

# **Evaluation of the Effectiveness of a Low-Cost Statistical Methodology to Target Services to Participants of a Local Welfare-to-Work Program**

**Randall Eberts**

W.E. Upjohn Institute for Employment Research

November 2015

Paper prepared for presentation at the

2015 APPAM Fall Research Conference

Session: Lessons Learned from Public Workforce Program Experiments

## **Abstract**

The Obama Administration has emphasized the use of low-cost interventions based on insights from behavioral economics to increase access to social programs and improve government efficiency. Several initiatives are underway but few directly target workforce programs. This paper reports on a USDOL-funded pilot conducted in the late 1990s to illustrate how low-cost interventions in workforce programs can improve employment outcomes of participants. The pilot is relevant for the current interest in low-cost RCT trials by demonstrating how a simple improvement in the referral of participants to services can improve outcomes, how RCT can be embedded in the existing program, and how administrative data can be used to minimize the cost and disruption of the evaluation.

## Introduction

The purpose of this paper is to show how low-cost interventions can be integrated into the operations of existing workforce programs. Recent interest in using lessons from behavioral economics to improve the participation and engagement in social programs has led to a growing number of initiatives that have attempted to use random control trials (RCT) experiments to improve program design, particularly in the way information is presented to participants. The Obama administration has made this approach a priority in how it administers social programs. In 2014 the Administration created the Social and Behavioral Sciences Team (SBST), dubbed the Nudge Squad presumably after Richard Thaler's and Cass Sunstein's influential book entitled *Nudge*. Their book documents the use of behavioral science in improving the participation in social programs and thus the effectiveness of the programs. Even before the creation of the SBST, the administration used lessons from behavioral economics in designing certain programs in the American Recovery and Reinvestment Act (ARRA) so that consumers would respond more quickly and effectively to the economic stimulus initiatives.

The UK has also pursued lessons from the insights of behavioral economics. In 2010 the UK Cabinet Office established the Behavioural Insights Team (BIT) with the purpose of finding "intelligent ways to encourage, support and enable people to make better choices for themselves."<sup>1</sup> One of its first interventions was to work with Jobcentre Plus office staff to redesign the process individuals went through when they signed on to receive benefits and began their job search process. Since then, BIT has run over 150 randomized control trials evaluating interventions in a wide variety of social areas.<sup>2</sup>

According to a survey paper by DellaVigna (2009), "behavioral economics stresses empirical findings of behavior that are partially at odds with standard economic assumptions. The key empirical findings from field research in behavioral economics imply that individuals can make systematic errors or be put off by complexity, that they procrastinate, and that they hold non-standard preferences and non-standard beliefs."<sup>3</sup> Therefore, insights from behavioral economics focus on ways to simplify complex decision-making processes that may tax the ability of individuals to effectively navigate government programs. Thus, the SBST projects and the UK BIT initiatives are designed to address primarily the behavioral barriers that affect how people engage with programs.<sup>4</sup> While the expected results may be modest, so are the costs, resulting in potentially large returns on investment.

The Obama Administration has formalized the use of behavioral insights by directing federal agencies to initiate and test such procedures. In 2013 the Administration sent a memo to the heads of federal agencies stating that "many innovative companies use rapidly conducted randomized field trials to identify high impact innovations and move them quickly into

---

<sup>1</sup> UK Cabinet, Behavioural Insights Team, Annual Update, 2010-2011, p. 3.

<sup>2</sup> Recently, the Behavioural Insights Team has evolved into a social purpose company and is no longer embedded in the Cabinet Office

<sup>3</sup> DellaVigna S (2009) Psychology and Economics: Evidence from the Field. *J Econ Lit* 47:315-372.

<sup>4</sup> Social and Behavioral Sciences Team 2015 Annual Report, p. xii, September 2015.

production.”<sup>5</sup> Random control trials are definitely not new to evaluating social programs. However, in the past most RCT evaluations were hugely expensive and took years to conduct and analyze. The approach advocated by the Administration is to try to streamline the evaluation process by embedding the process within the programs receiving the interventions. This is possible if agencies already collect data that records participant outcomes and characteristics and participants can be easily randomly selected into control and treatment groups.

For example, a conference on RCT held in the summer of 2014 and sponsored by the Office of Technology Policy and the Coalition for Evidence-Based Policy explored effective ways to embed low-cost RCTs in government social programs. Participants asserted that the following steps should be taken: 1) greater research access to government administrative data, such as UI wage records for workforce programs, with appropriate privacy protections, 2) increased government funding opportunities that specifically focus on low-cost RCTs, and 3) more high-profile competitions and challenges for low-cost RCTs, such as those launched by the Coalition for Evidence-based Policy.<sup>6</sup>

As previously mentioned, the Obama Administration institutionalized the use of behavioral insights by establishing the Social and Behavioral Sciences Team in 2014. SBST is a cross-agency team, housed in the White House Office of Science and Technology Policy, with the purpose of translating findings and methods from the social and behavioral sciences into improvements in Federal policies and programs.<sup>7</sup> During its first year of operation, its team has focused on executing proof-of-concept projects where behavioral insights could be embedded directly into programs at a low cost and lead to quantifiable and immediate improvements in program outcomes. The team pursued two areas where behavioral science could play a significant role: improving access to programs and improving government efficiency. Seventeen projects are listed in the SBST September 15, 2015 Annual Report, which included promoting retirement savings, improving college access, increasing medical insurance coverage, reducing delinquent debt repayments, and several others. Many of the projects included simple ways to communicate with individuals to improve their engagement in Federal programs.

On September 15, 2015, behavioral insights were further codified into Federal social policy when President Obama signed an Executive Order that encourages Federal agencies to “design its policies and programs to reflect our best understanding of how people engage with, participate in, use, and respond to those policies and programs. It specifically directs agencies to:

1. Identify opportunities to help qualifying individuals, families, communities, and businesses access public programs and benefits...by removing administrative hurdles, shortening wait times, and simplifying forms;
2. Improve how information is presented to consumers ... by considering how the content, format, timing, and medium by which information is conveyed affects comprehension and action by individuals, as appropriate;

---

<sup>5</sup> “Memorandum to the Heads of Departments and Agencies,” Office of Management and Budget, July 2013.

<sup>6</sup> Maya Shanka, “How Low-cost Randomized Controlled Trials Can Drive Effective Social Spending,” White House, July 30, 2014 at 2:05 pm.

<sup>7</sup> Social and Behavioral Sciences Team 2015 Annual Report, p. iii, September 2015.

3. Identify programs that offer choices and carefully consider how the presentation and structure of those choices, including the order, number, and arrangement of options, can most effectively promote public welfare, as appropriate, giving particular consideration to the selection and setting of default options; and
4. Review elements of their policies and programs that are designed to encourage or make it easier for Americans to take specific actions...”<sup>8</sup>

Despite the intense interest by the Administration in using low-cost RCTs to evaluate the effectiveness of behavioral insights in Federal programs, few initiatives have been directed at workforce programs, although the USDOL is sponsoring a few at this time.<sup>9</sup> The recently released Annual Report of the SBST lists 17 projects, but none involves Federal workforce programs. Therefore, the purpose of this paper is to provide an example of how simple but effective low-cost interventions and evaluations can be integrated into workforce programs.

The pilot program described here illustrates several aspects of the Administration’s concept of low-cost RCTs. First, the pilot focuses on two of the four directives to Federal agencies in the President’s executive order: improve how information is presented and improve how choices of programs are presented to customers. According to Babcock et al. (2012), job search assistance and employment services should be simplified and streamlined by making tools available that gather information on the education and employment opportunities pursued by others like the participant, list job openings that may interest the participant, and provide information on the projected growth in occupations (p. 8). The pilot streamlines the intake process by reducing the number of times participants must fill out registration forms and tries to match participants with providers that are better suited to respond to their specific needs. Second, part of the setup of this intervention was based on establishing an employability score, which was derived from statistical procedures similar to the profiling score required under Worker Profiling and Reemployment Services (WPRS). Babcock et al. (2012) advocate using profiling for job search assistance in order to discourage workers from procrastinating from receiving the services that might be helpful in finding employment. Third, the pilot embeds an RCT experiment directly in the intake process by randomly assigning participants to one of three service providers, stratified by three levels of employability. Fourth, the RCT uses administrative data generated by the program to record participant characteristics and employment outcomes, which provides a low-cost evaluation instrument which can yield results in a short period of time.

While the program described here is not new (the pilot was conducted in the late 1990s and early 2000s) and has been reported in previous publications, it is still instructive in providing

---

<sup>8</sup> The White House, Executive Order: Using Behavioral Science Insights to Better Serve the American People, September 15, 2015.

<sup>9</sup> One of the USDOL-sponsored “Nudge” programs directed at workforce programs is being conducted at the Upjohn Institute through its division that administers WIOA programs for a four-county area in southwest Michigan. The program was developed by front-line staff with the assistance of Ideas42 and Mathematica. It is being evaluated using an embedded RCT experiment.

an example that may guide the implementation of future initiatives.<sup>10</sup> Furthermore, since the program has already been evaluated using RCT, which was embedded in the intervention, the outcomes of the intervention are available.

## **Work First Pilot**

### *Overview*

The purpose of the pilot was to improve the employment outcomes of participants of the state welfare-to-work program, by streamlining the referral process so that services could be tailored to best meet the needs of participants. Funded by the USDOL and developed by the Upjohn Institute, the pilot referred welfare-to-work participants to one of three service providers based on a statistical algorithm that used administrative data to determine which provider offered services that were shown to be most effective for customers possessing specific characteristics and employment backgrounds. The referral system also streamlined the intake process by having information about each participant already loaded in the computer system before they reported for orientation. This approach avoided the need for participants to fill out forms with information they had already provided to providers in previous visits. Each provider offered different services and different approaches to delivering those services. Before the pilot was established, the local Workforce Investment Board (LWIB) where the pilot took place randomly referred participants to three different providers. Therefore, the relationships between different types of services and employment outcomes for groups of participants with different characteristics were based on a randomized sample. Using this sample, the observed employment outcomes were regressed against personal characteristics of the participants, and these relationships were then used to refer new enrollees to providers based on the enrollees' personal characteristics and the recent experience of similar participants. The initiative demonstrated that customizing services based on participant characteristics could increase the effectiveness and efficiency of the intervention. A random assignment evaluation of the pilot showed that targeting services in this way could significantly increase the 90-day employment retention rate of participants by 20 percentage points, yielding a benefit-cost ratio of greater than three. The paper draws lessons from this experiment on how low-cost interventions have the potential to significantly improve the effectiveness of workforce programs.

### *Detailed Description of the Work First Pilot*

The purpose of Michigan's Work First Program was to move welfare recipients into jobs as quickly as possible. The program provided welfare recipients reemployment skills, support, and opportunities to obtain employment, and it offered instruction in the proper techniques for writing resumes, completing applications, and interviewing for jobs. All enrollees received similar services regardless of their needs, although the three service providers varied in scope and intensity of services. More intensive skill training was available only to those who held a

---

<sup>10</sup> Randall W. Eberts, "Using Statistical Assessment Tools to Target Services to Work First Participants," in Randall W. Eberts, Christopher O'Leary, and Stephen A. Wandner (eds.) Targeting Employment Services, Kalamazoo MI: W.E. Upjohn Institute for Employment Research, 2002. The description of the pilot included in this paper will draw heavily from this publication.

job or those who had repeatedly failed to find employment. After clients completed the core services, they were expected to search intensively for work and accept offers that provided at least 20 hours of work per week at or above minimum wage.<sup>11</sup> Customers employed for 90 consecutive days in a qualified job were considered a successful outcome, and they were exited from the program. As an incentive for finding work, participants were permitted to keep the first \$200 earned each month and 20 percent over that without a reduction in benefits. Participants also received transportation, child care, and Medicaid for a limited time.

Michigan's local Workforce Development Boards (WDBs) worked closely with the local offices of the state's social service agency, the Family Independence Agency (FIA), to administer the Work First program. FIA determined welfare eligibility, issued welfare payments, and referred welfare recipients to Work First programs, while the Work First agency provided welfare recipients with employment services, through intermediaries. At the time of the pilot, FIA referred all applicants for public assistance to Work First, with some exceptions.

FIA referred all qualified applicants to Work First within 10 days of their applying for cash assistance. Applicants were notified of the date and time they are to enroll in the program and attend orientation. Orientation included an introduction to the Work First program, specification of the roles and responsibilities of the program and client, and a brief assessment of the client's situation and immediate needs, including the possible need for supportive services.<sup>12</sup> In-depth assessment and counseling were offered only to those in considerable need. In most cases, all those referred by FIA to Work First were required to participate in the same job search and job readiness workshops regardless of their past work histories or qualifications. Job search/job club workshops provided training in appropriate skills in seeking, locating, applying for, and obtaining employment. Job search training was typically conducted in group settings. Each person was expected to develop a résumé and to understand the proper techniques for completing applications and interviewing for jobs.

In Michigan, intermediaries, not the local WDBs, provided employment services to Work First participants. During the time the pilot was conducted, the Kalamazoo/St. Joseph WDB subcontracted with three local service organizations to provide employment services to participants included in the pilot. The formal agreements required that the organizations provide a set of basic services. Nevertheless, the number of hours in which customers participated in these activities varied, and in some cases, service providers offered additional services beyond those prescribed by the contract or they provided assistance using different approaches. Therefore, the level and intensity of services varied across service providers, although they all were in compliance with their contracts and state requirements.

---

<sup>11</sup>Allowable work activities included 1) unsubsidized employment; 2) subsidized private sector employment; 3) subsidized public sector employment; 4) on-the-job training; 5) job search and job readiness training and activities up to six weeks; 6) community service programs; and 7) no more than 12 months of vocational educational training.

<sup>12</sup>The initial assessment that was performed at orientation before the pilot began was minimal, and staff at the Kalamazoo-St. Joseph WDB did not consider this assessment adequate to be used as a basis of referring customers to service providers. Therefore, prior to the pilot customers were randomly assigned to providers.

### *Targeting Services to Work First Participants*

Michigan's Work First program, as well as many other welfare-to-work programs, provided basic instruction in job search techniques and minimal assistance in contacting employers. All Work First participants, regardless of their qualifications and work experience, were required to participate in these services. Research, however, has shown that the benefits from these basic services vary across the welfare population and that this variation depends to a large extent on an individual's characteristics, past work experience, and welfare dependence.<sup>13</sup> Therefore, targeting services to the specific needs of participants instead of pursuing a one-size-fits-all approach opened the possibility of improving the effectiveness of welfare-to-work programs and of helping states make more efficient use of their resources.

The idea of targeting services to would-be welfare recipients who can benefit most from the assistance predated the welfare reform movement. Ellwood (1986) explored the possibility of using statistical means to identify individuals who are most likely to be long-term welfare recipients. He estimated recidivism rates and exit rates using the characteristics of individuals and their previous employment and welfare histories as predictors. Based upon his ability to identify those at risk of becoming long-term welfare recipients, he concluded that the effectiveness of welfare programs (in particular AFDC) could be enhanced by targeting services to welfare recipients with specific characteristics. Support for targeting is further found in the evaluations of previous programs and demonstrations that targeted the welfare population. Gueron and Pauly (1991) reviewed the evaluations of a host of programs, both broad-coverage and small and selective voluntary programs, in order to discern whether the effectiveness of the service components within these programs vary among participants. They concluded that the impacts do vary among participants and that they are larger for more disadvantaged recipients. For example, they cited an analysis of the Supported Work program that shows that services were more effective for women who had never worked and had been on welfare longer. The same pattern emerged from a reanalysis of Supported Work and the quasi-experimental studies of WIN and CETA on AFDC recipients in which welfare recipients with little or no recent work experience benefitted substantially more than did those with some recent work experience (Grossman, Maynard, and Roberts, 1985). Friedlander (1988), in an analysis of five selected welfare employment programs, identified additional characteristics that affected program effectiveness, including marital status, education, and the number and ages of children. While still important, they were less strongly related to future employment and welfare receipt than past employment and welfare experience.

The concept of targeting services can also be placed within current insights from behavioral economics. Babcock et al. (2012) point out that the tendency of individuals to have imperfect self control creates behavioral barriers to reemployment. These factors generate a situation where stronger work and job search requirements that appear overly paternalistic in the usual model of homo economics may actually benefit the program participants themselves by helping them overcome self-control problems (p. 4). The targeting approach is consistent with

---

<sup>13</sup>See Bloom and Michalopoulos (2001) for a synthesis of research on the effectiveness of welfare-to-work programs.

insights of behavioral economics by referring participants to service providers that offer services and a delivery-of-service philosophy that more closely matches the way in which specific individuals process information and respond to guidance (paternalistic approach versus a hands-off approach).

Once it was established that the effectiveness of programs varied by participants, the next issue regarding targeting services was the method by which various subgroups are identified. Identification methods vary widely, from using the subjective evaluations of staff to a much more objective assessment based on statistical techniques. The effectiveness of those approaches varied as well. Gueron and Pauly (1991) cited two studies that used the perceptions and knowledge of staff about their clients to refer them to services. The first case was a study of AFDC recipients in Louisville who were randomly assigned to participate in job club activities. Before they entered the job club, staff counselors rated them on job-readiness based upon their perception of the client's motivation and skills. The study analyzed the relationship between these initial ratings and participants' performance in the program, such as attending job club or dropping out, and finding a job during job club. The study found that there was no relationship between the job-readiness ratings and those measures of performance. In a second study cited by Gueron and Pauly, intake workers in the Homemaker-Home Health Aide Demonstrations rated the job readiness of clients by supplementing their own perceptions with quantitative information, such as a client's education, work experience, and other personal attributes. Even though this additional information was considered, the study found that the staff based their ratings primarily on perceptions, with only a weak relationship between a client's intake information and the ratings. The study further reported that although the ratings were correlated with post-program performance outcomes they did not help to distinguish the success of program participants from those in the control group.

## **Components of the Intervention**

### ***The Work First Pilot Project's Statistical Assessment and Referral System***

The Kalamazoo-St. Joseph pilot incorporated the statistical assessment and referral system into the initial intake and orientation process. Each welfare recipient who enrolled in Work First was immediately assigned a score indicating his or her probability of finding employment. The score provided an assessment of each participant's need for services, based upon the past experience of local Work First participants like themselves. A high score indicated that a person had little need of services, since past participants with the same set of characteristics had a high probability of finding a job without much if any intervention. Those with a low score required more services, since past recipients with similar attributes had less success in finding and retaining employment. Each participant was then referred to one of three subcontractors based on their employability scores.

This statistical assessment model was based on the outcomes of participants entering the program during 1996. Table 1 displays the characteristics of Work First participants who enrolled in the program in 1996. Participants were predominantly single parents who had not completed high school and who had been on welfare for less than 36 months during the last five



years. Some of the participants had completed a general equivalency diploma (GED), but few received vocational training.

Data were obtained from the intake forms and the tracking system developed and maintained by the Kalamazoo/St. Joseph WDB. For most participants, multiple activities were recorded. The type of activity, the number of hours engaged in each activity, and the starting and ending dates of each activity were included in the files. Consequently, it was possible to piece together a sequence of activities between the time participants entered and left the program.

**Table 1 Variables Used in the Work First Statistical Assessment Model**

| Name         | Description   | Mean  |
|--------------|---|-------|
| sglprnt      | =1 if single parent   | 0.827 |
| age          | Age at time of enrollment   | 29.7  |
| age2         | Age squared   |       |
| noschl       | no formal schooling   | 0.038 |
| grlt9        | grade level completed less than 9 <sup>th</sup> grade   | 0.056 |
| gr9          | completed 9 <sup>th</sup> grade   | 0.056 |
| gr10         | completed 10 <sup>th</sup> grade  | 0.089 |
| gr11         | completed 11 <sup>th</sup> grade  | 0.191 |
| gr12         | completed 12 <sup>th</sup> grade (omitted from analysis, thus reference)  | 0.387 |
| post1        | completed one year of postsecondary   | 0.012 |
| post2        | completed two years of postsecondary  | 0.016 |
| post3        | completed three years of postsecondary  | 0.004 |
| post4        | completed four years of postsecondary   | 0.001 |
| ged          | earned GED certification  | 0.161 |
| YOU          | Youth Opportunities Unlimited   | 0.189 |
| goodwill     | Goodwill Industries   | 0.179 |
| foundat      | Behavioral Foundation   | 0.303 |
| comstock     | Comstock  | 0.045 |
| sturgis      | Sturgis   | 0.040 |
| rivers3      | Three Rivers  | 0.240 |
| voced        | attended postsecondary vocational education program   | 0.014 |
| notarget     | not a target group, which includes AFDC received any 36 of preceding 60 months, youngest child 16–18, or custodial parent under 24 and who has not completed high school or with little or no work experience | 0.528 |
| AFDC36       | received AFDC any 36 of preceding 60 months   | 0.343 |
| code20_1     | qualified unsubsidized employment prior to assignment   | 0.190 |
| code20_2     | qualified unsubsidized employment prior to assignment in previous enrollment  | 0.003 |
| nocmpl       | terminated as noncompliant in previous enrollment (code 59, 60, or 61)  | 0.057 |
| employed     | terminated as employed in qualified unsubsidized job  | 0.427 |
| Observations |   | 1,546 |

SOURCE: Author's calculations of Kalamazoo-St. Joseph Work First Administrative data, 1996-1997.

Results of the logit estimation are shown in Table 2. Focusing on the signs of the statistically significant coefficients, Work First participants are more likely to complete 90 consecutive days of employment if they had completed 12<sup>th</sup> grade (the omitted variable in the equation), were older, were employed prior to first assignment, enrolled in the program earlier in

the year rather than later, and were not out of compliance if they had previously enrolled in Work First.<sup>14</sup>

**Table 2. Logit Estimates of the Basic Statistical Assessment Model**

| Logit Estimates             |             | Number of obs=1,546    |        |       |                 |           |
|-----------------------------|-------------|------------------------|--------|-------|-----------------|-----------|
| Log Likelihood = -948.47621 |             | $\chi^2(23)=213.10$    |        |       |                 |           |
|                             |             | Prob > $\chi^2=0.0000$ |        |       |                 |           |
|                             |             | Pseudo $R^2=0.1010$    |        |       |                 |           |
| Employed                    | Coefficient | Standard Error         | z      | P> z  | [95% Confidence | Interval] |
| sglprnt                     | 0.223       | 0.156                  | 1.429  | 0.153 | -0.083          | 0.528     |
| age                         | 0.115       | 0.041                  | 2.790  | 0.005 | 0.034           | 0.196     |
| age2                        | -0.002      | 0.001                  | -2.602 | 0.009 | -0.003          | -0.000    |
| noschl                      | -1.801      | 0.555                  | -3.244 | 0.001 | -2.889          | -0.713    |
| grlt9                       | -0.454      | 0.304                  | -1.495 | 0.135 | -1.049          | 0.141     |
| gr9                         | -0.167      | 0.252                  | -0.662 | 0.508 | -0.661          | 0.327     |
| gr10                        | -0.775      | 0.218                  | -3.553 | 0.000 | -1.203          | -0.348    |
| gr11                        | -0.431      | 0.157                  | -2.744 | 0.006 | -0.739          | -0.123    |
| ged                         | 0.174       | 0.162                  | 1.074  | 0.283 | -0.143          | 0.492     |
| voced                       | -0.591      | 0.487                  | -1.212 | 0.225 | -1.546          | 0.364     |
| post1                       | 0.079       | 0.501                  | 0.159  | 0.874 | -0.903          | 1.062     |
| post2                       | 0.162       | 0.438                  | 0.371  | 0.711 | -0.695          | 1.020     |
| post3                       | 0.011       | 0.884                  | 0.013  | 0.990 | -1.721          | 1.744     |
| goodwill                    | -0.463      | 0.187                  | -2.485 | 0.013 | -0.829          | -0.098    |
| foundat                     | -0.560      | 0.164                  | -3.406 | 0.001 | -0.883          | -0.238    |
| sturgis                     | 0.005       | 0.300                  | 0.017  | 0.986 | -0.582          | 0.593     |
| comstock                    | 0.127       | 0.302                  | 0.421  | 0.673 | -0.465          | 0.719     |
| rivers3                     | -0.454      | 0.172                  | -2.641 | 0.008 | -0.791          | -0.117    |
| notarget                    | 0.064       | 0.116                  | 0.555  | 0.579 | -0.163          | 0.292     |
| addate                      | -0.003      | 0.001                  | -5.424 | 0.000 | -0.004          | -0.002    |
| code20_1                    | 1.107       | 0.144                  | 7.683  | 0.000 | 0.825           | 1.390     |
| code20_2                    | -0.393      | 1.055                  | -0.373 | 0.709 | -2.46           | 1.674     |
| nocmpl                      | -0.750      | 0.281                  | -2.672 | 0.008 | -1.301          | -0.200    |
| _cons                       | 36.921      | 7.260                  | 5.086  | 0.000 | 22.693          | 51.150    |

SOURCE: Author's calculations of Kalamazoo-St. Joseph Work First Administrative data, 1996-1997.

The only variable that may need an explanation for its inclusion in the model is the date of admission into Work First. The coefficient on this variable is negative and statistically significant. Therefore, those who enrolled in Work First in more recent periods experienced a lower probability of finding and maintaining employment for 90 consecutive days. The

<sup>14</sup>These results are consistent with previous studies that examine the employment prospects of welfare recipients. Estimates based on the national SIPP survey found that education and prior employment history were important determinants of the likelihood of leaving welfare for employment (see Eberts 1997, Appendix). A study for the State of Texas also found these factors to be important (Schexnayder, King, and Olson 1991). The Texas study also found that the number of children, the age of the welfare recipient, the duration on welfare, and the use of the employment service and participation in job training programs also affected the likelihood of employment in the expected direction. The employment- and training-related results from Texas are consistent with our results from Work First that prior employment and compliance with previous Work First enrollment positively affect the likelihood of qualified employment.

percentage of Work First participants reaching this status steadily declined from the first quarter of 1996, when the sample began. During the first and second quarters of 1996, 53 percent of participants in the sample were employed for 90 days, after which the percentage dropped to 50 percent during the third quarter, 31 percent during the fourth quarter, and 24 percent during the first quarter of 1997. The admission date variable can be interpreted as a proxy for attributes of Work First participants that are not captured in the characteristics included in the model. Work First staff observed that as the pool of welfare recipients going through the program diminished, enrollees were increasingly less qualified to find and hold jobs. The variable may also capture changes in the program and changes in local labor market conditions over time

Applying the estimated coefficients to the characteristics associated with each Work First participant yields predictions of the probability of employment for each individual. Consequently, each Work First enrollee can be ranked according to this estimated probability.<sup>15</sup> For heuristic purposes, one can view the distribution of employability scores as representing participants lined up to enter the Work First program according to their probabilities of finding employment. If the door is envisioned to be on the left side of the graph in Figure 1, those with the least propensity to find a job are at the front of the line, and the participants with the highest propensity are at the end of the queue. According to our model, the estimated probabilities of employment range from a low of 0.02 to a high of 0.90. Therefore, the person at the head of the line has almost no chance of finding a job and would need considerably more assistance than the person at the end of the line, who is almost certain to find employment without much help. Although 43 percent of the Work First participants in the sample found employment, the model did not assign anyone a probability of 100 percent. However, the spread is quite large, spanning most of the range from zero to one.

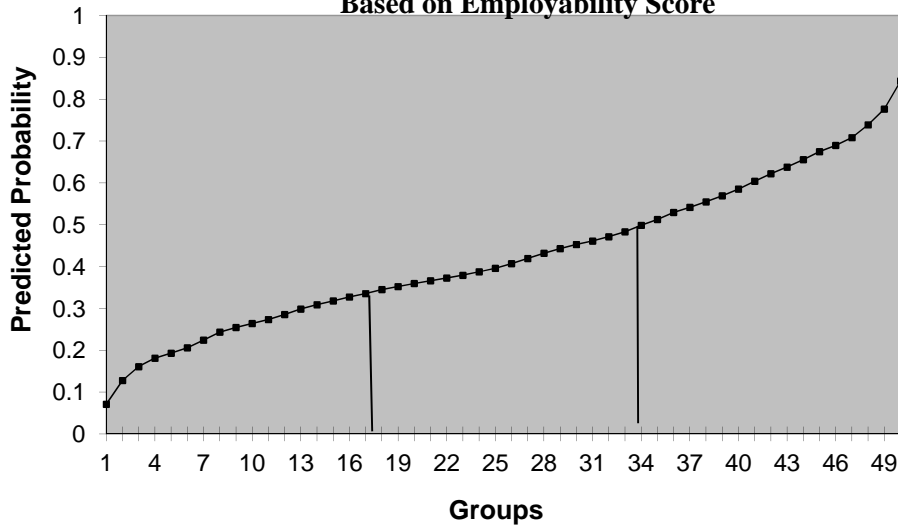
The employability score is used to categorize participants according to their perceived difficulty in finding employment. Participants are divided into three groups—low-level employability, middle-level employability, and high employability—as shown in figure 1. Within each of the three employability categories, individuals are then randomly assigned to one of the three providers. This approach allows the assessment of which of the three providers yields the best employment outcomes for participants of different employability levels.<sup>16</sup>

---

<sup>15</sup>Several criteria can be used to judge the utility of the model in its ability to distinguish among Work First participants as to their likelihood of finding employment. Two measures are considered here: 1) the relative steepness of the distribution of each individual's employment probabilities; and 2) the width of the confidence intervals. The model satisfies both criteria, as described in Eberts (2001).

<sup>16</sup>There is another component of the evaluation that was included in the original study but for simplicity will not be included in this description. In addition to using random assignment to determine the optimal combination of service providers across the three employability categories, the WDB staff, prior to the start of the pilot, came up with their own determination, based on their own assessment, of the optimal combination of services. This was run through the random assignment evaluation. It turned out that the staff chose the optimal combination, as determined by the random assignment results.

**Figure 1**  
**Referral of Participants to Providers**  
**Based on Employability Score**



*Accounting for Differences in Service Activities Among Providers*

The Kalamazoo/St. Joseph WDB contracted with three organizations to provide employment services to participants of the Work First program. The providers delivered services that met state and federal requirements regarding content and duration. However, there was some flexibility within the requirements. WDB staff observed that providers differed in their style and philosophies in delivering services and in the number of hours in which participants were engaged in specific activities. These observed differences were critical to the pilot by providing the opportunity to refer participants to the provider, and thus the mix and style of services, which best met their needs. Prior to the pilot, participants were assigned to the three service providers on a random basis, since staff had no meaningful way to assign customers.

Work First participants engaged in a variety of activities as part of their requirement for successfully participating in the program. Most participants began with assessment and employability planning (code 12). As shown in Table 3, 83 percent of all participants received those services in 1996. The percentage was higher for those who were not employed prior to entering Work First, about 90 percent. Around half the participants engaged in group or individual job-search assistance, which included counseling, job-seeking skills training, and may have included support on a one-to-one basis (code 13). Fifty-three percent were employed in a job (code 1) that paid minimum wage or higher and the employment was for 20 hours or more per week (or 35 hours if a working spouse). Another six percent were employed in unsubsidized employment that did not meet the requirements of code 1. Nineteen percent of the participants were in unsubsidized employment when referred, obtained subsidized employment meeting the requirements of code 1 prior to reporting, or obtained the appropriate employment prior to

reporting to the first activity. Only a handful of participants (two percent) were referred to community service programs or vocational educational training.

**Table 3 Selected Activities of Work First Programs**

| Activity                              | Code | Mean | Standard deviation | Minimum | Maximum |
|---------------------------------------|------|------|--------------------|---------|---------|
| Unsubsidized employment               | (01) | 0.53 | 0.50               | 0       | 1       |
| Job readiness                         | (10) | 0.09 | 0.28               | 0       | 1       |
| Assessment and employability planning | (12) | 0.83 | 0.37               | 0       | 1       |
| Job search                            | (13) | 0.55 | 0.50               | 0       | 1       |
| Part-time employment                  | (19) | 0.06 | 0.24               | 0       | 1       |
| Employment prior to assignment        | (20) | 0.19 | 0.39               | 0       | 1       |
| Community service                     | (33) | 0.01 | 0.11               | 0       | 1       |
| Voc. ed. training                     | (34) | 0.01 | 0.09               | 0       | 1       |

SOURCE: Author's calculations of Kalamazoo-St. Joseph Work First Administrative data, 1996-1997.

The length of time Work First enrollees engaged in activities varied by type of activity and by subcontractor. For example, as shown in Table 4, 38.1 percent of the participants spent two hours in the assessment and employability planning activity, while 39.6 percent spent 20 hours in the same activity. Of the three subcontractors within the Kalamazoo area, YOU averaged 7.3 hours, Behavioral Foundation 11.2 hours, and Goodwill 16.0 hours in this activity. The higher average for Goodwill results from a much larger percentage of participants spending time in the services than those assigned to other providers. Over three-quarters of those going to Goodwill spent 20 hours in this service. Only 27 percent of the participants receiving services from either YOU or the Foundation received 20 hours of this service. For those going to YOU, two-thirds of the participants received two hours or less of assessment and planning. Hours spent in this activity for those receiving services from the Foundation were split between 2, 15 or 16, and 20 hours. The wide distribution may indicate that these individuals have more discretion in how much time they spend in various activities.

Providers also differed in their approaches to delivering services. For instance, one provider stressed a goal-oriented approach to job search, requiring that participants call a given number of employers each day until they found a job. Another provider offered more assistance to customers in conducting phone inquiries and interviewing for jobs. Staff would work directly with customers to show them how to find employment postings and telephone numbers, how to inquire about the job posting, and how to present themselves during interviews. This same organization would also provide more intensive training at times to those who were not able to find a job during their initial several weeks in the program.

**Table 4 Distribution of Hours Engaged in Assessment and Employability Planning**

| Hours | Percent |            |          |      |
|-------|---------|------------|----------|------|
|       | All     | Foundation | Goodwill | YOU  |
|       | 5.9     | 1.9        | 1.9      | 14.6 |
| 2     | 38.1    | 38.3       | 19.0     | 52.8 |
| 3     | 0.2     | 0.5        | 0.0      | 0.0  |
| 4     | 0.4     | 0.5        | 0.5      | 0.0  |
| 5     | 0.1     | 0.0        | 0.5      | 0.0  |
| 6     | 0.1     | 0.0        | 0.0      | 0.4  |
| 7     | 0.0     | 0.0        | 0.0      | 0.0  |
| 8     | 0.1     | 0.0        | 0.0      | 0.4  |
| 9     | 0.0     | 0.0        | 0.0      | 0.0  |
| 10    | 0.1     | 0.0        | 0.0      | 0.4  |
| 11    | 0.7     | 0.0        | 0.5      | 1.9  |
| 12    | 0.2     | 0.5        | 0.0      | 0.0  |
| 13    | 0.0     | 0.0        | 0.0      | 0.0  |
| 14    | 0.2     | 0.5        | 0.0      | 0.0  |
| 15    | 4.8     | 11.1       | 0.0      | 0.0  |
| 16    | 9.3     | 19.6       | 0.9      | 1.9  |
| 17    | 0.0     | 0.0        | 0.0      | 0.0  |
| 18    | 0.0     | 0.0        | 0.0      | 0.0  |
| 19    | 0.0     | 0.0        | 0.0      | 0.0  |
| 20    | 39.6    | 26.9       | 76.8     | 27.7 |

SOURCE: Author's calculations of Kalamazoo-St. Joseph Work First Administrative data, 1996-1997.

## **Evaluation of the Kalamazoo/St. Joseph Work First Profiling Pilot**

### *Design of the Evaluation*

The Kalamazoo/St. Joseph Work First profiling pilot was evaluated using a random assignment approach. The evaluation included participants who entered the program from March 1998 to March 2000. During the two-year period, nearly 3,600 welfare recipients who were single parents were assigned to the three providers serving the Kalamazoo area.<sup>17</sup>

The computerized intake process was designed so that welfare recipients referred to Work First from The Family Independence Agency (FIA) were randomly assigned to various groups. The random assignment procedure took place in three steps. First, participants were divided into one of three groups, depending upon their employability score. Assignment of participants to the three employability groups was based on their relative ranking in the distribution of employability scores of those who enrolled in Work First at that session. It was not based on a predetermined cutoff value. Those participants with employability scores in the lowest 40 percent of the distribution were assigned to the low employability group (L), the next 20 percent were assigned to the middle group (M), and the highest 40 percent were assigned to

<sup>17</sup>About half the participants went through the program at least twice. For purposes of the evaluation, we included only the last time the person appeared in the program, if they appeared more than once. We adopted this approach to avoid biasing the evaluation toward multiple enrollees. One could argue that including the same person more than once in the evaluation overweights that person's experience relative to those who entered the program only once. More will be said about this approach in a subsequent section.

the high group (H). The middle group included only 20 percent of the participants because the treatment provider for that group, YOU, could accommodate only that percentage because of capacity constraints.<sup>18</sup> Second, those within each employability group were randomly assigned to one of the three providers. The number of participants in each group is displayed in Table 5.

---

<sup>18</sup>The actual assignment of employability scores was slightly different from the way in which the statistical assessment model was originally estimated. The model was estimated based on the entire set of individuals who participated in and completed the program during a year's time. The computation of the employability score, based on the coefficients from the model, was done at each intake and orientation session. These sessions took place twice a week. Obviously, only a small number of people who participate in the program each year attend each session.

Because of the small number of participants at each session, it may be the case that individuals in attendance on any given day were not fully representative of the Work First population. In examining the distribution of employability scores for each session, we found that on some days the employability scores would cluster on the high side, while on other days they would center on the low side of the distribution. Since the cutoffs were determined by dividing the distribution of scores of individuals who showed up on a given day, it could be the case that individuals with lower than average employability scores were assigned to the "high" employability group while on another day individuals with higher than average employability scores were assigned to the "low" employability group. It depends upon who was referred to a particular session.

Another difference between the employability scores as originally estimated and those assigned to participants during the pilot was the magnitude of the score. We recognized that the employability scores declined over the year in which the statistical assessment model was estimated. This relationship was consistent with the general observation by the WDB staff that as an increasing number of Work First participants found jobs, those remaining would have lower skills and be harder to place into jobs and more difficult to serve. To account for this trend, we included in the model the date that the participant enrolled in the program. The coefficient on this variable (adate), as shown in Table 4, was relatively large and highly statistically significant. The value of the coefficient (-0.003) was large relative to the mean of the variable (approximately 14460, which is the date expressed in machine language).

However, it turns out that as time increased from the date in which the model was estimated to when it was used to assign the employability scores, the coefficient played a much larger role in determining the size of the predicted value. The mean value of the employability score fell from about 0.30 in the original model to 0.05 in the evaluation. Most of the difference is due to the more advanced date. When the date is rolled back to its average value during the period in which the model was estimated, the mean employability score for the sample used in the evaluation increases to 0.46.

Further investigation shows that the rank ordering of employability scores computed with and without the adjustment for the time is highly correlated. The correlation coefficient of the actual employability score assigned to participants during the evaluation and the hypothetical one when the date of enrollment is rolled back by two years is 0.82.

**Table 5 Number of Participants Assigned to Each Provider**

## Employability Groups

| Provider   | Low | Middle | High | Total |
|------------|-----|--------|------|-------|
| Goodwill   | 144 | 73     | 164  | 381   |
| Foundation | 177 | 83     | 211  | 471   |
| YOU        | 59  | 26     | 54   | 140   |
| Total      | 380 | 183    | 429  | 992   |

SOURCE: Author's calculations of Kalamazoo-St. Joseph Work First

The primary outcome measure for the evaluation is the retention rate; that is, whether or not the participant was employed 90 consecutive days. Table 6 shows the retention rates of those in the control and treatment groups by employability group and provider. In this case, there is considerable variation both across groups and within groups. Note that the actual retention rate averaged for each group increases from the lowest employability group to the highest. For the control group, it increases from 11.6 percent for the lowest group to 21.7 percent for the highest employability group. The treatment group also follows the pattern of increasing retention rates from low to high employability groups. The same monotonic increase is exhibited for each provider except YOU. However, as shown in Table 6, the upper and lower bounds of the 95 percent confidence intervals overlap across the various groups.<sup>19</sup>

**Table 6 Upper and Lower Bounds of the 95 Percent Confidence Intervals for the Retention Rates of Each Provider**

|            | Employability Groups |       |       |        |       |       |       |       |       |
|------------|----------------------|-------|-------|--------|-------|-------|-------|-------|-------|
|            | Low                  |       |       | Middle |       |       | High  |       |       |
|            | Lower                | Mean  | Upper | Lower  | Mean  | Upper | Lower | Mean  | Upper |
| Goodwill   | 0.094                | 0.153 | 0.212 | 0.124  | 0.219 | 0.314 | 0.162 | 0.226 | 0.290 |
| Foundation | 0.039                | 0.079 | 0.119 | 0.069  | 0.145 | 0.221 | 0.167 | 0.223 | 0.279 |
| YOU        | 0.049                | 0.136 | 0.223 | 0.188  | 0.370 | 0.552 | 0.068 | 0.167 | 0.266 |

SOURCE: Author's calculations of Kalamazoo-St. Joseph Work First Administrative data, 1996-1997.

<sup>19</sup>The overlap is not as great between the low and middle employability groups as it is between the middle and high groups. The difference in the average retention rates for the low and middle employability groups is statistically significant at the 95 percent significance level. On the other hand, the difference in the average retention rates for the middle and high employability groups is not.



## ***Retention Rates by Various Combinations of Providers***

In order to determine whether different combinations of assignments of employability groups to service providers yield different outcomes, we examined six combinations.<sup>20</sup> The effects of the various combinations are measured by computing the number of participants within each employability group who retained their jobs if everyone in that group received services from the same provider. To illustrate this approach, consider the first combination listed in Table 7. The designation “gyk” refers to the combination in which all participants in the low employability group (the left-most group in Table 5) is hypothetically assigned to Goodwill (g); all participants in the middle employability group are assigned to YOU (y); and all participants in the high employability group are assigned to Behavioral Foundation (k). Since participants in the control group were randomly assigned to each of the providers within each of the three employability groups, using the subgroup assigned to a particular subcontractor to represent the effects for everyone in that employability group is a sound approach.

Using this approach, the appropriate retention rate for each employability group is multiplied by the total number of participants in the control group to compute the number of participants within that group who retained their job for 90 consecutive days. For instance, for the first combination, the retention rate of 0.153 for Goodwill is multiplied by 380 (see Table 5), the size of the control for the low employment group. This yields 58, which indicates that 58 participants in the control group of the low employability group would have retained their jobs if all were assigned to Goodwill. The same calculation is performed for the middle group, multiplying 0.380 by 183 which yields 68, and for the high group, multiplying 0.223 by 429, which yields 96. Summing these three numbers yields the total number of participants in the three control groups who retained their jobs, 222. Dividing by the total number of participants in the control groups results in the hypothetical retention rate if the combination “gyk” were used to assign participants.

**Table 7 Number of Participants Employed 90 Consecutive Days by Combination of Providers**

| Combination<br>of providers | Employability group |        |      | Total | Ranking |
|-----------------------------|---------------------|--------|------|-------|---------|
|                             | Low                 | Middle | High |       |         |
| 1 gyk                       | 58                  | 68     | 96   | 222   | 1       |
| 2 gky                       | 58                  | 26     | 72   | 156   | 5       |
| 3 ygk                       | 52                  | 40     | 96   | 188   | 3       |
| 4 ykg                       | 52                  | 26     | 97   | 175   | 4       |
| 5 kyg                       | 30                  | 68     | 97   | 195   | 2       |
| 6 kgy                       | 30                  | 40     | 72   | 142   | 6       |

NOTE: Providers are designated as letter: “g” = Goodwill; “k” = Foundation; “y” = YOU. The combination “gyk” refers to the low employability group assigned to Goodwill, the middle employability group to YOU, and the high employability group to the Foundation.

SOURCE: Author’s calculations of Kalamazoo/St. Joseph Work First administrative data, 1996-1997.

<sup>20</sup>More than six combinations are possible with three providers and three groups by assigning more than one employability group to a provider. However, we adhered to the WDB’s contractual arrangement during the pilot that all three providers delivered services. Therefore, we eliminated from consideration combinations that assigned two or three groups to one service provider.

Performing these calculations for all six combinations provides a convenient measure of the effectiveness of the various combinations. As shown in Table 7, the number of retentions ranges from a high of 222 for the combination “gyk” to a low of 141 for “kgy”. The difference between the highest and lowest is 79 retentions, or 56 percent. The difference between the highest number and the average is 44, or 25 percent. The results indicate that using the statistical tool to assess and refer Work First participants can increase the effectiveness of the program without increasing cost. The optimal combination of providers “gyk” yields a 25 percent higher retention rate than if the participants were randomly assigned to the providers.

Differences between any of the various pairs of combinations are statistically significant at the 95 percent significance level. Table 8 displays the difference in the retention rates and the  $t$  statistics for each pair of combinations. For instance, the difference between the retention rate for combination “gyk” and for combination “gky” is 0.066 (e.g.,  $65 \div 992$ ). The  $t$  statistic for this pair is 5.26, which is much greater than the critical value of 1.96 for a 95 percent significance level. Note that 10 out of the possible 15 pairs are statistically significant. Only those with differences in the retention rates of less than two percentage points (approximately 20 participants out of 992) are not statistically significant.

Based upon the analysis of the effectiveness of the combinations of providers, it appears that Goodwill had a comparative advantage in serving low employability participants, YOU in serving middle employability participants, and Behavioral Foundation in serving high employability customers. This combination of assignments was the same as the treatment group, which was determined by staff knowledge of the approaches taken by each provider and an analysis of welfare recipients who had participated in the program before the pilot began. However, it is beyond the scope of the pilot to determine the specific aspects of each provider’s approach that led to this outcome.<sup>21</sup>

---

<sup>21</sup>As previously noted, the retention rate for those in the middle employability control group assigned to YOU is higher than the rate for the treatment group assigned to YOU. If, as intended, individuals were randomly assigned to the treatment and control groups, and those within the control group were randomly assigned to the providers, one would expect the two retention rates to be similar. We tried two alternative approaches of deriving retention estimates for the different combinations that may mitigate the problem. The first approach controlled for factors that could be responsible for the significant difference between the treatment and control groups assigned to a specific provider. One possible factor is the date in which participants enter the program. It could be the case that because of the small number enrolled during each session and the nonrandom nature of referrals from FIA, the time of enrollment may lead to these differences. The second method combined the outcomes of both control and treatment groups. In this way, we reduced the effect of the timing of enrollment by considering outcomes from both groups. Both approaches yield results that are similar to the original approach.

**Table 8 Differences in Retention Rates between Pairs of Combinations of Providers**

| Providers   |     | 1 | 2     | 3      | 4      | 5      | 6     |
|---|-----|---|-------|--------|--------|--------|-------|
| Differences in retention rates                        |     |   |       |        |        |        |       |
| 1   | gyk | — | 0.066 | 0.034  | 0.046  | 0.026  | 0.080 |
| 2   | gky |   | —     | -0.031 | -0.019 | -0.039 | 0.014 |
| 3   | ygk |   |       | —      | 0.012  | -0.008 | 0.045 |
| 4   | ykg |   |       |        | —      | -0.020 | 0.033 |
| 5   | kyg |   |       |        |        | —      | 0.053 |
| 6   | kgy |   |       |        |        |        | —     |
| <i>t</i> -Statistics of difference in retention rates |     |   |       |        |        |        |       |
| 1   | gyk | — | 5.260 | 2.671  | 3.654  | 2.028  | 6.487 |
| 2   | gky |   | —     | -2.603 | -1.618 | -3.245 | 1.244 |
| 3   | ygk |   |       | —      | 0.986  | -0.644 | 3.842 |
| 4   | ykg |   |       |        | —      | -1.630 | 2.860 |
| 5   | kyg |   |       |        |        | —      | 4.481 |
| 6   | kgy |   |       |        |        |        | —     |

NOTE: Standard deviation derived according to the following formula:

$$\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

where  $\hat{p} = \frac{x_1+x_2}{n_1+n_2}$ ;  $\hat{q} = 1 - \hat{p}$ ; and  $x_1$  and  $x_2$  are the number of successes in the samples of size  $n_1$  and  $n_2$ .

SOURCE: Author's calculations of Kalamazoo/St. Joseph Work First Administrative data, 1996-1997.

### **Benefit/Cost Analysis of the Statistical Assessment and Referral System<sup>22</sup>**

The benefits of using the statistical assessment and referral system can be quantified by taking into account the earnings received by those additional participants who retained their jobs. As shown in the previous section, the optimal assignment rule yielded a net increase of 47 participants who retained their jobs 90 consecutive days over the number retaining their jobs in the group created by random assignment. Consequently, the net effect of the statistical assessment and referral system is computed by considering the difference in retention rates and earnings of the two groups. A benefit-to-cost ratio is then calculated by dividing the net effect by the cost of the pilot.<sup>23</sup>

The earnings are comprised of two components: the number of participants who retained their jobs (R) and the average weekly earnings of participants in that group during the 90 days (calculated here as 13 weeks) of employment (E). As shown in Table 9, the average weekly earnings of those in the optimal assignment group (denoted by subscript T) is \$192 and of those

<sup>22</sup>I wish to thank Kevin Hollenbeck and Jeff Smith for suggestions on conducting the benefit/cost analysis.

<sup>23</sup>The social value of the new system may be less than the value computed here because of displacement effects among the welfare population. It is conceivable that the additional retention by participants of the program with the new system may displace other welfare recipients from their existing jobs or preclude new Work First participants from finding jobs since the additional retentions reduce the job vacancies.

in the randomly assigned group (denoted by subscript C) is \$195, which suggests that the increase in employment retention from the optimal combination of service providers did not reduce significantly the earnings obtained. The difference in earnings of the two groups ( $B_T - B_C$ ) can be decomposed in the following way, using the control group as the base of comparison:

$$B_T - B_C = [(R_T - R_C) * E_C] + [(E_T - E_C) * R_C] + [(R_T - R_C) * (E_T - E_C)]$$

**Table 9 Average Weekly Earnings by Different Combinations of Providers**

| Combination of providers     | Average weekly earnings (\$) |
|------------------------------|------------------------------|
| gyk (optimal retention rate) | 192                          |
| gky                          | 211                          |
| ygk                          | 181                          |
| ykg                          | 175                          |
| kyg                          | 165                          |
| kyg                          | 189                          |
| Randomly assigned            | 195                          |

NOTE: Providers are designated by letters: “g” Goodwill; “k” Foundation; and “y” YOU. The combination “gyk” refers to the low employability group assigned to Goodwill, the middle employability group to YOU, and the high employability group to Foundation.

This decomposition yields the net effect in terms of additional earnings to program participants as a result of the statistical assessment and referral system. It is assumed here that the earnings difference continues for eight quarters, with two possible scenarios considered. The first scenario assumes that the difference in the number of participants retaining their jobs for 90 days persists throughout the 8 quarters. The second scenario assumes that the difference in job retention narrows throughout the eight-quarter period until the two series are equal. In both scenarios, wages are assumed to grow by 3 percent per year, and a 10 percent annual discount rate is used when computing the net present value of the earnings streams. As shown in Tables 10 and 11, under the first scenario, the net present value of the difference in the earnings streams of the optimal combination and the randomly assigned is \$840,827; under the second scenario, it is \$471,054.

**Table 10 Difference in Earnings between Treatment and Control Groups and Benefit-to-Cost Ratio of the System**

| Quarters after leaving program | Optimal group earnings minus randomly assigned group earnings (\$) |                           |
|--------------------------------|--|---------------------------|
|                                | No narrowing of earnings gap                                       | Narrowing of earnings gap |
| 1                              | 112,179  | 112,179                   |
| 2                              | 113,666  | 98,706                    |
| 3                              | 115,165  | 85,073                    |
| 4                              | 116,675  | 71,279                    |
| 5                              | 118,197  | 57,321                    |
| 6                              | 119,730  | 43,197                    |
| 7                              | 121,274  | 28,906                    |
| 8                              | 122,830  | 14,445                    |
| Net present value (\$)         | 840,827  | 471,054                   |
| Program cost (4)               | 145,000  | 145,000                   |
| Benefit-to-cost ratio          | 5.8  | 3.3                       |

NOTE: The first column of earnings assumes that the retention rates remain the same throughout the eight-quarter period while the average weekly earnings converge. The second column of earnings assumes that they converge until they are equal in the ninth quarter. Wages are assumed to increase 3 percent per year, and a 10 percent discount rate is assumed for the net present value calculation.

**Table 11 Difference in Earnings between Treatment and Control Groups and Benefit-to-Cost Ratio of the System, Assuming Retention Rates Converge**

| Quarters after leaving the program | $B_T - B_C$ | $R_T$ | $R_C$ | $E_T$    | $E_C$    |
|------------------------------------|-------------|-------|-------|----------|----------|
| 1                                  | \$112,179   | 222   | 175   | \$192.00 | \$195.00 |
| 2                                  | \$98,706    | 216   | 175   | \$193.44 | \$196.08 |
| 3                                  | \$85,073    | 210   | 175   | \$194.89 | \$197.18 |
| 4                                  | \$71,279    | 204   | 175   | \$196.35 | \$198.28 |
| 5                                  | \$57,321    | 198   | 175   | \$197.83 | \$199.39 |
| 6                                  | \$43,197    | 193   | 175   | \$199.31 | \$200.51 |
| 7                                  | \$28,906    | 187   | 175   | \$200.80 | \$201.63 |
| 8                                  | \$14,445    | 181   | 175   | \$202.31 | \$202.77 |
| Net present value                  | \$471,054   |       |       |          |          |
| Program Cost                       | \$145,000   |       |       |          |          |
| Benefit-to-cost Ratio              | 3.3         |       |       |          |          |

Note: This calculation of net impact and benefit-to-cost ratio assumes that the retention rates and the average weekly earnings converge during the eight-quarter period. Wages are assumed to increase 3 percent per year and a 10 percent discount rate is assumed for the net present value calculation.

The additional costs incurred to develop and operate the statistical assessment and referral system for the two-year life of the pilot totaled \$145,000. This expense included designing and integrating the system into the existing Work First program, which cost roughly \$105,000, and hiring a part-time person to administer the system during the intake and orientation process, which amounted to another \$40,000 during the two-year period. Dividing the net present value for each scenario by the program costs of \$145,000, yields a benefit-to-cost ratio for the first scenario of 5.8 and a ratio for the second scenario of 3.3.

## **Summary**

The Obama Administration has emphasized the use of low-cost interventions based on insights from behavioral economics to increase access to social programs and improve government efficiency. Several initiatives are underway but few directly target workforce programs. This paper reports on a USDOL-funded pilot conducted in the late 1990s to illustrate how low-cost interventions in workforce programs can improve employment outcomes of participants. The pilot is relevant for the current interest in low-cost RCT trials by demonstrating how a simple improvement in the referral of participants to services can improve outcomes, how RCT can be embedded in the existing program, and how administrative data can be used to minimize the cost and disruption of the evaluation.

The purpose of the Work First pilot was to determine the benefits of using a statistical assessment tool to target employment services to meet the needs of Work First participants more effectively. This approach is consistent with insights of behavioral economics by referring participants to service providers that offer services and a delivery-of-service philosophy that more closely matches the way in which specific individuals process information and respond to guidance (paternalistic approach versus a hands off approach). The statistical assessment tool estimated the probability that a participant would be employed for 90 consecutive days by relating this outcome to the personal characteristics and work history of former Work First participants. Estimates were based on administrative records of welfare recipients who had participated in the Work First program prior to the time of the pilot.

The evaluation yielded the following results. First, the statistical model exhibited sufficient precision to distinguish among participants according to their likelihood of working 90 consecutive days. Second, there was considerable variation in the retention rates among the various combinations of providers offering services to participants in the three employability groups, as identified by the assessment tool. The retention rate of the combination of providers that yielded the highest rate was 56 percent higher than the combination yielding the lowest rate, and 27 percent higher than if the participants were randomly assigned to providers. Third, the benefit-to-cost ratio of the pilot project ranged from 3.3 to 5.8, depending upon assumptions regarding the persistence over time of the earnings differences between the treatment and control groups.

## References

- Babcock, Linda, William J. Congdon, Lawrence F. Katz, and Sendhil Mullainathan. 2012. "Notes on Behavioral Economics and Labor Market Policy." *IZA Journal of Labor Policy* 1(1): 1-14.
- Eberts, Randall W. 1997. "The Use of Profiling to Target Services in State Welfare-to-Work Programs: An Example of Process and Implementation." Working paper, W.E. Upjohn Institute for Employment Research, Kalamazoo, MI.
- Eberts, Randall W. 2001. "Design, Implementation, and Evaluation of the Work First Profiling Pilot Project." Draft report prepared by the W.E. Upjohn Institute for Employment Research for the Employment and Training Administration, U.S. Department of Labor, March.
- Freedman, S., D. Friedlander, W. Lin, and A. Schweder. 1996. The GAIN Evaluation: Five-Year Impacts of Employment, Earnings, and AFDC Receipt. New York: Manpower Demonstration Research Corporation.
- Grossman, Jean, Rebecca Maynard, and Judith Roberts. *Reanalysis of the Effects of Selected Employment and Training Programs for Welfare Recipients*. Princeton, NJ: Mathematic Policy Research, Inc., 1985.
- Gueron, Judith, and Edward Pauly. 1991. *From Welfare to Work.*, New York: Russell Sage Foundation.
- Ellwood, D. 1986. "Targeting the 'Would-Be Long-Term Recipient: Who Should Be Served.'" Princeton: Mathematica Policy Research.
- Friedlander, Daniel. 1988. *Subgroup Impacts and Performance Indicators for Selected Welfare Employment Programs*. New York, NY: Manpower Demonstration and Research Corporation. "Notes on Behavioral Economics and Labor Market Policy."
- Schexnayder, Deanna, Christopher King, and Jerome Olson. 1991. "A Baseline Analysis of the Factors Influencing AFDC Duration and Labor Market Outcomes." Report to the Center for the Study of Human Resources and the Bureau of Business Research, The University of Texas at Austin.