What Do Welfare-to-Work Demonstrations Reveal to Welfare Reformers?

John V. Pepper
Department of Economics
University of Virginia
Charlottesville, VA 22903
jvp3m@virginia.edu

August 25, 1999

This research has been supported in part by the 1998-99 Joint Center for Poverty Research/ASPE Small Grants Program. I have benefitted from the comments of Stephen Bell, Leora Friedberg, Carolyn Heinrich, Matt Lyon, and Charles Manski. I have also benefitted from the opportunity to present this work at the 1999 JCPR Small Grants Conference and in seminars at Mathematica and the University of Virginia. I thank Jessica Sautner for research assistance. The data used in this paper are derived from files made available to researchers by the MDRC. The author remains solely responsible for how the data have been used and interpreted.

Abstract

What Do Welfare-to-Work Demonstrations Reveal to Welfare Reformers?

John V. Pepper

Under the new welfare system, states must design and institute programs that both provide assistance and encourage work, two objectives that have thus far appeared incompatible. Will states meet these new requirements? For many innovative programs, the randomized welfare-to-work experiments conducted over the last three decades may be the only source of observed data. While these experiments yield information on the outcomes of mandated treatments, the new regime permits states and localities much discretion. Using data from four experiments conducted in the mid-1980s, this study examines what welfare-to-work demonstrations reveal about outcomes when the treatments are heterogenous. In the absence of assumptions, these data allow us to draw only limited inferences about the labor market outcomes of welfare recipients. Combined with prior information, however, data from experimental demonstrations are informative, suggesting either that the long run federal requirements cannot be met or that these standards will only be met under special circumstances.

1. Introduction

In August 1996, the open-ended federal entitlement program of Aid to Families with Dependent Children was replaced with the Temporary Assistance for Needy Families (TANF) block grant for states. The Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (PRWORA) gives states broad discretion in defining eligibility and benefits for welfare programs. TANF funds, however, are not guaranteed. For state governments to be assured of continued full funding, at least 25 percent of single parent recipients must be participating in work activities in 1997, with this proportion increasing by 5 percent each year until it reaches 50 percent in 2002.[1]

Thus, under the new law states must design and institute welfare programs that both provide assistance and encourage work, two objectives that have so far appeared incompatible. Will states meet the federal requirements? To date, there is almost no empirical evidence to address this question.

Experimental analyses, however, are thought to provide relevant data. Over the last three decades, numerous randomized social experiments have examined whether work and training programs enable welfare recipients to participate in the economic mainstream. Typically, a sample of welfare recipients is randomly placed in either a treatment group which receives job training or a control group which receives the standard benefits. At some point in the future, we then observe the labor market outcomes of each individual. For many innovative programs, these randomized experiments may be the only source of observed data.

What do these data reveal? With random assignment, the sampling process identifies the fraction of welfare recipients who would work if they were all to receive training or if instead they were all to receive the standard benefits. By comparing these probabilities, researchers measure the effectiveness of the training program. In general, this literature suggests that basic training programs slightly increase labor force participation, hours worked, and wages, and decrease the likelihood of prolonged participation on welfare.

While experiments yield information on the outcomes of mandated treatments, the new regime permits states and localities much discretion. In fact, with the budget constraints imposed by block grant funding, state governments are unlikely to mandate training for all welfare recipients. Rather, some will be assigned to training, while others will receive standard benefits.

Still, these experiments might be informative. Using data from four well-known social experiments conducted by the Manpower Demonstration Research Corporation (MDRC), I examine what experimental evaluations reveal about outcomes under the new regime. In particular, I evaluate the labor force participation probability two years after the treatment is assigned. This estimated employment probability is then compared to the long-run federal requirement that half of single parent recipients participate in the labor force.[2]

---

[1] States failing to meet the minimum participation rates are subject to a penalty of 5 to 7 percent of their annual federal block grant. There are also requirements placed on recipients. After no more than 24 months, recipients are required to work in exchange for benefits, and a recipient can receive benefits for no more than 60 months in total.

[2] This comparison might be best thought of as illustrative. In practice, the estimated labor force participation probability reported in this paper may not be the measure which is compared against the federal benchmark. The reported estimates neither account for state specific welfare regulations nor the welfare status of the recipients. Under the PRWORA, however, state governments are given some power to define both the numerator (i.e., work

After describing the four experiments in Section 2, Section 3 evaluates what these data can and cannot reveal. Abstracting from the well-known critiques of randomized experiments, I maintain the best-case assumption that these MDRC evaluations identify the effects of the training programs. In particular, the experiment reveals the labor force participation probability if all recipients are assigned to a single treatment. These data cannot, however, identify the expected outcomes if some recipients are trained while others receive the standard benefits.

One might speculate that the outcome under the new regime will necessarily lie between the outcomes under the mandated training treatment and the mandated standard benefit treatment, with the precise location depending on what fractions of participants are assigned to training. This conjecture is true if being assigned to training never reduces the chances of participating in the labor force, but is not necessarily true if some persons benefit from training, while others do not.

Arguably, in fact, the modest training programs evaluated by the MDRC do not benefit everyone in this diverse population. Rather, it seems likely that the effects of treatment are heterogenous, with some fraction unaffected, some fraction employed only if assigned to training, and some fraction employed only if given standard benefits. Being assigned to training appears to impose significant net costs on some recipients such that their time and resources might otherwise be used more effectively.

Thus, the labor force participation probability that will be observed if the effect of treatment may be heterogenous is germane. Applying a nonparametric bounding method developed by Manski (1995; 1997), Sections 4 and 5 provide alternative estimates of the labor force participation probability when the treatments assigned are mixed.[3] I begin in Section 4 by asking the logical question of what the data alone reveal. In the absence of assumptions, these experiments only allow us to draw limited inferences about the labor market outcomes of welfare recipients. For the most part, the data do not reveal whether or not the federal standards will be met if a training program is adopted. In one case, however, these no-assumption estimates suggest that the program cannot meet the long run federal standards.

Rather than take the extreme position of either making no assumptions or assuming that training never hurts, I then explore a middle ground. In Section 5 predictions are made using alternative models that each reflect different *a priori* beliefs about the treatment assignment process. For instance, planners may train recipients who are most likely to benefit from training. While this outcome optimization model does not generally lead to point identification, informative nonparametric bounds are derived. In particular, the estimated bounds suggest that at least one-quarter and at most three-quarters of the caseload will participate in the labor force two years after the program is implemented.

Alternatively, budget constraints might limit the fraction of the caseload assigned to training. Under the restriction that no training is offered, the data identify the labor force participation probability. If instead, treatments are heterogenous, this budget constraint model narrows the no-assumption bound derived in Section 4. If, for instance, 10 percent of the caseload is trained, the estimated bounds suggest that regardless of the assignment policy, less than half of the welfare recipients can be expected to enter the labor market two years after the initial treatment.

---

activities) and the denominator (i.e., the caseload and eligibility criteria) when computing labor force participation probabilities.

[3] Dehijia (1999) examines a similar question using data from the California GAIN experiment combined with a parametric Bayesian model which identifies the unknown distributions of interest. While his results identify the labor market outcomes of interest, they rest on admittedly strong assumptions. Here, I examine what can be learned under an arguably weaker yet more plausible set of assumptions.

It might be that under the new regime planners will optimize the employment probability given a budget constraint. New theoretical results formalizing the implications of this constrained optimization model are developed in Section 6 and Appendix A. Intuitively, the employment probability under this constrained model can be no larger than under the unconstrained outcome optimization model. That is, if planners are optimizing outcomes, a constraint cannot improve the employment probability.

In Section 7, I draw conclusions. Two substantive findings emerge. First, some programs cannot possibly meet the federal labor force participation requirement. Second, others may meet the requirement if there is enough heterogeneity in treatment response and if state/local officials act in way that optimizes the labor force participation probability.

2. The Data

This analysis exploits four well known work and training demonstrations conducted in the 1980s by the Manpower Demonstration Research Corporation (MDRC). In particular, the four programs are the Arkansas WORK Program, the Baltimore Options Program, the San Diego Saturation Work Initiative Model (SWIM), and the Virginia Employment Services Program (ESP). For each program, welfare recipients were randomly selected to either participate in a basic work or training activity, or to receive the standard benefits package. Each of these evaluations were broad coverage programs with at least two years of follow-up data. All single parent families receiving AFDC and whose children were at least 6 years of age were mandated to participate. For the Arkansas WORK program, coverage was extended to families with children at least 3 years of age.

These MDRC experiments appear particularly well-suited for evaluating the types of welfare and training programs which might be adopted under TANF. For the most part, these experiments stressed labor force participation rather than human capital development. Thus, the training programs were modest, focusing on supervised job search and unpaid work assignments. Educational activities were only offered in limited cases.

The MDRC welfare-to-work programs were also mandatory. As a result, the "treatment" did not only consist of participation in program assignments, but also consisted of sample members hearing a message that they had to do something while on welfare. Furthermore, noncompliance with the assigned treatment led to sanctions or possible expulsion from the program. Over 50 percent of the caseload complied with the assigned training. Those that did not comply either exited the welfare system, were sanctioned, or were excused for reasonable cause.[4] As a result, all sample members assigned to training experience a treatment. That is, I evaluate the intention to treat, rather than the treatment itself.[5]

Table 1 provides additional summary information on these programs, as reported in Gueron and Pauly (1991, Table 3.1, pp. 85-92). To evaluate these experimental programs, the MDRC selected

---

[4] Most of those who failed to comply either left or became ineligible for AFDC. Sanctions, which eliminated the parent's portion of the grant for between 3 and 6 months, were applied in less than 10 percent of the cases (Friedlander and Burtless, Table 3-1, 1995).

[5] Note that under the intention to treat framework, all respondents comply with their assigned treatment. This definition also eliminates any bias arising from control group respondents obtaining the treatment. Respondents in the control group are not, by definition, assigned to training. If, instead, one wants to isolate the effects of the training program, as opposed to being assigned to training, noncompliance and substitution may become central concerns.

samples of size 1127, 3150, 2757 and 3211, in Arkansas, Baltimore, San Diego and Virginia, respectively. The data reveal the outcomes of all respondents, regardless of whether or not they complied with their assigned treatment. For each respondent, the data reveal the treatment received – training or the standard welfare benefits - and numerous labor market and welfare participation outcome measures. In this paper, the outcome variable of interest is whether or not the respondent participated in the labor force two years after treatment. For additional details on these data see Gueron and Pauly (1991) and Friedlander and Burtless (1995).

3. What Experimental Data Can Reveal

This section describes a primary identification problem that arises when using data from experimental evaluations to infer the outcomes that will prevail when treatments are heterogenous. Properly designed and implemented social experiments identify the distribution of outcomes that would occur if all individuals are given the same treatment. These data, however, cannot reveal the distribution of outcomes that will occur under the new regime. A mixing problem results in that data from experiments with mandated treatments cannot identify the outcome that would occur if treatment assignment varies across welfare recipients (Manski, 1995;1997). Section 3.1 formalizes what experimental data can reveal, while Section 3.2 describes what these data cannot reveal. While the data alone cannot identify the expected outcomes under the new regime, it might be conjectured that the employment probability must lie between those observed in the control and treatment groups. In Section 3.3, I evaluate this conjecture, concluding that the assumptions required to support this bound are not credible.

3.1 What Do Data from Welfare-to-Work Experiments Reveal?

To formalize the mixing problem, I begin by distinguishing between the outcome that would occur were a welfare recipient to have been assigned to training $y(1)$, and the outcome that would occur were she to have received the standard benefits, $y(0)$. In particular, let $y(\cdot)$ equal one if the individual would have participated in the labor force after the treatment period and zero otherwise. Let z denote the actual treatment received, where $z = 1$ if assigned to training and 0 otherwise. To simplify the exposition, I leave implicit variables used to condition on observed characteristics of individual respondents (e.g., single parents with children at least 6 years of age).

What do the data reveal? For those who were assigned to training ($z = 1$) the employment indicator $y(1)$ is observed but $y(0)$ is latent, while for those who received the standard benefits ($z = 0$) the outcome $y(0)$ is observed but $y(1)$ is latent. Thus, the data reveal the employment probability for those who were assigned to training, $P[\, y(1) = 1 \mid z = 1]$, and for those who merely received standard benefits, $P[\, y(0) = 1 \mid z = 0]$. In social experiments, the actual treatment received is randomly assigned so that the treatment, z, is statistically independent of the labor force participation indicators, $y(1)$ and $y(0)$. That is, the labor force participation probability of those who were assigned to training, $P[\, y(1) = 1 \mid z = 1]$, reveals the outcome that would occur if everyone where to receive training, $P[\, y(1) = 1]$. Likewise, the employment probability of those assigned to the control group, $P[\, y(0) = 1 \mid z = 0]$, reveals the outcome that would occur if the entire caseload received the standard benefits package, $P[\, y(0) = 1]$. Thus, the data from social experiments identify the employment probability if all welfare recipients are assigned to training, $P[\, y(1) = 1]$, or if instead all recipients receive the standard benefits, $P[\, y(0) = 1\,]$. Sample averages are used to consistently estimate these probabilities.

Table 2 displays the estimated employment probability for the treatment and control groups, along with the corresponding 90 percent bootstrapped confidence intervals.[6] In each case, the results suggest that job training programs slightly increase the probability of employment. Under the Virginia Employment Service Program (ESP), for instance, the labor force participation probability if all welfare recipients receive training is estimated to be 39.0 percent. If instead, all welfare recipients receive the standard benefits, these data suggest that 33.9 percent would be working after two years. Thus, the ESP increases the probability of employment by 0.051.

In practice, whether these data actually reveal the distribution of outcomes under mandatory

[6] See Pepper (1998) for a more detailed description of the mechanics of the bootstrapping method.

treatment policies is of considerable disagreement. There are many well known and important critiques of randomized social experiments.[7] Some concerns are procedural. The program may not be properly implemented, so that outcomes y(1) and y(0) are dependent upon the realized treatment z. Even if properly run, a more fundamental question is whether the demonstration program operates in the same fashion as it would if it were actually implemented. Certainly, concerns about external validity are germane (Campbell and Stanley, 1966; Hotz, Imbens, and Mortimer, 1998) : both the economy and the welfare system have undergone major changes since the mid 1980s. More generally, macro-feedback effects (Garfinkel, Manski and Michalopolous, 1992), Hawthorne Effects, and entry effects (Heckman, 1992; Moffitt, 1992) all suggest that small scale demonstration projects may not reveal the outcomes that would occur if the program were instituted on a larger scale.

In this paper, I abstract from these concerns by assuming that the MDRC demonstrations identify the effects of the various job training programs. That is, up to sampling error, the data are assumed to reveal the labor force participation probability if all recipients are assigned to training, P[y(1) = 1], and if all recipients are given the standard benefits, P[y(0) = 1]. While there are shortcomings with the MDRC evaluations, there is some support for this assumption. These evaluations are generally recognized as well designed and implemented social experiments (see, for example, Greenberg and Wiseman (1992), and Wiseman (1991)). To the extent, however, that this assumption is violated, the data from these experiments will be less informative (otz, Imbens, and Mortimer, 1998).

3.2 The Mixing Problem

The purpose of this paper is to examine the distribution of outcomes that will occur if treatment assignment varies across the caseload. Assume that a treatment selection policy, m, assigns some welfare recipients to be trained, and others to receive standard benefits. Let y_m be the realized outcome under policy m and let z_m be the realized treatment. In particular, I code $y_m = 1$ if an individual participates in the labor force after two years and $y_m = 0$ otherwise, and $z_m = 1$ if a person were to be assigned to job training, and 0 otherwise. Thus, I am interested in learning the labor force participation probability for individuals given the treatment policy m, $P[\ y_m = 1\ ]$.

In the absence of assumptions restricting the relationships between the treatment selection policy, m, and the employment indicators, y(1) and y(0), data from experimental evaluations cannot identify the distribution of outcomes when some individuals are trained and others are not. To see this use the law of total probability to write

$$(1) \qquad P[\ y_m = 1\ ] = P[\ y(1) = 1 \mid z_m = 1] * P[z_m = 1\ ] + P[\ y(0) = 1 \mid z_m = 0] * P[z_m = 0\ ].$$

The experimental evidence does not reveal how treatments will be assigned under policy m, nor how the latent outcomes y(1) and y(0) are related. Thus, a "mixing problem" results in that data from these experiments cannot identify the probability a recipient receives treatment t, $P[\ z_m = 1\ ]$, nor the labor force participation probabilities among those persons who will receive treatment t, $P[\ y(t) = 1 \mid z_m = t\ ]$ (see Manski 1995; 1997).

3.3 Outcome Optimization Bounds

[7] See Hausman and Wise (1985), Manski and Garfinkel (1992), and Manski (1996) for general critiques of the experimental methodology. Wiseman (1991) and Greenberg and Wiseman (1992) critically examine the MDRC demonstrations.

Given the mixing problem, one might speculate that the employment probability under the new regime will necessarily lie between the outcomes under the mandated training and standard benefit treatments, with the precise location depending on what fraction of participants are assigned to training. This hypothesis is true if being assigned to training never reduces the likelihood of participating in the labor force (Manski, 1997, Proposition 4).[8] After all, under this ordered outcomes assumption, the planner can do no better in terms of maximizing the employment probability than assigning everyone to training and no worse than assigning everyone standard benefits.

Thus, if outcomes are ordered, the estimates displayed in Table 2 imply that the training programs cannot achieve the long-run objectives of the federal reform. Under the Virginia ESP, for instance, the data imply that at least 33.3 percent and at most 39.3 percent of welfare recipients will be employed. Similarly, the labor force participation probability will not exceed 23.8 percent under the Arkansas Work program, 37.7 percent under the Baltimore Options Program, and 35 percent under the SWIM program. All of these upper bounds fall far short of the 50 percent long run benchmark.

If instead the effects of treatment are heterogenous, with some fraction unaffected, some fraction employed only if assigned to training, and some fraction employed only if given standard benefits, the ordered outcomes bounds displayed Table 2 do not apply. Planners can do better than assigning everyone to training and can do worse than assigning everyone to receive the standard benefits. Thus, the employment probability under the new regime is not in general constrained to lie between the outcomes under mandated treatments.

Arguably, in fact, the effects of treatment are heterogenous. Given that the welfare caseload is diverse, reflecting a broad range of skills, backgrounds and challenges, it seems unlikely that a basic skills and job search program will benefit this entire population (Pavetti, L. et al., 1999). A program which enables recipients with low levels of human capital to transition into the labor force may not benefit those with different skills or impediments. Rather, for some fraction of the caseload, the participants' time and resources devoted to training might otherwise be used more effectively. In fact, nearly half of the recipients assigned to training choose to either leave the program or incur sanctions. While some fraction leave welfare for work, the majority simply incur the costs of noncompliance (Fraker *et al.*,1997; Friedlander and Hamilton ,1993; Pavetti, L. et al., 1999). For these and others, the net costs of training may be substantial.

Although data from randomized experiments cannot reject the ordered outcomes assumption, the credibility of this restriction is tenuous. In Sections 4-6, I examine what the data reveal about the employment probability under the new regime if the effects of treatment might be heterogenous.

4. No Assumption Bounds

With limited budgets and the potentially prohibitive cost of work and training programs, state governments are unlikely to mandate a single job training program for all welfare recipients. Instead, states might systematically select certain recipients to be trained, others to participate in community work projects and others to receive short term cash assistance. Thus, rather than making the extreme assumption that all recipients are assigned to job training, a logical first step is to examine what these data reveal in the absence of assumptions. The general result is that knowledge of the employment probabilities under homogenous treatment policies yield a one-sided bound on the labor force participation

[8] Equation (1) reveals that this hypothesis is also true in the improbable case that the treatment assignment process under the new regime is random (Manski, 1997, Proposition 5).

probability.

To formalize these restrictions, it is useful to first explore how selection policies might affect outcomes. Treatment only affects some individuals. In particular, a fraction P[y(1) = 1 1 y(0) = 0] of the population "benefits" from job training, while a fraction P[y(1) = 0 1 y(0) = 1] "benefits" from the standard program. The remainder are unaffected by the treatment policy, with some fraction P[y(1) = 1 1 y(0) = 1] participating in the labor force regardless of the treatment and some fraction P[y(1) = 0 1 y(0) = 0] not working.

Thus, the employment probability will be maximized if those who benefit from training are assigned to training, and everyone else is given standard benefits. Regardless of the treatment assignment policy, however, the unemployment rate will at least equal P[y(1) = 0 1 y(0) = 0]. In contrast, the labor force participation probability will be minimized if all who benefit from training receive standard benefits and everyone else is assigned to training. Regardless of the treatment assignment policy, however, the employment rate must at least equal P[y(1) = 1 1 y(0) = 1]. Thus, the joint distribution of labor force participation indicators, y(1) and y(0), implies that

(2.) P[y(1) = 1 1 y(0) = 1] # P[y_m = 1] # 1 - P[y(1) = 0 1 y(0) = 0].

Notice that the width of the bound in Equation (2) equals the fraction of individuals who are affected by the treatment selection probability, P[y(1) = 1 1 y(0) = 0] + P[y(1) = 0 1 y(0) = 1]. Thus, in the absence of assumptions used to address the mixing problem, the joint distribution of y(1) and y(0) will only identify the labor force participation probability if treatments have no effect on outcomes. Still, these bounds are informative. The lower bound will generally be greater than zero while the upper bound will generally be less than one.

However, since the experimental data do not reveal the joint distribution of the labor force participation indicators, y(1) and y(0), the bounds in Equation (2) are not identified. Instead, the data reveal the labor force participation probability if all recipients are assigned training, P[y(1) = 1], or if all recipients receive the standard benefits, P[y(0) = 1]. These observed distributions imply informative restrictions on the joint distribution in Equation (2).

To illustrate how the marginal distributions on the labor force participation probabilities under mandated treatments are used to bound the joint distribution, consider the employment probabilities for the Virginia ESP. The estimates displayed in Table 2 imply that the labor force participation probability would be 39.0 percent under mandatory training and 33.9 percent under standard benefits.

Many joint distributions are consistent with these marginal probabilities. Table 3A displays one extreme case where the outcomes under training and standard benefits exhibit the strongest positive correlation. In this case, the treatment has almost no influence on outcomes. Only 5.1 percent of respondents benefit from training, and no one benefits from the standard program. Given this distribution, at most 39.0 percent of the caseload will participate in the labor force while at least 33.1 percent will participate.

Table 3B displays the other extreme where the outcomes exhibit the strongest negative correlation. In this case, the treatment has the largest possible effect on outcomes. Regardless of the selection method, at least 27.1 percent of the caseload will be unemployed. The labor force participation outcomes of the remaining 72.9 percent of the caseload depend upon the assignment process. It might be that all 72.9 percent will participate in the labor force, while it might be the entire caseload will be unemployed.

The joint distributions displayed Tables 3A and 3B, and others, are consistent with the ESP outcomes revealed by the experiments. In the absence of additional information, the distribution with the strongest negative association between the outcomes – i.e., Table 3B – defines the range of possible labor force participation probabilities under the new regime. This distribution maximizes the fraction of the caseload affected by the assignment process. Thus, if a state were to adopt the ESP program, the labor force participation probability must lie between 0 and 72.9. Some of the uncertainly reflected in this bound results from not knowing the assignment process (i.e., see the bounds in Equation 2), while some results from not knowing what fraction of the caseload is affected by the assignment process.

To formalize these restrictions, consider the fraction of the caseload who will be employed regardless of the treatment selection process, $P[\, y(1) = 1 \cap y(0) = 1 \,]$. Observe that

(3) $$P[\, y(1) = 1 \cap y(0) = 1 \,] = P[\, y(1) = 1 \,] + P[\, y(0) = 1 \,] - P[\, y(1) = 1 \cup y(0) = 1 \,].$$

Although the data cannot reveal the probability that a person might be employed, $P[\, y(1) = 1 \cup y(0) = 1 \,]$, this unknown probability is bounded. In particular, this probability can be no smaller than the labor force participation probability under either uniform treatment policy and no larger than the minimum of $P[\, y(1) = 1] + P[\, y(0) = 1 \,]$ and one. Thus, a sharp bound on the fraction who will be working regardless of the selection process equals

(4) $$\max\{ 0, P[\, y(1) = 1 \,] + P[y(0) = 1 \,] - 1 \} \leq P[\, y(1) = 1 \cap y(0) = 1 \,] \leq \min\{ P[\, y(1) = 1 \,], P[\, y(0) = 1 \,] \}$$

(Frechet, 1951). The upper bound is realized if the outcomes exhibit the strongest possible positive association (see, for example, Table 3A) in which case treatments have relatively little effect on outcomes. The lower bound is realized in the case where the outcomes exhibit the strongest possible negative association (see, for example, Table 3B) in which case treatments have the largest effect on outcomes. Similar reasoning can be used to bound the fraction of the caseload who will be unemployed regardless of the assignment process, $P[\, y(1) = 0 \cap y(0) = 0 \,]$.

Combining the bounds in Equations (2) and (4) reveals that

(5) $$\mathrm{Max}\{ 0, P[\, y(1) =1 \,] + P[\, y(0) = 1 \,] - 1 \} \leq P[\, y_m = 1 \,] \leq \mathrm{Min}\{ 1, P[\, y(1) =1 \,] + P[\, y(0) = 1 \,] \}$$

(Manski, 1997, Proposition 1). As in Equation (2), the upper (lower) bound is only realized if all who benefit from training are assigned to training (standard benefits). Additional uncertainty is introduced in that the data cannot reveal what fraction of the caseload is influenced by the treatment selection process.

Since the experimental data reveal the employment probabilities under a uniform treatment policy, the bounds for an unknown training assignment rule m are nonparametrically identified. The sample means in Table 2 are used to estimate these bounds.

Table 4 presents the no-assumption estimates for each of the four programs. Notice that these bounds are only informative on one side. In the absence of data, we know that the employment probability lies between zero and one. For each of the four programs, the data narrow the upper bound while the lower bound remains at zero.

Still, these bounds are informative. We learn, for instance, that no matter what assignment policy is used, the employment probability under the Arkansas WORK program will not meet the long run labor force participation requirements to assure full TANF funding. At most, just over 44 percent of the caseload will be participating in the labor force two years after the program is implemented. In contrast, the data are inconclusive about the labor force participation probabilities under the Virginia, San Diego, and Baltimore programs. In the absence of additional assumptions, the data cannot reveal whether or not adopting these programs will achieve the long run labor force participation requirements.[9]

5. Alternative Assumptions

A logical place to impose restrictions is on the treatment assignment policy m. One possibility is that all welfare recipients are trained. A second possibility is that administrators randomly assign welfare recipients to receive training or standard benefits. In these cases, as revealed by Equation (1), the experimental data bound the possible employment probabilities to lie between the labor force participation probabilities displayed in Table 2.

There are many other possible assumptions. Here, I examine the implications of two easily understood and commonly suggested restrictions on the assignment process. The first model assumes administrators assign job training to maximize the chance of employment. The second model assumes that a certain fraction of recipients are assigned to training but makes no assumption about the selection rule.

5.1 Outcome Optimization

Assume that administrators observe the employment indicators{ y(1), y(0) }and select the treatment, job training or standard benefits, that maximizes the labor force participation probability.[10] That is, $y_m = \max [y(1), y(0)]$. Under this outcome optimization model, the treatment selection policy will maximize the labor force participation probability so that $P[y_m = 1] = 1 - P[y(1) = 0$ 1 $y(0) = 0]$. Regardless of the assignment process, some fraction of the caseload will remain unemployed. If we observed the joint distribution of outcomes, this probability would be identified. Given the experiments, however, there remains uncertainty in that the data cannot reveal what fraction of the caseload is affected by the assignment process.

Intuitively, under this assignment rule administrators can do no worse in terms of maximizing the employment probability of welfare recipients than what would have occurred if all recipients were assigned to job training, and no better than the no-assumption upper bound in Equation (5). Formally, the

---

[9] These estimates apply for all single parent recipients, whether or not they are on welfare two years after the treatment. If restricted to those on welfare, the estimated lower bound remains at zero while the upper bound is less than 50 percent in all four experiments.

[10] This model imposes the strong assumption that administrator know the latent labor force participation outcomes for each individual. As such, the resulting estimates can be viewed as providing information on what might happen in the best case scenario. Similar results will apply under a weaker version of rational expectations where, given the available information s, the planner knows the employment probabilities P[y(1) = 1*s] and P[y(0) = 1*s]. If the information s uniquely identifies each individual, then the planner observes the latent outcomes { y(1), y(0) }. Otherwise, the planner only knows the distribution of outcomes among those with characteristics s. See Manski and Nagin (1998) for further details. Dehijia(1999) also examines a version of the outcome optimization model.

Frechet (1951) bounds in Equation (4) imply the sharp restriction that

$$(6) \qquad \max\{ P[y(1) = 1], P[y(0) = 1] \} \leq P[y_m = 1] \leq \min\{ P[y(1) = 1]+ P[y(0) = 1] , 1 \}$$

(Manski, 1997, Proposition 6).

While the upper bound in Equation (6) coincides with the no-assumption upper bound in Equation (5), the lower bound is informative. In particular, the estimates in Table 2 suggest that under the Baltimore, San Diego and Virginia programs, at least one-third of the caseload will work. However, even in this best-case model, where planners with rational expectations maximize the expected outcome, there remains much uncertainty. As before, the Arkansas WORK program will not meet long run federal requirements. At most, only 44.2 percent of the caseload will be working after two years. The Virginia, Baltimore, and San Diego programs, however, may or may not meet the long-run standard that at least half of the caseload participates in the labor force. Consider, for instance, Virginia's ESP program. If planners combine this program with an outcome optimization assignment rule, the estimated bounds imply that at least 39.0 percent and at most 72.9 percent of welfare recipients will be employed after two years. Where the realized labor force participation probability will lie depends upon the association between the latent labor force participation indicators, y(1) and y(0). If these outcomes have a strong positive association (e.g., Table 3A), the realized probability will lie closer to the lower bound. In contrast, if the association is strongly negative (e.g., Table 3B), the realized probability will approach the upper bound.

5.2 Budget Constraint Model

Restrictions on the treatment assignment process might arise from the costs of treatment. Table 1 shows that the net cost of the four training programs evaluated by the MDRC ranges from $118 per recipient for the Arkansas WORK program to nearly $1,000 per recipient for the Baltimore Options program. Arguably, planners operate under a strict budget constraint specifying the fraction of recipients assigned to training and to standard benefits.

Suppose that a certain known fraction of welfare recipients receive job training, while the remainder receive standard benefits. Let $p = P[z_m = 1]$ be the fraction of recipients who will receive training, and $1-p = P[z_m = 0]$ be the fraction who will receive standard benefits.[11] The experiment and this budget constraint restriction imply sharp bounds on the labor force participation probability.

To illustrate how the budget constraint model restricts the outcomes, consider the distribution of employment outcomes under the ESP displayed in Table 3B. Suppose that a planner, under this extreme distribution, is restricted to assign training to 10 percent of the caseload and standard benefits to the remaining 90 percent. In the best case, training would be assigned to recipients who benefit from training so that the labor force participation probability would be 43.9 percent. In the worst case, training is assigned to those who benefit from the standard program, so that the labor force participation probability would equal 23.9 percent. Thus, the labor force participation probability under this budget constraint

---

[11] An alternative model assumes that planners are restricted to training no more than a certain fraction p of the caseload, with no constraints imposed on the fractions assigned to standard benefits. That is, $p \geq P[z_m = 1]$. The bounds under this weak budget constraint model are derived in Appendix A, Proposition 3. While in principle, these bounds can differ from those estimated under the strict constraint, the results found when p = 0.10 and p = 0.25, the two cases examined below, are identical.

model lies between 23.9 and 43.9 percent.

To formalize this bound, consider the labor force participation probability under assignment policy m. Given a known fraction of recipients p will be assigned to training, Equation (1) can be rewritten as

$$(1')\qquad P[\ y_m = 1\ ] = P[\ y(1) = 1 \mid z_m = 1] * p + P[\ y(0) = 1 \mid z_m = 0] * (1-p).$$

The data do not identify the employment probabilities for those who will be trained $P[\ y(1) = 1 \mid z_m = 1]$, nor for those who will receive standard benefits, $P[\ y(0) = 1 \mid z_m = 0]$. After all, the way in which recipients will be assigned to treatments is unknown.

As before, however, information on the observed outcomes under uniform treatment policies can be used to bound the labor force participation probability under some unknown assignment policy m. The experiment reveals the probability a recipient works if everyone is assigned to training $P[y(1) = 1\ ]$ while our interest is in learning the labor force participation probability for those who will be assigned to training. The relationship between these two probabilities is highlighted using the law of total probability to write

$$(7)\qquad P[y(1) = 1\ ] = P[y(1) = 1^* z_m = 1]*p + P[y(1) = 1^* z_m = 0]*(1-p).$$

Since the unknown probability $P[y(1) = 1^* z_m = 0]$ lies in the interval [0, 1] we can bound the labor force participation probability for recipients who will be assigned to training. In particular,

$$(8)\qquad \text{Max}[\ 0, (P[y(1) = 1 \mid x] - 1 + p)/p\ ]\ \#\ P[y(1) = 1^* x, z_m = 1]\ \#\ \text{Min}[\ 1, P[y(1) = 1 \mid x] / p\ ]$$

Analogous bounds can be derived for the labor force participation probability for those who will receive standard benefits, $P[y(0) = 1^* z_m = 0]$.

From Equations (1') and (8) it follows that

$$(9)\qquad \max\{\ 0, P[\ y(1) = 1\ ] - (1-p)\ \} + \max\{\ 0, P[\ y(0) = 1\ ] - p\}$$
$$\#\ \ P[\ y_m = 1 \mid x]\ \ \#$$
$$\min\{\ p, P[\ y(1) = 1\ ]\ \} + \min\{\ 1-p, P[\ y(0) = 1\ ]\ \}$$

(Manski, 1997, Proposition 7). Notice that as the fraction trained approaches one, the bounds will center around the outcome that would be observed if all recipients are assigned to training, $P[\ y(1) = 1\ ]$, while as the fraction approaches zero the bounds center around the outcome that would be observed if all recipients receive standard benefits, $P[\ y(0) = 1\ ]$. More generally, restricting the fraction of recipients who are trained narrows the no-assumption bounds in Equation (5).

Table 5 displays estimated bounds under the constraint that either 10 percent or 25 percent of the caseload will be assigned to training. If restricted to training one-tenth of the caseload, the Baltimore will meet the short run federal labor force participation standard of 25 percent, and the Arkansas WORK, San Diego SWIM, and Virginia ESP programs might meet this standard. In all cases, however, employment outcomes under these programs fall short of the long-run requirement.

In contrast, if planners train one-quarter of the caseload, the estimated bounds are too wide to infer whether or not the programs will meet any of the federal labor force participation standards. With increased flexibility planners might either improve or degrade the decision making process. Thus, as the constraint become less restrictive the upper bound increases and the lower bound decreases.

Notice that constraining the administrator to assign one-quarter of the caseload to training does not affect the estimated bounds associated with the Arkansas WORK program. That is, the budget constraint bound of [0, 0.442] is equivalent to the no-assumption bound displayed in Table 4. The MDRC data reveal that no more than 23.8 percent of the caseload can benefit from training and no more than 20.5 percent can benefit from the standard assistance (see Table 2). To maximize the labor force participation probability, those that benefit from training are assigned to training, while to minimize this probability those who benefit from training receive standard benefits. Neither of these extreme policies is ruled out by the budget constraint model.

6. Joint Outcome Optimization and Budget Constraint Model

Arguably, both the budget constraint and outcome optimization models apply. That is, the objective of the planner might be to maximize the employment probability with a constraint specifying the fraction of recipients assigned to both job training and standard benefits. Under the outcome optimization model, administrators assign training to those who benefit from training, and standard benefits to everyone else. Given this assignment rule, a budget constraint is binding if the fraction assigned to training, p, falls below the proportion who benefit from training or if the fraction receiving standard assistance, 1-p, falls below the proportion who benefit from the standard program.

Consider the labor force participation probabilities under the ESP. In the extreme case displayed in Table 3A, the assignment process affects 5.1 percent of the caseload. Assuming the fraction assigned to training exceeds 5.1 percent, the outcome optimization model ensures that the labor force participation probability will equal 39.0 percent, the outcome that would be observed if all recipients were to receive training. If the budget constraint restricts the fraction assigned training to be less than one-twentieth of the caseload, then the observed labor force participation rate would equal 0.339 + p. In contrast, Table 3B depicts the case where the largest fraction of the caseload is affected by the assignment process. Here, if the administrator is restricted to train less than 39.0 percent of the caseload the budget constraint will be binding. Likewise, if less than 33.9 percent of the caseload is assigned to standard benefits, planners cannot optimize outcomes.

Intuitively, under this constrained optimization model planners can do no worse in terms of maximizing the employment probability than what would have occurred if all recipients were assigned to receive standard benefits and no better than the upper bound in the budget constraint model formalized in Equation (9). In fact, the bounds under this constrained optimization model turn out to be:

<u>Proposition 1</u>: For Bernoulli random variables y(1) and y(0), let P[y(1) = 1] and P[y(0) = 1] be known. Subject to the constraint that $P[z_m = 1] = p$, let $y_m = \max\{ y(1), y(0) \}$. Assume, without loss of generality, that $P[y(1) = 1] \geq P[y(0) = 1]$. Then

$$\min\{ P[y(0) = 1] + p, P[y(1) = 1] \} \leq P[y_m = 1] \leq \min\{ p, P[y(1) = 1] \} + \min\{ 1- p, P[y(0) = 1] \}. \quad (10)$$[12]

While the upper bound in Equation (10) coincides with the budget constraint upper bound in Equation (9),

[12] The proof of Proposition 1 as well as for the bound under the weaker condition that $p \geq P[z_m = 1]$ are in Appendix A. For the two cases examined here (p = 0.10 or p = 0.25), the estimated bounds are identical under the strict and weak constraints. In general, however, this does not have to be the case. Intuitively, under the outcome optimization model, the bounds found using a strict budget constraint are subsets of the bounds under the weak constraint.

the lower bound is more informative.

Table 6 displays the estimated outcome optimization bounds under the constraint that either 10 percent or 25 percent of the caseload will be trained. Under the Baltimore, San Diego or Virginia programs over one-third of the caseload will be working two years after the treatment is assigned. Thus, the short run employment standards will be met. Whether the long run requirements are achieved depends on both the constraint and the association between the outcomes. If 10 percent of the caseload is assigned to training, the programs cannot meet the long run employment standard. If instead one-fourth of the caseload is assigned to training, the upper bound exceeds the 50 percent benchmark in all cases. Thus, if the labor force participation is optimized and if outcomes are negatively associated (see, for instance, Table 3B), the long run federal benchmark will be achieved. If these outcomes have a strong positive association (e.g., Table 3A), however, planners cannot meet the federal benchmark.

6. Conclusion

Understanding the outcomes that can be expected from implementing a job training program is a central concern to academics, policy makers, and program administrators. What will the post welfare reform world look like? What fraction of the caseload will work? How will poverty and welfare spells change? What will happen to the teenage pregnancy rate? Almost no empirical evidence has been brought to bear on these and other fundamental questions.

Under the new federal regulations, state and local governments must design and implement welfare programs that meet minimum labor force participation requirements. In 1997, at least 25 percent of the single parent families on welfare must be working. By the year 2002, federal regulations require labor force participation rates in excess of 50 percent.

In many cases, the only available information on innovative programs comes from welfare-to-work demonstrations conducted since the War on Poverty. What do these experiments reveal to welfare reformers? Two general findings emerge. First, some programs cannot meet the federal standards. Second, other programs may meet the requirements if there is both sufficient heterogeneity in the treatment response and administrators optimize outcomes, or something reasonably close. While achieving this latter requirement depends upon the objective of the planners, the former requirement depends upon the fraction of the caseload affected by treatment.

Consider, for example, the Virginia Employment Services Program. In the absence of any restrictions to address the mixing problem, the employment probability under the ESP falls within [0, 0.726]. Thus, the data suggest that program may meet the federal objectives. The experimental data alone, however, reveal almost nothing about whether instituting the ESP will achieve the federal employment standards.

Prior information substantially narrows the no-assumption bound. If, for instance, administrators assign treatments to optimize the employment probability, the experiments imply that the program will achieve the short run federal standards. At least of 39.3 percent of the caseload will work. Even under the outcome optimization model, however, the data do not reveal whether outcomes under the ESP program can meet the long run federal labor force participation standards. After all, the experimental data cannot reveal the fraction of the caseload affected by the assignment process. If instead, we assume that the planners are restricted to train 10 percent of the caseload, the labor force participation probability will lie between [0.239, 0.439]. Under this restriction, we learn that the caseload may meet the short run federal standards but cannot achieve the long-run objectives.

Clearly, even with these experimental data, there remains much uncertainty. Without prior information on the association between the outcome variables under the different treatments, the data do not reveal whether the states will meet the federal standards. What modes of data collection would yield information beyond what can be inferred from the MDRC experiments? One possibility would be a set of experiments where the randomization is at the site level rather than at the individual level. Here a treatment would be a set of rules under which case managers would operate, rather than a mandate. An experiment of this type would reveal how case managers would make treatment assignments. In principle, it would directly reveal the outcome of policy m.

Appendix A: Outcome Optimization Budget Constraint Model

In this appendix, I formally derive nonparametric bounds under various restrictions on the fractions assigned to treatment. Three cases are examined. First, I prove Proposition 1 which bounds the outcome distribution under the optimization model with a fixed fraction p assigned to treatment 1 and 1-p assigned to treatment 0. The second maintains the optimization model but relaxes the constraint so that no more than p recipients are assigned to treatment 1. Finally, under this weak budget constraint, I formalize the bounds in the absence of the outcome optimization assumption.

A1. Proposition 1: Outcome Optimization with a Strict Budget Constraint

The outcome optimization model implies that treatment is assigned to maximize outcomes. That is, $y_m = \max\{ y(1), y(0)\}$. Thus, for a binary outcome variable

A1.) $P[y_m = 1] = 1 - P[y(1) = 0 \cap y(0) = 0] = P[y(1) = 1 \cup y(0) = 1].$

To highlight the treatment assignment process, it is useful to decompose the union in Equation (A1) into three distinct intersections. In particular, we know that

A2.)
$$\begin{aligned} P[y(1) = 1 \cup y(0) = 1] = \; & P[y(1) = 1 \cap y(0) = 1] + \\ & P[y(1) = 1 \cap y(0) = 0] + \\ & P[y(1) = 0 \cap y(0) = 1] \\ = \; & P[y(1) = 1 \cap y(0) = 1] + \\ & P[y(1) = 1] - P[y(1) = 1 \cap y(0) = 1] \} + \\ & P[y(0) = 1] - P[y(1) = 1 \cap y(0) = 1] \}. \end{aligned}$$

Notice that under the unconstrained outcome optimization model, $P[y_m = 1]$ monotonically decreases with $P[y(1) = 1 \cap y(0) = 1]$.

If outcomes are optimized, treatment 1 is assigned for all agents with y(1) = 1 and y(0) = 0, and treatment 0 is given to those with y(1) = 0 and y(0) = 1. Thus, $P[z_m = 1] \geq P [y(1) = 1 \cap y(0) = 0]$ and $P[z_m = 0] \geq P[y(1) = 0 \cap y(0) = 1]$. Suppose that the budget constraint model is imposed such that a known fraction p are assigned to treatment 1 and (1-p) are assigned to treatment 0. With this constraint, it might be that some who benefit from a particular treatment are assigned to the alternative. That is,

A3.)
$$\begin{aligned} P[y_m = 1] = \; & P[y(1) = 1 \cap y(0) = 1] + \\ & \min\{ p, P[y(1) = 1] - P[y(1) = 1 \cap y(0) = 1] \} + \end{aligned}$$

min(1-p, P[y(0) = 1] - P[y(1) = 1 1 y(0) = 1] }.

Notice that the outcome distribution P[y_m = 1] under this constrained optimization model monotonically decreases in P[y(1) = 1 1 y(0) = 1]. This follows from the fact that the constraints on the fraction assigned to treatment 1 and the fraction assigned to treatment 0 cannot both be binding. To see this, let q = P[y(1) = 1 1 y(0) = 1] and suppose that both p < P[y(1) = 1] - q and (1-p) < P[y(0) = 1] - q. In this case, P[y_m = 1] = q + 1. Thus, for both constraints to be binding it must be the case the q = 0. The Frechet bound in Equation (4) shows that a necessary condition for the joint probability q to equal zero is that P[y(1) = 1] + P[y(0) = 1] < 1. Yet, if this condition holds, then it cannot be the case that p < P[y(1) = 1] and (1-p) < P[y(0) = 1].

Thus, from the Frechet (1951) bound in Equation (4), a sharp lower bound on P[y_m = 1] is found by setting P[y(1) = 1 1 y(0) = 1] = min{ P[y(0) = 1], P[y(1) = 1]}, while the upper bound sets this unobserved probability equal to max{ 0,P[y(0) = 1] + P[y(1) = 1] -1 }. When P[y(0) = 1] + P[y(1) = 1] - 1 > 0, the upper bound on the outcome distribution can be written as

A4.) P[y_m = 1] # P[y(0) = 1] + P[y(1) = 1] - 1 +
min{ p, P[y(0) = 0] } +
min(1-p, P[y(1) = 0]}

= min{ p, P[y(1) = 1] } +
min(1-p, P[y(0) = 1]}.

The first equality follows from fact that P[y(t) = 1] + P[y(t) = 0] = 1 for t = 0,1. The second equality follows from the fact that both constraints cannot be binding. Thus, either p > P[y(0) = 0]] (1- p) < P[y(0) = 1] or (1-p) > P[y(1) = 0]] p < P[y(1) = 1]. If neither constraint is binding then the upper bound equals one. If the restriction on the fraction assigned to treatment 1 is binding, then the upper bound is { p + P[y(0) = 1] }. If the restriction on the fraction assigned to treatment 0 is binding, then the upper bound is { (1-p) + P[y(1) = 1] }.

Thus, we have

<u>Proposition 1</u>: For Bernoulli random variables y(1) and y(0), let P[y(1) = 1] and P[y(0) = 1] be known. Subject to the constraint that P[z_m = 1] = p, let y_m = max{ y(1), y(0) }. Assume, without loss of generality, that P[y(1) = 1] $ P[y(0) = 1]. Then,

A5.) min{ P[y(0) = 1] + p, P[y(1) = 1] }

P[y_m = 1]

min{ p, P[y(1) = 1] } + min{ 1- p, P[y(0) = 1] }.

A2. <u>Proposition 2:</u> Outcome Optimization with a Weak Budget Constraint

Suppose a weak version of the budget constraint model applies such that an administrator is constrained to assign no more than a known fraction p to treatment 1. That is, assume that p $P[z_m = 1].

No restrictions are imposed on the fraction of recipients assigned to treatment 0. Now

$$\text{A6.)}\quad P[\, y_m = 1\,] = P[\, y(1) = 1 \cap y(0) = 1] + \min\{\ p, P[\, y(1) = 1 \cap y(0) = 0]\ \} + P[\, y(1) = 0 \cap y(0) = 1]$$

$$= P[\, y(0) = 1\,] + \min\{\ p, P[\, y(1) = 1\,] - P[\, y(1) = 1 \cap y(0) = 1]\ \}$$

Again, the outcome distribution monotonically decreases in $P[\, y(1) = 1 \cap y(0) = 1]$. The second proposition follows:

<u>Proposition 2</u>: For Bernoulli random variables y(1) and y(0), let P[y(1) = 1] and P[y(0) = 1] be known. Subject to the constraint that $P[\, z_m = 1] \leq p$, let $y_m = \max\{\, y(1), y(0)\,\}$. Assume $P[\, y(1) = 1\,] \geq P[\, y(0) = 1\,]$. Then,

$$\text{A7.)}\quad \min\{\, P[\, y(0) = 1\,] + p, P[\, y(1) = 1\,]\,\} \leq P[\, y_m = 1\,] \leq \min\{\, 1, P[\, y(0) = 1\,] + P[\, y(1) = 1\,], P[\, y(0) = 1\,] + p\}.$$

Notice that the lower bounds in Proposition 1 and 2 are identical. Under these optimization models, the outcome distribution monotonically deceases in $P[\, y(1) = 1 \cap y(0) = 1\,]$. Thus, given the Frechet bound in Equation (4), the lower bound in Equation (A7) is realized if $P[\, y(1) = 1 \cap y(0) = 1\,] = P[\, y(0) = 1]$. Equivalently, since $P[\, y(1) = 1 \cap y(0) = 1\,] + P[\, y(1) = 0 \cap y(0) = 1\,] = P[\, y(0) = 1]$, the lower bound is realized when $P[\, y(1) = 0 \cap y(0) = 1\,] = 0$. Thus, removing the restrictions on the fraction assigned to treatment 0 has no effect on the lower bound.

The upper bounds, however, may differ. In particular, if $p > P[\, y(1) = 1]$ the upper bound under the weak budget constraint equals the no-assumption upper bound, $\min\{\, P[\, y(1) = 1] + P[\, y(0) = 1\,], 1\}$. Under the strict budget constraint, the upper bound equals $\min\{\, P[\, y(1) = 1] + P[\, y(0) = 1\,], P[\, y(1) = 1\,] + (1-p)\,\}$.

A3. <u>Proposition 3</u>: Weak Budget Constraint Model

Suppose the weak constraint that $p \geq P[\, z_m = 1\,]$ applies but the outcome optimization model does not hold. The upper bound on the employment probability coincides with the upper bound found under the outcome optimization model in Proposition 2. If, in contrast, outcomes are minimized, $z_m = 0$ if y(1) = 1 and y(0) = 0 and $z_m = 1$ if y(1) = 0 and y(0) = 1. Thus, in the absence of the constraint, $P[\, y_m = 1\,] \geq P[\, y(1) = 1 \cap y(0) = 1]$. With the constraint

$$\text{A8.)}\quad P[\, y_m = 1\,] \geq P[\, y(1) = 1 \cap y(0) = 1] + \max(\, 0, P[\, y(1) = 0 \cap y(0) = 1] - p\,)$$

$$= P[\, y(1) = 1 \cap y(0) = 1] + \max(\, 0, P[\, y(0) = 1\,] - P[\, y(1) = 1 \cap y(0) = 1] - p\,\}.$$

$$= \max(\, P[\, y(1) = 1 \cap y(0) = 1], P[\, y(0) = 1\,] - p\,\}.$$

Notice that if outcomes are minimized, $P[\ y_m = 1\ ]$ monotonically increases in $P[\ y(1) = 1 \cap y(0) = 1]$. Thus, from the Frechet bound in Equation (4), a sharp lower bound on $P[\ y_m = 1\ ]$ is found by setting $P[\ y(1) = 1 \cap y(0) = 1] = \max\{\ 0, P[\ y(0) = 1] + P[\ y(1) = 1] - 1\ \}$. A third proposition follows:

> <u>Proposition 3</u>: For Bernoulli random variables y(1) and y(0), let $P[\ y(1) = 1]$ and $P[\ y(0) = 1\ ]$ be known. Assume that $P[\ z_m = 1] \leq p$ and that $P[\ y(1) = 1\ ] \geq P[\ y(0) = 1\ ]$. Then,

A9.)

$$\begin{aligned} &\text{Max}\{\ 0, P[\ y(0) = 1\ ] + P[\ y(1) = 1\ ] - 1, P[\ y(0) = 1] - p\ \} \\ &\qquad \leq P[\ y_m = 1\ ] \leq \\ &\text{Min}\{\ 1, P[\ y(0) = 1\ ] + P[\ y(1) = 1\ ]\ , P[\ y(0) = 1\ ] + p\}. \end{aligned}$$

If $p > P[\ y(0) = 1]$, the constraint is not binding and the no-assumption lower bound, $\text{Max}\{\ 0, P[\ y(0) = 1\ ] + P[\ y(1) = 1\ ] - 1\ \}$, may be realized. In contrast, Equation (A8) and the Frechet bound imply that the constraint will be binding if $p < P[\ y(1) = 0 \cap y(0) = 1] \leq \min\{\ P[\ y(0) = 1, y(1) = 0\ ]\ \}$. In this case, the lower bound of $P[\ y(0) = 1] - p$ will be strictly greater than the no-assumption bound.

References

Campbell, D. and J. Stanley. (1966). Experimental and Quasi-Experimental Designs for Research. Boston: Houghton Mifflin.

Dehijia, Rajeev H. (1999). "Program Evaluation as a Decision Problem," National Bureau of Economic Research Working Paper, No. 6954.

Fraker, Tom M., Lucia A. Nixon, Jan A. Losby, Carol S. Prindle and John F. Else. 1997. *Iowa's Limited Benefit Plan*. Washington, D.C.: Mathematica Policy Research, Inc.

Frechet, M. (1951). "Sur Les Tableaux de Correlation Donte les Marges Sont Donnèes." *Annals de Universitè de Lyon A*, ser. 3, 14: 53-77.

Friedlander, D. and G. Burtless. (1995). Five Years After: The Long Term Effects of Welfare-to-Work Programs. New York: Russell Sage Foundation.

Friedlander, Daniel and Gayle Hamilton. 1993. *The Saturation Work Model in San Diego: A Five-Year Follow-Up Study*. New York: Manpower Demonstration Research Corporation

Garfinkel, I., C. Manski, and C. Michalopolous. (1992). "Micro-Experiments and Macro Effects." In C. Manski and I Garfinkel, eds., Evaluating Welfare and Training Programs. Cambridge, MA: Harvard University Press.

Greenberg, D. and M. Wiseman. (1992). "What did the OBRA Demonstrations Do?" In C. Manski and I Garfinkel, eds., Evaluating Welfare and Training Programs. Cambridge, MA: Harvard University Press.

Gueron, J.M. and E. Pauly. (1991). From Welfare to Work. New York: Russell Sage Foundation.

Hausman, J. and D. Wise, eds. (1985). Social Experimentation. University of Chicago Press.

Heckman, J. (1992). "Randomization and Social Policy Evaluation." In C. Manski and I Garfinkel, eds., Evaluating Welfare and Training Programs. Cambridge, MA: Harvard University Press.

Hotz, V.J., G.W. Imbens, and J.H. Mortimer. (1998). "Predicting the Efficacy of Future Training Programs Using Past Experiances," University of California - Los Angeles Working Paper.

Manski, C. (1995). Identification Problems in the Social Sciences. Cambridge, MA: Harvard University Press.

Manski, C. (1996), "Learning About Treatment Effects from Experiments with Random Assignment of Treatments," *Journal of Human Resources*, 31, 707-733.

Manski, C. (1997). "The Mixing Problem in Program Evaluations." *The Review of Economic Studies*, 64, 537-553.

Manski, C. and I. Garfinkel, eds. (1992). Evaluating Welfare and Training Programs. Cambridge, MA: Harvard University Press.

Manski, C. and D. Nagin. (1998). "Bounding Disagreements about Treatment Effects: A Case Study of Sentencing and Recidivism," Sociological Methodology 1998.

Moffitt, R. (1992). "Evaluation Methods for Program Entry Effects." In C. Manski and I Garfinkel, eds., Evaluating Welfare and Training Programs. Cambridge, MA: Harvard University Press.

Pavetti, LaDonna, Krista Olson, Demetra Nightingale, Amy-Ellen Duke, and Julie Isaacs. (1997). *Welfare-to-Work Options for Families Facing Personal and Family Challenges: Rationale and Program Strategies,* Urban Institute Press, Washington, D.C.

Pepper, J.V. (2000). "The Intergenerational Transmission of Welfare Receipt: A Nonparametric Bounds Analysis," *The Review of Economics and Statistics*, forthcoming.

Wiseman, M. (1991). Ed. "Research and Policy: A Symposium on the Family Support Act of 1988," *The Journal of Policy Analysis and Management, 10(4), pp.588-666.*

Table 1: Selected Characteristics of the MDRC Social Experiment

	Arkansas	Baltimore	San Diego	Virginia
Start Date	1983	1982	1985	1983
Major Program	Sequence of job search, unpaid work.	Choice of job search, unpaid work, education	Sequence of job search, unpaid work, education	Sequence of job search, unpaid work, education
Age of Children	3 or older	6 or older	6 or older	6 or older
Length of Follow up	36 months	36 months	27 months	30 months
Fraction Complying With Training	58.3%	38.0%	45.0%	64.4%
Net Cost per Person Trained[1]	$118	$953	$919	$430
Observed Outcomes	Employment, earnings, welfare receipt and payments	Employment, earnings, welfare receipt and payments	Employment, earnings, welfare receipt and payments	Employment, earnings, welfare receipt and payments
Sample Size	1,127	2,757	3,211	3,150

1. Net cost are computed as the average cost per enrolled recipient minus the average cost for recipients in the control group. The net cost estimates include all expenditures by the operating agency as well as services provided by other agencies that were considered part of the treatment program.

Source: Gueron and Pauly, 1991 and Friedlander and Burtless, 1995..

Table 2: Estimated Probability of Employment Eight Quarters After Treatment

	Control Group: P[y(0) = 1]	Treatment Group: P[y(1) = 1]
Arkansas	20.5% (0.175, 0.230)	23.8% (0.213, 0.265)
Baltimore	37.7% (0.255, 0.392)	38.8% (0.373, 0.412)
San Diego	28.5% (0.265, 0.302)	35.0% (0.331, 0.370)
Virginia	33.9% (0.315, 0.362)	39.0% (0.373, 0.408)

Note: Bootstrapped 90 percent confidence interval are in parenthesis below the estimates.

Table 3A: A Joint Distribution Consistent with the Labor Force Participation Probabilities for the Virginia Employment Services Program:
Strongest Positive Association Between the Outcomes

Treatment		Training	
	Outcomes	y(1) = 0	y(1) = 1
Standard Benefit	y(0) = 0	61.0%	5.1%
	y(0) = 1	0.0%	33.9%

Table 3B: A Joint Distribution Consistent with the Labor Force Participation Probabilities for the Virginia Employment Services Program:
Strongest Negative Association Between the Outcomes

Treatment		Training	
	Outcomes	y(1) = 0	y(1) = 1
Standard Benefit	y(0) = 0	27.1%	39.0%
	y(0) = 1	33.9%	0.0%

Table 4: Estimated No Assumption Bounds on the Probability of Employment Eight Quarters After Treatment Under Assignment Policy m, P[y_m = 1]

	Lower Bound	Upper Bound
Arkansas	0.0%	44.2% (0.474)
Baltimore	0.0%	76.5% (0.796)
San Diego	0.0%	63.5% (0.661)
Virginia	0.0%	72.9% (0. 759)

Note: The 0.95 quantiles of bootstrapped upper bound are in parenthesis below the estimates.

Table 5: Estimated Bounds on the Probability of Employment Eight Quarters After Treatment Under the Budget Constraint Model

10 Percent Trained, 90 Percent Standard benefits

	Lower Bound	Upper Bound
Arkansas	10.5% (0.075)	30.5% (0.334)
Baltimore	27.7% (0.256)	47.7% (0.499)
San Diego	18.5% (0.165)	38.5% (0.402)
Virginia	23.9% (0.215)	43.9% (0.462)

25 Percent Trained, 75 Percent Standard benefits

	Lower Bound	Upper Bound
Arkansas	0.0% (0.000)	44.2% (0.474)
Baltimore	12.7% (0.106)	62.7% (0.649)
San Diego	3.5% (0.015)	53.5% (0.552)
Virginia	8.9% (0.065)	58.9% (0.612)

Note: The 0.95 quantiles of bootstrapped upper bound are in parenthesis below the upper bound estimates and the 0.05 quantiles of the boostrapped lower bound are in parenthesis below the lower bound estimates.

Table 6: Estimated Bounds on the Probability of Employment Eight Quarters After Treatment Under the Joint Outcome Optimization Budget Constraint Model

10 Percent Trained, 90 Percent Standard benefits

	Lower Bound	Upper Bound
Arkansas	23.8% (0.213)	30.5% (0.334)
Baltimore	38.8% (0.373)	47.7% (0.499)
San Diego	35.0% (0.331)	38.5% (0.402)
Virginia	39.0% (0.373)	43.9% (0.462)

25 Percent Trained, 75 Percent Standard benefits

	Lower Bound	Upper Bound
Arkansas	23.8% (0.213)	44.2% (0.474)
Baltimore	38.8% (0.373)	62.7% (0.649)
San Diego	35.0% (0.221)	53.5% (0.552)
Virginia	39.0% (0.373)	58.9% (0.612)

Note: The 0.95 quantiles of bootstrapped upper bound are in parenthesis below the upper bound estimates and the 0.05 quantiles of the boostrapped lower bound are in parenthesis below the lower bound estimates.