requires the assumption that the effect of education and labor market region are constant over time, which we believe to be reasonable.

The resulting samples were matched using PSM. In practice, there are many ways to use propensity scores to match treatment and comparison groups. PSM can match individuals one-to-one, one-to-many, or many-to-one. Nearest neighbor matching matches all individuals to their closest PSM fit and can have multiple non-participants for every treatment individual or find the best matches only using non-participants once. In any matching procedure, tradeoffs exist between sample size and match quality. Thus the researcher determines how close of a match is close enough to balance these conflicting interests.

To create the most appropriate comparison group, we use an approach called Kernel density matching, which uses the propensity scores as weights and calculates a weighted average of matched non-participants for each participant.[13] Thus multiple individuals combined create a 'comparison match' that balances characteristics and predicts wages based on that balance. In this way, better matches are given higher 'weight'

---

[13] One important matching characteristic is pre-program wages. Characteristics like motivation cannot be directly measured, so wages are used as an instrument. Training participants' wages have been found to dip in the periods leading up to training due, a phenomena called the 'Ashenfelter dip' (Heinrich, Mueser, & Toske, 2008) and are often followed by quarters of stagnant wages due to the opportunity cost in decreased wages of dedicating time to training (often called 'lock in' effects). Thus for matching purposes we have used eight quarters of data, from twelve to four quarters prior to program start date to avoid the Ashenfelter dip.

and therefore more influence in the comparison group. The propensity score is estimated using our observable variables, including past employment and earnings, gender, age, race, veteran status, highest level of education attained, geography, and past enrollment in public benefits programs. Past employment and earnings are particularly important because they shed light on factors that are not directly recorded in the data, such as individual-level productivity and motivation.

Estimation

Acknowledging that our matching may not completely control for unobserved characteristics, we then subject the matched individuals to the Difference-in-Difference approach adapted from Imbens and Wooldridge (2009). Difference-in-Difference is especially appropriate in program evaluation because it eliminates both time effects and group-specific effects from the 'noise' and leaves the main program effect. Assuming matches are similar between the two groups, the treatment group's post-program wages minus the comparison group's post-program wages result in net program impacts. Thus, one main advantage of using Difference-in-Differences is that even if the two groups are not similar in ways that are consistent across time, it eliminates group impacts and leaves only the effect of treatment on the treated.

Our guiding equation is adapted from Imbens and Wooldridge (2009) and is implemented with Villa's (2011) *diff* command in Stata.[14] Assume that we have N individuals indexed i=1...N for whom we observe (G_i, w_{i0}, w_{i1}) where:

G_i=Group membership

$\overline{w}_{i0}$=Average wages for the first four quarters pre participation

$\overline{w}_{i1}$=Average wages for the last x quarters after enrollment

---

[14] Villa, "DIFF: Stata Module to Perform Differences in Differences Estimation."

Thus the estimation equation is:

$$\overline{w}_{i1} - \overline{w}_{i0} = \beta + \widehat{\tau DID} * G_i + \epsilon_i$$

The coefficient of interest is $\widehat{\tau DID}$ or the predicted change in average wages pre-program participation to post-program participation for the treatment group.

The number of quarters over which to estimate program effects depends on average program participation length. Some programs are more intensive than others and will have a larger primary negative 'lock in' effect on wages for the first quarters after participation (Card et al., 2010). Increasing the follow up period can improve overall impact accuracy. Wage changes are important, but just as critical to the analysis is how long increases are sustained. If an individual sees a sustained increase of $500 a quarter for the rest of their career, then the impacts of that program would be much larger over time than for a program where impacts taper off after three or four quarters.

Estimation Results: Earnings

Table 3 contains the net impacts on average annual earnings for both cohorts in the short, medium, and long term. Over the medium term (five to eight quarters after program entrance) the WIA Adult program was responsible for an increase in annual earnings of 37 percent. In other words, program participants had earnings 37 percent higher than they would have if they not participated in the program. This percentage increase translates into positive impacts of $4,333 and $3,611 in annual earnings for the 2007-2008 and 2009-2010 cohorts, respectively. Note that the first cohort experiences much larger wage gains over the short term but also much more fade out across time.

There is no clear trend across all cohorts with regard to whether net impacts on earnings grow or dissipate over time. For the 2007-2008 cohort, we observe that impacts on earnings decrease as time goes on from the short-term to the long-term. This is consistent with Figure 1 in the sense that the unemployment

rises sharply in their longer term. For the 2009-2010 cohort, impacts stayed fairly constant as time went on, which is consistent with the falling unemployment rates found in Figure 1. The 2007-2008 cohorts exited their programs as the Great Recession was starting (the recession officially began December 2007), while the 2009-2010 cohorts exited as the recession was ending (the recession officially ended June 2009).

The estimated results for various populations over the medium term are shown in Table 4. A few trends stand out. First, gender matters. The results for both cohorts show much larger, statistically significant impacts for men. Second, the results for African Americans seem greater than for whites. No clear patterns seem to emerge for differences across age or education, either across levels or across cohorts.

Estimation Results: Quarterly Cash Benefits (MFIP and SNAP)

Our analysis of net impacts on the receipt and average amount of quarterly cash benefits (MFIP and SNAP) received yields far fewer statistically significant results. Tables 5 and 6 contain the results. In terms of cash benefit receipt, both cohorts tend to show small, negative impacts. The exception to this is the 2009-2010 cohort, which shows a positive impact in the short term. Restricting the analysis to those receiving cash benefits, the 2007-2008 cohort shows a negative impact in terms of average amount of cash benefits received (though the results are not statistically significant) and the 2009-2010 cohort shows a positive impact (though only the short-term impact is statistically significant).

Alternative Estimation Approaches

The richness of our data allows us to explore the sensitivity of our results with several different estimators.[15] In particular, we compare the results from several estimators that are often applied to cross-

---

[15] As long as we have the right set of covariates, we do not expect our results to be sensitive to estimation approach. See for example Mueser et al. (2007).

section data. Unlike our data, these data usually lack observations about workers prior to treatment. Given the lack of treatment, it is important to model the selection process. Methods commonly applied to modeling the selection process include regression adjustment, inverse probability weighting, augmented inverse probability weighting, nearest neighbor matching, and the combination of regression adjustment with inverse probability weighting. We briefly describe the results from each method when applied to the post-evaluation observations in our data.

Regression adjustment

The regression adjustment approach models the outcome variable as a function of observables. We follow this approach by modeling wages as a function of gender, age, race (white v. nonwhite), veteran status, education level, geographic region (in or out of main metropolitan area), and benefits received. The differences in the conditional means represent the average treatment effects. Both the average treatment effects and the average treatment effects on the treated for our two samples are shown in Table 6. The results are qualitatively and quantitatively similar to those presented earlier, although slightly smaller in magnitude. Additionally, the results demonstrate a significant difference across the two periods, with the effects for the 2009-2010 cohort smaller than those for the 2007-2008 cohort.

(Augmented) Inverse probability weighting

Rather than model the outcome variable as a function of observables, we could instead assume that program participation was a function of observables. The inverse probability weighting approach is one way to model the treatment selection. To consider the possibility that treatment selection is a function of observed demographic variables, we fit a logit model of treatment as a function of gender, age, race (white v. nonwhite), veteran status, education level, geographic region (in or out of main metropolitan area), and benefits received. The results are also shown in Table 6 and are very similar to the regression adjustment estimates presented above.

The IPW estimator with a bias correction is known as the augmented inverse probability weighting estimator. The bias correction comes from modeling both the outcome and the treatment. Another approach that models both the outcome and the treatment is the regression adjustment with inverse probability weighting. Both of these generate similar results, as shown in Table 6.

Nearest neighbor matching

The nearest neighbor matching estimator is the most similar to the propensity score matching estimator used above. The NNM estimator does not use a functional form and therefore is much more flexible. We match on the same set of observable covariates listed above: gender, age, race (white v. nonwhite), veteran status, education level, geographic region (in or out of main metropolitan area), and benefits received.[16] These results are also shown in Table 6.

Conclusions

Evaluating Minnesota's Workforce Investment Act during two different phases of the business cycle and applying several different estimation approaches generates several important results. The first is that the kernel density PSM DiD results that rely on panel data are qualitatively and quantitatively similar to some newer techniques that rely purely on post-treatment cross-section data. Second, the results from Minnesota adds to the list of states that have WIA evaluations, and suggests that Minnesota's results are larger than those found in WIA evaluations from other states. Minnesota's unemployment rates are lower than the national average, which is consistent with the third main result of this study: phases of the business cycle may play a significant role in explaining differences in WIA evaluations across different time periods. Consistent

---

[16] In addition, we specified a version which requires exact matches of discrete covariates and a version which adds a large-sample bias adjustment as suggested by Adadie and Imbens (2006,2011). Both models produced similar estimates and are not shown.

with Kaplan et al. (2005) and Couch et al. (2011), participants during low-unemployment times experience larger wage gains than workers in higher-unemployment times.

Finally, it adds benefits to the usual outcome variables of income and employment. Our results suggest that benefit use is not statistically significantly affected by participation in the WIA program. This result is important for those computing return on investment estimates for WIA Adult programs who want to know if participants reduce government expenditures (if training reduces benefit use) or increase government expenditures (if training brings workers into the government "system" and thus increases benefit use).

References

Abadie, Alberto and Guido W Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica*, 74(1), 235-67.

Abadie, Alberto and Guido W Imbens. 2011. "Bias-Corrected Matching Estimators for Average Treatment Effects." *Journal of Business & Economic Statistics*, 29: 1–11.

Abadie, Alberto. 2005. "Semiparametric Difference-in-Differences Estimators." *The Review of Economic Studies*, 72(1), 1-19.

Abbring, Jaap H., Gerard J. van den Berg, Pieter A. Gautier, A. Gijsbert C. van Lomwel, Jan C. van Ours, and Christopher J. Ruhm. 2002. "Displaced Workers in the United States and the Netherlands." In *Losing Work, Moving On: International Perspectives on Worker Displacement*, Peter J. Kuhn, ed. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 105–194. http://dx.doi.org/10.17848/9781417505333.ch2

Andersson, Fredrik; Harry J Holzer; Julia I Lane; David Rosenblum and Jeffrey Smith. 2013. "Does Federally-Funded Job Training Work? Nonexperimental Estimates of WIA Training Impacts Using Longitudinal Data on Workers and Firms". National Bureau of Economic Research working paper 19446.

Ashenfelter, Orley. (1978) "Estimating the Effect of Training Programs on Earnings" *Review of Economics and Statistics* 6, no. 1 (1978): 47-57.

Athey, Susan and Guido W Imbens. 2006. "Identification and Inference in Nonlinear Difference-in-Differences Models." *Econometrica*, 74(2), 431-97.

Azemati, Hanna; Michael Belinsky; Ryan Gillette; Jeffrey Liebman; Alina Sellman and Angela Wyse. 2013. "Social Impact Bonds: Lessons Learned So Far." *Community Development Investment Review*, (01), 023-33.

Basu, Anirban. 2014. "Estimating Person-Centered Treatment (Pet) Effects Using Instrumental Variables: An Application to Evaluating Prostate Cancer Treatments." *Journal of Applied Econometrics*, 29(4), 671-91.

Bender, Stefan; Christian Dustmann; David Margolis; and Costas Meghir. 2002. "Worker Displacement in France and Germany." In *Losing Work, Moving On: International Perspectives on Worker Displacement*, Peter J. Kuhn, ed. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 375–470.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. (2004) "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics* 119(2), 249-275.

Burda, Michael C and Antje Mertens. 2001. "Estimating Wage Losses of Displaced Workers in Germany." *Labour Economics*, 8(1), 15-41.

Card, David; Jochen Kluve and Andrea Weber. 2010. "Active Labour Market Policy Evaluations: A Meta-Analysis." *The Economic Journal*, 120(548), F452-F77.

Carrington, William J. 1993. "Wage Losses for Displaced Workers: Is It Really the Firm That Matters?" *Journal of Human Resources*, 28(3), 435-62.

Cattaneo, Matias D. 2010. "Efficient Semiparametric Estimation of Multi-Valued Treatment Effects under Ignorability." *Journal of Econometrics*, 155(2), 138-54.

Cattaneo, Matias D; David M Drukker and Ashley D Holland. 2013. "Estimation of Multivalued Treatment Effects under Conditional Independence." *Stata Journal*, 13(3), 407-50.

Chrisinger, Colleen K. 2013. "Earnings Progression and the Workforce Investment Act: Evidence from Washington State." *Industrial Relations: A Journal of Economy and Society*, 52(4), 853-77.

Clark, Don P; Henry W Herzog and Alan M Schlottmann. 1998. "Import Competition, Employment Risk, and the Job-Search Outcomes of Trade-Displaced Manufacturing Workers." *Industrial Relations*, 37, 182-206.

Couch, Kenneth A. 2001. "Earnings Losses and Unemployment of Displaced Workers in Germany." *Industrial & Labor Relations Review*, 54(3), 559-72.

Couch, Kenneth A. & Jolly, Nicholas A. & Placzek, Dana W., 2011. "Earnings losses of displaced workers and the business cycle: An analysis with administrative data" Economics Letters, Elsevier, 111(1), 16-19, April.

Decker, Paul T and Jillian A Berk. 2011. "Ten Years of the Workforce Investment Act (Wia): Interpreting the Research on Wia and Related Programs." *Journal of Policy Analysis and Management*, 30(4), 906-26.

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

Donald, Stephen G and Kevin Lang. 2007. "Inference with Difference-in-Differences and Other Panel Data." *The Review of Economics and Statistics*, 89(2), 221-33.

Drukker, David (2015) "Introduction to Treatment Effects in Stata: Part 2: The STATA Blog, 24 August, http://blog.stata.com/2015/08/24/introduction-to-treatment-effects-in-stata-part-2/.

Eberts, Randall W; Wei-Jang Huang and Jing Cai. 2013. "A Methodology for Setting State and Local Regression-Adjusted Performance Targets for Workforce Investment Act Programs." Upjohn Institute Working Paper 13-189. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Greenberg, David H; Charles Michalopoulos and Philip K Robin. 2006. "Do Experimental and Nonexperimental Evaluations Give Different Answers About the Effectiveness of Government-Funded Training Programs?" *Journal of Policy Analysis and Management*, 25(3), 523-52.

Heckman, James and Jeffrey Smith. (1999) "The Pre-Programme Earnings Dip and the Determinants of Participation in a Social Programme. Implications for Simple Programme Evaluation Strategies" *The Economic Journal* 109(457), 313-348.

Heckman, James and Jeffrey Smith. (2003) "The determinants of participation in a social program: Evidence from a prototypical job training program" National Bureau of Economic Research, no. w9818 (2003).

Heckman, James J; Hidehiko Ichimura and Petra E Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *The Review of Economic Studies*, 64(4), 605-54.

Heckman, James; Hidehiko Ichimura; Jeffrey Smith and Petra Todd. 1998. "Characterizing Selection Bias Using Experimental Data." *Econometrica*, 66(5), 1017-98.

Heinrich, Carolyn et al. "New Estimates of Public Employment and Training Program Net Impacts: A Nonexperimental Evaluation of the Workforce Investment Act Program" IZA DP No. 4569 (2009).

Heinrich, Carolyn J; Peter Mueser; K Troske and Jacob M Benus. 2008. "Workforce Investment Act Non-Experimental Net Impact Evaluation." *IMPAQ International*.

Heinrich, Carolyn J; Peter R Mueser; Kenneth R Troske; Kyung-Seong Jeon and Daver C Kahvecioglu. 2013. "Do Public Employment and Training Programs Work?" *IZA Journal of Labor economics*, 2(1), 1-23.

Heinrich, Carolyn J; Peter R Mueser; Kenneth R Troske; Kyung-Seong Jeon and Daver C Kahvecioglu. 2011. "A Nonexperimental Evaluation of Wia Programs." In *The Workforce Investment Act: Implementation Experiences and Evaluation Findings,* eds. Douglas J. Besharov and Phoebe H. Cottingham, Kalamazoo, MI: WE Upjohn Institute for Employment Research, 371-404.

Helwig, Ryan T. 2001. "Worker Displacement in a Strong Labor Market." *Monthly Labor Review*, 13.

Hollenbeck, Kevin and Wei-Jang Huang. "Net Impact and Benefit-Cost Estimates of the Workforce Development System in Washington State." Upjohn Institute (2006).

Hollenbeck, Kevin. 2009. "Workforce Investment Act (Wia) Net Impact Estimates and Rates of Return," European Commission-sponsored meeting," What the European Social Fund Can Learn from the WIA Experience," Washington, DC. http://research. upjohn. org/confpapers/2. Citeseer,

Hollenbeck, Kevin. Workforce Investment Act (WIA) net impact estimates and rates of return. Upjohn Institute Conference Papers, 1-23 (2009) .

Hollenbeck, Kevin. 2011. "Short-Term Net Impact Estimates and Rates of Return." *The Workforce Investment Act: Implementation Experiences and Evaluation Findings*, 347-70.

Hollenbeck, Kevin; Daniel Schroeder; Christopher King; and Wei-Jang Huang "Net Impact Estimates for Services Provided through the Workforce Investment Act" Employment and Training Administration Occasional Paper 2006-06, U.S. Department of Labor (2005).

Hollenbeck, Kevin; Daniel Schroeder; Christopher T King; Wei-jang Huang and Wei-lang Huang. 2005. "Net Impact Estimates for Services Provided through the Workforce Investment Act." *Department of Labor, Employment and Training Administration.* Citeseer,

Howland, Marie and George E Peterson. 1988. "Labor Market Conditions and the Reemployment of Displaced Workers." *Industrial & Labor Relations Review*, 42(1), 109-22.

Huber, Chuck (2015) "Introduction to Treatment Effects in Stata: Part 1" *The STATA Blog* 7 July. http://blog.stata.com/tag/regression-adjustment/

Imbens, Guido W and Jeffrey M Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature*, 47(1), 5-86.

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan. 1993b. *The Costs of Worker Dislocation.* Kalamazoo: W. E. Upjohn Institute for Employment Research.

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan. 1993a. "Earnings Losses of Displaced Workers." *American Economic Review* 83(4): 685–709.

Kaplan, David S., Gabriel Martínez, and Raymond Robertson. 2005. "What Happens to Wages after Displacement?" *Economia: Journal of the Latin American and Caribbean Economic Association* 5.2: 197-234.

King, Christopher and Carolyn Heinrich. "How Effective are Workforce Development Programs? Implications for U.S. Workforce Policies." Association for Public Policy Analysis and Management's Fall Research Conference (2011): 1-26.

King, Christopher and Carolyn Heinrich. 2011. "How Effective are Workforce Development Programs? Implications for U.S. Workforce Policies." Association for Public Policy Analysis and Management's Fall Research Conference.

King, Christopher et al. "Returns from Investments in Workforce Services: Texas Statewide Estimates for Participants, Taxpayers, and Society." Ray Marshall Center for the Study of Human Resources.

King, Christopher T. 2004. "The Effectiveness of Publicly Financed Training in the United States: Implications for WIA and Related Programs." In *Job Training Policy in the United States*, eds. Christopher J. O'Leary, Robert A. Straits, and Stephen A. Wandner, W.E. Upjohn Institute, 57-99.

King, Christopher T; Ying Tang; Tara Smith; Daniel Schroeder and Burt S Barnow. 2008. *Returns from Investments in Workforce Services: Texas Statewide Estimates for Participants, Taxpayers, and Society*. Ray Marshall Center for the Study of Human Resources, Lyndon B. Johnson School of Public Affairs, University of Texas at Austin.

King, Christopher. The effectiveness of publicly financed training in the United States: Implications for WIA and related programs, Upjohn Institute Press Book Chapters (2004): 57-99.

LaLonde, Robert. (1986) "Evaluating the Econometric Evaluations of Training Programs" *American Economic Review* 76(4), 604-620.

Liebman, Jeffrey et al. "Social Impact Bonds: Lessons Learned So Far." John F. Kennedy School of Government, Harvard University. Community Development Investment Review.

Miles, Marty, and Stacy Woodruff-Bolte. "Apples to Apples: Making Data Work for Community-Based Workforce Development Programs." Corporation for a Skilled Workforce (2013).

Miles, Marty, and Stacy Woodruff-Bolte. 2013. "Apples to Apples: Making Data Work for Community-Based Workforce Development Programs." Corporation for a Skilled Workforce.

Moore, Richard W and Philip C Gorman. 2009. "The Impact of Training and Demographics in Wia Program Performance: A Statistical Analysis." *Human Resource Development Quarterly*, 20(4), 381.

Mueser, Peter R; Kenneth R Troske and Alexey Gorislavsky. 2007. "Using State Administrative Data to Measure Program Performance." *The Review of Economics and Statistics*, 89(4), 761-83.

Mueser, Peter, Kenneth Troske, and Alexey Gorislavsky (2007) "Using state administrative data to measure program performance." *The Review of Economics and Statistics*, 89(4), 761-783.

Nichols, Austin. (2007) "Causal Inference with Observational Data" *The Stata Journal* 7(4), 507-541.

Office of the Legislative Auditor, State of Minnesota. 2010. "Evaluation Report: Workforce Programs".

Plesca, Miana and Jeffrey Smith. "Evaluating multi-treatment Programs: Theory and Evidence from the U.S. Job Training Partnership Act Experiment" *Empirical Economics* 32 (2007): 491-528.

Plesca, Miana and Jeffrey Smith. 2007. "Evaluating Multi-Treatment Programs: Theory and Evidence from the Us Job Training Partnership Act Experiment." *Empirical Economics*, 32(2-3), 491-528.

Ruhm, Christopher J. 1991a. "Are Workers Permanently Scarred by Job Displacements?" *American Economic Review* 81(1): 319–24.

Ruhm, Christopher J. 1991b. "Displacement Induced Joblessness." *Review of Economics and Statistics* 73(3): 517–22.

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

Smith, Jeffrey and Petra Todd (2005) "Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics* 125(1-2), 305-53.

Specktor, Carole. 2004. "Using Return-on-Investment Measures for Workforce Development Programs in Minnesota." Hubert H. Humphrey Institute of Public Affairs.

Stevens, Ann Huff. 1995. "Long-Term Effects of Job Displacement: Evidence from the Panel Study of Income Dynamics." Working paper 5343. Cambridge, Mass.: National Bureau of Economic Research.

Stevens, Ann Huff. 1997. "Persistent Effects of Job Displacement: The Importance of Multiple Job Losses." *Journal of Labor Economics* 15(1, part 1): 165–88.

Stevens, Ann Huff. 2001. "Changes in Earnings Instability and Job Loss." *Industrial and Labor Relations Review* 55(1): 60–78.

U.S. Department of Labor. 2015. "Training and Employment Guidance Letter No. 30-14: Negotiating or Extending Performance Goals for the Workforce Investment Act (WIA) Title IB Programs and Wagner-Peyser Act Funded Activities for Program Year (PY) 2015" http://wdr.doleta.gov/directives/attach/TEGL/TEGL_30-14.pdf

Villa, Juan. 2015. "Diff: Stata Module to Perform Differences in Differences Estimation." Statistical Software Components. Boston College Department of Economics.

Workforce Investment Act of 1998. Department of Labor. August 7, 1998. http://www.doleta.gov/usworkforce/wia/wialaw.pdf

TABLE 1: DEMOGRAPHIC CHARACTERISTICS OF WIA ADULT COHORTS

	2007-2008		2009-2010	
	N	%	N	%
OVERALL TOTAL	745	100%	1,093	100%
AGE				
18-24	97	13.0%	230	21.0%
25-44	308	41.3%	548	50.1%
45-54	128	17.2%	201	18.4%
55-64	38	5.1%	65	5.9%
Other or No Data	174	23.4%	49	4.5%
GENDER				
Female	305	40.9%	569	52.1%
Male	268	36.0%	475	43.5%
Other or No Data	172	23.1%	47	4.3
RACE				
African American	115	15.4%	219	20.0%
White	414	55.6%	735	67.2%
Other or No Data	216	29.0%	139	12.7%
GEOGRAPHY				
Greater Minnesota	536	71.9%	687	62.9%
Seven-County Metro	209	28.1%	406	37.1%
EDUCATIONLEVEL				
Less than HS	113	15.2%	127	11.6%
HS Diploma or Equivalent	313	42.0%	466	42.6%
Some Postsecondary	259	34.8%	386	35.3%
AA, BA, and Above	60	8.1%	114	10.4%
AVERAGE ANNUAL EARNINGS (at Baseline)	$18,566		$17,841	

Table 2: Pre-Treatment Match Diagnostics

		Mean Weighted Variables							
Cohort	Cohort Size (N)	Log Annual Earnings	% Female	Age	% White	% Veteran	Education Level	% Seven-County Metro	Quarterly Cash Benefits
WIA Adult 2007-2008	745	9.264	53.2%	36.9	72.3%	4.5%	3.30	28.1%	$237
Comparison Group	1,751	9.234	54.3%	37.5	72.4%	4.7%	3.20	34.9%	$208
Difference		**0.03**	**-1.1%**	**-0.6**	**-0.1%**	**-0.2%**	**0.10**	**-6.9%**	**$29**
p-value		0.5095	0.6074	0.2607	0.9473	0.858	0.0114**	0.0004***	0.2394
WIA Adult 2009-2010	1,093	9.183	54.5%	35.3	70.4%	3.5%	3.44	37.1%	$312
Comparison Group	2,093	9.136	55.6%	35.1	70.0%	3.5%	3.38	38.0%	$296
Difference		**0.047**	**-1.1%**	**0.2**	**0.4%**	**0.0%**	**0.06**	**-0.9%**	**$15**
p-value		0.2258	0.5707	0.7074	0.7929	0.9828	0.0848*	0.599	0.6005

Notes: Means and t-tests are estimated by linear regression. Inference: *** p<0.01; ** p<0.05; * p<0.1. Education level is a weighted average using the following coding: 0 - other and/or missing; 1 - grade school or less; 2 - incomplete high school; 3 - high school diploma or equivalent; 4 - some postsecondary education; 5 - associates or bachelor degree; 6 - masters or doctorate. "% white" is calculated based on those providing information.

Table 3: NET IMPACTS ON AVERAGE ANNUAL EARNINGS

Cohort	Baseline	Short-Term	Medium Term	Long Term
		Follow-Up Periods		
	Log Annual Earnings	Log Annual Earnings	Log Annual Earnings	Log Annual Earnings
WIA Adult 2007-2008	9.375	9.651	9.553	9.503
	(0.056)	(0.042)	(0.052)	(0.059)
Comparison Group	9.237	9.084	9.102	9.314
	(0.016)	(0.032)	(0.037)	(0.035)
Difference	0.138	0.567	0.451	0.189
	(0.058)	(0.053)	(0.064)	(0.069)
Difference-in-Difference		**0.429 *****	**0.313 *****	**0.052**
		(0.078)	(0.086)	(0.090)
Net Impact (Percent Change in Annual Earnings)		**+53.6%**	**+36.8%**	**+5.3%**
Net Impact (Change in Annual Earnings, in Dollars)		**+$6,316**	**+$4,333**	**+$629**
WIA Adult 2009-2010	9.189	9.253	9.455	9.674
	(0.042)	(0.037)	(0.039)	(0.039)
Comparison Group	9.135	8.927	9.087	9.341
	(0.013)	(0.029)	(0.032)	(0.028)
Difference	0.054	0.326	0.368	0.333
	(0.044)	(0.047)	(0.050)	(0.048)
Difference-in-Difference		**0.272 *****	**0.314 *****	**0.279 *****
		(0.065)	(0.067)	(0.065)
Net Impact (Percent Change in Annual Earnings)		**+31.3%**	**+36.9%**	**+32.2%**
Net Impact (Change in Annual Earnings, in Dollars)		**+$3,060**	**+$3,611**	**+$3,150**

Notes: Timeframes: Baseline 4-8 quarters before entrance; short-term 1-4 quarters post-entrance; medium-term 5-8 quarters; long-term 9-12 quarters. Robust Standard Errors are in parentheses. Means and Standard Errors are estimated by linear regression. Inference: *** p<0.01; ** p<0.05; * p<0.1

**Table 4: Disaggregated Net Impacts on Average Earnings
over the Medium Term**

	2007-2008		2009-2010	
	Log Annual Earnings Diff-in-Diff	Net Impact (% Change in Earnings)	Log Annual Earnings Diff-in-Diff	Net Impact (% Change in Earnings)
OVERALL	0.313***	**+36.8%**	0.314***	**+36.9%**
	(0.086)		(0.067)	
BY AGE				
18-24	0.318	**+37.4%**	0.527***	**+69.4%**
	(0.197)		(0.129)	
25-44	0.347***	**+41.5%**	0.307	**+35.9%**
	(0.111)		(0.091)	
45-54	0.298	**+34.7%**	0.448***	**+56.5%**
	(0.197)		(0.127)	
55-64	0.348*	**+41.6%**	0.495***	**+64.0%**
	(0.199)		(0.129)	
BY GENDER				
Female	0.361***	**+43.5%**	0.306***	**+35.8%**
	(0.118)		(0.080)	
Male	0.549***	**+73.2%**	0.52***	**+68.2%**
	(0.122)		(0.106)	
BY RACE				
African American	0.661***	**+93.7%**	0.687***	**+98.8%**
	(0.180)		(0.163)	
White	0.453***	**+57.3%**	0.313***	**+36.8%**
	(0.092)		(0.068)	
BY GEOGRAPHY				
Greater Minnesota	0.377***	**+45.8%**	0.334***	**+39.7%**
	(0.109)		(0.073)	
Seven-County Metro	0.349***	**+41.8%**	0.438***	**+55.0%**
	(0.129)		(0.115)	
BY EDUCATION LEVEL				
Less than HS	0.513***	**+67.0%**	0.55***	**+73.3%**
	(0.196)		(0.184)	
HS Diploma or Equiv.	0.357***	**+42.9%**	0.233**	**+26.2%**
	(0.134)		(0.104)	
Some Postsecondary	0.348**	**+41.6%**	0.508***	**+66.2%**
	(0.140)		(0.097)	
AA, BA, and Above	0.585**	**+79.5%**	0.457**	**+57.9%**
	(0.264)		(0.200)	

Table 5: Net Impacts on Quarterly Cash Benefit Receipt (MFIP and SNAP)

Cohort	Baseline	Follow-Up Periods		
		Short-Term	Medium Term	Long Term
	Benefits Receipt	Benefits Receipt	Benefits Receipt	Benefits Receipt
WIA Adult 2007-2008	0.280	0.317	0.277	0.295
	(0.019)	(0.020)	(0.019)	(0.019)
Comparison Group	0.183	0.236	0.228	0.255
	(0.011)	(0.012)	(0.012)	(0.012)
Difference	0.097	0.081	0.049	0.039
	(0.022)	(0.023)	(0.022)	(0.023)
Difference-in-Difference		**-0.016**	**-0.049**	**-0.058**
		(0.031)	(0.031)	(0.031)
Net Impact (Percent Change in Likelihood of Receiving Benefits)		**-1.6%**	**-4.9%**	**-5.8%**
WIA Adult 2009-2010	0.294	0.456	0.383	0.372
	(0.014)	(0.015)	(0.015)	(0.015)
Comparison Group	0.211	0.293	0.313	0.292
	(0.011)	(0.012)	(0.012)	(0.012)
Difference	0.083	0.163	0.070	0.080
	(0.018)	(0.019)	(0.019)	(0.019)
Difference-in-Difference		**0.080 ***	**-0.014**	**-0.003**
		(0.026)	(0.026)	(0.026)
Net Impact (Percent Change in Likelihood of Receiving Benefits)		**+8.0%**	**-1.4%**	**-0.3%**

Notes: Timeframes: Baseline 4-8 quarters before entrance; short-term 1-4 quarters post-entrance; medium-term 5-8 quarters; long-term 9-12 quarters. Robust Standard Errors are in parentheses. Means and Standard Errors are estimated by linear regression. Inference: *** p<0.01; ** p<0.05; * p<0.1.

Table 6: Net Impacts on Average Quarterly Cash Benefits (MFIP and SNAP)

| Cohort	Baseline	Follow-Up Periods		
		Short-Term	Medium Term	Long Term
	Log Quarterly Benefits	Log Quarterly Benefits	Log Quarterly Benefits	Log Quarterly Benefits
WIA Adult 2007-2008	6.273	6.701	6.719	6.996
	(0.116)	(0.113)	(0.138)	(0.117)
Comparison Group	5.773	6.204	6.387	6.541
	(0.086)	(0.081)	(0.094)	(0.074)
Difference	0.5	0.497	0.332	0.455
	(0.144)	(0.139)	(0.167)	(0.138)
Difference-in-Difference		**-0.003**	**-0.168**	**-0.045**
		(0.200)	(0.221)	(0.200)
Net Impact (Percent Change in Quarterly Benefits)		**-0.3%**	**-15.5%**	**-4.4%**
Net Impact (Change in Quarterly Benefits, in Dollars)		**-$2**	**-$82**	**-$23**
WIA Adult 2009-2010	6.28	7.125	6.956	6.88
	(0.080)	(0.065)	(0.076)	(0.079)
Comparison Group	5.988	6.544	6.45	6.467
	(0.079)	(0.111)	(0.137)	(0.291)
Difference	0.291	0.581	0.506	0.413
	(0.112)	(0.129)	(0.157)	(0.113)
Difference-in-Difference		**0.29 ***	**0.215**	**0.122**
		(0.171)	(0.193)	(0.159)
Net Impact (Percent Change in Quarterly Benefits)		**+33.6%**	**+24.0%**	**+13.0%**
Net Impact (Change in Quarterly Benefits, in Dollars)		**+$180**	**+$128**	**+$69**

Notes: Timeframes: Baseline 4-8 quarters before entrance; short-term 1-4 quarters post-entrance; medium-term 5-8 quarters; long-term 9-12 quarters. Robust Standard Errors are in parentheses. Means and Standard Errors are estimated by linear regression. Inference: *** p<0.01; ** p<0.05; * p<0.1.

Table 7: Alternative Estimation Methods

	2007-2008	2009-2010
Kernel Density PSM DiD (Medium Term)	**0.313**	**0.314**
	(0.086)	(0.067)
Regression Adjustment ATE	0..259	0.192
	(0.052)	(0.039)
Regression Adjustment ATET	0.289	0.222
	(0.050)	(0.037)
Inverse-probability weights ATE	0.260	0.205
	(0.048)	(0.038)
Inverse-probability weights ATET	0.267	0.191
	(0.050)	(0.037)
Augmented IPW ATE	0.250	0.206
	(0.048)	(0.037)
IPW Regression Adjustment ATE	0.253	0.211
	(0.048)	(0.037)
Nearest Neighbor Matching	0.175	0.229
	(0.074)	(0.060)

Notes: Standard errors in parentheses. The Kernel Density PSM DiD estimates are from Table 4. Each regression approach is defined in the text and is applied to the post-treatment period observations only.

Figure 1: Unemployment Rates and Cohort Program Exit Date Windows
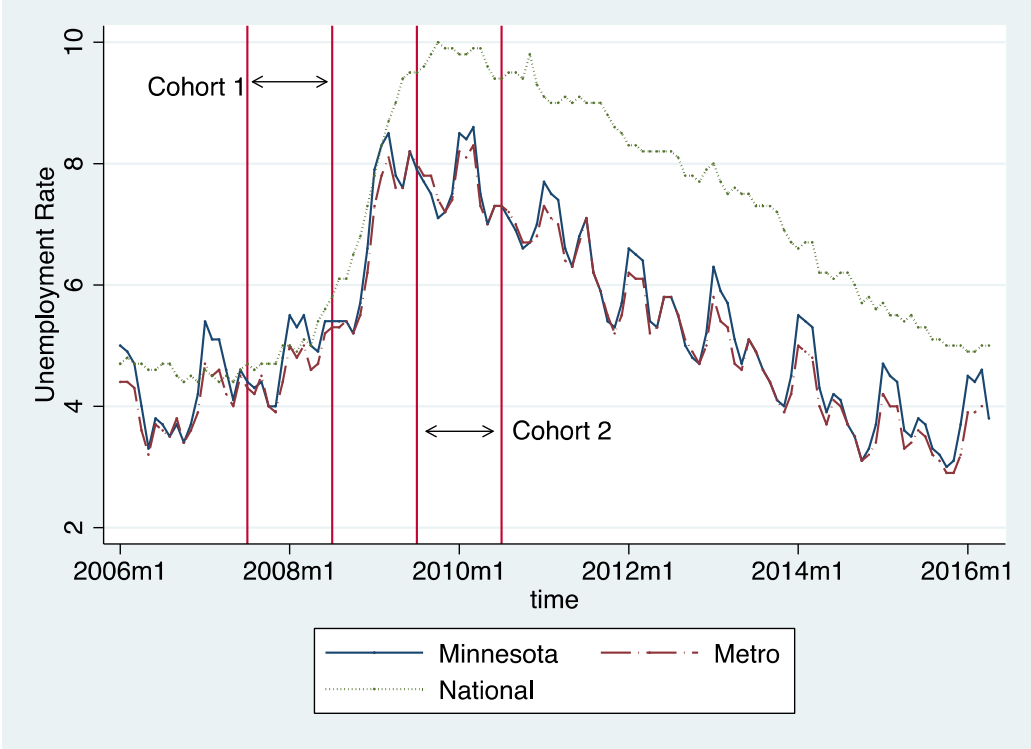
Appendix: Project Background

Over the last decade, Minnesota's leaders have increasingly noted the need for more rigorous and standardized approaches to evaluating workforce programs. In 2009, Minnesota enacted a new set of "Uniform Program Accountability Measures"—known as UPAM—for Minnesota's Department of Employment and Economic Development (DEED) economic development and workforce development programs.[17] One of the measures stipulated by the UPAM law was return on investment. In response to this law, the GWDC launched its Return on Investment Initiative[18] and convened an advisory group to study the issue and develop a standardized measure, in alignment with its statutory role to:

Advise the governor on the development and implementation of statewide and local performance standards and measures relating to applicable federal human resource programs and the coordination of performance standards and measures among programs.[19]

Made up of a broad range of partners, including key staff from relevant state agencies, workforce development service providers, business members, community-based organizations, and data evaluation experts, the GWDC advisory group set out to develop a standard return on investment methodology that could be applied to workforce programs administered or funded with public dollars.

To guide their work, the advisory group agreed to a number of shared values and goals for the methodology, namely that it should be transparent and credible, adaptable and sensitive to change, relatively simple to administer, and yield timely and relevant results.

Shift of Focus to Net Impact Evaluation

The advisory group chose to estimate return on investment through net impact evaluation, which takes a scientific approach to estimating and attributing program impacts, limiting the use of broad assumptions. The advisory group studied a number of net impact/return on investment evaluations in other states and at the federal level.

In early 2010, the Office of the Legislative Auditor released an evaluation report on workforce programs that provided some initial net impact findings, and recommended that, "DEED should adopt a set of standard approaches for assessing workforce program outcomes, including periodic comparisons of workforce program participants and non-participants."[20] This report further emphasized the importance of developing and implementing a standardized net impact framework for Minnesota's workforce development programs—a foundation upon which return on investment estimates could be built.

The advisory group developed the broad parameters of the framework, such as which outcomes to measure, how to measure them, and how to estimate associated costs and benefits

---

[17] Minnesota Statutes § 116J.997, which have since been repealed and replaced.
[18] This initiative has since been re-named the *Net Impact Initiative.*
[19] See Minnesota Statute 116L.665 Subd. 3c, available at: https://www.revisor.mn.gov/statutes/?id=116L.665.
[20] Office of the Legislative Auditor, State of Minnesota, "Evaluation Report: Workforce Programs." Available at: http://www.auditor.leg.state.mn.us/ped/pedrep/workforce.pdf.

for the purpose of return on investment analyses. In doing so, the group balanced the varied interests and perspectives of its members to develop a consensus-based framework ready for implementation.
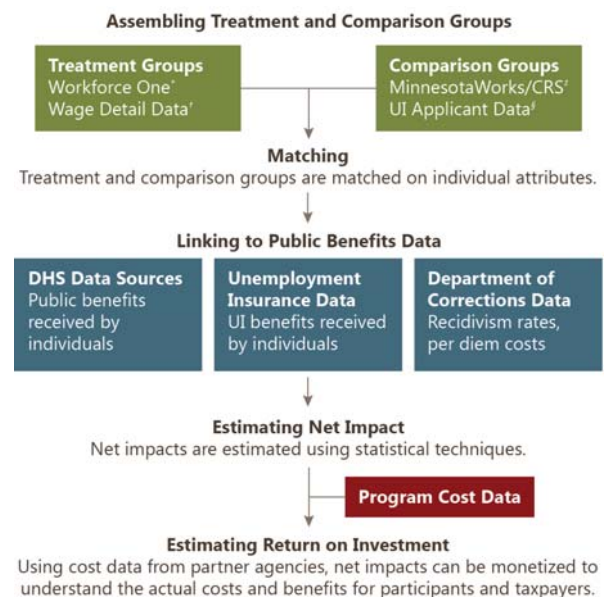
Pilot Study: Purpose and Scope

In 2012, the advisory group began to lay the groundwork for a relatively small-scale pilot evaluation. The goal of the pilot has been to inform the development of a high-quality net impact evaluation framework, and to establish the necessary data procedures to make way for the implementation of a larger, ongoing performance management effort.

The work of the pilot has involved building partnerships with various state agencies, studying state agency programs, and establishing data-sharing procedures across various data systems. Along the way, we have addressed numerous challenges, including a lack of precedent for the type and scope of cross-agency data sharing required, instituting an infrastructure for storing and sharing large datasets, and determining exactly which data are needed and how to properly link them. Inconsistencies in how (and whether) data are reported and defined across state data systems were also a barrier, requiring a great deal of work to standardize the data and prepare it for analysis.

Challenges notwithstanding, a major benefit of the approach has been that it uses data that are already available; it does not require service providers to collect additional data or program participants to self-report. The pilot has provided a clear pathway forward for future analyses, minimizing much of the upfront cost in large part due to the voluntary participation of the evaluator.

The pilot evaluation analyzes the net impacts of two major workforce programs, the Workforce Investment Act (WIA) Adult Program and the Dislocated Worker Program (both WIA- and state-funded) operating between 2007 and 2010 in the midst of the Great Recession and its after-effects. The evaluation focuses on the impacts of these programs on employment, earnings, and participant use of cash benefits (namely MFIP and SNAP).[21] Workforce program participant outcomes are analyzed against the outcomes of comparison groups constructed from similar non-participants who either applied for unemployment insurance benefits or who registered at a WorkForce Center or online at MinnesotaWorks.net.[22]

---

[21] MFIP is the Minnesota Family Investment Program, the state's TANF program; SNAP is the Supplemental Nutrition Assistance Program, formerly known as Food Stamps.
[22] See page 8 for details on the evaluation design.

Collaborative Approach

Along the way, the advisory group, evaluator, and staff have worked collaboratively with other evaluation experts in Minnesota to build consensus around a common framework for net impact and return on investment analysis. This includes Wilder Research, the Greater Twin Cities United Way, and Invest in Outcomes. The goal is that by achieving consensus, we may encourage further evidence-based policy-making efforts, foster the credibility and transparency of the framework, and save time and resources by avoiding unnecessary duplication. Thus another goal of the pilot project (and this report) is to provide clear guidance to state agencies and evaluators for reproducing and improving this work in the future.