Institute for Research on Poverty
Discussion Paper no. 1096-96

Are Welfare Employment Programs Effective?

Lawrence M. Mead
New York University

July 1996

Abstract

Employment programs meant to place welfare adults in work or training became an important part of Aid to Families with Dependent Children starting in the 1980s. These programs are effective if one means that they have positive impacts in evaluations, less so if one expects them to make a large and visible change in the welfare problem. In programs evaluated by the Manpower Demonstration Research Corporation, impacts on employment, earnings, and dependency are small in absolute terms but somewhat larger as a percentage of the control group mean. Impacts are understated in some studies because randomization occurred only after enrollment in the work program or because control group members had access to equivalent services. Results are also depressed by the failure of many experimentals to participate in the tested program. Programs raise the activity of experimentals in work-related activities much more than they raise earnings or employment. Effects on dependency are understated because evaluations do not capture deterrence effects. The sharp decline in AFDC in Wisconsin in recent years suggests the power of work requirements to drive the rolls down.

Are Welfare Employment Programs Effective?

The employment programs that states have implemented for welfare recipients since the 1980s constitute a promising development in social policy, but how promising is disputed. Some experts say the programs are effective. Some do not. Much of the debate turns on the shifting meanings of "effective." "Welfare" here largely means Aid to Families with Dependent Children (AFDC).[1]

Congress first established mandatory work programs for AFDC adults in 1967, when it enacted the Work Incentive (WIN) program, but WIN never seriously affected more than a small part of the welfare caseload, due to weak funding and implementation. After allowing states to experiment for much of the 1980s, Congress in 1988 replaced WIN with the Jobs Opportunities and Basic Skills Training Program, or JOBS. This program was implemented during fiscal years 1990–95, and some states obtained waivers of normal federal rules to continue to run their own experimental work programs, which differed in some details.[2] In 1995, the new Republican Congress decided to dismantle JOBS in favor of block-granting welfare to the states. But it kept stringent work requirements, which will require states to keep running welfare employment programs of some kind. Thus, the issue of what these programs achieve will live on.

In the following sections, I first describe various meanings of program effectiveness. I then argue that welfare employment programs are more successful than assessments based on the evaluation literature suggest. Impacts tend to be understated for technical reasons, and they vary with how fully programs are implemented. Evaluations also fail to capture effects on client activity and much of the effect on welfare caseloads. The potential of work requirements to deter entry to welfare appears to be substantial. The case of Wisconsin suggests as much. Reforming welfare through work is really an administrative challenge. The new block grant is premised on the idea that work programs can drive the rolls down dramatically.

MEANINGS OF EFFECTIVENESS

When we say a social program is "effective" we mean, at the least, that it achieves something.[3] Evaluators have taught the policy world that it is not enough that money be spent on doing good. The criterion of success should not be mere *inputs* to a program. Nor should it be mere *outputs* in the sense of services delivered to clients. Nor should it even be *outcomes*, meaning the changes in the clients' condition or status, such as income or employability, that occur during or after a program treatment. Rather, it should be *impact*, meaning the *change* in outcomes that is actually *due to* the program.

Outcome versus Impact

We cannot take positive change in outcomes as evidence of impact. Improved results such as job entries may have occurred for reasons unrelated to the program. Suppose a work program has managed to place its clients in jobs. Perhaps it succeeded only because it chose motivated clients or because the economy was favorable. That is, the jobs would have been obtained even without the program, and thus it should get no credit for them. The need to define the counterfactual, to measure program results against what would otherwise have occurred, is especially great in the case of welfare. People usually go on aid because they have lost a job or a spouse. Many rebound into employment or remarriage and thus leave aid, whether or not anything is done to help them. A welfare employment program should not get credit for this turnover. It achieves impact only if it causes *more* recipients to take jobs and leave aid then would have done so *without* the program.

The rationale for experimental evaluations is that they isolate the true from the apparent effects of a program. A random sample of clients is drawn from the population to be served, and then this sample is randomly allocated to two groups, one of which receives the experimental treatment while the other—the "control" group—does not. Due to random assignment, the program cannot pick out the most motivated people to serve. The experimental clients will probably not differ systematically in any attribute, either measured or unmeasured, from the controls. If they experience better outcomes, therefore, that gain can truly be attributed to the program. The difference is an unbiased measure of impact.[4] The Manpower Demonstration Research Corporation (MDRC), the main evaluator of welfare employment programs, has built its reputation on this sort of study.

Is the Impact Test Demanding?

From one point of view, the impact criterion is demanding. It means that programs must "really" have effects, not only appear to. A program does not succeed because it generates success stories, let alone because of good intentions. One of the ironies of the War on Poverty of the 1960s is that the experts who designed new programs for the poor also insisted on evaluating them rigorously. Faced with the impact test, however, many popular programs, such as Head Start, proved disappointing. The effect of this sort of social science was conservatizing—against ambitious government programming (Aaron, 1978). The experience gave rise to "Rossi's Law": "The expected value of any net impact assessment of any large-scale social program is zero" (Rossi, 1987).[5]

From another viewpoint, however, the impact criterion is undemanding, for it requires *only* impact. That is, it asks that a program show effects on a social problem, but not that it fully overcome that problem. Commonly, experimental evaluations find that a program has some impact on skills or earnings, but not a large one. But even a small impact is enough to justify the program's expense in a cost-benefit analysis, especially if consideration is given to far-in-the-future reductions in crime or welfare by the clients. The program is then judged to be a rational investment, and its designers advocate that society spend more on it.[6] But the effects on the problem in question, for example school failure by disadvantaged students, may still be too small to be apparent to anybody but the evaluators. Thus, conventional evaluation can declare programs successful, yet they may have no effects on social problems that the public can see. This disjuncture is especially likely when an evaluation uses a large sample of clients, because it may then detect quite small differences in outcome between experimentals and controls.

These conflicting notions of effectiveness underlie much of the debate over the value of antipoverty programs. The public clearly wants government to solve problems, not only to affect them. Inability to invent programs that could reduce crime or welfarism or elevate school performance in visible ways contributed as much as anything to the image of incompetence in domestic policy that came to dog the federal government in the 1970s and 1980s. This in turn fomented the rightward political trend of those years (Mead, 1992: chs. 2, 10–11).

Republicans took power in the White House with Ronald Reagan in 1981 and in Congress with Newt Gingrich in 1995 promising to cut back antipoverty programs. As their rationale, they had only to cite the long series of disappointing evaluations—mostly done by liberal evaluators—in education, training, criminal justice, and other areas. That record gave credibility to the further charge that federal programs, particularly welfare, had actually exacerbated poverty by rewarding the poor for unwed pregnancy, nonwork, and other dysfunctional behaviors (Murray, 1984).

But on the other side, liberals retort that the effect of tested programs is seldom actually zero. Most do show some impacts, even if they are smaller than we would wish. Some programs, notably the Perry Preschool project in Ypsilanti, Michigan, have shown long-term benefits that make the programs look more successful after years than they did at first (Berrueta-Clement et al., 1984). Thus, the view that government always fails is ignorant and unfair. Good programs can make some headway against poverty, provided government is prepared to fund them long-term. They are “good investments” that society should be patient with (Schorr, 1988; Mayer and Jencks, 1995). But again, effectiveness in this sense does not satisfy the public desire actually to solve social problems.

Effectiveness in Welfare Employment

A persuasive definition of effectiveness ought to include both impact and what I will call sufficiency. That is, the effective program is one which has clear effects on the problem it addresses, *and* which really solves the problem, or could do so if it were fully implemented.

In welfare employment the chief goals are to raise the earnings and employment of welfare adults and to reduce the dependency of their families. These goals are not usually in conflict, since for the adults to go to work usually raises incomes and often takes the families off welfare, if not out of poverty.[7] Therefore, to achieve impact and sufficiency, welfare work programs should clearly improve employment and reduce dependency, *and* they should be able to do this for the bulk of welfare adults.

That would imply achieving and sustaining employment at least for the recipients who are judged to be employable. Experts judge that the employable include perhaps four-fifths of the women who ever go on welfare and about two-thirds of those on the rolls at a given time, who tend to be the more disadvantaged (Maynard, 1994: 114–15, 114 n. 2; Mead, 1992: 124–27). Alternatively, effectiveness might mean driving enough of the employable off welfare so that the rolls were radically reduced, even if the remaining cases consisted largely of the unemployable.

UNDERSTATED IMPACTS

Are welfare employment programs effective in these terms? There is at least hope. If anything has repealed Rossi's Law, it is the welfare work programs evaluated by MDRC since the early 1980s. These projects showed more consistent impacts in randomized evaluations than any previous class of social program serving large populations. The first of the projects, in San Diego, showed as early as 1984 that it could raise the employment of welfare applicants and reduce their dependence on aid (Goldman et al., 1984), and from that point the whole debate in Washington about welfare and poverty took a more hopeful turn. The perception grew that these programs worked, and for this reason the

Family Support Act of 1988 (FSA) was devoted mainly to the creation of JOBS, the most ambitious welfare work structure yet (Wiseman, 1991a).

Small Differences

Analysts at the political extremes, both right and left, tend to dismiss welfare employment impacts as inconsequential. This serves their political agenda, which is to argue that reformism must fail and radical change is unavoidable. They say, respectively, that poverty is rooted in the evils of the welfare state or the capitalist economy (e.g., Murray, 1994: 27–28; Piven and Cloward, 1993: 387–91). From a certain point of view, effects are indeed small. Table 1 describes the average impacts from the nine primary evaluations of welfare employment programs that MDRC completed from 1985 to date. The middle column shows the effects expressed as absolute differences between the experimental and control groups. It is difficult to believe that a rise in earnings of $329 or in employment of 3 percentage points could transform the welfare system.

For several reasons, however, impacts expressed this way are understated.

Percentages of Baseline

It is reasonable to ask, not only how large impacts are in absolute terms, but how large they are relative to where recipients begin. If impacts are expressed as percentages of the control-group mean, as in the right-hand column of Table 1, they appear more substantial. Perhaps the average program raises employment or reduces welfare by only 3 points, but as a percentage of *baseline* the typical improvement is two to four times as great.

In the most notable programs evaluated by MDRC, the Saturation Work Initiative Model (SWIM) in San Diego and California's Greater Avenues for Independence (GAIN) in Riverside County, the effects were sizable by any standards. The earnings impacts were $658 and $1,010 a

TABLE 1
Average Results from Nine Main MDRC Welfare Employment Evaluations

Impact	Control Mean	Experimental-Control Difference	Percentage Change
Average earnings	$1,992	$329	16
Employed at end of final year	28.4%	3.0%	11
Average AFDC payments	$2,409	-$175	-7
On welfare at end of final year	55.2%	-3.1%	-6

Note: For details of individual studies and sources, see Appendix Table A.

year, respectively, representing 29 and 40 percent of the control-group mean. The programs also reduced annual welfare payments by $553 and $584, or 14 and 17 percent of the baseline, respectively. Over three years, Riverside raised the share of its clients who worked by a quarter and their earnings by close to half, compared to controls (Riccio, Friedlander, and Freedman, 1994: 120).[8]

Late Randomization

Impacts are understated, furthermore, because the baseline against which effects are measured is higher than many people suppose. The way we think of experimental evaluation is that the experimental group gets the new treatment and the control group gets nothing at all. But in practice, the control clients often get treatment much like that given to the experimentals. It is not always possible to keep the controls entirely out of the experiment. In MDRC's Florida study, 20 percent of the control group was inadvertently exposed to various aspects of the experimental program, a mistake that reduced the measured impacts of the program by as much as 20 percent (Kemple, Friedlander, and Fellerath, 1995: 55–57, 101–4).

More subtly, in several of the MDRC projects, the controls received some elements of the welfare employment treatment by design. An evaluation that tests all dimensions of a work test would separate the experimental and control clients at the point where they first applied for welfare or were found to be mandatory for the work requirement.[9] But of those required to enter welfare work programs, only two-thirds or less actually do so,[10] in part because they never complete their applications or leave welfare quickly. One reason may be that they heard a message about the need to work and avoid dependency. Not facing the same requirements, the controls hear such a message less and are more likely to go on aid and stay there. The difference is a deterrent effect that represents part of a work program's impact. If randomization is at application, the experiment will get credit for it, just as it will for any success in moving clients into jobs after they enter the program. In MDRC's studies in Baltimore, Florida, San Diego, and Virginia, this was the research design.

However, in the five other MDRC projects (Arkansas, California GAIN, Cook County, San Diego SWIM, and West Virginia), randomization was done only among recipients who had been assigned to work *and* who had reported to the work program for registration and orientation. Only then were some admitted to the program and some consigned to a control group. This meant that the program could get credit only for impacts produced *after* program entry, typically by moving its clients toward work. It could not get credit for deterrence occurring between application and registration. Also, by attending registration and orientation, the controls got some of the message about the need to work intended for the experimentals. For both reasons, the achievements of the programs will appear smaller than if randomization were at application.[11]

Controls in Other Programs

Further, clients in the control groups may well be exposed to suasions unrelated to the tested program. By the time MDRC began its projects in the early 1980s, welfare mothers whose youngest children were at least 6 were already supposed to enroll in the WIN program, as were other welfare adults without family or school responsibilities. Clients relegated to control groups in the MDRC studies escaped the experimental work program but were still subject to WIN. The top panel of Table 2 shows activity levels in studies where MDRC compared participation by the experimentals in the tested program to participation by the controls in WIN. This comparison is only rough, because studies differed in many respects, including the point of randomization.[12] Activity was much lower in WIN than in the new programs, chiefly because WIN had less funding and was more leniently administered. The WIN levels, though, are not zero. While WIN generally was ineffective, any effect it did have reduced the measured impact in these experiments.

TABLE 2
Activity Levels of Experimentals and Controls in MDRC Evaluations (percentage)

		Experimentals	Controls
Studies using WIN data	In any employment and training activity	40.4	5.5
	In job search or job club	34.6	1.4
Studies using other data	In any employment and training activity	55.6	28.5
	In job search or job club	40.3	6.4
	In education or training	29.8	22.7

Note: Intervals covered vary. For details of individual studies and sources, see Appendix Table B.

Controls were also able to obtain education and training services outside welfare employment entirely. In several projects where MDRC had information from community college records or surveys of clients, it was able to show participation in these activities as well. The lower panel of Table 2 shows average activity levels for experimentals and controls in these studies. Again, the comparison is only rough. However, participation by controls is much higher than when only WIN was considered. Twenty-nine percent of the controls had some involvement in employment-related activities, mostly in education or training. If these services were of any benefit, the effect again was to reduce the measured impact of the experiment.

The problem of the nonzero baseline would be slight if the experiment being tested were the only program of its kind. Welfare work programs, however, were instituted in the 1970s and 1980s, a time when many other federal and local training and education programs already existed. Recipients not chosen for an MDRC experiment could go and get similar training from local agencies operating under the Job Training Partnership Act (JTPA), vocational education, or local community colleges. Often in a locality, all these programs compete for clients, and not enough eligibles remain outside the system to construct a truly unserved control group to evaluate any of them.

IMPACTS AND IMPLEMENTATION

Likewise, one may presume for several reasons that the evaluated impacts of welfare employment programs are not fixed. They vary with the degree of program implementation. As work requirements are more fully implemented, effects may be expected to rise. That gives hope that work programs could show a sufficient impact on the welfare problem, not only a palpable one.

The Averaging Problem

Just as the control group can include people who receive services, so the experimental group can include people who do not, and again measured impact is reduced. In experimental method, the impact is taken to be the difference in *average* outcomes between experimentals and controls, with the means for both groups computed over all their members, whatever happened to them.[13] But as noted above, not all experimental clients enter the tested program. This means that the experimental average includes many people who did not actually receive the treatment.

Nonparticipation is minor in voluntary programs, where it is assumed that clients apply for benefits and will receive them if admitted. It is major, however, for mandatory programs, which serve a population much of which does not come forward on its own.[14] In both WIN and JOBS, the vast majority of clients participate only because they are required to as a condition of welfare eligibility.[15] This ensures that many will drop out and not be served, depressing recorded impact.

It makes sense to test a program's effects on the entire experimental group if it is meant to serve an entire population, as is true for mandatory work programs. It is also necessary methodologically. Since subjects are randomly assigned to the experimental and control group, the impact estimate is true only if it considers all individuals in both groups—whether or not they actually participated actively. One cannot single out the clients who were active and calculate an impact just for them. They might differ, in measured and unmeasured ways, from the control group, and the estimate would no longer be untainted by selection bias.

The effect, however, is to understate considerably how much welfare adults may gain from employment. Impacts can look small, in part, because the group averages merge the improvements in earnings, or reductions in dependency, realized by clients who go to work with a lot of zeros for the nonworkers. In SWIM, for example, the average experimental earned $2,903 during the program's second year, but since only 49 percent of this group worked during that year, the average *employed* experimental earned about twice that much (Gueron and Pauly, 1991: 145). If one focuses on the gains of the working recipients and not the average impacts, the potential of work to solve the welfare problem looks greater.

The point here is not that the impacts of work programs are really twice what they seem. We cannot assume that the recipients who work and realize large gains do so because of the work program. Nor is it reasonable to compare results for working experimentals to the averages for controls. An averaging process also occurs for the control group, and there too individuals with earnings will do considerably better than the average. The difference in average impact is still a fair test of the program.

Rather, the point is that impact is likely to vary positively with participation rates. If work programs have any good effect, they have it by causing more recipients to work and thus realize the gains typical of the workers. The more clients actually partake of the program, the more are likely to have earnings and other favorable outcomes. Then a higher proportion of experimentals will contribute something to the averages, and measured impact will rise (Greenberg and Wiseman, 1992: 39–40, 45–46). This will happen even if the gains per employed experimental do not improve at all, and perhaps even if they drop.

Conversely, if programs fail to show effect, a major reason is simply that so many recipients referred for work never actually participate. The difference between typical gains and working gains was enormous even for SWIM, which achieved one of the highest participation rates ever achieved in welfare employment. It would be even greater in a typical program with lower participation. The evaluations of welfare employment programs are tests more of the participation level than the efficacy of services. One cannot construe an evaluation of a program with low participation as a fair test of what that program, if fully implemented, might achieve.

Reaching Below the Cream

Impacts will rise for another reason as well. As a higher proportion of the eligibles participate in a work program, the additional clients tend to be those for whom impact will be higher.

When a work program first serves a welfare caseload, the clients who most readily participate are typically the most employable. They tend to be motivated and high school graduates, and they often have a prior work history. They enter the program as volunteers or with little suasion, and they quickly get jobs, often at good wages, and leave welfare. Some experts think that welfare employment programs should be voluntary, on the view that clients must be self-motivated to succeed (Lynn, 1989). However, clients like this would typically succeed even without the program. Their fortunes will be little different from those of equivalent controls who get work on their own, so the program generates little impact by serving them. These easy-to-serve cases make up about a third of the adults on welfare, and it was this third that welfare employment chiefly served prior to the experiments evaluated by MDRC. That largely explains why WIN had little palpable effect on the caseload (Mead, 1992: 167–68, 171–72).[16]

To maximize impact, a program must reach deeper down in the employability barrel. A close study of the MDRC projects revealed that most of their effects on employment and welfare did not come from fresh applicants to welfare, who tend to be the most employable. They came, rather, from applicants returning to welfare after a break, and in some instances from recipients already on the rolls—that is, from groups that are less employable (Friedlander, 1988). This jibes with the estimate cited earlier that perhaps two-thirds of adults on the rolls are employable. It implies that, to maximize impact, welfare employment programs must expand their participation so that they reach at least the top two-thirds of the caseload.[17] Since programs cannot achieve these levels voluntarily, this in turn implies that they must be mandatory, and must enforce participation stringently (Mead, 1990).[18]

It is quite possible that impacts would eventually fall as the bottom of the barrel was reached. This is because the most disadvantaged clients have too many problems to benefit much from conventional employment and training treatments. They might be given assignments short of actual work, such as participation in community activities (Herr, Halpern, and Conrad, 1991), but these would not produce much gain in employment or lower dependency. However, participation rates in most work programs are still well short of this point of diminishing returns.

Impacts Have Already Grown

One can argue that the impacts recorded by programs serving the disadvantaged have already grown as those programs have become more demanding. Most would say that the welfare employment programs evaluated, as a group, show stronger effects on employment and earnings than voluntary government training programs run under the auspices of the Comprehensive Employment and Training Act (CETA) of 1973–82 or the Job Training Partnership Act (JTPA) since 1982. While most of the welfare work programs at least show effects, JTPA's impacts in its latest evaluation were mostly nonsignificant and, for some groups, negative (Bloom et al., 1992). One reason might be is that the welfare employment programs reach more of the ambivalent clients where the potential to produce change compared to controls is greater.

Among the welfare employment projects studied by MDRC, one can say that impacts have improved as it has become clearer how to run effective programs. Considering all the studies, there is no general trend for impacts to improve over time, but there is for projects within California, the state that has had the largest influence on national developments.[19] The initial San Diego program of the early 1980s started the vogue for welfare employment. Its successor, SWIM, was then able to improve on its impacts primarily by enforcing participation more stiffly. By emphasizing job search as well as participation, Riverside GAIN then improved on SWIM. The GAIN evaluation as a whole obtained impacts comparable to the initial San Diego results, even though, of the six county programs tested, only those in San Diego and Riverside were conspicuously well-run. The other counties are now under pressure to run JOBS more as Riverside does. In 1995, the California legislature amended GAIN to weaken the strong presumptions of remediation contained in the original legislation and to require all recipients to look for work up front.

Recently, MDRC released preliminary results from a nationwide evaluation of JOBS which look even more encouraging. The data come from three local sites (Atlanta, GA, Riverside, CA, and Grand Rapids, MI) that are testing two different versions of JOBS, a “labor force attachment” model stressing job search and immediate work, and a “human capital development” approach that encourages clients to enter education or training in preference to work in hopes of getting better jobs later. Table 3 shows the impacts for both models after two years. The labor force attachment results are much the stronger, in part because impacts from remediation may take more than two years to appear.

The labor force attachment results are also considerably better than the averages in Table 1. They are comparable to Riverside’s, and they are now produced through a common program philosophy in three different localities.[20] That again suggests that as work programs are better implemented, they will show higher impacts around the country.

IMPACTS ON LIFESTYLE

There is additional reason to call welfare employment programs effective, if we look beyond orthodox evaluation studies such as MDRC’s. Economic effects capture vital dimensions of program performance, but not the noneconomic dimensions that are, perhaps, politically more important.

One would imagine from the MDRC studies that the goals of the programs were simply to raise recipients’ earnings so that more of them can leave welfare and money is saved. But it is clear from studies of welfare politics that the main controversies surrounding AFDC are not economic. The

TABLE 3
Preliminary Impacts from MDRC National JOBS Evaluation

	Outcome	Control Mean	Absolute Impact	Percentage Change
Labor force attachment model	Average earnings	$2,712	$696	25.8
	Percent with earnings	34.4	8.1	23.5
	Average AFDC payments	$3,312	-$732	-21.9
	Percent on welfare	68.3	-11.1	-16.2
Human capital development model	Average earnings	$2,508	-$24	-0.8
	Percent with earnings	32.4	2.6	8.1
	Average AFDC payments	$3,420	-$456	-13.5
	Percent on welfare	68.8	-4.2	-6.1

Source: Freedman and Friedlander (1995).
Note: Dollar figures are annualized from data for months.

public and politicians are not exercised about welfare primarily because it costs money. Its expenses are well below those of larger social insurance programs such as Social Security or Medicare.

The Lifestyle Issue

Rather, the chief controversy is about the lifestyle of welfare adults. While about half of welfare families go on the rolls for two years or less, the other half stays on longer, and the rolls at a point in time are dominated by these chronic cases. The long-term recipients remain dependent largely because they do not work regularly and often have children out of wedlock. Who is responsible for these behaviors? Could the adults work more than they do? Are they to be seen as "deserving" of aid, or as "undeserving"? These are the issues that dominate controversy about welfare, not the cost of the program.

Polls show that the public wants to enforce good behavior on the adults, but also to help people in need. That combination makes voters uneasy at proposals simply to spend more or less on welfare. Spending more does not change behavior, and spending less might put the needy at risk. Most people would prefer to turn AFDC into a work program, in which families get an income but the adults have to work in return for it, just as taxpayers do. Then good behavior is enforced while families are assured of income. If a work-oriented welfare costs more money, in order to guarantee jobs or child care to the recipients, voters accept that, provided the work demand is serious.

Politically, putting the adults to work is a lot more important than getting them off welfare, although the one should often accomplish the other. The public wish is not so much to save money as to challenge the defeated, dependent culture characteristic of the long-term poor. People perceive that mindset as contrary to the individualist ethos of the nation. To mobilize the dependent so that they honor that ethic is, in the public mind, the innermost goal of reform.[21]

From this perspective, the key aim of welfare employment is simply to have welfare adults do more to help themselves. It is first of all to have them work, but also to look for work or take other steps that will eventually lead to work. Actually working or leaving aid may be the final goal, but to show effort is also an end in itself. The Family Support Act (FSA) was crafted only in part to maximize economic impacts. States were required to spend 55 percent of their federal JOBS money on disadvantaged groups of clients likely to benefit most, and federal officials were told to draft performance measures for the program that would promote impacts.

Raising Activity Levels

However, the main controversy surrounding FSA was not about this but about whether participation levels would be high enough in JOBS to satisfy public pressures for work effort. The main demand of Congressional Republicans and the Reagan White House was for standards much higher than in WIN, which was required to serve only 15 percent of its eligibles annually. FSA required states to involve 20 percent of their employable recipients in JOBS on a monthly basis by 1995, a much stiffer standard. And regulations written by the Bush Administration defined participation to require assigning clients to activities of at least 20 hours a week, a standard many localities protested as excessive (Mead, 1992: chap. 9).

Thus, the activity levels indicated in Table 2 have great importance in themselves, quite aside from the influence activity may have on impacts. While controls are active enough to understate the measured effects of experimental programs, there is still an enormous difference in level between them and the experimentals. To take the more conservative numbers from the lower panel, *almost twice as many* recipients do something to help themselves when they are subject to a work program as when they are not. And while experimentals participate in education and training only slightly more often than controls, they enter job search *over six times* more often. And looking for work is the activity the public most wants to see from dependent adults.

A related idea is "coverage." The broadest goal of a work program is to make the caseload accountable for lifestyle in various ways. This it may do by causing more recipients to go to work or look for work, but also by ensuring that those who do not comply absorb a sanction (a reduced grant) or leave welfare. A caseload that must do one of these things is “covered” in a sense meaningful to the public, even if only part of it goes to work. In seven studies, MDRC estimates that the evaluated programs achieved coverage of 75 to 97 percent of their eligibles over nine to twelve months (Gueron and Pauly, 1991: 133–34). In SWIM, the most stringent, only 3 percent of clients managed to remain on welfare and in SWIM without participating in the program, undertaking education or training on their own, working, or incurring a sanction (Hamilton, 1988: 140–42).

These effects on activity dwarf any of the economic impacts. Politically, they are probably at least as important, and more should be made of them as a justification of welfare employment programs. They provide one of the main reasons to think these programs could have a sufficient impact on the welfare problem.

IMPACTS ON DEPENDENCY

It would be difficult to argue for sufficiency, however, unless work programs show more effect on the welfare rolls than they have seemed to thus far. As Table 1 suggests, impacts on dependency or on welfare payments have generally been smaller, both absolutely and relative to the baseline, than effects on employment or earnings.

It is understandable, in part, that employment impacts should exceed welfare impacts. Not all welfare adults who go to work earn enough money to leave the rolls, although their grants should be reduced. Welfare rules allow recipients who work to keep some of their earnings rather than reduce their grants by the full amount, a “work incentive” meant to promote employment. The more generous such incentives, the more possible it is to work without leaving welfare.[22] Meanwhile, many controls leave the rolls for a variety of reasons (work, remarriage, ineligibility). Against a baseline with such turnover, it is harder for work programs to show an impact on dependency than on employment (Gueron and Pauly, 1991: 149, 152).

Understated Impacts

However, welfare impacts are also understated, for reasons beyond those mentioned above. One of these is that not all the jobs recipients achieve, in part due to work programs, are reported to welfare. Some clients work "off the books" so they can receive their earnings without any reduction in their welfare grant. When surveyed by the government, only 6 percent of AFDC mothers report that they are working, even part-time (U.S. House, 1994: 402). But according to analyses of the Panel Study of Income Dynamics (PSID), an academic database, the real level is probably a quarter or more, and 51 percent of welfare mothers work at some time while on welfare (Moffitt, 1983: 1029; Harris, 1993: 329–30).[23] According to an interview study of 50 welfare mothers, 19 percent of their income came from various forms of employment, legal or illegal (Edin and Jencks, 1992: 208). If welfare agencies could detect all this work, they would close many more cases and the welfare impacts of work programs would look larger.

Even when welfare caseworkers know about jobs, they may not adjust grants or close cases because of the paperwork involved. According to some accounts, caseworkers actively discourage recipients from working or reporting earnings, for fear that the case will become "error-prone" and trigger quality-control sanctions (Bane and Ellwood, 1994: 6–7, 126–27). Failure to reduce grants for earnings is no doubt one reason why recipients who work do not have their benefits "taxed" away at the rate legislators intend. Between 1967 and 1981, AFDC adults who took jobs were supposed to lose about two-thirds of their earnings in reduced grants, but the real tax rate was less than a third (Fraker, Moffitt, and Wolf, 1985: 255–57).

Deterrence

Welfare impacts are also understated because conventional evaluations capture mostly the "exit" effect of a work program—its ability to cause people already on the rolls to leave welfare faster. They miss much of the "entry" effects—the impact that the presence of a work program may have on whether people go on welfare in the first place. Entry effects might be both negative or positive. A work requirement might keep some people off welfare. Knowing they will have to work, they prefer to avoid aid or to get a job on their own. But it may also attract people to welfare if it removes the stigma of aid and/or makes clients more employable (Moffitt, 1996: 32–39).

One negative entry effect—the withdrawal from welfare that occurs between initial application and registration for a work program—was mentioned above. This is an effect not credited to the program if randomization occurs only after work registration. By the same argument, even randomization at initial application will fail to capture deterrence effects that occur before people even apply for aid. A serious work or child support requirement might change the whole climate surrounding welfare, making dependency seem less acceptable and thus reducing the rolls. This would, however, presume a greater degree of implementation than most work programs have achieved.[24]

It is difficult to study such effects experimentally because they occur in advance of welfare application, and there is no clearly defined program clientele from which one could draw a sample for random assignment. It has not appeared to the MDRC evaluators that the institution of more serious work programs caused applications to welfare to drop off, but these studies produce no systematic information on the question (Gueron and Pauly, 1991: 77, 183–84).

Yet entry effects can be substantial. One simulation using PSID found that a program like JOBS that was implemented for the whole caseload would cause a variety of entry and exit effects. Under varying assumptions, the caseload might grow by as much as 8 percent—or drop by as much as 25 percent. A demanding, well-implemented JOBS program would reduce the rolls (Moffitt, 1996).[25]

As it happens, JOBS was implemented in the midst of a substantial run-up of the AFDC caseload. Between 1989 and 1993, the rolls grew 29 percent nationally. This certainly implies that JOBS did not in any general sense restrain dependency. However, a statistical analysis suggests that states that implemented JOBS forcefully suffered markedly less welfare growth than others. States that had more recipients active in JOBS in the early 1990s suffered less welfare growth, even allowing for social and economic differences among the states (Mead, 1995b: 1–8).

These simulation and statistical results are not directly comparable to the impact findings from the MDRC evaluations. But they are in some rough sense stronger. They suggest that the evaluations do not capture all that welfare employment might do to drive caseloads down, especially if it is fully and forcefully implemented.

WELFARE DECLINE IN WISCONSIN

The case of Wisconsin may well illustrate the capacity of work programs to reduce dependency in ways the public can see.[26] In recent years, this state has reduced its welfare rolls remarkably. During the 1989–93 national welfare boom, when on average, state AFDC rolls grew 35 percent, they fell 3 percent in Wisconsin. While Louisiana and Mississippi also had small declines, these are Southern states with a tradition of minimal social services and very low welfare benefits. Wisconsin somehow cut its rolls despite paying the twelfth-highest benefits in the country.

Over the longer period from 1986 to 1994, the state cut AFDC cases a remarkable 23 percent. Dependency fell 42 percent in the average county, over 60 percent in a few rural counties, and 5 percent even in Milwaukee, the site of the state's largest and most nonwhite caseload, according to state figures (State of Wisconsin, 1994: 14–15). The state's hot economy in the late 1980s was one probable cause. So was a 6 percent cut in the benefit level when Governor Tommy Thompson first took office in 1987, since with lower payment levels fewer families qualify for aid. But these forces are not the whole story, since the caseload fall continued after the cut and during the early 1990s when Wisconsin went through a recession, albeit a mild one compared to elsewhere in the country.

JOBS in Wisconsin

Besides favorable economic conditions, the factor Wisconsin officials mention most is tougher work enforcement. The state decided in 1986 to build up its welfare employment programs using its own money, and starting in 1987 many counties received increased funding. This enabled them to hire the staff necessary to cover more of the caseload. For the first time the majority of adults could assume that they would have to do something to help themselves if they went on welfare. According to officials, this administrative event was what chiefly precipitated the caseload fall, although the wide availability of jobs also allowed more recipients to find work once they were in the programs. Further, starting in 1989, increased federal funding for the new JOBS program came on-line, and yet more of the caseload was served. Behind the buildup stood the long-term maturation of Wisconsin’s excellent welfare bureaucracy, whose well-funded staff and computer systems enable it to track clients more closely than in most states.

No work program in Wisconsin has yet received an experimental evaluation, so there is no strict proof that JOBS has helped drive the caseload down.[27] However, the state has run JOBS in a manner similar to that of the programs that have demonstrated the most impact in MDRC evaluations done elsewhere. The counties, which run the local programs, typically enforce participation stringently, the policy that probably has the largest effect on impacts. Although many programs have traditionally favored education and training as a service strategy, the programs most noted for high work levels and caseload reductions emphasize job search and actual work. The Riverside program, top-ranked in MDRC’s GAIN evaluation, also emphasized these activities. Of the successful counties, the most exemplary is probably Kenosha, where most recipients have to work at least part time before they qualify for training and 40 percent of clients are already working, either in private or government jobs, a figure much higher than in the typical JOBS program. In 1993, the National Alliance of Business chose Kenosha as the best JOBS program in the country.

The Thompson Experiments

In addition to implementing JOBS, the Thompson Administration has launched its own experiments in welfare in a number of counties. Typically, these programs are tougher versions of JOBS. They require working or looking for work earlier or more definitely than the regular work program. The most important is Work First, which requires that would-be recipients look for work even before their applications for aid are approved, and Work Not Welfare, which limits cash aid to 24 months and requires recipients to “work off” benefits by working or other assigned activities (Wiseman, 1996).

Local officials credit these efforts too with helping to drive the caseload down. They say that the rhetoric of work obligation and self-reliance surrounding the programs has “sent a message” to the recipients that employment is now going to be demanded of them, and this has persuaded many to leave the rolls for work even without further suasions. While the experiments are too recent to explain much of the caseload decline since 1987, the officials may have a case. According to state figures, in the eighteen counties running Work First, the caseload fell an average of 40 percent between January 1987 and October 1994, the period prior to the experiment. That advance is probably due mostly to the benefit cut and the regular JOBS program backed by a good economy. But after Work First was instituted, the rolls then fell another 13 points from the new base by July 1995.

By imposing a work requirement up front, the Thompson programs aim to produce exactly the deterrent effect that other evidence suggests does much to reduce dependency. Work First and Work Not Welfare station employment caseworkers at welfare intake offices to try to dissuade people from applying for aid. They advise the needy to get work immediately or to seek aid from their families.

They point out that people entering welfare will immediately face demands to work, and that in Work Not Welfare cash aid is time-limited.

By these means a substantial proportion of applicants are in fact diverted from welfare. According to state data, in Pierce and Fond du Lac, the two counties running Work Not Welfare, the caseload fell 47 percent between January 1987 and December 1994, prior to the experiment. In January 1995, Work Not Welfare was instituted, and from the new base caseloads then fell another 39 percent by July 1995. According to Fond du Lac officials, much of that decline reflects diversion. Fresh intake to the rolls has virtually ceased, and as departures continue, the rolls drop.

Is Diversion Denial?

Some might ask how diversion differs from simply denying people access to aid; Wisconsin might as well withdraw from the AFDC program or reduce benefits to subsistence levels. This is a criticism welfare liberals have levied against AFDC ever since conservatives started building up fraud and abuse detection and work and child support requirements in the 1970s (Lipsky, 1984; Brodkin, 1986; Sosin, 1986). But diversion does not simply close the door to welfare. Applicants, if eligible, can still apply for aid and get it, and those who decide not to apply are not simply abandoned by government. Welfare officials help them arrange alternatives to AFDC, including referring them to other programs, and government still accepts some responsibility for good outcomes. The direction the new policies give the needy about how to avoid dependency is itself a service—one missing if welfare just goes out of business.

There need be no opposition between the poor helping themselves and government assuming responsibility. The new programs, like JOBS, are paternalistic. They insist that the responsibility for coping with poverty be divided between the needy and government, with both sides shouldering some of the burden. Diversion, like enforcing work within welfare, is one way to accomplish that.

REFORM AS ADMINISTRATION

Experience to date suggests that the impacts of work requirements are limited, not so much by the nature of the programs, as by the administrative capacities of the governments doing the implementing. In order to show sufficiency, work programs must reach the vast bulk of the adult recipients, and clients have to be supervised closely to be sure they enroll and then fulfill their assignments. That requires that support services such as child care be arranged. More critically, it requires energetic case managers to track down people who drop out, and it requires backup systems such as computer monitoring and adjudication. It requires creating financial incentives for staff and contractors to achieve results. It requires inspiring the whole organization with the work mission. All this takes leadership and money, although in the best programs the cost is more than offset by lower grant payments as people leave welfare (Bardach, 1993; Mead, 1995a).

The programs that garnered the highest impacts in the MDRC evaluations—San Diego's SWIM and Riverside GAIN—were well-funded and conspicuously well-organized and managed. San Diego drew upon a long prior experience with welfare employment programs (Hamilton, 1988: 15–17). Equally, the programs that proved most disappointing—Cook County and Florida—suffered from significant administrative problems. Cook County attempted to impose a work requirement on its whole caseload but lacked the resources to run job clubs (a successful work search strategy involving groups of clients) or to supervise individual job searches closely (Quint and Guy, 1986). Florida suffered a growing caseload and serious funding cuts in the middle of the program, leading to child care shortages and an inability to monitor activities closely (Kemple, Friedlander, and Fellerath, 1995: ES6-ES8).[28]

One reason to think that Wisconsin's programs have impact is that they, like SWIM and Riverside, are unusually well-run. In most if not all Wisconsin counties, JOBS staff are able and motivated, have adequate funding for child care and other support services, and enjoy effective backup systems. The state has been running welfare work programs well for nearly 30 years, and that prowess carries over to JOBS and the current Thompson experiments.[29] As a result, Wisconsin achieves among the highest participation rates of any state in JOBS. In 1994, 32 percent of its employable welfare adults met the JOBS participation standard, twice the rate required for that year and the tenth-highest figure in the country.[30] In 1990, Kenosha had already achieved participation rates by the JOBS measure of more than 50 percent, or over twice the rate mandated for all states in 1995 (Wiseman, 1991b).

The need for capable administration is an important constraint on work programs. Not all localities are yet able to implement them well enough to show impact, let alone sufficiency. But the same is even more true of most other social programs that show positive impacts. Many compensatory programs judged successful by evaluators, including Perry Preschool, are hothouse plants—implemented only for small caseloads by specially dedicated staffs operating outside government. Among compensatory programs, welfare employment programs are unique in having shown impact on large populations even when run by regular government agencies.

The Wisconsin experience suggests, if it does not prove, that work programs can achieve sufficiency. The Wisconsin programs not only have probable impact on welfare, they have enough effect so that the dependency problem is really changed. They reach the vast bulk of the caseload, and they have driven enough cases off the rolls so that dependency is fundamentally reduced, if not overcome. They have done this across much of the state, although not yet in Milwaukee. It is likely that, forcefully implemented, work requirements have a power to change the expectations surrounding welfare and make it much more of a last resort for the needy than it is today. That is exactly what the public, in every state, wants welfare reform to achieve.

THE NEW WELFARE SYSTEM

Very likely, the use of welfare work programs to deter entry to welfare will become even more important in the future, under the new welfare structure just enacted in Washington. The Republican

Congress has recast welfare as a block grant with few federal controls. Broadly, states now can set the eligibility rules for federal assistance, not only the benefit level as they did under AFDC.

JOBS was disestablished and its funding folded into the block grant. There is now no federally mandated welfare employment structure. Yet Congress is still banking heavily on programs like JOBS to reform welfare. It believes they can attain sufficiency and then some. Under the block grant, recipients must work within two years and cannot receive aid for more than five years, policies that presume states can move most current dependents toward employment.

Furthermore, 15 percent of adults receiving aid under the block grant are supposed to be engaged in “work activities” by fiscal 1996, rising to 50 percent in the year 2002 and thereafter. The weekly hours of effort demanded to be a participant rise from the current 20 in 1996 to 35 in 2002 and thereafter. Even tougher standards are set for two-parent AFDC cases. “Work activities” are defined to mean primarily actual work in private or subsidized jobs, although some job search and vocational education are allowed.

These standards are vastly more demanding than those in JOBS. The participation rates apply to *all* welfare adults (except mothers with children under 1, at state option), not only to the 43 percent now defined as employable (U.S. House, 1994: 358).[31] Therefore, to meet them will be over twice as difficult as comparable rates were in JOBS. The 35-hour standard is virtually full-time, as against the half-time standard in JOBS. On their face, such demands look impracticable. They call for participation levels not yet realized anywhere in the country. It is doubtful whether even well-run programs could achieve such levels by 2002—*assuming caseloads as they are now*.

But Congress is assuming that they will not remain as they are now, that the enforcement of such demands even early on in the implementation process will cause large deterrence effects. The intent, in short, is not to serve the current caseload but to “tip” it and drive the employable adults off the rolls, as Wisconsin is already doing. Then states would actually have to realize the participation standards only for a much smaller caseload. As an incentive to do this, the participation rates that states would have to meet will be *reduced* by as many percentage points as they have cut their caseloads since the new regime went into effect.

The intention is so radical that localities may frustrate it. If they want federal assistance funding, they face a Hobson's choice between enforcing work with unprecedented toughness or denying federal aid to much of the current caseload, which would then have to be supported on local funding or not at all. However, the penalty for failing to meet the standards is only 5 percent of the federal grant. Many states may prefer to take that penalty rather than face the alternatives. Or they may go to the courts or Congress and try to have the standards suspended or changed.

But the public will to enforce work is clearcut. If states attempt to take the penalty and opt out, Congress may increase it so that states will have no choice but to enforce work if they want federal welfare aid. Then programs like JOBS would have to be used, as Congress and voters intend, truly to break the back of dependency in America. Work would be enforced with a seriousness never before seen, and employable adults would leave welfare in droves.

If this happens, the question surrounding welfare work programs will change radically. The issue will no longer be whether they have impact, and the focus of the existing evaluations on exit effects—on whether clients already on welfare are helped—will seem almost irrelevant. The effect of work requirements will be undoubted, and it will come overwhelmingly through reducing entry to the rolls. The question will shift from whether there are effects to whether they are salutary. Is it a good idea to drive most of the employable away from the welfare system? Do the unemployable also lose aid? What happens to those who are diverted from aid? Are they working or not, better off or not?

As noted, these are questions that experimental evaluations cannot address. To answer them will require different methods, for example surveys of the population eligible for welfare, or of people formerly on welfare. Very little research of this kind exists.[32] In short, if welfare employment is used to drive caseloads down, research, as well as social policy, will enter a new world.

APPENDIX TABLE A
Results of Main MDRC Welfare Employment Evaluations

Study	Outcome in Final Year	Control Mean	Experimental Control Difference	Percentage Change
Arkansas	Average earnings	$1,085	$337	31%
1985	Employed at end of final year	18.3%	6.2%	34
	Average AFDC payments	$910	-$168	-18
	On welfare at end of final year	40.1%	-7.3%	-18
Baltimore	Average earnings	$2,989	$511	17
1985	Employed at end of final year	40.3%	0.4%	1
	Average AFDC payments	$1,815	-$31	-2
	On welfare at end of final year	48.4	-0.2%	0
California GAIN	Average earnings	$2,523	$636	25
1994	Employed at end of final year	24.1%	4.5%	19
	Average AFDC payments	$4,163	-$331	-8
	On welfare at end of final year	55.5%	-3.0%	-5
(Riverside GAIN)	(Average earnings)	($2,552)	($1,010)	(40)
1994	(Employed at end of final year)	(24.6%)	(6.6%)	(27)
	(Average AFDC payments)	($3,448)	(-$584)	(-17)
	(On welfare at end of final year)	(45.8%)	(-5.2%)	(-11)
Cook County	Average earnings	$1,217	$10	1
(job search/	Employed at end of final year	21.4%	1.3%	6
work experience)	Average AFDC payments	$3,146	-$40	-1
1987	On welfare at end of final year	80.8%	-1.9%	-2
Florida	Average earnings	$3,138	$80	3
1995	Employed at end of final year	37.8%	0.4%	1
	Average AFDC payments	$1,945	-$113	-6
	On welfare at end of final year	53.6%	-2.4%	-4
San Diego	Average earnings	$1,937	$443	23
(applicants in	Employed at end of final year	36.9%	5.5%	15
job search/work	Average AFDC payments	$2,750	-$226	-8
experience)	On welfare at end of final year	47.9%	-2.0	-4
1986				

(table continues)

APPENDIX TABLE A, continued

Study	Outcome in Final Year	Control Mean	Experimental Control Difference	Percentage Change
San Diego SWIM 1989	Average earnings	$2,246	$658	29
	Employed at end of final year	29.3%	5.4%	18
	Average AFDC payments	$3,961	-$553	-14
	On welfare at end of final year	58.7%	-7.4%	-13
Virginia 1986	Average earnings	$2,356	$268	11
	Employed at end of final year	34.1%	4.6%	13
	Average AFDC payments	$1,295	-$111	-9
	On welfare at end of final year	39.3%	-2.6%	-7
West Virginia 1986	Average earnings	$435	$16	4
	Employed at end of final year	13.1%	-1.0	-8
	Average AFDC payments	$1,692	$0	0
	On welfare at end of final year	72.5%	-1.5	-2
Averages (excludes Riverside)	Average earnings	$1,992	$329	16
	Employed at end of final year	28.4%	3.0%	11
	Average AFDC payments	$2,409	-$175	-7
	On welfare at end of final year	55.2%	-3.1%	-6

Sources: Gueron and Pauly (1991), pp. 142–43; Kemple, Friedlander, and Fellerath (1995), p. ES-16; Riccio, Friedlander, and Freedman (1994), pp. 120, 122.

Notes: Studies are weighted equally. Averages exclude Riverside. Results are for AFDC Basic, not AFDC-UP, and are for the last full year of the evaluations, which varied from 1 to 3 years in follow-up. "Percentage change" means the absolute impact as a percentage of the control-group mean. Averages for the percentage changes are means of the individual percentage change figures, not calculated from the averages of the control means and experimental-control differences.

APPENDIX TABLE B

Experimental vs. Control Activity Levels in MDRC Evaluation Studies (Percentage)

	Ever Active in	Experimentals	Controls
Studies using WIN data			
Baltimore[a]	Any employment and training activity	44.5	3.4
(Over 12 months)	Job search or job club	25.0	2.0
	Education and training	16.6	1.0
San Diego	Any employment and training activity	46.4	6.1
(Over 9 months)	Job search or job club	44.1	0.8
West Virginia	Any employment and training activity[b]	30.2	6.9
(Over 9 months)			
Average	Any employment and training activity	40.4	5.5
	Job search or job club	34.6	1.4
Studies using other data			
Florida	Any employment and training activity	63.9	40.1
(Over 2 years)	Job search or job club	42.7	18.7
	Education and training	41.5	30.5
Cook County	Any employment and training activity	47.2	16.9
(Over 9 months)	Job search or job club	36.1	2.2
	Education and training	16.9	15.9
California GAIN	Job search or job club	28.5	3.9
(Over 2–3 years)	Education and training[c]	26.4	23.1
San Diego SWIM	Job search or job club	53.7	0.7
(Over 2–3 years)	Education and training[d]	48.2	31.2
Virginia	Education and training	15.8	12.7
(Over 15–28 months)[e]			
Average	Any employment and training activity	55.6	28.5
	Job search or job club	40.3	6.4
	Education and training	29.8	22.7

APPENDIX TABLE B, continued

Sources:

Baltimore: Friedlander et al. (1985), p. 62.
California GAIN: Riccio, Friedlander, and Freedman (1994), p. 41.
Cook County: Friedlander et al. (1987), p. 45.
Florida: Kemple, Friedlander, and Fellerath (1995), pp. 51, 54.
San Diego: Goldman, Friedlander, and Long (1986), p. 194.
San Diego SWIM: Hamilton and Friedlander (1989), p. 38.
Virginia: Riccio et al. (1986), p. 82.
West Virginia: Friedlander et al. (1986), p. 73.

[a]Figures cover AFDC Basic and AFDC-UP.
[b]Means Community Work Experience Programs plus other activity.
[c]Means vocational plus postsecondary education.
[d]Means community college plus JTPA.
[e]Interval length is given in Hamilton and Friedlander (1989), p. 39.

Endnotes

[1]Adult recipients of Food Stamps are also subject to work requirements, and some pressures to enforce work have also appeared in Supplemental Security Income (SSI). But the AFDC work programs are much the most well-developed.

[2]For the politics and policy developments surrounding WIN and JOBS, see Mead (1986, 1992).

[3]"Social programs" here means compensatory service programs such as education or training that attempt to enlarge the skills or earnings of their clients, not transfer programs that distribute income, health care, or other direct benefits. The effects of transfers are more measurable, and less disputed, than those of the service programs.

[4]I set aside here other ways that experimental trials may be biased, for example because they fail to test the conditions that would prevail if the experimental program were permanently and universally instituted.

[5]The social programs of the 1930s, such as the Works Progress Administration, are fondly remembered as having helped overcome the Great Depression. But they did not undergo experimental impact evaluations. If they had, they might not have fared any better than recent antipoverty programs.

[6]Such claims were common at the "Conference on Social Programs that Really Work," where this paper was first presented, at the Institute of Government and Public Affairs, University of Illinois at Chicago, Chicago, IL, Oct. 20, 1995.

[7] Some appreciations by liberal analysts suggest that the goals are in conflict, because when recipients leave welfare they may lose more in income and other benefits, such as child and health care, than they gain. See Danziger, Sandefur, and Weinberg (1994), especially chapters 7 and 8. But research and experience show that the vast majority of welfare recipients who work are better off, at least in income terms, than without work, and few fail to work for lack of child care and health care. See Mead (1992: 71–72, 115–24).

[8]For details about SWIM and Riverside, see Appendix Table A.

[9]Clients may become mandatory for work, not only when they first enter welfare, but later. Recipients have to recertify their eligibility every six months. At one of these reassessments, it may be discovered that they now meet the requirements for the work program, perhaps because their youngest child has passed some threshold age. They will then be referred to the program, and could enter an evaluation, just like someone else who was referred on first application.

[10]In seven MDRC studies, the percentage of clients referred to work programs who ever participated in them ranged from 38 to 64 percent. See Gueron and Pauly (1991: 128–34).

[11]Admittedly, the reported impacts of the four studies with early random assignment were not larger than those of the five studies with later randomization. Indeed, they were smaller. The late-randomizers included the formidable SWIM and GAIN studies. There was, however, no statistically significant difference between the two groups.

[12]The closer to program activities randomization is, the higher participation rates will be, because there are fewer stages prior to participation for clients to drop out.

[13]In practice, evaluators do not subtract the raw group means but estimate the difference using multivariate statistical models. With this method, impacts are more likely to be statistically significant, and they are adjusted for the slight demographic differences that may exist between the experimental and control groups. Due to random assignment, such models can be used without fear of selection bias.

[14]The Family Support Act does not explicitly require JOBS to be mandatory. However, it gives states authority to compel participation by recipients, on pain of reducing their grants, and it levies participation standards on states that, in practice, force most of them to use that authority.

[15]In WIN in 1979, at the height of its implementation, 17 percent of clients were volunteers (Mead, 1988: 274 n. 33). I know of no equivalent estimate for JOBS.

[16]The third was served in the sense of participating in some form, at some time in the course of a year. In contrast, the very tough SWIM program realized its impacts by achieving participation rates

by this standard of 52 percent (Hamilton, 1988: xviii, 155–58, 162, 187). Analyses of participation in the MDRC studies usually use this standard. It is more lenient than in JOBS, which requires that participants be scheduled for 20 hours a week and attend 75 percent of these hours.

[17]In one demonstration from 1979–81, five local WIN programs that raised their participation rates from 53 to 74 percent in these years also raised their job entries by 58 percent, in the face of declining national economy. See CSR, Inc., and Osora & Associates (1982).

[18]Statistical analyses also suggest that higher participation in WIN or JOBS is associated with stronger outcomes, particularly job entries and departures from welfare, controlling for demographic features of the clients and the local labor market. See Mead (1988, 1995a). Due to the possibility of selection bias, these studies measure results less definitively than experimental evaluations, but they reveal more about which administrative assignments produce results.

[19]The correlation between the year a study began and ended and its impacts is in fact slightly inverse, because of the strong results that Baltimore and San Diego recorded early in the sequence.

[20]According to Freedman and Friedlander (1995), Riverside did not dominate the national results.

[21]For summaries of studies bearing on the public opinion and psychology of welfare, see Mead (1992: chaps. 3, 7).

[22]Congress increased work incentives in 1962 and 1967 but cut them back sharply in 1981. The main criticism was that they did not raise work levels but did make it harder to close cases when recipients went to work.

[23]In the first, but not the published, version of this paper, Harris estimates that about a third of the mothers work at a given time.

[24]Discussions of deterrence and the research problems are as yet rare. See Manski and Garfinkel (1992: chaps. 1, 6–7) and Moffitt (1996).

[25]Moffitt found that the rolls would grow 4 percent if the work program were voluntary, and 2 percent if it were mandatory and required mothers to work half the hours normal for female family heads. The rolls would fall 3 percent if the requirement were three-quarters of those hours. Since Moffitt says the average female family head works 30 hours a week and JOBS requires 20, the latter estimate comes closest to capturing the potential of JOBS.

[26]The following discussion is based largely on Mead (1995a, 1995c).

[27]An exception was an experimental test of a pre-JOBS work program in Rock County (Pawasarat and Quinn 1993: chap. 11). I am now doing a statistical analysis, modeled on Mead (1995b), to try to connect variations in caseload change around the Wisconsin counties to JOBS and other factors. Preliminary results suggest that counties that effectively enforced work and child support payments also had sharper declines in dependency, controlling for other variables.

[28]For details on the performance of these programs, see Appendix Table A.

[29]The state was a model of the efficient WIN organization as set out in Mitchell, Chadwin, and Nightingale (1980), the pioneering study of welfare employment implementation.

[30]Other than Washington, all the states that outperformed Wisconsin had smaller and less urban welfare caseloads. Data from the Office of Family Assistance, U.S. Department of Health and Human Services.

[31]JOBS exempts several groups, most importantly welfare mothers whose youngest child is under age 3, although states can lower this to age 1.

[32]The only example I know is Danziger and Kossoudji (1994–1995), a study of former general assistance (GA) recipients in Michigan who were denied coverage following program cuts in 1991. Results show limited employment and substantial hardship among the former recipients. They are, however, of limited relevance to the likeliest form of AFDC reform—tougher work requirements rather than elimination of the program. And AFDC mothers are, on average, more employable than GA clients, so the hardship produced by cuts might be less.

References

Aaron, Henry J. 1978. *Politics and the Professors: The Great Society in Perspective.* Washington, D.C.: Brookings Institution.

Bane, Mary Jo, and David T. Ellwood. 1994. *Welfare Realities: From Rhetoric to Reform*. Cambridge, Mass.: Harvard University Press.

Bardach, Eugene. 1993. *Improving the Productivity of JOBS Programs.* New York: Manpower Demonstration Research Corporation.

Berrueta-Clement, John R., et al. 1984. *Changed Lives: The Effects of the Perry Preschool Program on Youths Through Age 19.* Ypsilanti, Mich.: High/Scope Press.

Bloom, Howard S., et al. 1992. *The National JTPA Study: Title II-A Impacts on Earnings and Employment at 18 Months*. Bethesda, Md.: Abt Associates.

Brodkin, Evelyn Z. 1986. *The False Promise of Administrative Reform: Implementing Quality Control in Welfare*. Philadelphia: Temple University Press.

CSR, Inc., and Osora & Associates. 1982. *Final Report: Evaluation of the WIN Total Registrant Involvement Project*. Washington, D.C.: CSR, Inc.

Danziger, Sandra K., and Sherrie A. Kossoudji. 1994–95. “What Happened to General Assistance Recipients in Michigan?” *Focus* 16, no. 2 (Winter): 32–34.

Danziger, Sheldon H., Gary D. Sandefur, and Daniel H. Weinberg, eds. 1994. *Confronting Poverty: Prescriptions for Change*. Cambridge, Mass.: Harvard University Press.

Edin, Kathryn, and Christopher Jencks. 1992. “Reforming Welfare.” In Christopher Jencks, *Rethinking Social Policy: Race, Poverty, and the Underclass*, chap. 6. Cambridge, Mass.: Harvard University Press

Fraker, Thomas, Robert Moffitt, and Douglas Wolf. 1985. “Effective Tax Rates and Guarantees in the AFDC Program, 1967–1982.” *Journal of Human Resources* 20, no. 2 (Spring): 251–63.

Freedman, Stephen, and Daniel Friedlander. 1995. *The JOBS Evaluation: Early Findings on Program Impacts in Three Sites*. New York: Manpower Demonstration Research Corporation.

Friedlander, Daniel. 1988. *Subgroup Impacts and Performance Indicators for Selected Welfare Employment Programs*. New York: Manpower Demonstration Research Corporation.

Friedlander, Daniel, et al. 1985. *Maryland: Final Report on the Employment Initiatives Evaluation.* New York: Manpower Demonstration Research Corporation.

Friedlander, Daniel, et al. 1986. *West Virginia: Final Report on the Community Work Experience Demonstrations.* New York: Manpower Demonstration Research Corporation.

Friedlander, Daniel, et al. 1987. *Final Report on Job Search and Work Experience in Cook County*. New York: Manpower Demonstration Research Corporation.

Goldman, Barbara, Daniel Friedlander, and David Long. 1986. *Final Report on the San Diego Job Search and Work Experience Demonstration*. New York: Manpower Demonstration Research Corporation.

Goldman, Barbara, et al. 1984. *Preliminary Findings from the San Diego Job Search and Work Experience Demonstration*. New York: Manpower Demonstration Research Corporation.

Greenberg, David, and Michael Wiseman. 1992. “What Did the OBRA Demonstrations Do?” In *Evaluating Welfare and Training Programs,* ed. Charles F. Manski and Irwin Garfinkel, chap. 1. Cambridge, Mass.: Harvard University Press.

Gueron, Judith M., and Edward Pauly, with Cameran M. Lougy. 1991. *From Welfare to Work*. New York: Russell Sage Foundation.

Hamilton, Gayle. 1988. *Interim Report on the Saturation Work Initiative Model in San Diego.* New York: Manpower Demonstration Research Corporation.

Hamilton, Gayle, and Daniel Friedlander. 1989. *Final Report on the Saturation Work Initiative Model in San Diego.* New York: Manpower Demonstration Research Corporation.

Harris, Kathleen Mullan. 1993. “Work and Welfare among Single Mothers in Poverty.” *American Journal of Sociology* 99, no. 2 (September): 317–52.

Herr, Toby, Robert Halpern, and Aimée Conrad. 1991. *Changing What Counts: Rethinking the Journey Out of Welfare*. Evanston, Ill.: Northwestern University, Center for Urban Affairs and Policy Research.

Kemple, James J., Daniel Friedlander, and Veronica Fellerath. 1995. *Florida’s Project Independence: Benefits, Costs, and Two-Year Impacts of Florida’s JOBS Program.* New York: Manpower Demonstration Research Corporation.

Lipsky, Michael. 1988. “Bureaucratic Disentitlement in Social Welfare Programs.” *Social Service Review* 58, no. 1 (March): 3–27.

Lynn, Laurence E., Jr., ed. 1989. “Symposium: The Craft of Public Management.” *Journal of Policy Analysis and Management* 8, no. 2 (Spring): 284–306.

Manski, Charles F., and Irwin Garfinkel, eds. 1992. *Evaluating Welfare and Training Programs.* Cambridge, Mass.: Harvard University Press.

Mayer, Susan, and Christopher Jencks. 1995. “War on Poverty: No Apologies, Please.” *New York Times,* November 9, p. A29.

Maynard, Rebecca A. 1994. “Subsidized Employment and Non-Labor Market Alternatives for Welfare Recipients.” In *The Work Alternative: Welfare Reform and the Realities of the Job Market,* ed.

Demetra Smith Nightingale and Robert H. Haveman, chap. 5. Washington, D.C.: Urban Institute Press.

Mead, Lawrence M. 1986. *Beyond Entitlement: The Social Obligations of Citizenship.* New York: Free Press.

Mead, Lawrence M. 1988. “The Potential for Work Enforcement: A Study of WIN.” *Journal of Policy Analysis and Management* 7, no. 2 (Winter): 264–88.

Mead, Lawrence M. 1990. “Should Workfare Be Mandatory? What Research Says.” *Journal of Policy Analysis and Management* 9, no. 3 (Summer): 400–404.

Mead, Lawrence M. 1992. *The New Politics of Poverty: The Nonworking Poor in America.* New York: Basic Books.

Mead, Lawrence M. 1995a. “The New Paternalism in Action: Welfare Reform in Wisconsin.” Milwaukee: Wisconsin Policy Research Institute.

Mead, Lawrence M. 1995b. “Taming Welfare Growth: The Role of the JOBS Program.” Paper presented at the annual conference of the Association of Public Policy Analysis, Washington, D.C., November 3, 1995.

Mead, Lawrence M. 1995c. “The Change in Wisconsin Welfare.” *WI: Wisconsin Interest* 4, no. 2 (Fall/Winter): 17–24.

Mitchell, John J., Mark L. Chadwin, and Demetra S. Nightingale. 1980. *Implementing Welfare-Employment Programs: An Institutional Analysis of the Work Incentive (WIN) Program*, R&D Monograph 78, U.S. Department of Labor. Washington, D.C.: U.S. Government Printing Office.

Moffitt, Robert 1983. “An Economic Model of Welfare Stigma.” *American Economic Review* 73, no. 5 (December): 1023–35.

Moffitt, Robert A. 1996. “The Effect of Employment and Training Programs on Entry and Exit from the Welfare Caseload.” *Journal of Policy Analysis and Management* 15, no. 1 (Winter): 32–50.

Murray, Charles. 1984). *Losing Ground: American Social Policy, 1950–1980.* New York: Basic Books.

Murray, Charles. 1994. “What to Do about Welfare.” *Commentary* (December): 26–34.

Pawasarat, John, and Lois M. Quinn. 1993. *Wisconsin Welfare Employment Experiments: An Evaluation of the WEJT and CWEP Programs.* Milwaukee: Employment and Training Institute, University of Wisconsin–Milwaukee.

Piven, Frances Fox, and Richard A. Cloward. 1993. *Regulating the Poor: The Functions of Public Welfare.* Updated ed. New York: Vintage.

Quint, Janet, and Cynthia Guy. 1986. *Interim Findings from the Illinois WIN Demonstration Program in Cook County.* New York: Manpower Demonstration Research Corporation.

Riccio, James, Daniel Friedlander, and Stephen Freedman. 1994. *GAIN: Benefits, Costs, and Three-Year Impacts of a Welfare-to-Work Program.* New York: Manpower Demonstration Research Corporation.

Riccio, James, et al. 1986. *Final Report on the Virginia Employment Services Program.* New York: Manpower Demonstration Research Corporation.

Rossi, Peter H. 1987. "The Iron Law of Evaluation and Other Metallic Rules." In *Research in Social Problems and Public Policy*, ed. Joann H. Miller and Michael Lewis, vol. 4, pp. 3–20. Greenwich, Conn.: JAI Press.

Schorr, Lisbeth B., with Daniel Schorr. 1988. *Within Our Reach: Breaking the Cycle of Disadvantage.* New York: Doubleday.

Sosin, Michael R. 1986. "Legal Rights and Welfare Change, 1960–1980." In *Fighting Poverty: What Works and What Doesn't*, ed. Sheldon H. Danziger and Daniel H. Weinberg, chap. 11. Cambridge, Mass.: Harvard University Press.

State of Wisconsin, Department of Health and Social Services. 1994. *JOBS Annual Report: State of Wisconsin.* Madison: Department of Health and Social Services.

U.S. House of Representatives, Committee on Ways and Means. 1994. *Overview of Entitlement Programs: 1994 Green Book.* Washington, D.C.: U.S. Government Printing Office.

Wiseman, Michael, ed. 1991a. "Research and Policy: A Symposium on the Family Support Act of 1988." *Journal of Policy Analysis and Management* 10, no. 4 (Fall): 588–666.

Wiseman, Michael. 1991b. "Sample Family Support Act Job Opportunity and Basic Skills Training (JOBS) Participation Data (Revised)." Madison: La Follette Institute of Public Affairs, University of Wisconsin.

Wiseman, Michael. Forthcoming. "Some Strategies for Welfare Reform: The Wisconsin Story." *Journal of Policy Analysis and Management* 15, no. 4 (Fall).

U. S. Welfare Programs Antipoverty Programs. U.S. Welfare Programs fit into 13 large categories which are listed and described below. All U.S. Welfare Programs provide benefits to low-income individuals and families. The programs represent entitlements to all Americans but benefits are only paid to individuals and families with low income. Negative Income Tax â€“ Two tax credit programs are administered by the Internal Revenue Service (IRS) to distribute money to low-income Americans. The tax credits include a â€œrefundableâ€ portion which is paid to individuals and families that owe no income tax