National Evaluation of Welfare-to-Work Strategies

What Works Best for Whom: Impacts of 20 Welfare-to-Work Programs by Subgroup

U.S. Department of Health and Human Services
Administration for Children and Families
Office of the Assistant Secretary for Planning and Evaluation

U.S. Department of Education
Office of the Under Secretary
Office of Vocational and Adult Education

January 2001

Prepared by:
Charles Michalopoulos
Christine Schwartz

with
Diana Adams-Ciardullo

Manpower Demonstration Research Corporation

The Manpower Demonstration Research Corporation (MDRC) is conducting the National Evaluation of Welfare-to-Work Strategies under a contract with the U.S. Department of Health and Human Services (HHS), funded by HHS under a competitive award, Contract No. HHS-100-89-0030. Child Trends, as a subcontractor, is conducting the analyses of outcomes for young children (the Child Outcomes Study). HHS is also receiving funding for the evaluation from the U.S. Department of Education. The study of one of the sites in the evaluation, Riverside County (California), is also conducted under a contract from the California Department of Social Services (CDSS). CDSS, in turn, is receiving funding from the California State Job Training Coordinating Council, the California Department of Education, HHS, and the Ford Foundation. Additional funding to support the Child Outcomes Study portion of the evaluation is provided by the following foundations: the Foundation for Child Development, the William T. Grant Foundation, and an anonymous funder.

The findings and conclusions presented herein do not necessarily represent the official positions or policies of the funders.

Contents

Tables and Figures

Table

Figure

Acknowledgments

The foundation of this report is the commitment and hard work of hundreds of people and dozens of organizations over more than 15 years, from the beginning of the evaluation of SWIM in 1985 through the ongoing NEWWS Evaluation. Critical to the evaluations on which the present study is based was the support and assistance of administrators and staff in the 8 states, 14 sites, and 20 programs discussed. For the original and five-year follow-up studies of SWIM, the evaluators received assistance from state staff in California and county staff in San Diego County; for GAIN, state staff in California and county staff in Alameda, Butte, Los Angeles, Riverside, San Diego, and Tulare Counties; for the NEWWS Evaluation, state staff in California and county staff in Riverside County, state staff in Georgia and county staff in Fulton County, state staff in Michigan and county staff in Kent and Wayne Counties, state staff in Ohio and county staff in Franklin County, state staff in Oklahoma and county staff in Oklahoma, Cleveland, and Pottawatomie Counties, and state staff in Oregon and county staff in Multnomah and Washington Counties; for FTP, state staff in Florida and county staff in Escambia County; and for MFIP, state staff in Minnesota and county staff in Anoka, Dakota, Hennepin, Mille Lacs, Morrison, Sherburne, and Todd Counties.

The willingness of these staff to allow their programs to be studied using an elaborate research design including random assignment, to provide insights into how their programs were implemented, and to allow detailed data collection was of crucial importance. Their contributions will help shape policy for many years to come.

Executive Summary

In 1996, Congress radically transformed the nation's cash assistance welfare program when it passed the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA). The legislation replaced the 60-year-old Aid to Families with Dependent Children (AFDC) entitlement program with Temporary Assistance for Needy Families (TANF), a funding mechanism that provides states with block grants and considerable flexibility in designing their welfare programs. In addition to making other changes, many states responded by expanding their employment and training programs or changing the focus of their existing programs. A number of states replaced voluntary welfare-to-work programs that emphasized education and training with mandatory programs that stressed quick employment. While many aspects of the 1996 legislation and the state policies that followed were untested, the use of mandatory welfare-to-work programs was not. During the ten years prior to PRWORA, large-scale rigorous studies of welfare-to-work programs were launched in many states and counties. This report investigates results from 20 of these programs to determine who has benefited from welfare-to-work programs (and who has not) and whether some practices appear more effective than others at increasing the employment and earnings of single-parent welfare recipients.

The programs studied in this report share two key characteristics. They all required some portion of the welfare caseload to participate in a welfare-to-work program or risk losing some or all of their welfare benefits through sanctions. And they were all studied by the Manpower Demonstration Research Corporation (MDRC) using a rigorous experimental research design in which individuals were assigned at random either to a *program group,* which was required to participate in an employment or training program, or to a *control group,* which did not have access to the program.

In other ways, the 20 programs are quite diverse (see Table 1 for a summary of the programs). They operated in many states and counties across the country, with programs in Atlanta, Georgia; Columbus, Ohio; Detroit, Michigan; Grand Rapids, Michigan; Oklahoma City, Oklahoma; Escambia County (Pensacola), Florida; Portland, Oregon; six counties in California (Riverside, Los Angeles, San Diego, Alameda, Butte, and Tulare); and seven counties in Minnesota. While all began operating prior to the passage of PRWORA, the earliest began in 1985 and the latest are still in operation. The programs also vary in origin; most were part of state welfare-to-work programs funded under the Job Opportunity and Basic Skills Training (JOBS) program of the Family Support Act of 1988; however, one was a federal demonstration to test how high participation could be among individuals who were supposed to enroll in the program, and two were begun under waivers of the AFDC program when it was still in place. Finally, the programs vary in their approach to helping welfare recipients find work; five programs encouraged or required nearly all individuals to look for work, seven focused on basic education for most participants, and eight used a mix of the two approaches, encouraging or requiring more job-ready participants to look for work but allowing others to build skills through basic education. Although welfare-to-work programs have changed in response to welfare reform, these programs are relevant to the current policy debate; many of the 20 programs are still being operated, two contain

Table 1

Brief Descriptions of 20 Welfare-to-Work Programs

San Diego's Saturation Work Initiative Model (SWIM) operated between July 1985 and September 1987. SWIM provided a fixed sequence of services: job-search workshop, unpaid work experience, and education and training for those still jobless.
Alameda, Butte, Los Angeles, Riverside, San Diego, and Tulare Counties, California, ran versions of the Greater Avenues for Independence (GAIN) program beginning in the mid 1980s. Operated statewide, GAIN directed individuals considered "in need of basic education" to basic education, but required others to enroll in a job search activity.
Atlanta, Georgia; Grand Rapids, Michigan; and Riverside, California, operated two welfare-to-work programs each as part of the JOBS program authorized by the Family Support Act of 1988. In each site, some individuals were assigned to a Labor Force Attachment (LFA) program that required most participants to initially look for work; other individuals were assigned to a Human Capital Development (HCD) program that placed most participants in basic education.
Columbus, Ohio, tested two approaches to case management as part of the state's JOBS program. In the Columbus Traditional program, two different workers handled income maintenance and employment and training case management. In the Columbus Integrated program, one staff member handled both. Both programs were education-focused, placing most participants into basic education and some into post-secondary programs.
Detroit, Michigan, and Oklahoma City, Oklahoma, ran education-focused JOBS programs that assigned most individuals to basic education. These are the only programs studied in this report in which the mandate to participate was not strongly enforced.
Portland, Oregon's JOBS program was employment-focused; staff told clients that their goal should be to get a job. Participants were told to wait until they found a "good" job and those in need of more skills were encouraged to enroll in short-term education or training initially and look for a job later.
The Minnesota Family Investment Program (MFIP) was begun in 1994 in seven rural and urban counties in Minnesota. The MFIP policy combined a mandatory employment and training program for long-term welfare recipients with financial incentives to encourage them to work. MFIP's welfare-to-work program was an employment-focused program that encouraged participants to take a job quickly.
Florida's Family Transition Program (FTP) was implemented in 1994 in Escambia County, Florida. Participants who were considered not job-ready were allowed to participate in education and skills development; others were required to look for work. In addition to the welfare-to-work program, FTP offered financial incentives to work and imposed a time limit on receipt of welfare benefits. As a result, FTP has the key components of many states' TANF policies.

other features of states' TANF programs such as financial incentives and time limits, and most enforced the mandate to participate in their programs by using tough sanctions (although most sanction policies were not as tough as those used by many states today).

The results analyzed in this report may be particularly important at this time. In addition to giving states flexibility in designing their welfare programs, PRWORA also required a growing percentage of the welfare caseload to be working or participating in work-related activities and it imposed a five-year time limit on how long most families could receive federal support. States may be better able to meet their obligation and help welfare recipients become self-sufficient before they reach the time limit if they understand what has worked in the past and if they know which groups may require more or different types of help because they have not benefited from previous efforts.

I. The Findings in Brief

As mentioned above, people in each site were assigned at random to either a program group or a control group. Since random assignment ensured that the groups were similar at the time of random assignment, any differences that emerged between them could reliably be attributed to the mandatory welfare-to-work programs. Comparing outcomes for the program and control groups therefore reveals the effects of the program. The key findings follow.

- **For most subgroups, people in the program groups had higher earnings and lower welfare payments than people in the control groups, but generally had the same combined income from earnings, AFDC, and Food Stamps.** When samples from the 20 programs were combined, effects on annual earnings were similar for most subgroups; they exceeded $1,000 per year for only one group and were close to zero for only one group. The programs also reduced annual AFDC payments by similar amounts for all groups, with the effects ranging between $200 and $600. As a result of increased earnings and reduced welfare payments, the programs generally neither increased nor decreased combined income from earnings, welfare, and Food Stamps.

- **Measures of psychosocial well-being and barriers to work were typically not strongly related to impacts on earnings.** Private Opinion Survey data were used to define subgroups based on risk of depression, mastery, work-related parental concerns, preference for work, health or emotional problems, child care problem, and transportation problems, all measured at the time of random assignment. In general, there was little relationship between these measures and impacts. The one exception was risk of depression. The programs did not affect earnings for people at high risk of depression when they entered the study, and the programs had smaller effects for those at high risk than for those at low risk.

- **The programs increased earnings about as much for the more disadvantaged groups as for the less disadvantaged groups. Nevertheless, the more disadvantaged groups earned much less than others.** The programs increased earnings for long-term recipients, high school nongraduates, families with three children or more, and people with no recent work experience. In particular, the programs increased earnings for the most disadvantaged group: long-term recipients who did not have a high school diploma and had not worked in the year prior to random assignment. Although the programs increased earnings across the board, they typically increased earnings no more for the more disadvantaged groups than for the less disadvantaged groups. As a result, earnings for the more disadvantaged groups remained far below earnings for other groups even after participating in these programs.

- **Employment-focused programs tended to be more effective than education-focused programs for the more disadvantaged groups. Programs that provided a mix of first activities tended to help the broadest range of people.** For the more disadvantaged groups, most of the programs with the largest effects on earnings were employment-focused. Programs with an education focus rarely had large effects for these groups. In a rigorous comparison of employment-focused and education-focused programs that magnified the differences between these two types of models, programs that required nearly all participants initially to look for work had larger effects on earnings for the more disadvantaged groups than programs that enrolled most people initially into basic education. However, the two program models had similar effects for the less disadvantaged groups. A number of programs that provided a mix of first activities (some of which were employment-focused) produced large earnings gains for the more disadvantaged groups *and* the less disadvantaged groups. Thus, programs with a mix of first activities were effective for the broadest range of individuals.

II. Research Questions

This report tries to answer the question of "what works best for whom" in mandatory welfare-to-work programs for single-parent welfare recipients. Implicit in this question are three broad research issues.

- **Which groups were affected the most and the least?**

To answer the "for whom" part of the question, the report examines subgroups of single-parent families based on a number of characteristics, including educational attainment; work and welfare history; race, ethnicity, and sex; number and age of children; barriers to work because of child care, transportation, and health or emotional problems; preference for work over welfare; parental concerns about leaving family for work; and depression and feeling of mastery over life circumstances. To investigate results for a group of individuals expected to be especially hard to help, a most disadvantaged subgroup was defined to include long-term recipients (those who had ever been on welfare two years or more prior to random assignment) who had not graduated from high school and who had no earnings in the year prior to random assignment. Likewise, a least disadvantaged group was defined as individuals with none of these barriers, while individuals were considered moderately disadvantaged if they had one or two barriers. To search for an even more disadvantaged group, the most disadvantaged group was further divided by some of the psychosocial measures and barriers to work, such as risk of depression, mastery, and child care problems.

Understanding what happened to various groups will require looking at both *outcomes* — how much groups earned on average or what their average income was, for example — and *impacts* — how much average earnings or other outcomes increased or decreased because of the programs. Some groups with low earnings may not have benefited from the programs studied in this report. Likewise, some groups may be benefiting from welfare-to-work programs, but still be left without enough earnings to move completely off welfare. For those groups, policymakers may need to use new strategies such as offering post-employment services or help in overcoming substance abuse or domestic violence.

- **In what dimensions are the programs succeeding?**

In studying outcomes and impacts, the report investigates three dimensions: earnings, welfare benefits, and income. Policymakers may want to encourage welfare recipients to work; for them, the "best" program may be the one that increases employment and earnings the most. Other policymakers may be primarily interested in reducing spending on welfare; for them, the best program may be the one that reduces cash assistance the most. Welfare recipients and policymakers concerned about child and family poverty may care most about total income; for them, the best program may be the one that increases income the most.

- **Which programs or program models work best?**

These programs vary in a number of ways, including how they helped clients make the transition from welfare to work, who was enrolled in the programs, how the programs were implemented, where the programs were implemented, and the economic conditions under which they were implemented. If programs with one set of characteristics consistently outperformed others for some subgroups, policymakers might want to repeat those programs for some welfare recipients.

III. Pooled Results Across Subgroups

Published results show that most of these programs increased earnings and reduced welfare receipt overall, but led to no change in combined income from earnings, welfare, and Food Stamps. This study produced similar results for a wide range of subgroups. Overall, the programs increased earnings and reduced welfare payments for most subgroups, an encouraging finding that suggests that few groups were left behind. Table 2 summarizes these impacts for a variety of subgroups when samples from the 20 programs are combined.

- **If the objective of welfare-to-work programs is to increase earnings, this set of programs worked well for almost every group.**

The primary purpose of welfare-to-work programs is to help recipients go to work and increase their earnings. Overall, the 20 programs studied in this report succeeded in this regard. On average, they increased annual earnings by about $500 per person; that is, program group members earned about $500 more per year on average than control group members.[1] Moreover, the programs increased earnings by a similar amount across a wide range of subgroups (see Table 2). Only for new applicants did the effect on earnings exceed $1,000 and only for the group at high risk of depression did the programs not significantly increase earnings. (See the accompanying box for a discussion of statistical significance.)

[1]All dollar amounts were inflation-adjusted to 1997 dollars.

National Evaluation of Welfare-to-Work Strategies

Table 2

Impacts on Earnings, AFDC Payments, and Income Pooled Across Welfare-to-Work Programs, by Selected Characteristics at the Time of Random Assignment

		Average Total Earnings per Year, Years 1-3 ($)		Average Total AFDC Payments per Year, Years 1-3 ($)		Average Total Income per Year, Years 1-3 ($)	
Subgroup at Baseline	Sample Size	Control Group	Impact	Control Group	Impact	Control Group	Impact
Baseline characteristic							
Total earnings in past 12 months							
No earnings	41,434	1,754	571 ***	4,675	-416 ***	8,082	41
Less than $5,000	20,554	3,425	399 ***	3,696	-359 ***	8,707	-58
$5,000 or more	9,944	6,957	548 ***	2,967	-305 ***	11,200	143
Welfare history[a]			†		†		††
Long-term recipient	43,339	2,480	544 ***	4,791	-433 ***	9,027	4
Short-term recipient	21,333	3,708	534 ***	3,400	-337 ***	8,463	94
New applicant	6,853	3,025	1,106 ***	2,611	-218 **	6,819	773 ***
High school credential			††				††
No high school diploma/GED	31,139	1,867	430 ***	4,708	-395 ***	8,282	-66
High school diploma/GED	40,793	3,751	627 ***	3,749	-389 ***	8,989	123 *
Number of children			†††		†		
3 or more	18,179	2,523	682 ***	5,604	-458 ***	10,412	93
2	22,950	2,957	663 ***	4,185	-408 ***	8,769	128 *
1	30,562	3,196	328 ***	3,268	-326 ***	7,589	-65
Level of disadvantage[b]			††				
Most disadvantaged	14,393	983	404 ***	5,570	-411 ***	8,426	-116
Moderately disadvantaged	47,113	2,955	599 ***	4,066	-414 ***	8,591	79
Least disadvantaged	10,019	5,664	421 ***	2,677	-282 ***	9,558	41

(continued)

Table 2 (continued)

Subgroup at Baseline	Sample Size	Average Total Earnings per Year, Years 1-3 ($) Control Group	Impact	Average Total AFDC Payments per Year, Years 1-3 ($) Control Group	Impact	Average Total Income per Year, Years 1-3 ($) Control Group	Impact
Psychosocial indicator							
Risk of Depression Scale[c]			†		†		
High	2,507	3,071	289	3,308	-392 ***	8,281	-236
Moderate	4,157	3,138	460 ***	3,381	-442 ***	8,510	-121
Low	10,588	3,049	769 ***	3,496	-596 ***	8,529	-22
Work-Related Parental Concerns Scale					†		
High	4,786	1,973	748 ***	4,012	-551 ***	8,067	-76
Low	15,796	3,524	588 ***	3,099	-407 ***	8,424	51
Mastery Scale							
Low	7,680	2,554	623 ***	3,480	-378 ***	8,016	100
High	12,911	3,503	672 ***	3,196	-465 ***	8,495	33
~~Barrier to work or participation~~							
Health or emotional problem[d]							
Yes	5,507	2,097	552 ***	3,518	-355 ***	7,619	28
No	15,181	3,525	663 ***	3,220	-468 ***	8,563	36
Transportation problem							
Yes	7,212	2,026	725 ***	3,736	-457 ***	7,753	110
No	13,252	3,742	616 ***	3,078	-430 ***	8,616	20
Child care problem							
Yes	12,478	2,827	648 ***	3,591	-452 ***	8,397	9
No	7,832	3,666	651 ***	2,863	-439 ***	8,222	71

(continued)

Table 2 (continued)

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, Food Stamp records, Private Opinion Survey (POS) data, and Background Information Forms (BIFs).

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

[a]Sample members were classified as new applicants if they responded on the BIF that they had never received welfare in the past; as short-term recipients if they had received welfare before on their own case or their spouse's case for a total of less than two years; and as long-term recipients if they had received welfare for two years or more prior to random assignment.

[b]Individuals were classified as most disadvantaged if they had no earnings in the year prior to random assignment, did not have a high school diploma or GED at random assignment, and had received welfare for two years or more prior to random assignment; and as least disadvantaged if they had none of these characteristics. All other sample members were classified as moderately disadvantaged.

[c]Risk of depression subgroups include sample members from four NEWWS sites only: Atlanta, Grand Rapids, Riverside, and Portland.

[d]Sample members in the "has barrier" category on this measure could have had a health or emotional problem themselves, which they reported as a barrier to work or participation at random assignment, or a family member could have had such a problem.

Defining Statistical Significance

Statistical significance is used to determine whether estimated differences between two groups are real or due to chance. Usually, statistical significance is defined at a certain level. Thus, if a difference is statistically significant at the 5 percent level, the implication is that there is only a 5 percent chance that the difference is due to chance. In this report (which follows generally accepted practices), the minimum acceptable level of statistical significance is 10 percent. Any difference with a significance level less than or equal to 10 percent is described as being *statistically significant* (or not likely to be due to chance). Any difference with a significance level greater than 10 percent is described as *not statistically significant* (or possibly due to chance).

- **Measures of psychosocial well-being and barriers to work were typically not strongly related to impacts on earnings.**

Private Opinion Survey (POS) data from some of the programs were used to define subgroups based on risk of depression, mastery, work-related parental concerns, preference for work; and health or emotional, child care, and transportation barriers to work, all measured at the time of random assignment. In general, there was little relationship between these measures and impacts (see Table 2). The one exception was risk of depression. The programs did not affect earnings for people at high risk of depression when they entered the study, and had significantly smaller effects for those at high risk than for those at low risk. These results are consistent in some ways with the programs that were studied. While most provided assistance with child care and transportation, few explicitly tried to address psychological problems.

- **If the objective of welfare-to-work programs is to reduce welfare payments, this set of programs succeeded for most subgroups.**

A second objective of welfare-to-work programs is to reduce the use and cost of welfare programs. This may occur directly through sanctioning or by creating a burden that makes people want to leave welfare. However, the primary mechanism for reducing welfare payments is the work that results from the programs' services. In all programs studied in this report, an individual's welfare benefit was reduced by some amount if she earned above a threshold known as the earnings disregard. Since the programs significantly increased earnings, they should also have reduced welfare benefit amounts, and they did. On average, they reduced annual welfare payments by nearly $400 and reduced Food Stamp payments by another $100 (not shown in Table 2).

Just as the programs increased earnings by about the same amount for a broad range of subgroups, they tended to reduce welfare payments by similar amounts for most subgroups (see Table 2). In fact, impacts on welfare payments were, if anything, more similar across subgroups than were impacts on earnings. For no subgroup did the annual impact on welfare payments fall below $200 or rise above $600.

- **If the objective of welfare-to-work programs is to increase income from earnings and public assistance, welfare-to-work programs succeeded for few groups, but were more likely to have increased income for the less disadvantaged groups.**

As described above, the programs' effects on earnings were about the same as their effects on welfare plus Food Stamps. As a result, the programs did not significantly increase combined income from earnings, welfare, and Food Stamps. A few subgroups were exceptions to this result, although all of the exceptions occurred for the less disadvantaged subgroups (see Table 2). The programs increased annual income by nearly $800 for new applicants but barely changed income for long-term recipients, and they increased income by more than $100 for high school graduates but did not significantly change income for nongraduates.

Although the programs did not increase income for most subgroups, they also did not decrease income for most subgroups. This might be viewed as a positive result for two reasons. First, the programs might have reduced income because individuals were either sanctioned or lost their job and decided not to reapply for welfare benefits. Although this probably happened for some *individuals*, there is no evidence that it occurred so frequently that the average income of entire groups was reduced. Second, the income amounts shown in Table 2 reflect only welfare, Food Stamps, and earnings. In particular, they exclude income from the federal Earned Income Credit (EIC), a source of considerable income for working poor families, and the programs' impacts on income would have been bigger if the EIC had been included.[2] At the same time, the calculation of income also ignores a number of work-related expenses, such as payroll and income taxes, child care costs, and transportation costs.

IV. Impacts for the More Disadvantaged Subgroups

All but one of the programs being studied met the provisions of the JOBS program, which were designed to benefit those most likely to be long-term recipients. An important question, therefore, is whether the programs succeeded for their targeted groups. The broad answer is that they did. As discussed above, the programs increased earnings for most groups, including the more disadvantaged groups. However, several important results warrant further discussion.

- **The programs increased earnings about as much for the most disadvantaged groups as for the moderately and least disadvantaged groups.**

As discussed above, impacts on earnings were spread fairly evenly across subgroups. Earnings gains due to the programs were as large for long-term recipients as for short-term recipients; almost as large for high school graduates as for nongraduates; slightly larger for families with three children or more than for families with one child; and larger for people with no recent work experience than for those with some recent work experience. An especially encouraging finding is that impacts on earnings for the group classified as the most disadvantaged were about as large as those for the least disadvantaged group and almost as large as those for the moderately disadvantaged group.

[2]This measure of income also excludes other income sources and income from other household members. In the studies in which the information has been collected through surveys, however, the impact on other income sources has generally been small.

- **The programs reduced welfare payments more for the more disadvantaged groups than for the less disadvantaged groups.**

As discussed above, reductions in welfare payments were fairly similar across subgroups. However, there is a hint that reductions were slightly greater for the more disadvantaged groups. For example, welfare payments were reduced by twice as much for long-term recipients as for new welfare applicants even though the programs' impact on earnings was twice as large for new applicants as for long-term recipients. Likewise, welfare reductions were nearly identical for high school graduates and nongraduates, even though high school graduates had significantly larger earnings impacts. Welfare reductions were also almost twice as much for the most disadvantaged sample members as for the least disadvantaged sample members; however, earnings impacts were also higher for the most disadvantaged group.

- **The programs did not increase earnings for sample members at high risk of depression but increased earnings substantially for those at low risk.**

Welfare-to-work programs have been designed to help people with few job skills and little work experience. However, a disproportionate number of welfare recipients also exhibit symptoms of depression, and depression may keep them from taking advantage of welfare-to-work programs and from working. As indicated above, this report finds reason to be concerned. Overall, the programs did not increase the earnings of sample members at high risk of depression, but increased the earnings of those at low risk by a substantial amount. At the same time, the programs decreased welfare payments to those at high and at low risk by a similar amount. Regardless of risk of depression, however, the programs neither significantly increased or decreased combined income from earnings, AFDC, and Food Stamps.

- **The effects of the programs depended on the kind of disadvantage an individual suffered from.**

In an analysis not shown in Table 2, individuals who were receiving welfare at the time of random assignment were divided into eight groups according to whether they were long-term recipients, whether they had graduated from high school, and whether they had recent work experience. Earnings impacts were larger for more disadvantaged groups if the disadvantages included lack of prior work experience, but smaller if the disadvantages included lack of a high school diploma. They were about the same for long-term recipients as for others. This analysis suggests impacts are related not to the number but to the kind of disadvantage.

- **Measures of psychosocial well-being did not help define a new group of the hard to serve who were not being helped by the programs.**

As welfare rolls decline, states are being left with a caseload that is harder to serve than the individuals who were randomly assigned in these programs. To try to define an extremely disadvantaged group, the most disadvantaged group shown in Table 2 was further divided according to the psychosocial measures described above (risk of depression, mastery, and so on.). In general, the psychosocial measures did not help define a new group of the extremely disadvantaged who were not benefiting from the programs. For example, the programs significantly increased earnings for members of the most disadvantaged group who were also at high risk of depression. Moreover, this impact on earnings was about as large for the most disadvantaged sample members at low risk of depression. (Although the programs did not significantly increase earnings for the group at high risk of depression overall, this was due to low earnings impact for the *least* disadvantaged sample members at high risk of depression.)

V. Outcomes for the More Disadvantaged Subgroups

One objective of welfare-to-work programs is to increase the earnings of welfare recipients. A related objective is to help welfare recipients earn enough to end their reliance on public assistance. This is an especially important goal under time-limited welfare. Even if welfare-to-work programs increase earnings levels, those levels might remain too low to eliminate a family's need for welfare. For families who eventually reach the time limit and lose their welfare benefits, their income might then be insufficient to meet even basic needs such as food and housing.

- **Despite positive effects on earnings for the more disadvantaged welfare recipients, absolute levels of earnings remained particularly low for these groups.**

During the three-year follow-up period studied in this report, the more disadvantaged members of the control group earned substantially less on average than others (see Table 2). Individuals with no earnings in the year prior to random assignment earned only one-fourth as much as those with $5,000 or more in prior-year earnings.[3] The same was true for other subgroups. Sample members who had not graduated from high school earned only half as much as those who had graduated. Long-term recipients also earned substantially less than short-term recipients. The most troublesome outcome, however, is the average earnings level for the most disadvantaged group (long-term recipients who have not graduated from high school and who have no recent work experience). For control group members in this subgroup, average annual earnings over the three-year follow-up period were less than $1,000 compared with almost $6,000 for the least disadvantaged group. Although the welfare-to-work programs increased earnings across the board, they typically increased earnings no more for the more disadvantaged groups than for the less disadvantaged groups. As a result, earnings for the more disadvantaged groups were as far below earnings for other groups *after* participating in these programs as they were before, and new policies may be needed to raise their earnings.

- **The sample members at high risk of depression were financially as well off as those at low risk.**

As described above, individuals at high risk of depression were one of the few subgroups that did not have significant earnings impacts from these mandatory welfare-to-work programs. In terms of economic well-being, however, depression might not be as important as work experience, education, and welfare history. Although the programs did not increase earnings for those at high risk of depression, Table 2 shows that the average annual earnings and income were similar for control group members at high and at low risk. In contrast, earnings for high school nongraduates fell far below earnings for

[3]Since average earnings includes zero earnings for people who are not working, some of the differences across subgroups are due to lower employment rates. For example, people with no earnings in the year prior to random assignment were only half as likely to work as those with $5,000 or more in prior-year earnings (not shown in Table 2). Even among those who worked, however, people with no earnings in the year prior to random assignment earned about half as much as those with $5,000 or more in prior-year earnings (not shown in Table 2).

graduates, and earnings for people with no recent work experience were much lower than earnings for people with substantial recent work experience.

VI. Evidence on Which Approaches Work Best

The previous sections argued that the welfare-to-work programs as a group increased earnings for the more disadvantaged and the less disadvantaged groups by similar amounts. Although the pooled results show few differences across subgroups, it is possible that some program models performed better than others for some subgroups. The four categories shown in Table 3 provide one means of classifying the program models. Although program model is an important dimension on which to compare the programs, it is important to remember that the programs differed in a number of other dimensions, including who was enrolled, when and where programs took place, and the economic conditions at the time they took place.

The largest of the four categories shown in Table 3 contains the education-focused programs which sought to place most participants initially in basic education (the three HCD programs, the two Columbus programs, Detroit, and Oklahoma City). At the other extreme are the four employment-focused programs with job search as the first activity for most participants (the three LFA programs and SWIM). Four other programs (Riverside GAIN, Portland, FTP, and MFIP) were also employment-focused, but they used a mix of first activities by enrolling more job-ready individuals in job search and allowing or directing others to enroll in basic education. Finally, the remaining five GAIN sites used a mix of activities without an employment focus. Even though the six GAIN sites followed the same policy, Riverside differed from the other five in that nearly all staff emphasized quick employment to participants; in the other five sites, most staff did not.

National Evaluation of Welfare-to-Work Strategies

Table 3

Summary of Self-Sufficiency Approaches of 20 Welfare-to-Work Programs

Education-Focused	Mix of First Activities Without Employment Focus	Employment-Focused With Mix of First Activity	Employment-Focused With Job Search as First Activity
Atlanta HCD	Alameda GAIN	Riverside GAIN	Atlanta LFA
Grand Rapids HCD	Butte GAIN	Portland	Grand Rapids LFA
Riverside HCD	Los Angeles GAIN	Florida FTP	Riverside LFA
Columbus Integrated	San Diego GAIN	Minnesota MFIP	San Diego SWIM
Columbus Traditional	Tulare GAIN		
Detroit			
Oklahoma City			

- **Employment-focused programs tended to be more effective than education-focused programs for the more disadvantaged groups. Portland and River-**

side GAIN, two of the employment-focused programs that allowed some individuals to build skills through basic education, were especially effective.

Over the three-year follow-up period, employment-focused programs produced four of the five largest earnings impacts for individuals with no earnings in the year prior to random assignment, for long-term welfare recipients, and for the most disadvantaged group and three of the five largest earnings impacts for high school nongraduates (see Table 4). Programs with an education focus are listed only once. Even in the third year of follow-up (not shown), after individuals initially enrolled in basic education had time to gain some skills and then find work, most of the programs with the largest effects on earnings were employment-focused, and education-focused programs barely made the list of the most effective programs for the more disadvantaged groups. Two programs in particular stand out from the rest. Riverside GAIN produced the second or third largest average earnings impact for each group of the more disadvantaged people shown in the upper part of Table 4. Portland's JOBS program likewise produced some of the largest impacts for each group. Both programs were employment-focused, but both also used a mix of job search and basic education as first activities.

- **Programs with a mix of activities tended to help the widest range of individuals.**

Programs with a mix of activities dominate the list of the most effective programs for the *less* disadvantaged participants (the lower part of Table 4). GAIN programs were especially effective for the less advantaged participants, but FTP and Portland's JOBS program were also effective for some of these groups. Programs with a mix of first activities were also frequently effective for the more disadvantaged participants. This is largely because Riverside GAIN and Portland were so successful — two programs that were also employment-focused — but MFIP and the GAIN program in Butte also produced large earnings impacts for these groups (as did FTP and the GAIN program in San Diego in the third year of follow-up; not shown in Table 4). Thus, programs with a mix of first activities were effective for the broadest mix of individuals.[4]

It is interesting that programs with a mix of first activities did better than education-focused programs for the more disadvantaged groups even though both emphasized basic education for the more disadvantaged. Likewise, it is interesting that they did better than job search programs for the less disadvantaged groups even though both emphasized job search for job-ready participants. The broad success of the mixed programs may indicate that determining whether individuals need basic education is more difficult than determining whether they have graduated from high school or worked recently. In fact, the programs with a mix of first activities used other criteria, such as scores on tests of basic skills and English proficiency. Thus, programs with a mix of first activities

[4]A number of programs did not randomly assign new applicants (including Los Angeles and Tulare in GAIN, and most of the programs evaluated as part of NEWWS). In addition, this report includes only long-term welfare recipients from MFIP because others in MFIP were not immediately required to participate in employment and training services. Therefore, only 8 of the 20 programs being studied were among the most effective for new applicants.

National Evaluation of Welfare-to-Work Strategies

Table 4

Programs with Largest Impacts on Earnings in Years 1-3
Across 20 Welfare-to-Work Programs,
for More Disadvantaged and Less Disadvantaged Groups

	More Disadvantaged Groups			
Ranking by Impact Size	No Earnings in Year Prior to Random Assignment	Without High School Diploma or GED	Long-Term Welfare Recipients	Most Disadvantaged
Largest impact	Portland ($1,476)	Butte GAIN ($1,257)	Butte GAIN ($1,445)	Minnesota MFIP ($1,115)
2nd largest impact	Riverside GAIN ($1,262)	Riverside GAIN ($1,029)	Riverside GAIN ($1,296)	Grand Rapids LFA ($1,035)
3rd largest impact	Minnesota MFIP ($1,074)	Grand Rapids LFA ($838)	Portland ($1,222)	Riverside GAIN ($1,026)
4th largest impact	Riverside LFA ($782)	Columbus Integrated ($808)	Riverside LFA ($742)	Portland ($701)
5th largest impact	Alameda GAIN ($659)	Portland ($767)	Atlanta LFA ($586)	Riverside LFA ($668)

	Less Disadvantaged Groups			
Ranking by Impact Size	Earnings of $5,000 or More in Year Prior to Random Assignment	With High School Diploma or GED	Short-Term Welfare Recipients	Least Disadvantaged
Largest impact	Butte GAIN ($3,670)	Riverside GAIN ($1,780)	Riverside GAIN ($1,409)	Riverside GAIN ($1,976)
2nd largest impact	Riverside GAIN ($1,917)	Alameda GAIN ($1,203)	San Diego GAIN ($1,022)	Butte GAIN ($1,593)
3rd largest impact	San Diego GAIN ($1,471)	Portland ($1,202)	Portland ($1,012)	San Diego GAIN ($1,549)
4th largest impact	Detroit ($1,260)	San Diego GAIN ($1,030)	Butte GAIN ($885)	SWIM ($1,504)
5th largest impact	Grand Rapids HCD ($970)	Florida FTP ($833)	Florida FTP ($840)	Florida FTP ($698)

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records and Background Information Forms (BIFs).

NOTE: The parenthetical numbers are the programs' three-year earnings impacts.

may have been more effective at increasing earnings because they effectively determined who would benefit from job search and who would benefit from basic education.

- **Programs that required most individuals to immediately look for work increased earnings faster than programs that directed most toward basic education, but those differences dissipated over time. Nevertheless, for the more disadvantaged groups, programs that emphasized job search increased earnings overall more than programs that emphasized basic education.**

Post-AFDC welfare-to-work programs have primarily used a "work-first" approach that encourages recipients to look for work immediately. However, many welfare recipients and advocates for welfare recipients decry the lack of opportunities to augment skills through education. Atlanta, Grand Rapids, and Riverside provide the best comparison of the two approaches. In each site, two programs operated side by side. While one program emphasized quick job entry (labor force attachment, or LFA) by requiring most participants to initially look for work, the other emphasized basic education (human capital development, or HCD) and enrolled most individuals initially in basic education. People were randomly assigned to one of the two programs, so that any differences in impacts of the programs were due to differences in the programs themselves, particularly the different emphases.

For several subgroups that were examined, the LFA programs initially produced larger earnings impacts than the HCD programs (see Table 5), but differences in earnings impacts were no longer statistically significant for any of the subgroups by the third year of the follow-up period. Over the three-year period, however, the LFA programs produced significantly higher earnings impacts than the HCD programs for four groups of the more disadvantaged recipients: those without a high school diploma or GED, those at high risk of depression, those with no earnings in the year prior to random assignment, and those considered the most disadvantaged. In comparison, the LFA and HCD programs produced essentially the same earnings impacts over the three-year period for the less disadvantaged counterparts of these groups. Five years of follow-up information will eventually be available for people in all of these programs, and it will be interesting to see how the two approaches compare over a longer period.

VII. Policy Implications

For a policymaker or program administrator, the results in this report yield several important implications.

- **It is possible to help the most disadvantaged participants if resources are targeted toward them and programs are developed to meet their needs.**

The Family Support Act of 1988 required states to target welfare-to-work programs toward welfare recipients who were the most likely to have a very long stay on welfare and the least likely to work. States were also required to offer a mix of services that were thought most likely to benefit this hard-to-serve group and to subsidize child care, transportation, and work-related expenses for participants in their welfare-to-work programs. Most of the programs studied in this report were either oper-

National Evaluation of Welfare-to-Work Strategies

Table 5

Impacts on Earnings in the LFA and HCD Programs for Selected Subgroups

	Year 1			Year 3			Years 1-3		
Program and Subgroup	LFA	HCD	Difference	LFA	HCD	Difference	LFA	HCD	Difference
By high school credential ($)									
No high school diploma or GED	658 ***	160 *	498 ###	625 ***	504 ***	121	636 ***	319 ***	317 ###
High school diploma or GED	415 ***	183	232 #	314 *	431 **	-116	366 **	373 **	-7
By risk of depression ($)									
High risk	675 **	-193	869 ###	175	-369	544	417	-201	618 ##
Moderate risk	476 **	110	366 #	553 *	667 **	-114	462 **	363 *	99
Low risk	547 ***	271 **	276 ##	499 ***	730 ***	-231	540 ***	536 ***	4
By earnings in past 12 months ($)									
No earnings	522 ***	158 **	364 ###	566 ***	522 ***	44	535 ***	377 ***	157 #
Less than $5,000	540 ***	36	504 ###	445 **	264	182	468 ***	166	302 #
$5,000 or more	525	272	253	-135	341	-477	284	289	-6
By level of disadvantage ($)									
Most disadvantaged	500 ***	120	380 ###	690 ***	493 ***	198	602 ***	316 ***	286 ##
Moderately disadvantaged	624 ***	161	463 ###	449 ***	478 ***	-29	540 ***	374 ***	167
Least disadvantaged	n/a	n/a		n/a	n/a		n/a	n/a	

SOURCE: MDRC calculations from unemployment insurance (UI) earnings records.

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

A two-tailed t-test was applied to differences between outcomes for the two programs. Statistical significance levels are indicated as # = 10 percent, ## = 5 percent, and ### = 1 percent.

N/a = not applicable.

ated under the Family Support Act or anticipated the key requirements of the act. As described above, the programs did increase earnings for the more disadvantaged groups.

In studying a group of mandatory but lower-cost welfare-to-work programs from the early 1980s, Daniel Friedlander (*Subgroup Impacts and Performance Indicators for Selected Welfare-to-Work Programs.* New York: MDRC, 1988) found, in contrast, that earnings impacts were small for the more disadvantaged. Since the programs studied by Friedlander preceded the FSA in both time and character, the comparison suggests that the approach of the FSA was more successful in increasing earnings of the more disadvantaged. More broadly, it suggests that it is possible to help the more disadvantaged participants.

- **A mix of job search and education increases earnings the most for the broadest range of individuals.**

Most of the programs with the largest effects on earnings used a mix of job search and basic education as first activities. People who appeared to be ready to work were required to look for work, but participants who lacked basic skills were allowed to enroll in basic education. For the more disadvantaged groups, programs with a mix of first activities were especially effective if they were also employment-focused, suggesting that program administrators may want to build programs that have a mix of services. Some caution should be used in interpreting this result, however. There has been no direct, rigorous comparison of a program with a mix of first activities with a program that emphasized primarily job search or basic education. The success of the mixed programs could stem from other factors such as the state of the economy or program location (most of the programs that used a mix of first activities were in California, for example).

- **Job search rather than education increases earnings quickly**.

If resources limit a program to one activity for most participants, that activity should be job search if the objective is to increase employment and earnings quickly. This makes sense, since people who are in school have less time to work and earn. By the third year of follow-up, for example, the two approaches were about equally effective at increasing earnings. Over a three-year period of time, however, job search appeared to increase earnings more than basic education for the more disadvantaged participants (but not for the less disadvantaged participants).

- **Psychological problems may still be an impediment to the success of welfare-to-work programs.**

This report investigated the impact of welfare-to-work programs by risk of depression and feelings of self-efficacy. Although individuals at high risk of depression in the control group fared as well in the labor market as those at low risk, the former group was less able to capitalize on the ability of welfare-to-work programs to increase earnings. These results suggest that welfare administrators may need to implement different or more intensive interventions for the depressed. It also suggests that further research is needed to understand whether other psychological problems limit the effectiveness of welfare-to-work programs.

Chapter 1

Introduction

The passage of the federal Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) in 1996 allowed states to make substantial changes in their welfare policies. Although policy changes regarding time limits have received the greatest attention, most states also changed their policies in several ways to more directly help or encourage welfare recipients to find work. Most have introduced financial incentives to make work pay and to increase income for families in which the parent does work. Many have expanded their supports for the working poor — for example, by increasing child care subsidies. In addition, many states have altered or expanded their welfare-to-work programs. While welfare-to-work programs prior to PRWORA were often de facto voluntary programs designed to increase skills through basic education and training, most states now operate mandatory employment-focused programs that emphasize quick employment and that enforce their mandates through sanctions.

This report examines one aspect of the new policies — welfare-to-work programs — by studying 20 mandatory welfare-to-work programs implemented prior to the passage of PRWORA. The programs share two distinguishing features. They all required some portion of the welfare caseload to participate in a welfare-to-work program, and they were all studied by the Manpower Demonstration Research Corporation (MDRC) using a rigorous experimental research design in which individuals were randomly assigned either to a program group, which was required to participate in a welfare-to-work program, or to a control group, which did not have access to the program. Although welfare-to-work programs have changed in response to welfare reform, these pre-PRWORA programs are still relevant; many of the 20 programs are still being operated, and two contain other features of states' Temporary Assistance for Needy Families (TANF) programs such as financial incentives and time limits.[1]

This report tries to answer the question of "what works best for whom" among these welfare-to-work programs. Implicit in this question are three issues. For which identifiable subgroups of individuals did the programs have the largest and smallest impacts? Did successful programs affect only earnings and welfare benefits, or did they also increase income from earnings and public assistance? Which programs or program models had the most promising effects either for a broad range of subgroups or for some particularly important groups?

Welfare-to-work programs have existed for several decades, and earlier studies tried to determine who does and does not benefit from such programs. The Work Incentive Program (WIN), which began in 1967, required single parents who were receiving welfare to participate in a welfare-to-work program if they did not have children under age 6. A number of state and local WIN programs were

[1]While these programs cover many of the practices used in current welfare-to-work programs, they were not intended to be and should not be considered representative of welfare-to-work programs in general. Many of the welfare-to-work programs included in this report were chosen because of their promising or innovative practices. In addition, programs were chosen for other reasons that may make them unlike the "average" welfare-to-work program. For example, programs had to have data systems capable of organizing and transmitting data to MDRC and willing to meet research requirements.

evaluated during the early 1980s using random assignment and results were both encouraging and cautionary. In summarizing results from these studies, Gueron and Pauly (1991) and Friedlander and Burtless (1995) found that the programs generally increased employment and earnings and were cost-effective. In comparing impacts for various groups in five of these studies, however, Friedlander (1988) found that earnings gains were concentrated not among the most disadvantaged, who are expected to have a hard time finding work and leaving welfare, nor among the least disadvantaged, who are most likely to work without assistance from a welfare-to-work program. Instead, Friedlander found the largest earnings gains among a middle group of welfare applicants who had spent some but not a great deal of prior time on welfare. In contrast, welfare savings came primarily from long-term recipients, especially those without a high school diploma or with little recent work experience.

Partly in response to these findings, the Family Support Act of 1988 (FSA) created the Job Opportunities and Basic Skills Training (JOBS) program, which required states to target resources toward welfare recipients who were the most likely to have a very long stay on welfare and the least likely to work. States were also required to offer services that were thought most likely to benefit this hard-to-serve group, with an emphasis on basic education for those without a high school diploma. Most of the programs in this report were either operated under FSA or anticipated the key requirements of the act. Therefore, one of the report's objectives is to determine whether this approach worked, that is, whether programs successfully increased the employment and earnings of the more disadvantaged participants by concentrating resources on them and by changing the nature and intensity of services to reflect their needs.

The JOBS program was ended by TANF, the 1996 block grant program. To prevent reductions in their block grants, states must meet federal "participation standards" by engaging large portions of TANF recipients in work or work-related activities. These standards have prompted states and counties to increase the breadth of the caseload required to enroll in welfare-to-work programs. In addition, some states have increased their use of sanctioning and the penalties imposed by sanctions to enforce the expanded mandates (Quint et al., 1999). By broadening the caseload covered by welfare-to-work programs, states have most likely increased the number of the most disadvantaged recipients participating in welfare-to-work programs. Likewise, with the advent of welfare time limits and increasingly mandatory programs many of the least disadvantaged recipients have left the welfare rolls. Therefore, the typical welfare recipient subject to welfare-to-work programs may now be more disadvantaged than previously and the most disadvantaged may not have even been part of welfare-to-work programs in the past. This report attempts to determine whether this most disadvantaged subgroup is also helped by welfare-to-work programs. While no program in the study enrolled as broad a cross section of the welfare caseload as states are currently enrolling, the number of the most disadvantaged individuals is large enough that we may reliably assess the impacts of these programs on this group. Moreover, a number of the studies included survey questions that allow us to ask whether psychosocial factors such as depression help define a new group of extremely disadvantaged individuals who might not have benefited from these programs.

In addition to enrolling more of the caseload, many post-AFDC welfare-to-work programs have used a primarily "work-first" approach that encourages recipients to take a job quickly. However, many welfare recipients and advocates for welfare recipients decry the lack of opportunities to augment meager skills through education and training. This report also seeks to inform the debate by determining whether programs that encouraged basic education helped the most disadvantaged more or less than programs that were more focused on immediate employment. Of the 20 programs included in the report, five sought to place individuals quickly into a job by requiring most to look for work initially. Another seven were more focused on basic education (for example, Adult Basic Education, English as a Second Language, and GED preparation), though not exclusively so, and the ultimate objective of these education-focused programs was to help participants obtain a job by increasing their marketable skills. The remaining eight programs used a mix of the two approaches by requiring individuals deemed able to find a reasonable job to look for a job but encouraging or requiring those who were judged to need immediate improvement in their skills to enroll in basic education. Especially informative are comparisons of programs in three sites that operated both education- and employment-focused programs. In these sites, people were assigned at random to one of the two programs, providing a rigorous comparison of the two approaches.

The "most disadvantaged" can be defined in a number of different ways. Long-term welfare recipients who have received welfare for a number of years seem the most likely to run into TANF's five-year lifetime limit. People with little recent work experience may have the fewest relevant work skills and therefore have the hardest time finding work. A host of family problems such as health or emotional problems of a child or other family member, and lack of affordable, acceptable, or accessible child care are also likely to be relevant. The depressed, those who feel they do not have control over their destiny, and those with poor attitudes about work may all struggle in the labor market. All 20 studies include baseline information on demographic characteristics such as work history, welfare history, age, education, and number of children. In addition, in nearly half of the programs studied in this report a Private Opinion Survey (POS) asked participants at random assignment to reveal information about such characteristics as risk of depression, barriers to work, preference for work, and family attachment. Not only will the report be able to provide a systematic examination of the role of these characteristics across a number of programs, it will also explore whether these characteristics helped define "disadvantaged."

The remainder of the report is organized as follows. Chapter 2 describes the 20 programs being studied with an eye toward trying to understand similarities and differences that may affect subgroup impacts. To help the reader understand how subgroup comparisons are made, Chapter 3 examines in depth one measure of current interest: risk of depression. Chapter 4 describes outcomes for the control group. The next chapters, the heart of the report, examine subgroup impacts based on demographic information (Chapter 5), composite subgroups defined by several characteristics (Chapter 6), and subgroups based on psychosocial indicators and barriers to work (Chapter 7). Chapter 8 presents a preliminary attempt to understand variation in impacts across subgroups and programs.

Chapter 2

Characteristics of the Programs

This report presents results for single-parent families from 20 different welfare-to-work programs operated in eight states and more than a dozen counties over a period of more than 10 years. Despite their diverse settings, the programs share two important features. First, they all required some part of the welfare caseload to participate in an employment or training program with the ultimate objective of helping clients leave welfare for work. Some programs tried to accomplish this primarily through job search, some assigned most clients initially to basic education and training, others used a mix of first activities, allowing participants who lacked basic skills enroll in basic education but requiring others to start by looking for work. Two programs also contained financial incentives to encourage work, and one program placed a time limit on welfare receipt.

Second, all of these programs were evaluated rigorously using random assignment. When people entered a study they were assigned at random either to a *program group,* which was subject to the rules and requirements of the welfare-to-work program, or to a *control group,* which did not have access to the program.[1] Because people were assigned at random, any differences that developed after random assignment could reasonably be attributed to the different policies they faced.

This chapter provides some background information on the 20 programs, including the program models and the characteristics of the sample members and the sites. The programs include:[2]

- San Diego's Saturation Work Initiative Model (SWIM)
- Six programs from California's Greater Avenues for Independence (GAIN)
- Eleven programs from the National Evaluation of Welfare-to-Work Strategies (NEWWS)
- Minnesota's Family Investment Program (MFIP)
- Florida's Family Transition Program (FTP)

Table 2.1 summarizes the 20 programs, and a fuller description is provided in the remainder of this section.

[1]As discussed below, in four of the sites people were randomly assigned to a control group or one of two program groups with somewhat different interventions.

[2]All 20 programs have three years or more of information on outcomes for sample members after random assignment. Three other programs studied since 1990 by MDRC — Florida's Project Independence, Los Angeles's Jobs First GAIN, and Connecticut's Jobs First — also had mandatory welfare-to-work services, but did not have three years of follow-up data when this report was written.

National Evaluation of Welfare-to-Work Strategies

Table 2.1

Summary of Programs

Program	Program Activities	Coverage and Mandatoriness	Sample and Site Characteristics	Other Features	Outcome	Annual Impacts ($)
SWIM (San Diego)	Two-week job search workshop followed by Employment Work Experience Program (EWEP) and job club; if no job after 13 weeks, assessed for education and training	Broad coverage Mandatory for parents with no child under age 6	Majority applicants Broad racial/ethnic mix Welfare grant: $617 (1986) Unemployment rate: 5.0%		Earnings Years 1-3 Year 3 AFDC Years 1-3 Year 3	 652*** 677** -694*** -661***
GAIN						
Alameda	ABE, GED preparation, or ESL if no high school diploma, lacked basic reading and math skills, or non-English-speaking; job search for others	Enrolled only long-term recipients Mandatory for parents with no child under age 6	Long-term recipients only Majority African-American Welfare grant: $694 (1989) Unemployment rate: 4.4%		Earnings Years 1-3 Year 3 AFDC Years 1-3 Year 3	 545* 839** -202 -316
Butte	ABE, GED preparation, or ESL if no high school diploma, lacked basic reading and math skills, or non-English-speaking; job search for others	Broad coverage Mandatory for parents with no child under age 6	Majority applicants Primarily white Rural county Welfare grant: $694 (1989) Unemployment rate: 8.0%	Delayed enrolling many participants for several months to keep cases per worker low	Earnings Years 1-3 Year 3 AFDC Years 1-3 Year 3	 908** 1,108** -90 -70
Los Angeles	ABE, GED preparation, or ESL if no high school diploma, lacked basic reading and math skills, or non-English-speaking; job search for others	Enrolled only long-term recipients Mandatory for parents with no child under age 6	Long-term recipients only 50% Hispanic, 35% African-American 50% with child under age 6 Welfare grant: $694 (1989) Unemployment rate: 4.6%		Earnings Years 1-3 Year 3 AFDC Years 1-3 Year 3	 96 165 -354*** -277***

(continued)

Table 2.1 (continued)

Program	Program Activities	Coverage and Mandatoriness	Sample and Site Characteristics	Other Features	Outcome	Annual Impacts ($)
Riverside	ABE, GED preparation, or ESL if no high school diploma, lacked basic reading and math skills, or non-English-speaking; job search for others	Broad coverage Mandatory for parents with no child under age 6	Broad welfare history mix 50% white, 25% African-American Welfare grant: $694 (1989) Unemployment rate: 5.7%	Strongest employment focus of the six GAIN programs studied	Earnings Years 1-3 Year 3 AFDC Years 1-3 Year 3	 1,385*** 1,312*** -731*** -619***
San Diego	ABE, GED preparation, or ESL if no high school diploma, lacked basic reading and math skills, or non-English-speaking; job search for others	Broad coverage Mandatory for parents with no child under age 6	Broad welfare history and racial/ethnic mix Welfare grant: $694 (1989) Unemployment rate: 4.1%		Earnings Years 1-3 Year 3 AFDC Years 1-3 Year 3	 613*** 741*** -390*** -339**
Tulare	ABE, GED preparation, or ESL if no high school diploma, lacked basic reading and math skills, or non-English-speaking; job search for others	Broad coverage Mandatory for parents with no child under age 6	Majority long-term recipients 50% white, 40% African-American Agricultural county Welfare grant: $694 (1989) Unemployment rate: 10.3%		Earnings Years 1-3 Year 3 AFDC Years 1-3 Year 3	 96 568* 22 -147
NEWWS						
Atlanta LFA	Job search is first activity for almost all participants; if no job after job search, assessed for education and training	Broad coverage Mandatory for parents with no child under age 3	Majority long-term recipients 90% African-American Welfare grant: $280 (1993) Unemployment rate: 6.2%		Earnings Years 1-3 Year 3 AFDC Years 1-3 Year 3	490*** 532*** -207*** -180***

(continued)

Table 2.1 (continued)

Program	Program Activities	Coverage and Mandatoriness	Sample and Site Characteristics	Other Features	Outcome	Annual Impacts ($)
Atlanta HCD	Adult Basic Education is first activity for almost all participants	Broad coverage Mandatory for parents with no child under age 3	Majority long-term recipients 90% African-American Welfare grant: $280 (1993) Unemployment rate: 6.2%		Earnings	
					Years 1-3	302**
					Year 3	472**
					AFDC	
					Years 1-3	-171***
					Year 3	-155***
Grand Rapids LFA	Job search is first activity for almost all participants; if no job after job search, assessed for education and training	Broad coverage Mandatory for parents with no child under age 1	Majority long-term recipients 50% white, 40% African-American Welfare grant: $474 (1993) Unemployment rate: 5.5%		Earnings	
					Years 1-3	409**
					Year 3	350
					AFDC	
					Years 1-3	-698***
					Year 3	-556***
Grand Rapids HCD	Adult Basic Education is first activity for almost all participants	Broad coverage Mandatory for parents with no child under age 1	Majority long-term recipients 50% white, 40% African-American Welfare grant: $474 (1993) Unemployment rate: 5.5%		Earnings	
					Years 1-3	422**
					Year 3	444*
					AFDC	
					Years 1-3	-492***
					Year 3	-487***
Riverside LFA	Job search is first activity for almost all participants; if no job after job search, assessed for education and training	Broad coverage Mandatory for parents with no child under age 3	Few new applicants 50% white, 35% Hispanic Welfare grant: $624 (1993) Unemployment rate: 11.7%		Earnings	
					Years 1-3	571***
					Year 3	384**
					AFDC	
					Years 1-3	-662***
					Year 3	-598***
Riverside HCD	Adult Basic Education is first activity for almost all participants	Enrolled only those in need of basic skills Mandatory for parents with no child under age 3	Few new applicants 50% white, 35% Hispanic Welfare grant: $624 (1993) Unemployment rate: 11.7%		Earnings	
					Years 1-3	323**
					Year 3	485**
					AFDC	
					Years 1-3	-623***
					Year 3	-700***

(continued)

Table 2.1 (continued)

Program	Program Activities	Coverage and Mandatoriness	Sample and Site Characteristics	Other Features	Outcome	Annual Impacts ($)
Columbus Integrated	Education and training is first activity for almost all participants	Broad coverage Mandatory for parents with no child under age 3	Majority long-term recipients 50% white, 50% African-American Welfare grant: $341 (1993) Unemployment rate: 4.6%	Used integrated case management; one staff member handled both income maintenance and employment and training case management	Earnings	
					Years 1-3	361**
					Year 3	473**
					AFDC	
					Years 1-3	-362***
					Year 3	-381***
Columbus Traditional	Education and training is first activity for almost all participants	Broad coverage. Mandatory for parents with no child under age 3	Majority long-term recipients 50% white, 50% African-American Welfare grant: $341 (1993) Unemployment rate: 4.6%	Used traditional case management; different workers handled income maintenance and employment and training case management	Earnings	
					Years 1-3	300*
					Year 3	291
					AFDC	
					Years 1-3	-289***
					Year 3	-295***
Detroit	Long-term education and training encouraged for first half of study period; job search emphasized for second half of study period	Mandatory for parents with no child under age 1	Mainly long-term recipients Mostly African-American Welfare grant: $459 (1993) Unemployment rate: 8.0%	Implemented as a de facto voluntary program	Earnings	
					Years 1-3	374***
					Year 3	592***
					AFDC	
					Years 1-3	-182***
					Year 3	-281***
Oklahoma City	Long-term education and training encouraged in most cases instead of job search	Almost all applicants Mandatory for parents with no child under age 1	Almost all applicants 70% white, 30% African-American Welfare grant: $324 (1993) Unemployment rate: 5.6%	Implemented as a de facto voluntary program	Earnings	
					Years 1-3	28
					Year 3	-24
					AFDC	
					Years 1-3	-86**
					Year 3	-60

(continued)

Table 2.1 (continued)

Program	Program Activities	Coverage and Mandatoriness	Sample and Site Characteristics	Other Features	Outcome	Annual Impacts ($)
Portland	Less job-ready in basic education and training at discretion of case managers; job search for others	Broad coverage Mandatory for parents with no child under age 1	Mainly long-term recipients 80% white, 20% African-American Welfare grant: $460 (1993) Unemployment rate: 6.6%		Earnings Years 1-3 Year 3 AFDC Years 1-3 Year 3	 1,096*** 1,364*** -704*** -802***
FTP (Florida)	ABE, GED preparation, and ESL if no high school diploma or lacked basic reading and math skills or had trouble with English; job search for others	Mandatory for parents with no child under 6 months old	Half applicants 50% white, 50% African-American Welfare grant: $303 (1995) Unemployment rate: 5.2%	Time limits Financial incentives	Earnings Years 1-3 Year 3 AFDC Years 1-3 Year 3	 545*** 808*** -151*** -301***
MFIP (Minnesota)	Job search is initial activity for almost all participants	Mandatory only if received welfare for 36 of previous 60 months Mandatory for parents with no child under age 1	40% long-term recipients, 40% applicants 65% white, 35% African-American Welfare grant: $532 (1994) Unemployment rate: 4.2%	Financial incentives	Earnings Years 1-3 Year 3 AFDC Years 1-3 Year 3	 214 426 753*** 573***

I. The Program Models

A. The Saturation Work Initiative Model (SWIM)[3]

Operated between July 1985 and September 1987, the Saturation Work Initiative Model (SWIM) was an employment-focused program that was mandatory for most single-parent AFDC households with no child under age 6. To provide help in finding employment, SWIM started most participants off with a two-week job search workshop. Participants who did not find a job after job search were referred to the Employment Work Experience Program (EWEP), which required them to work 20 to 30 hours per week for 13 weeks in public or nonprofit agencies in exchange for their welfare benefits. Those who were still not working after EWEP were referred to community education and training programs.

B. Greater Avenues for Independence (GAIN)[4]

Implemented in the mid 1980s, Greater Avenues for Independence (GAIN) was California's welfare-to-work program. In six of the state's 58 counties, the effects of GAIN were studied using a random assignment evaluation begun in early 1988.[5]

Participants in the welfare-to-work program were placed in one of two tracks after an initial assessment. Individuals who had neither a high school diploma nor a General Educational Development (GED) certificate, who obtained low scores on either a basic reading or math test, or who were not proficient in English were considered "in need of basic education." Most entered a program of basic education, GED preparation, or English as a Second Language (ESL). Most other participants were required to enroll in a job search activity, primarily job club or supervised job search. If a participant in either track completed her first activity without finding a job, she may have been referred to on-the-job training, work experience, supported work, or other education and training.

Both tracks were somewhat flexible. A participant in need of basic education could elect to enroll in job search before education or training. If she failed to find a job after job search, however, she was required to enroll in the appropriate education program. Likewise, a participant not in need of basic education who was already engaged in an education or training program was usually allowed to complete that program before her job search.

In GAIN, deferrals were given to those who had a part-time job, a temporary illness, a family emergency, or other situations that would make attending an activity difficult. Moreover, a recipient was not required to register for GAIN if she had a full-time job, even if she was still on AFDC. Finally, GAIN was mandatory only if the family's youngest child was age 6 or over, although the evaluation studied a number of parents with younger children who were randomly assigned after volunteering to participate.

---

[3]For a more detailed discussion of the SWIM program, see Hamilton and Friedlander (1989).

[4]For a more complete description of the GAIN program design, see Riccio, Friedlander, and Freedman (1994), particularly Chapters 1 and 2.

[5]Although GAIN began before the Family Support Act was implemented, it met the provisions of the legislation and later became California's JOBS program.

Although the six GAIN counties studied shared a uniform program model, the characteristics of the counties and their implementation of the model differed somewhat.

- **<u>Alameda and Los Angeles</u>.** Both of these large urban counties chose to enroll only long-term recipients in GAIN during the period of random assignment and had much higher initial participation in basic education and training than in job search.

- **<u>Riverside and San Diego</u>.** These counties in southern California had a more diverse group of participants than the other four counties. Riverside stands out from the other five counties because nearly all staff emphasized quick employment, while no more than half in any other county did so. Although Riverside enrolled a substantial number of participants in education and training, its emphasis on quick employment was so strong that it is considered an employment-focused program. San Diego, in contrast, operated a typical GAIN program, ranking in the middle of the six counties in several of its implementation characteristics.

- **<u>Butte and Tulare</u>.** While these are the two smallest counties in the evaluation, they are otherwise quite different from each other. Butte is a rural northern county that had the greatest proportion of welfare applicants in its sample and the smallest proportion of people in need of basic education. Tulare is an agricultural county in which nearly two-thirds of participants were judged to need education and training.

C. <u>The National Evaluation of Welfare-to-Work Strategies (NEWWS)</u>

The National Evaluation of Welfare-to-Work Strategies (NEWWS) is a study of 11 welfare-to-work programs created or adapted to fit the provisions of the Job Opportunities and Basic Skills Training (JOBS) program of the Family Support Act of 1988 (FSA).[6] Under JOBS, all single-parent welfare recipients whose youngest child was age 3 or over (or age 1 or over at a state's discretion) were required to participate in a welfare-to-work program. Each state's program was required to offer adult education, job skills training, job-readiness activities, and job development and placement services. States were also required to provide at least two of the following services: job search, work supplementation, on-the-job training, and community work experience. To help welfare recipients take advantage of these services, states were required to provide subsidies for child care, transportation, and work-related expenses for JOBS participants. Transitional Medicaid and child care benefits were also to be provided to parents who left welfare for work.

The JOBS program was also designed to help states reach the hard-to-serve who sometimes fell through the cracks in earlier programs. To this end, states were required to spend at least 55 percent of JOBS resources on potential long-term recipients or among the more disadvantaged groups, including those who had received welfare in 36 of the prior 60 months, those who were custodial parents under age 24 without a high school diploma or GED, those who had

[6]For more information on the Family Support Act and the JOBS program, see Hamilton and Brock (1994), Chapter 1.

little work experience, and those who were about to lose eligibility for welfare because their youngest child was age 16 or over.

Under NEWWS, 11 welfare-to-work programs are being studied in seven sites: Atlanta, Georgia; Columbus, Ohio; Detroit and Grand Rapids, Michigan; Oklahoma City, Oklahoma; Portland, Oregon; and Riverside, California.[7] Although the Family Support Act had a number of requirements as described above, it allowed states considerable flexibility in program design and implementation. As a result, there is substantial variation in the 11 programs studied under NEWWS. Nevertheless, the programs fall into three broad categories.

- **Employment-focused with job search as the first activity.** Three sites — Atlanta, Grand Rapids, and Riverside — implemented "labor force attachment" (LFA) programs that required most participants to begin with job search activities.

- **Education-focused.** Seven programs emphasized education: "human capital development" (HCD) programs in Atlanta, Grand Rapids, and Riverside;[8] two programs in Columbus that tested different forms of case management;[9] and programs in Detroit and Oklahoma City. In each of these programs, most participants were initially enrolled in education and training programs. The programs in Atlanta, Grand Rapids, and Riverside emphasized basic education, while the programs in Columbus, Detroit, and Oklahoma City emphasized long-term or post-secondary education and training. Detroit and Oklahoma City enforced the participation mandate much less than other programs studied in NEWWS.[10]

- **Employment-focused with a mix of first activities.** The eleventh program evaluated under NEWWS was operated in Portland, Oregon. Like the three LFA programs, it emphasized to clients that the goal of the program was to get a job. Unlike the LFA programs, however, the Portland program encouraged participants to wait until they found a "good" job. In addition, Portland was similar to GAIN in that it encouraged those in need of more skills to enroll in education or training initially and look for a job later.

[7]Descriptions included here are adapted from Freedman et al. (2000), Chapter 3.

[8]In Atlanta and Grand Rapids, people were randomly assigned to the control group, the HCD program group, or the LFA program group. In Riverside, those in need of basic education according to the GAIN criteria described above were randomly assigned to one of these three groups, but those not in need of basic education were randomly assigned to either the control group or the LFA program group.

[9]In Columbus, people were randomly assigned to the control group, the traditional case management group (in which one caseworker verified eligibility for welfare and a second managed program participation), or the integrated case management group (in which one caseworker both verified eligibility and managed program participation).

[10]In addition, Detroit's program was education-focused only during the first half of the study period. It later became an employment-focused program that required nearly all participants to look for work.

D. The Minnesota Family Investment Program (MFIP)[11]

The Minnesota Family Investment Program (MFIP) was begun in 1994 to test whether financial incentives would encourage welfare recipients to work.[12] In addition to providing a financial incentive, however, MFIP required welfare recipients to participate in its welfare-to-work program after they had received welfare in 24 months over a three-year period. Individuals could avoid participating in the employment and training services if they had a child under age 1, if they were working 30 hours or more per week, or if they had other "good cause" reasons for not working. Since only long-term recipients were required to participate in the welfare-to-work services immediately after random assignment, the MFIP sample used in this report is limited to this group.

MFIP's welfare-to-work program was an employment-focused program that assigned more job-ready individuals to jobs search but allowed others to enroll initially in education programs. Nevertheless, many staff encouraged participants to take a job quickly, especially compared with Minnesota's JOBS program, STRIDE, for which the control group could volunteer. Almost half of MFIP case managers said they emphasized getting a job quickly, and about two in three long-term recipients in MFIP said they were encouraged to take a job quickly. The message is reflected in program participation rates: compared with STRIDE, MFIP increased the use of job search activities and enrollment in the career workshop, but reduced participation in post-secondary education.

Although this report is about the impacts of welfare-to-work programs, to understand MFIP's effects it is necessary to also understand its financial incentive.[13] Under MFIP, working welfare recipients essentially had their welfare guarantee increased by 20 percent. In addition, each dollar of earnings reduced the welfare grant by 62 cents under MFIP compared with 67 cents for the first four months of combining work and welfare under AFDC and 100 percent thereafter. As a result, a mother of two who worked 20 hours per week and earned $6 per hour would receive almost $250 more in income under MFIP than under AFDC (Figure 1.1 in Miller et al., 1997).

E. The Family Transition Program (FTP)

Florida's Family Transition Program (FTP) is in most respects quite different from the other programs studied in this report.[14] Like the other programs, it required participants to engage in employment and training services. Like MFIP, it included a financial incentive that made work pay more than it did under AFDC rules. Unlike any of the other 19 programs, however, the con-

[11]For more details on the MFIP program, see Miller et al. (1997), Chapter 1.

[12]MFIP is also the name of Minnesota's TANF program, which is a modified version of the MFIP program described here.

[13]To test the effects of financial incentives, individuals in MFIP were enrolled in a control group that neither received the financial incentive nor was mandated to participate in services, an incentives-only group that received the financial incentive but was not required to participate in services, or a full MFIP group that both received the financial incentive and was required to participate in services. In this report, we use information for only the control group and the full MFIP group.

[14]The description of the FTP program is adapted from Bloom, Kemple, and Rogers-Dillon (1997), Chapter 1; information on implementation of the program is from Chapter 3.

trol group in FTP was also required to participate in services through Project Independence, Florida's JOBS program. Perhaps more important, FTP imposed a time limit on receipt of welfare benefits. About 40 percent of the program group was considered more disadvantaged and allowed to receive welfare for 36 months in a 72-month period before reaching the program's time limit.[15] The remaining 60 percent of the program group was allowed to receive welfare for 24 months in a 60-month period before reaching the time limit.

Although both the control and program groups were required to participate in employment and training services, the mandate was different for the two groups. First, Project Independence was not fully funded during this period, so there remained a sizable difference in participation rates between the groups. Second, control group members with a child under age 3 were exempt from the participation mandate, whereas program group members with a child six months old or older had to participate. Third, mandates were much more strictly enforced for the program group than for the control group, with 30 to 40 percent of the program group being sanctioned at some point for not complying with the mandate. Fourth, participants who were considered not job-ready were allowed to participate in education and skills development, but the definition of job-ready was more inclusive in FTP than in Project Independence. While a participant in the control group was considered job-ready if she had a high school diploma or GED *or* had worked in 12 of the 24 months preceding random assignment, a participant in the program group was considered job-ready only if she had graduated from high school (or received a GED), worked in 12 of the 24 months preceding random assignment, *and* passed a literacy test (Bloom et al., 1998). In one period of random assignment, only about one-seventh of the program group was considered job-ready by this definition. As a result, FTP's primary effect on participation in services was to increase enrollment in education and training. Nevertheless, because FTP case managers said their job was primarily about helping people get off welfare and because FTP's other policies were clearly designed to encourage work, FTP is considered in this report to be an employment-focused program that used a mix of first activities.

II. <u>Characteristics of Sample Members and Sites</u>

In addition to varying in their self-sufficiency approach, the sites and programs varied substantially in the types of clients they served, the generosity of their welfare systems, and the robustness of the local economy. Table 2.1 summarizes some key aspects of the sample — welfare history, race and ethnic composition, and proportion with children under age 6. It also shows two key economic features of each site — the welfare guarantee for a single parent with two children during an early part of the follow-up period and the county unemployment rate for about the same period.

These features of the sample and sites may be important in understanding the subgroup impacts to follow. For example, as described earlier, Los Angeles GAIN enrolled only long-term welfare recipients while Butte GAIN enrolled a large number of welfare applicants. As a result, high school graduates in Los Angeles have at least one impediment to finding work: they have a long history of receiving welfare. Likewise, high school nongraduates have at least two impedi-

[15]A person was given a 36-month time limit if she had received welfare for at least 36 of the 60 months prior to random assignment or if she was a high school dropout under age 24 with little or no recent work history.

ments to finding work: in addition to relatively little education, they also have a long history of receiving welfare. Both high school graduates and nongraduates in Butte, on the other hand, are unlikely to have long welfare histories. If the number of barriers to work is key in determining a program's impacts, then high school graduates in Los Angeles may have impacts similar to high school *nongraduates* in Butte, even if the programs are equally effective for the same people. Differences between Los Angeles and Butte, therefore, may be due to differences in the way the GAIN program was implemented, but they might also be due to differences in welfare history, other demographics, or economic conditions.

In interpreting impact results, the state of the local economy and the generosity of welfare benefits might also be important. When the local economy is faring poorly, it may be difficult for many welfare recipients to find work, and the impacts of a welfare-to-work program might be small. Likewise, a welfare-to-work program might be unnecessary in a booming economy if many in the control group are able to find work on their own. On the other hand, a poor economy will tend to bring more high school graduates and more experienced workers into the welfare system, and these people may respond well to a welfare-to-work program.

Differences in welfare benefit levels may also be related to differences in the programs' impacts. It is relatively easier to find a job that pays better than welfare in low-grant states than in high-grant states. Those who are receiving welfare nonetheless are likely to have substantial problems that keep them from working. Welfare recipients with a given level of education, welfare history, work experience, and so on, might therefore be more disadvantaged in low-grant states than in high-grant states, and impacts might vary accordingly.

The programs and sites varied substantially on five key dimensions.

- **Welfare history.** The programs fall into three broad groups based on the number of long-term recipients, short-term recipients, and new welfare applicants who were randomly assigned. Some programs, either by design or because of the nature of their welfare caseloads, enrolled primarily or exclusively individuals who had received welfare for two years or more over the course of their life (long-term recipients). This group includes GAIN programs in Alameda County and Los Angeles County; JOBS programs in Atlanta, Columbus, Detroit, Grand Rapids, and Portland; and MFIP. At the other extreme, Butte and Oklahoma City enrolled more new applicants than either short-term recipients or long-term recipients. The remaining six programs — SWIM, FTP, and the other four GAIN sites — enrolled a broad mix of long-term recipients, short-term recipients, and welfare applicants.

- **Race and ethnicity.** In nine programs (MFIP, FTP, Alameda GAIN, and JOBS programs in Grand Rapids, Columbus, Oklahoma City, and Portland), there were large numbers of both white and African-American sample members. In four other programs, 85 percent or more of sample members were African-American (Atlanta and Detroit) or 85 percent or more were white (Butte). In the remaining seven programs at least 25 percent of sample members were Hispanic. All seven programs were run in California: the Los Angeles, Riverside, San Diego, and Tulare GAIN programs; the Riverside NEWWS programs, and SWIM.

- **Age of youngest child.** As discussed above, families whose youngest child was age 3 or over were required to participate in welfare-to-work programs by the Family Support Act. States were allowed to require participation for families with a child as young as age 1. MFIP, FTP, and JOBS programs in Detroit, Grand Rapids, Oklahoma City, and Portland had this requirement. In contrast, the pre-JOBS programs in SWIM and GAIN mandated participation only for families with no child under age 6. This key difference in program rules had predictable effects on the sample. In SWIM and five of the GAIN counties, fewer than 15 percent of sample members had a child under age 6. In the other sites, 40 percent or more had a child under age 6.

- **Local unemployment rate.** Most sites had moderate unemployment rates of 4 to 6 percent when their programs began, reflecting the strong state of the national economy over the last 15 years. Notable exceptions were Tulare and the Grand Rapids and Riverside NEWWS programs, all of which had unemployment rates exceeding 10 percent. In addition, Butte and Detroit had moderately high unemployment rates of about 8 percent.

- **Generosity of welfare benefits.** The eight states also fall roughly into three categories according to their welfare guarantees. At one end of the generosity scale is California, home of nine of the programs being studied, with benefits in 1993 of $624 per month for a single parent with two children. At the other end are Georgia, Florida, Oklahoma, and Ohio, all of which provided less than $350 per month for a family of three in 1993. The remaining three states — Michigan, Oregon, and Minnesota — fall in the middle, with welfare guarantees between $450 and $550 per month for about the same time period.

Other demographic characteristics. In many other ways, the samples were quite similar across the programs (results not shown in Table 2.1). Most sample members were women (more than 90 percent in each site). About half of the sample in each site had neither graduated from high school nor earned a GED at the time of random assignment. The average age of sample members was about 30 in all sites, and the average family had about two children. Even in these samples of welfare recipients and welfare applicants, most (more than 90 percent in some sites) had worked at some time in the past.[16]

III. Data Sources

All follow-up information for the studies comes from administrative records. Earnings information was taken from reports made by employers to the state unemployment insurance (UI)

[16]Reflecting the different welfare histories described earlier, however, there was substantial variation in whether sample members had worked in the year prior to random assignment. In sites such as Oklahoma City that enrolled many applicants, most sample members had recent work experience. In contrast, in sites such as Alameda and Los Angeles that enrolled only long-term recipients, few sample members had recent work experience.

system. AFDC and Food Stamp information comes from state or county welfare system administrative records.[17]

Some subgroups in this report are defined based solely on administrative records collected for the period prior to random assignment. Others are defined based on responses to a Background Information Form (BIF) filled out by a caseworker on behalf of each sample member at the time of random assignment. The BIF contains demographic information, such as the educational attainment, prior work experience, and welfare history of the sample member; marital status and number and ages of children; race and ethnicity, and sex.

Still other subgroups are defined based on responses to the Private Opinion Survey (POS), which was administered at the time of random assignment to sample members in MFIP, FTP, and the Atlanta, Grand Rapids, Portland, and Riverside NEWWS sites. The POS was designed to ascertain such information as sample members' risk of depression; mastery or locus of control; preference for work; barriers to work or program participation because of child care, transportation, and health or family problems; and degree of work-related parental concerns.

IV. Overall Impacts of the Programs

Although overall program impacts are not the focus of this report, it might be easier to find subgroup differences in effective programs than in ineffective programs since ineffective programs might have helped no group. Table 2.1 provides one measure of the effectiveness of the programs: their impacts on annual earnings and annual AFDC payments.[18]

Two programs stand out as the most successful: the Riverside GAIN and Portland JOBS programs each resulted in earnings gains exceeding $1,000 per year over the three years of follow-up. Butte likewise had quite large effects, increasing earnings by more than $1,100 in the third year of follow-up and more than $900 per year overall. At the other extreme, Los Angeles GAIN and Oklahoma City had virtually no effect on earnings. Most programs fell in the middle, increasing earnings between $200 and $800 per year.

MFIP stands out from the other programs as the only one to significantly increase public assistance amounts. Indeed, in each of the other 19 programs, impacts on welfare payments were negative, and in most cases there were large negative payments. MFIP's increase in welfare payments stemmed from the program's large financial work incentive that essentially increased welfare benefit amounts by 20 percent for working welfare recipients. These financial incentives resulted in welfare benefit payment amounts that were $750 per year higher on average for the pro-

[17]Because Food Stamp amounts were not collected for the evaluation of SWIM, public assistance amounts for SWIM include only AFDC. In MFIP, Food Stamps and General Assistance were included in the AFDC welfare check for members of the program group. As a result, public assistance amounts in MFIP for both the control and program groups represent the sum of AFDC, General Assistance, and the cash value of Food Stamps.

[18]For a number of reasons, the impacts shown in Table 2.1 differ from impacts shown in reports about the individual evaluations. Results in Table 2.1 were inflation-adjusted to 1997 dollars, while impacts in the individual reports were not. Results in the table were not regression-adjusted for baseline demographic characteristics, whereas results in individual reports were adjusted. In addition, not all long-term recipients in MFIP were included in the analysis because information was unavailable for some when it was conducted.

gram group than for the control group. In addition, MFIP increased combined income from earnings, AFDC, and Food Stamps by substantially more than any other program.

V. Summary of Program Features

Table 2.2 summarizes the self-sufficiency approaches used by the various programs.

National Evaluation of Welfare-to-Work Strategies

Table 2.2

Summary of Self-Sufficiency Approaches of 20 Welfare-to-Work Programs

Education-Focused	Mix of First Activity Without Employment Focus	Employment-Focused With Mix of First Activity	Employment-Focused With Job Search as First Activity
Atlanta HCD	Alameda GAIN	Riverside GAIN	Atlanta LFA
Grand Rapids HCD	Butte GAIN	Portland	Grand Rapids LFA
Riverside HCD	Los Angeles GAIN	Florida FTP	Riverside LFA
Columbus Integrated	San Diego GAIN	Minnesota MFIP	San Diego SWIM
Columbus Traditional	Tulare GAIN		
Detroit			
Oklahoma City			

Although all the programs included some education and some job search, they fall roughly into four categories. Seven of the JOBS programs focused on education, encouraging participants to enroll in basic education, typically in preparation for job search. Four programs — SWIM and the three LFA programs — were employment-focused and encouraged most participants to look for work regardless of their work experience, education level, or welfare history. Four programs — FTP, MFIP, Portland, and Riverside GAIN — were also employment-focused but used a mix of first activities that encouraged job-ready participants to look for work but allowed others to build skills through education and training with the ultimate objective of finding employment. The remaining five GAIN sites also used a mix of first activities, but were not as employment-focused as Riverside.

Programs also varied in the breadth of their enrollment, summarized in Table 2.3.

About half the programs enrolled a broad group of welfare recipients and welfare applicants. Most of the NEWWS programs focused on short- and long-term recipients, while the GAIN programs in Los Angeles and Alameda enrolled only long-term recipients. In contrast, the program in Oklahoma City enrolled a significant number of new applicants with no prior welfare receipt.

National Evaluation of Welfare-to-Work Strategies

Table 2.3

Program Enrollment

Evaluation or Program	Applicants	Short-Term Recipients	Long-Term Recipients	Notes
SWIM (San Diego)				
GAIN				
Alameda				
Butte				
Los Angeles				
Riverside				
San Diego				
Tulare				
NEWWS				
Atlanta LFA				
Atlanta HCD				
Grand Rapids LFA				
Grand Rapids HCD				
Riverside LFA				
Riverside HCD				Recipients in need of basic education
Columbus Integrated				
Columbus Traditional				
Detroit				
Oklahoma City				Primarily applicants
Portland				
FTP (Florida)				
MFIP (Minnesota)				

Chapter 3

An Example of Subgroup Comparisons: Risk of Depression

This report presents results for a number of outcomes, including earnings, welfare payment amounts, and combined income from earnings, AFDC, and Food Stamps. It presents these results for a number of subgroups — for example, by educational attainment, welfare history, work history, preference for work, and risk of depression. This chapter introduces some of the ideas and methods used throughout the report by presenting results for subgroups defined by an individual's risk of depression.

Sample members in the Portland, Atlanta, Riverside, and Grand Rapids NEWWS Evaluation were asked at the time of random assignment to fill out a Private Opinion Survey (POS) that included four items from the 20-item Center for Epidemiological Studies Depression (CES-D) Scale. In particular, each respondent was asked how often in the prior week she felt sad, how often she felt depressed, how often she felt lonely, and how often she had trouble shaking the blues. Answers to these questions were used to assess an individual's risk of depression because the CES-D Scale has been found to be correlated with clinical depression. That is, individuals who say they suffer from many of the symptoms or suffer from some symptoms frequently are more likely than others to be judged by a psychiatrist to be depressed.

Sample members were divided into three groups: those at high risk of depression, those at moderate risk, and those at low risk. (In the interest of brevity, those groups will often be referred to as the most depressed, the moderately depressed, and the least depressed.) Comparisons of outcomes and impacts for the three groups are summarized by three important findings.

- **The programs generally increased earnings for sample members at low risk of depression, but not for those at high risk. However, they reduced welfare payments for all three groups.**

The programs increased earnings by about $800 per year, on average, for the least depressed group and about $500 per year for the moderately depressed, but they did not significantly increase earnings for the most depressed. Nevertheless, they reduced welfare payments by about $600 per year for the least depressed and $400 per year for the most depressed. Only the Portland program deviated from this pattern. It increased earnings by more than $700 per year for the most depressed sample members and reduced their welfare payments by less than $300.

- **The programs did not significantly change combined income from earnings, AFDC, and Food Stamps for any of the groups.**

Average income neither significantly increased nor significantly decreased for any of the three subgroups. The fact that earnings gains did not translate into income gains reflects the tradeoff that welfare recipients face between earnings and cash assistance. Under the AFDC system in most states, an additional dollar of earnings usually reduced welfare payments by a dollar, leaving the family's income unchanged. The fact that average income did not decrease despite the use of sanctions, however, implies that the welfare-to-work programs did not financially hurt even the high-risk group, which is expected to have the most difficulty meeting the programs' mandates.

- **Earnings and income levels did not vary much with risk of depression levels.**

Despite the larger impacts on earnings for the least depressed sample members, earnings and income were similar for the three groups. Over a three-year follow-up period, the most depressed control group members earned $3,360 per year compared with $3,818 for the least depressed members. During the same period, income was also similar for the two groups: $8,045 for the most depressed and $8,507 for the least depressed.

The remainder of this chapter presents details on these findings.

I. Depression and Employment

We often assume that depressed people are less likely to work and more likely to remain on welfare than other people, and that welfare-to-work programs will consequently have smaller effects for them, but surprisingly little is known about the relationship between depression, welfare receipt, and employment, and particularly about whether depressed people respond well to interventions like the ones studied in this report.

A. Studies Linking Depression and Employment

The most relevant study linking employment outcomes to depression may be one by Danziger et al. (1998), who surveyed a sample of current and former welfare recipients to determine whether individuals suffering from an array of 14 problems were less likely than other people to be working. Among women in their sample who had had a bout of major depression in the year prior to being surveyed, 46.5 percent were working 20 hours or more per week; among others in their sample, 60.7 percent were working, a statistically significant difference.

Several studies of a group broader than welfare recipients have found that those who are more depressed are less likely to be employed, more likely to be poor, and likely to have lower income than those who are less depressed. In a sample of more than 2,500 Mexican-American and non-Hispanic white residents of Los Angeles without a chronic medical condition, Wells, Golding, and Burnam (1988) found that 72.8 percent of those without a psychiatric disorder were working compared with 64.0 percent of those with a psychiatric disorder. Studying an earlier group from Los Angeles, Aneshensel, Frerichs, and Huba (1984) found that the more depressed sample members had less income on average than the less depressed sample members. Studying a small group of residents in New Haven, Connecticut, Bruce, Takeuchi, and Leaf (1991) found that 7.9 percent of individuals in their sample who were poor had recently suffered a bout of major depression compared with 4.0 percent of those who were not poor. Interestingly, both studies assume that poverty or the lack of income caused people to become depressed, not the other way around.

Other studies have found less evidence linking depression to employment outcomes, but none of these other studies has looked explicitly at welfare recipients. Among recent high school graduates who were in the labor force, Winefield and Tiggemann (1990) found that those who were working were generally no less likely to be depressed as those who were unemployed. In conducting a longitudinal study of individuals aged 50 or over, Rohde, Lewinsohn, and Seeley (1990) found virtually no difference in employment rates among people who showed signs of depression in the first interview period and those who did not. More generally, they cited 14 papers that compared outcomes for people who had never been depressed with outcomes for people who were once depressed. Of these 14 papers, only 3 found much worse outcomes for people who had been depressed.

A study by Hock and DeMeis (1990) includes an intriguing result about mothers of infants. They found that working mothers were no less depressed on average than mothers who were not working. They did find, however, higher rates of depression for working women who preferred to be at home and for stay-at-home mothers who preferred to be working outside the home than for those who were doing what they wanted to do.

B. <u>Depression and Impacts of Welfare-to-Work Programs</u>

There is even less evidence on the effects of mandatory welfare-to-work programs for the more depressed and the less depressed welfare recipients. The most direct evidence comes from a report by Bos et al. (2001) on adult education in programs being studied as part of the NEWWS Evaluation. The report looked at impacts on earnings and AFDC payments for welfare recipients who had neither a high school diploma nor a GED at the time of random assignment, in three of the four NEWWS sites that will be discussed in this chapter.[1] The report found that the programs increased the earnings for individuals with the most symptoms of depression by less than they did for individuals with fewer symptoms.

This chapter looks at a broader section of the caseload to address the same question: did these welfare-to-work programs have larger effects on people at high risk of depression than on people at low risk of depression? Two hypotheses yield conflicting expectations. On the one hand, depressed people might have difficulty taking advantage of a welfare-to-work intervention. If this is true, the programs will not have increased the earnings and employment levels of the most depressed. Hocks and DeMeis's results might be consistent with this hypothesis. The most depressed sample members may be those who want to work but are not working. Even with the help of a welfare-to-work program, their depression may inhibit their ability to take advantage of the program. They will remain depressed and unable to fulfill their desire to go to work.

An alternative hypothesis is that the least depressed sample members are more likely to work and leave welfare on their own, that is, without the help of a welfare-to-work program. This would raise a high bar for the programs to hurdle, and the result may be little impact for the least depressed. If the literature cited above is correct, however, this hypothesis will be proved false. In the control group, which represents what the program group would have done without the welfare-to-work program, the most depressed sample members will be just as likely to work as the least depressed sample members. The bar would be consequently at the same height for both groups, and the least depressed group should have an easier time clearing it.

[1]The analysis also included about a quarter of the sample in Riverside's education-focused program who did have a high school diploma or GED at random assignment but who scored low on a math or reading appraisal test or had limited English, and thus were determined to be in need of basic education.

II. Results Across Seven Programs

A. Pooled Results for Years 1-3 After Random Assignment

Table 3.1 provides a useful overview of outcomes and impacts by risk of depression. Results shown are averaged across all seven programs that included CES-D Scale items.[2] Throughout the report such results will be referred to as "pooled."

1. Defining the Subgroups. Table 3.1 shows results for three mutually exclusive subgroups: individuals at low risk of depression, those at moderate risk, and those at high risk. As discussed above, the CES-D Scale is correlated with clinical depression, but is not the same as clinical depression. Nevertheless, the groups will often be referred to as least depressed rather than at low risk of depression, moderately depressed rather than at moderate risk, and most depressed rather than at high risk.

A person was considered at high risk of depression if her average response to the four questions described earlier indicated that she suffered from these four symptoms on most days. She was considered at low risk if she indicated on average that she suffered from these symptoms less than half the time, and she was considered at moderate risk if she was neither at high risk nor at low risk. Appendix A provides more details on how subgroups were defined.

As is true for all subgroups in this report, risk of depression was determined at the time of random assignment (when the Private Opinion Survey was administered). Random assignment ensures that control group members in a subgroup are about the same on average as program group members in that subgroup, except that one group was required to participate in a mandatory welfare-to-work program and the other was not. For example, the most depressed control group members will have about the same average work experience, welfare history, education, and so on, as the most depressed program group members at the time of random assignment. Because the groups were similar at random assignment, any differences that emerge later can be attributed to the effects of the welfare-to-work programs.

2. Outcomes for Program Group Members. Table 3.1 shows average outcomes for the program group during the three years after random assignment for five outcomes: annual earnings, employment, annual AFDC payment amounts, AFDC receipt, and combined annual income from earnings, cash assistance, and Food Stamps. In the program group, the least depressed earned more on average than the most depressed ($3,818 per year compared with $3,360 per year). Contributing to this difference in earnings is a difference in employment rates. While 76.5 percent of the least depressed program group members worked in this three-year period, only 73.7 percent of the most depressed worked.

Four thousand dollars is not a lot to earn in an entire year. One reason that average annual earnings were so low is that average dollar amounts include zeroes for sample members who had no earnings. In other words, the $3,360 average earnings for the most depressed program group

[2]The pooled results are a weighted average of the results for the seven programs, using weights that are proportional to the size of the program and control groups in each site. These results also correct for a change over time in the proportion of sample members assigned to the program group in Portland.

National Evaluation of Welfare-to-Work Strategies

Table 3.1

Selected Impacts Pooled Across Seven Welfare-to-Work Programs, by Risk of Depression at Random Assignment

	Years 1-3				Year 3			
Outcome and Subgroup	Program Group	Control Group	Program Impact		Program Group	Control Group	Program Impact	
Average annual earnings ($)				†				†
High risk of depression	3,360	3,071	289		4,225	4,039	186	
Moderate risk of depression	3,599	3,138	460	***	4,620	3,984	636	***
Low risk of depression	3,818	3,049	769	***	4,816	3,972	844	***
Ever employed (%)								
High risk of depression	73.7	67.4	6.3	***	55.8	53.2	2.5	
Moderate risk of depression	74.1	68.2	5.9	***	59.1	52.4	6.7	***
Low risk of depression	76.5	68.8	7.7	***	60.3	53.8	6.6	***
Average annual AFDC payments ($)				†				
High risk of depression	2,916	3,308	-392	***	2,009	2,352	-342	***
Moderate risk of depression	2,940	3,381	-442	***	2,115	2,582	-467	***
Low risk of depression	2,900	3,496	-596	***	2,056	2,634	-578	***
Ever received AFDC (%)								
High risk of depression	95.3	96.1	-0.9		50.7	58.3	-7.6	***
Moderate risk of depression	94.6	95.6	-1.1		53.6	58.9	-5.3	***
Low risk of depression	95.6	96.4	-0.8	**	52.8	60.5	-7.7	***
Average annual income from welfare and earnings ($)								
High risk of depression	8,045	8,281	-236		7,641	7,897	-256	
Moderate risk of depression	8,388	8,510	-121		8,256	8,236	20	
Low risk of depression	8,507	8,529	-22		8,322	8,276	46	

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, Food Stamp records, and Private Opinion Survey (POS) data.

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

Risk of depression subgroups include sample members from four NEWWS sites only: Atlanta, Grand Rapids, Riverside, and Portland.

members is an average of the positive earnings of the 73.7 percent of the sample who ever worked *and* the zero earnings of the 26.3 percent who never worked. Among the most depressed program group members who worked at some point during the year, average earnings were therefore $4,559 per year (or $3,360 average earnings divided by the 73.7 percent of those who worked). Among the least depressed program group members who worked, average earnings were $4,991 per year (or $3,818 divided by 76.5). Even after making this correction, however, annual earnings were very low. This reflects reality for many of the low-skilled people who entered these programs. With few marketable skills, most are likely to find a job that pays very little and many work only sporadically because they can find only temporary work.

Individuals who earn more typically receive smaller AFDC payments. Since the most depressed program group members earned less on average than the least depressed members, this may seem to imply that the most depressed should have received larger AFDC payments on average than the least depressed. Table 3.1 indicates that this was not so. People in the program group who were the most depressed received about the same AFDC amount as both the moderately depressed and the least depressed, and the three groups were about equally likely to have received welfare.

There are several potential explanations for these seemingly contradictory outcomes. First, different subgroups contain different people, and some groups may be eligible for smaller AFDC payments to start with. The more depressed sample members may have fewer children than the less depressed sample members, for example, or they may be more likely to live in a low-grant state (Georgia) than a high-grant state (California). Second, the relationship between earnings and cash assistance depends on how much an individual earns and how many months she has combined work and welfare. Finally, it is possible that the more depressed sample members were sanctioned more frequently than the less depressed sample members or that they were less likely to reapply for benefits when they lost their job, decreasing their welfare payments even when they did not work.

Although the objective of welfare-to-work programs is to reduce welfare dependence and increase employment and earnings, the bottom line for most welfare recipients is probably their total income. Table 3.1 indicates that average combined income from earnings, welfare, and Food Stamps was between $8,000 and about $8,500 for each of the three subgroups and was somewhat higher for the least depressed program group members than for the most depressed. This is not surprising, since the least depressed had the highest earnings, but welfare payment amounts were similar for all three subgroups.

The fact that income from earnings and AFDC was between $6,000 and $7,000 for each group and income from AFDC, earnings, and Food Stamps was between $8,000 and about $8,500 implies that the average family received about $1,500 per year in Food Stamps (not shown in Table 3.1). For the most depressed program group members, for example, earnings plus AFDC payments averaged $6,276 (or $3,360 from earnings plus $2,916 from AFDC payments). This is $1,769 less than the average income of $8,045 for this group. The difference represents the cash value of Food Stamps received by these families.

Just as $4,000 is not a lot to earn, $8,000 is not a lot for a family to live on for an entire year. Although many of these families were poor, this estimate of income somewhat overstates the problem because a number of sources of income were excluded. First, earnings and public assistance amounts reflect only what was shown in the administrative records for the states and counties in which the programs were being run. If an individual worked in an informal job, a federal government job, or another job not covered by the unemployment insurance system, her earnings are understated in Table 3.1. If she moved out of the state in which the program was being operated, her earnings and public assistance amounts might likewise be understated.

The income measure used in this report also understates total income because many types of income that are not available through administrative records were ignored. Mothers who leave welfare for work have an added incentive to make sure child support payments are made by noncustodial fathers of their children. Those who marry may benefit from the income of their spouse. Comparisons of other sources of income have typically found few differences between program and control groups, however. As a result, income levels in this report will be lower than incomes that families actually receive, but estimates of the impacts of the welfare-to-work programs on their income should be fairly accurate.

Perhaps a more significant exclusion from income is the Earned Income Credit (EIC), which currently provides nearly $4,000 per year for many working families. For a family with an income of only $10,000, this obviously represents a substantial portion of total income. At the same time, working families incur some expenses when they go to work. Most pay payroll taxes and some pay income taxes, while many incur costs from child care, transportation, clothing, and other goods related to working. While excluding the EIC understates families' disposable income, excluding work-related expenses overstates their income, and it is unclear which effect is greater.

3. Outcomes for Control Group Members. Table 3.1 also shows average outcomes for the control group. Although the average outcomes were similar for the two research groups, the patterns were somewhat different. First, the most depressed program group members were less likely to work and had lower average earnings than the least depressed; in contrast, control group members in the three subgroups were about equally likely to work and had about the same average earnings. In addition, employment rates for all subgroups were lower for the control group than for the program group. Second, in the program group the three subgroups had similar welfare payments and rates of welfare use; in the control group the least depressed sample members had slightly larger AFDC payment amounts than the most depressed members. Moreover, AFDC payment amounts were larger in the control group than in the program group for each of the three subgroups. Finally, average combined income was somewhat higher for the least depressed than for the most depressed in both the program and control groups, and income levels in the program group were similar to the levels in the control group.

4. Impacts of the Welfare-to-Work Programs. To determine the *impacts* of a program for a subgroup, average outcomes for program group members in the subgroup are compared with average outcomes for control group members. The results of this comparison for the three-year period after random assignment are shown in Table 3.1. For each of the three subgroups, earnings for the program group exceeded earnings for the control group. For the most depressed group, the difference was $289 per year. Thus, the programs' impact on earnings for this group was $289 per year. Similar comparisons reveal that the impact on earnings for the moderately depressed was higher, at $460 per year, and that the impact for the least depressed was higher still, at $769 per year. This pattern makes sense in light of the finding that the more depressed program group members earned less than the less depressed members, but that depression was unrelated to earnings for the control group.

Although random assignment and the way the subgroups were defined ensured that there were no systematic differences between the research groups prior to random assignment, chance would nevertheless have caused some small differences between the two research groups. To assess whether differences can confidently be attributed to the welfare-to-work programs or were likely due to chance, the concept of *statistical significance* is used. In this report, an impact is said to be statistically significant at the 10 percent level if there is less than a 10 percent chance that the estimated impact could have been generated by a program that had no effect; at the 5 percent level if there is less than a 5 percent chance; at the 1 percent level if there is less than a 1 percent chance.

Statistical significance does not directly indicate the magnitude or importance of an impact estimate, nor does it indicate that the program definitely had an effect; it indicates only whether differences in policies are likely to have affected outcomes. In some cases, results will be statistically insignificant because the program had little or no effect. In other cases, however, results will be statistically insignificant because a subgroup is too small to provide reliable estimates of the program's impacts. To reduce the number of subgroups for which the estimated impacts are unreliable, results throughout the report are presented only for subgroups that contain at least 200 sample members in the program and control groups combined.

In Table 3.1, stars indicate statistically significant impacts. The programs' impact on earnings for the three-year follow-up period for the least depressed sample members, for example, was $769 per year. The three stars next to the estimated impact indicate that this estimate is significantly different from zero at the 1 percent significance level. In comparison, the $289 impact on earnings for the most depressed sample members is not starred, indicating that the impact is not significantly different from zero. In other words, this $289 impact is consistent with the programs' having had no effect on earnings for this subgroup.

In comparing impacts across the three depression subgroups, there is a clear pattern: these seven welfare-to-work programs raised average earnings more for the least depressed than for the moderately depressed ($769 compared with $460) and more for the moderately depressed than for the most depressed ($460 compared with $289). Because differences across subgroups will exist even when the program has the same effect on all groups, statistical significance is also used to indicate whether differences across subgroups are likely due to the effects of the programs. To indicate statistical significance of impacts across subgroups, daggers are shown above the columns of stars. In this case, the dagger above the column of earnings impacts indicates that the welfare-to-work programs increased earnings significantly less for the most depressed than for the moderately depressed and the least depressed and that this difference across subgroups was significant at the 10 percent significance level.

Greater earnings gains for the least depressed also translated into greater reductions in AFDC payment amounts for this group. As shown in the table, the welfare-to-work programs reduced annual AFDC benefit amounts by $596 for the least depressed subgroup, but only $442 for the moderately depressed and $392 for the most depressed. The dagger accompanying these results indicates that these differences are also statistically significant.

As a result of these two opposing factors — larger increases in earnings due to the program accompanied by larger decreases in welfare payment amounts for the least depressed — income changed only slightly for each of the three subgroups. In addition, these impacts on income were small enough that they were likely due to chance for each subgroup. Furthermore, although the estimated impact is more negative for the most depressed than for the least depressed, differences in impacts on income across the three groups are small enough to be likely due to chance, not to subgroup responses to the program.

B. <u>Pooled Results for Year 3 Only</u>

Although impacts through the first three years of follow-up give a sense of the overall impact of the programs, for some purposes it is more instructive to look at impacts in specific years. For example, programs that stress human capital development by enrolling welfare recipients in education or training are not expected to increase earnings quickly. Such programs may even reduce the average program group member's earnings in an attempt to assist her in finding a higher-paying job with a larger long-term payoff. In addition, the CES-D Scale asks only about feelings in the past week; it might not be as accurate in predicting depression three years after random assignment as it is at predicting depression soon after random assignment.

Therefore, Table 3.1 shows pooled average outcomes and program impacts for the third year after random assignment. A comparison of results for years 1-3 with results for year 3 reveals several interesting patterns. Individuals in both the program and control groups were less dependent on welfare and more reliant on their own earnings in year 3 than over the entire three-year period. Average earnings for all subgroups was about $1,000 higher in year 3 than in the three-year period, and welfare payments were about $1,000 lower for all subgroups. Although nearly all sample members received welfare at some time during the three-year period, only about half received welfare in year 3. However, income was about the same in year 3 as in the full three-year period. In fact, for all subgroups, income was slightly lower in year 3 than in the three-year period. This may reflect the deficiencies of administrative records described above; that is, individuals who left the state or county where the program was being administered appear to have had no earnings or public assistance. It is also possible that individuals stopped receiving welfare because they married or their children became adults, but either did not replace the income lost from welfare or relied on income from a spouse or partner.

Although outcomes were different in year 3 than in the three-year period, the pattern of pooled impacts was similar for the two periods. Individuals at low risk of depression had the largest increases in earnings ($844 in year 3 and $769 per year over the three-year period); those at high risk of depression had the smallest increases in earnings ($186 in year 3 and $289 per year over the three-year period). Likewise, the least depressed had significantly larger earnings gains than the most depressed.

Reflecting their larger gains in earnings, the least depressed also had the largest decreases in welfare benefit amounts in year 3, and the most depressed had the smallest decreases. However, differences in welfare payments across the subgroups were not statistically significant in year 3. Likewise, differences in impacts on welfare receipt were not significantly different. While the programs reduced the proportion of the most depressed group receiving welfare by 7.6 percentage points, they reduced the proportion of the least depressed group by 7.7 percentage points.

In year 3, as in the entire follow-up period, combined income from welfare, earnings, and Food Stamps was changed little by these welfare-to-work programs, and differences in program impacts on income were not statistically significant across the three depression subgroups. In other words, even though the programs had much larger effects on the earnings of the least depressed sample members, they did not increase income for this group significantly more than they increased income for the most depressed.

III. Impacts by Individual Program

Table 3.2 compares the subgroups in more detail, showing impacts for the seven programs for which questions from the CES-D Scale were part of the Private Opinion Survey (POS). To conserve space, the table presents only the programs' impacts and indications of their statistical significance, but not average program or control group outcomes. As in Table 3.1, impacts were calculated as the difference in average outcomes between the program and control subgroups. Because these programs typically changed income very little, the table presents impacts only for earnings and AFDC payment amounts.

The pooled impacts show clearly that impacts on earnings were largest for the least depressed sample members and smallest for the most depressed. Results for the individual programs are not so clear-cut. Only in the Grand Rapids HCD program in year 3 were differences in earnings impacts across the three groups statistically significant. Only in Portland were impacts on AFDC payment amounts significantly different across the three groups.

On closer look, however, the individual program impacts on earnings broadly agree with the pooled impacts. For six of the seven programs, the most depressed sample members had the smallest earnings impacts. For all seven programs, earnings impacts were statistically significant and positive for the least depressed sample members. In comparison, earnings impacts were significantly different from zero for the moderately depressed in only two programs (Portland and Riverside LFA), and earnings impacts were significantly different from zero for the most depressed in only one program (Portland).

Because all programs had few sample members in the high-risk and moderate-risk groups, however, differences across the three subgroups were typically not statistically significant. This indicates the key advantage of pooled impacts: by combining results across a number of programs, they provide more powerful tests of which groups were affected and whether programs affected some groups more than others.

Although the pooled impacts on earnings were broadly consistent with the individual impacts, pooled impacts on welfare payments were strongly affected by the Portland program. In the pooled results of Table 3.1, the least depressed had significantly larger reductions in their AFDC payments than the most depressed. Only in Portland were the differences in impacts on AFDC payments large across

National Evaluation of Welfare-to-Work Strategies

Table 3.2

Impacts on Earnings and AFDC Payments Across Seven Welfare-to-Work Programs, by Risk of Depression at Random Assignment

	Sample	Impacts on Average Total Earnings per Year ($)		Impacts on Average Total AFDC Payments per Year ($)	
Program and Subgroup	Size	Years 1-3	Year 3	Years 1-3	Year 3
Atlanta LFA					
High risk	383	710	689	-166	-138
Moderate risk	762	217	539	-160	-143
Low risk	1,999	588 ***	569 **	-177 ***	-131 *
Atlanta HCD					
High risk	400	-19	-443	-155	-150
Moderate risk	826	394	948 **	-165 *	-155
Low risk	1,970	360 *	536 *	-139 **	-94
Grand Rapids LFA					
High risk	319	-15	-513	-800 ***	-626 **
Moderate risk	488	443	509	-842 ***	-684 ***
Low risk	1,148	512 **	475	-761 ***	-606 ***
Grand Rapids HCD			†		
High risk	304	-425	-940	-415 *	-359
Moderate risk	474	152	260	-503 ***	-493 **
Low risk	1,164	851 ***	1,039 **	-545 ***	-530 ***
Riverside LFA					
High risk	519	430	374	-949 ***	-806 **
Moderate risk	858	790 **	646	-804 ***	-752 ***
Low risk	2,425	716 ***	400	-705 ***	-647 ***
Riverside HCD					
High risk	270	-218	383	-394	-474
Moderate risk	444	529	577	-482	-699 *
Low risk	1,010	517 **	754 **	-654 ***	-714 ***
Portland				†††	†††
High risk	775	734 **	831 *	-279 *	-243
Moderate risk	1,174	771 **	1,004 **	-437 ***	-587 ***
Low risk	2,946	1,386 ***	1,754 ***	-990 ***	-1,075 ***

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, Food Stamp records, and Private Opinion Survey (POS) data.

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

Risk of depression subgroups include sample members from four NEWWS sites only: Atlanta, Grand Rapids, Riverside, and Portland.

the three groups. While the Portland program reduced welfare payments by about $1,000 per year for the least depressed, it reduced them by less than $300 per ear for the most depressed. In contrast, the LFA and HCD programs had similar impacts on welfare payments for the three subgroups. Not only were the differences not statistically significant, but they were typically small.

In addition to showing whether differences by subgroup are found across a wide range of programs, the results by individual program can show whether a wide range of programs benefit particular subgroups. According to Table 3.2, for example, all seven welfare-to-work programs increased earnings significantly for the least depressed sample members during the three-year follow-up period. These impacts range from a low of $360 per year in the Atlanta HCD program to a high of $1,386 per year in Portland. Thus, a wide range of welfare-to-work programs seem to benefit those at low risk of depression.

In comparison, only Portland increased earnings significantly for individuals at high risk of depression. Since Portland's welfare-to-work program is the only one of the seven to combine an employment focus with a broad mix of first activities, this could be an indication that such programs are the most effective at increasing earnings for both the more disadvantaged and the less disadvantaged groups. Portland is only one program, however, and such a broad conclusion requires more data to support it.

One possible explanation for the lack of significant impacts for the high-risk and moderate-risk groups is that few sample members are in those groups. In the six LFA and HCD programs, the number of high-risk sample members ranged from 270 (Riverside HCD) to 519 (Riverside LFA). The low-risk group, for whom earnings gains were typically statistically significant, had four to five times as many sample members. This again indicates the advantage of pooling results across programs: subgroups that are too small to yield valid inferences for one program may yield important results when combined.

If impacts for the high-risk group were insignificant because of small samples, however, most programs would have produced positive, but statistically insignificant, impacts for the high-risk group. They did not. In four of the LFA and HCD programs, the most depressed members of the program group earned (insignificantly) less than the most depressed members of the control group over the three-year period.

IV. <u>Implications</u>

One objective of subgroup comparisons is to help policymakers decide where to target resources and to help understand the relative advantages of different approaches to encouraging welfare recipients to work. The comparison of impacts by level of depression yielded several useful implications that are typical of the types of conclusions that will be drawn in the remainder of the report.

- **Welfare-to-work programs may need additional services if they are to increase self-sufficiency among welfare recipients at high risk of depression**.

Overall, these seven welfare-to-work programs did not significantly increase earnings for the group at high risk of depression. Only the most successful of the seven programs increased earnings significantly for the high-risk group, but even it increased earnings much more for others. If welfare-to-work programs are to be effective for the most depressed sample members, policymakers might have to augment their programs with interventions designed explicitly for depressed sample members.

- **Welfare-to-work programs are effective for the least depressed sample members in a wide variety of circumstances.**

All of the programs increased earnings significantly for the least depressed. Thus, the least depressed appear to benefit from education and from job search. They appear to benefit from programs in high-grant states (Riverside, California), moderate-grant states (Grand Rapids, Michigan, and Portland, Oregon), and low-grant states (Atlanta, Georgia). They appear to benefit from sites with weak economies (Riverside and Grand Rapids) and stronger economies (Atlanta and Portland). In short, policymakers may find it relatively easy work to increase self-sufficiency among the least depressed.

- **Only Portland's welfare-to-work program was effective regardless of risk of depression.**

The Portland program increased earnings significantly for the most depressed, least depressed, and moderately depressed sample members both in year 3 and over the three-year follow-up period. Portland was the only program to achieve this feat. In fact, Portland was the only program to increase earnings significantly for the most depressed and one of only two programs to increase earnings significantly for the moderately depressed. Although it was employment-focused, Portland was the only one of the seven programs that required job-ready individuals to look for work while allowing individuals in need of basic education to build skills before looking for work. Thus, Portland's results may provide a first indication that programs with a mix of first activities provide the most consistent impacts across a broad range of groups.

- **The most depressed are not financially far behind the least depressed.**

Welfare-to-work programs appear to increase earnings more for the least depressed sample members than for the most depressed members, but total income is similar across the groups. Without the welfare-to-work programs, in fact, there is little difference economically between the two groups. If policymakers are more concerned about income levels than about welfare receipt or self-sufficiency, resources might be better targeted to groups whose income is substantially below average.

Chapter 4

Control Group Outcomes by Subgroup

For policymakers, how much better off people in the programs were than those not in the programs is of primary importance. However, how well off people are without the program is equally important. Suppose, for example, that earnings were not affected by welfare-to-work programs for some subgroups. Should new services be implemented for these groups and additional resources expended to help them? It depends. If they had relatively high earnings or were relatively likely to have left welfare on their own, then new resources and services may not be warranted. If anything, resources might be directed away from the groups since they did well on their own and were not helped by the programs. Suppose that earnings were affected for other subgroups. Should administrators of welfare systems stop investigating new ways of helping these groups? Again, it depends. If the groups had relatively low earnings or were relatively likely to remain on welfare despite the welfare-to-work programs, then additional services or resources might be called for. In a world of time-limited welfare, in particular, additional services and resources might be directed their way if most individuals fared poorly with the current array of services, even though some benefited from those services. Finally, suppose that all subgroups within a grouping benefited equally from the programs. For example, individuals with child care problems benefited as much as those with no child care problems. This could imply that the programs addressed the concerns, but it could also imply that the measure of child care problems did not effectively discern real problems.

To explore what would have happened without the welfare-to-work programs, this chapter presents average earnings, AFDC payments, and combined income from earnings, AFDC, and Food Stamps for the control group. Outcomes are presented for seven types of subgroups defined using demographic information and administrative records on earnings and welfare: recent work history, welfare history, high school credential, number of children, age of youngest child, race and ethnicity, and sex. Outcomes are also presented for six types of subgroups defined using responses to the POS: preference for work, work-related parental concerns, mastery, health or emotional problems, child care problems, and transportation problems.

For each subgroup, the more disadvantaged group is expected to have lower earnings and higher welfare benefit amounts than the less disadvantaged group. For example, race, education, marital status, work experience, and disability status have all previously been found to be related to the likelihood of staying on welfare.[1] Thus, individuals with more earnings prior to random assignment should continue to earn more after random assignment; those with substantial welfare histories prior to random assignment should receive more welfare than others; high school graduates should earn more than high school nongraduates; parents with more children or younger children should earn less and receive more in welfare than other parents; individuals with high work-related parental concerns, low preference for work, low sense of mastery, and barriers to work and program participation should all earn less and receive more welfare than other parents.

[1]Bane and Ellwood, 1994.

In general, this is what we found. Three key findings summarize the relationship between characteristics and control group outcomes.

- **Recent work history, welfare history, and educational attainment were the best predictors of earnings and AFDC payment amounts.**

Individuals with the most recent work experience prior to random assignment earned several times more after random assignment than individuals with no recent work experience. Likewise, high school graduates earned twice as much as nongraduates. Long-term welfare recipients received twice as much welfare after random assignment as welfare applicants. In comparison, there were only modest earnings and welfare differences by age of youngest child, number of children, sex, and race and ethnicity.

- **Measures from the POS generally were related to outcomes after random assignment, but less so than work and welfare history.**

Parents with low preference for work, high work-related parental concerns, health and emotional problems, and child care and transportation barriers earned less than other sample members and received higher welfare benefit amounts. Differences in earnings between subgroups were generally less than $1,000 per year, substantially less than the differences associated with work history and welfare history.

- **Income was similar across a wide range of subgroups.**

Although earnings were much lower for some groups than for others, income from earnings, AFDC, and Food Stamps was similar across a wide range of subgroups. With very few exceptions, groups had average income ranging roughly from $7,500 to $9,000 per year. This reflects a basic tradeoff under the old AFDC program. Welfare recipients who worked often had their welfare benefits reduced one dollar for each additional dollar they earned. Several groups did have much higher income than the average. In particular, people who had earned $5,000 or more in the year prior to random assignment had average income exceeding $11,000 per year in the three years after random assignment, and parents with three children or more had average income of more than $10,400 over the three-year period.

I. Subgroups Defined Using Demographic Characteristics and Welfare and Work History

All 20 programs collected information from a Background Information Form (BIF) at the time of random assignment and from administrative records on earnings and welfare receipt prior to random assignment. This section presents control group outcomes for seven sets of subgroups defined using these data. These characteristics are recent work history, welfare history, high school credential, number of children, age of youngest child, race and ethnicity, and sex.

A. The Subgroups

1. Recent Work History. Three subgroups were defined based on prior earnings: individuals with no earnings in the year prior to random assignment; those with some earnings, but less than $5,000 (in 1997 dollars) in the prior year; and those with $5,000 or more in earnings. Reflecting the fact that welfare recipients are unlikely to work and that most sites did not randomly assign welfare applicants, about 58 percent of the sample had no earnings in the year prior to random assignment, while 29 percent earned less than $5,000. Only 14 percent of the sample earned $5,000 or more in the prior year.

2. Welfare History. Three subgroups were defined based on responses to a BIF question about prior welfare receipt and on administrative records. Individuals who had received welfare for two years or more in the past were designated *long-term recipients*. Those who had received welfare for less than two years were called *short-term recipients*. Those who had never received AFDC before were called *new applicants*.[2] About 59 percent of the sample were long-term recipients, 30 percent were short-term recipients, and 11 percent were new applicants.

3. High School Credential. Two subgroups were defined based on receipt of a high school credential: *high school graduates* had either a high school diploma or a GED certificate at random assignment; *high school nongraduates* did not. The sample was fairly evenly split between the two groups: 57 percent had a high school diploma or GED and 43 percent had neither.

4. Number of Children. Individuals were placed in subgroups depending on the number of children they had at the time of random assignment. About 25 percent had three children or more, 32 percent had two children, and about 43 percent had only one child.

5. Age of Youngest Child. Two subgroups were defined based on the age of the youngest child at random assignment: families with a child under age 6 and families with no child under age 6. Even though SWIM and the programs in the GAIN evaluation were not mandatory for families with a child under age 6, the overall sample is fairly evenly split between these two groups: 42 percent had a child under age 6 and 58 percent did not.

6. Race and Ethnicity. Responses to the BIF were used to assign individuals to one of the following groups: white, African-American, Hispanic, and other. The sample is about evenly split between white and African-American sample members, with about 40 percent in each group. About 15 percent of sample members were Hispanic, and the remainder, including Asian-Americans, Native Americans, and Alaskan natives, were in the other category.

7. Sex. Although most single-parent welfare recipients are female, pooling the sample from 20 experiments provides an opportunity to compare impacts for men and women. About 92 percent of the overall sample was female and 8 percent was male.

B. Outcomes

1. Earnings. Table 4.1 shows average annual earnings for control group members in these subgroups for two periods of time: the entire three-year period after random assignment and the third year alone.

[2]In most of the programs, those called new applicants were approved for benefits before random assignment and were therefore really welfare recipients. In Oklahoma City and FTP, however, applicants were randomly assigned before finding out whether they had been approved to receive benefits, and some never received benefits during the follow-up period.

National Evaluation of Welfare-to-Work Strategies

Table 4.1

Earnings and AFDC Payments Pooled Across 20 Welfare-to-Work Programs, for Control Group Members by Selected Characteristics at the Time of Random Assignment

Sample or Subgroup	Sample Size	Average Total Earnings per Year ($)		Average Total AFDC Payments per Year ($)		Average Total Income per Year ($)	
		Years 1-3	Year 3	Years 1-3	Year 3	Years 1-3	Year 3
Full sample	71,932	2,943	3,665	4,160	3,181	8,685	8,199
Total earnings in past 12 months							
No earnings	41,434	1,754	2,392	4,675	3,639	8,082	7,482
Less than $5,000	20,554	3,425	4,253	3,696	2,775	8,707	8,368
$5,000 or more	9,944	6,957	7,808	2,967	2,100	11,200	10,884
Welfare history[a]							
Long-term recipient	43,339	2,480	3,186	4,791	3,755	9,027	8,475
Short-term recipient	21,333	3,708	4,484	3,400	2,479	8,463	8,073
New applicant	6,853	3,025	3,276	2,611	1,729	6,819	5,938
High school credential							
No high school diploma or GED	31,139	1,867	2,269	4,708	3,728	8,282	7,519
High school diploma or GED	40,793	3,751	4,710	3,749	2,777	8,989	8,715
Number of children							
3 or more	18,179	2,523	3,179	5,604	4,427	10,412	9,590
2	22,950	2,957	3,709	4,185	3,206	8,769	8,295
1	30,562	3,196	3,930	3,268	2,405	7,589	7,286
Age of youngest child							
Under 6	30,153	2,766	3,573	3,875	3,047	8,574	8,271
6 or over	41,095	3,080	3,693	4,363	3,278	8,778	8,113

(continued)

Table 4.1 (continued)

Sample or Subgroup	Sample Size	Average Total Earnings per Year ($)		Average Total AFDC Payments per Year ($)		Average Total Income per Year ($)	
		Years 1-3	Year 3	Years 1-3	Year 3	Years 1-3	Year 3
Race and ethnicity							
White	29,634	3,044	3,822	3,628	2,603	8,047	7,533
African-American	28,042	3,058	3,889	4,061	3,214	9,059	8,791
Hispanic	10,384	2,849	3,285	5,708	4,593	9,872	9,134
Other	3,489	2,098	2,610	5,175	4,097	8,714	8,034
Gender							
Female	65,563	2,887	3,637	4,144	3,179	8,640	8,193
Male	5,919	3,518	3,909	4,417	3,275	9,204	8,249

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, Food Stamp records, and Background Information Forms (BIFs).

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

[a]Sample members were classified as new applicants if they responded on the BIF that they had never received welfare in the past; as short-term recipients if they had received welfare before on their own case or their spouse's case for a total of less than two years; and as long-term recipients if they had received welfare for two years or more prior to random assignment.

Economists argue that an individual's earnings ability reflects her skills. In that light, the earnings levels shown in Table 4.1 are not surprising. Two characteristics — prior earnings (a proxy perhaps for work experience) and high school credential — are closely related to skills, and those two characteristics are also responsible for the largest differences across subgroups. Individuals with no earnings in the year prior to random assignment earned only one-fourth as much after random assignment as those who had earned $5,000 or more in the year prior to random assignment. Although the difference is not as stark, high school graduates earned twice as much as nongraduates.

The remaining subgroups show only small differences, and most earned close to the average of about $3,000 over the three-year period and $3,665 in year 3. Earnings by welfare history show a surprising pattern, however. Although long-term recipients earned substantially less than short-term recipients, new applicants also earned substantially less than short-term recipients. This result is contrary to the general notion that the more disadvantaged groups (that is, short-term recipients) will earn less than the less disadvantaged groups (that is, new applicants) and reflects the fact that applicants came disproportionately from Oklahoma City, a site with very low earnings on average.[3]

2. AFDC Payments. Just as prior earnings were the best predictor of future earnings, welfare history was the best predictor of future welfare payments in the control group. According to Table 4.1, long-term recipients received nearly $5,000 per year in AFDC over the three-year follow-up period and nearly $4,000 in year 3 alone. New applicants received much less: about $2,600 per year overall and $1,700 in year 3. This reflects the fact that new applicants are unlikely to remain on welfare long, but that many long-term recipients will stay on welfare until the problems that kept them on welfare for such a long time are resolved.

Other characteristics also appear related to AFDC benefit amounts. Prior earnings predicted future AFDC payments almost as well as welfare history. Individuals with no prior earnings also received more than $4,500 per year in AFDC over the three-year period, and more than $3,600 in year 3. Those who had earned $5,000 or more in the prior year, because they were earning so much more, received much less AFDC: about $3,000 per year over the three-year period and about $2,000 in year 3. High school graduates received about $1,000 less per year than nongraduates. Families with three children or more received about $2,000 more per year than families with only one child.

AFDC payments to racial and ethnic groups highlight one of the difficulties in drawing conclusions based on pooled results: pooled comparisons reflect not only differences across the subgroups, but also different compositions of the caseload by site. In this case, white sample members and African-American sample members received substantially less AFDC than Hispanic sample members. Location is the most likely explanation. Nearly all of the Hispanic sample members came from California, the state with the highest welfare guarantees for the programs represented in this report. Thus, AFDC payments were most likely high for Hispanic sample members because of where they lived, not who they were. Other pooled results by race and ethnicity should be viewed with this in mind. More generally, pooled

[3]In other words, this is a consequence of weighting results by the number of sample members in each site when sites have different sample compositions. Appendix C discusses this issue in greater detail.

results will always confound to some extent the effects of welfare-to-work programs on groups of individuals and the composition of the more successful and less successful programs.

3. Combined Income from Earnings, AFDC, and Food Stamps. There is much less variation in average combined income from earnings AFDC, and Food Stamps across subgroups than in either earnings or welfare payments (see Table 4.1). On average, sample members had income between $8,000 and $9,000 per year. Only three groups — those who had earned $5,000 or more in the prior year, families with three children or more, and Hispanic sample members — had income above $9,500 per year on average. Only one group — new applicants — had income below $7,000 per year. This similarity in income reflects a basic tradeoff faced by recipients of public assistance under the old AFDC program. Many of those who increased their earnings lost one dollar of AFDC payments for every dollar of earnings, leaving them with no additional income.

Income for these sample members should generally increase over time because many went to work and left welfare entirely. Nevertheless, average combined income in the third year of follow-up was less than average combined income over the three-year period ($8,199 compared with $8,685). This may reflect the drawbacks of using administrative records: over time, people move and are no longer represented in a state or county's administrative records.[4]

II. Control Group Outcomes for Subgroups Defined Using the Private Opinion Survey

In 9 of the 20 programs — Portland, MFIP, FTP, and the LFA and HCD programs in Atlanta, Grand Rapids, and Riverside — sample members completed a confidential Private Opinion Survey (POS) at the time of random assignment. This section examines average outcomes for members of the control group for several subgroups defined using responses to the POS, including preference for work over welfare, work-related parental concerns, feelings of mastery, and barriers to work from child care, transportation, and health or emotional problems.

Plotnick, Klawitter, and Edwards (1998) used National Longitudinal Survey of Labor Market Experience, Youth Cohort (NLSY) data to examine the link between welfare receipt and several psychosocial measures, including self-esteem, locus of control, attitudes toward school, aversion to accepting welfare, work commitment, attitudes toward work and family roles for married women, and attitudes about nontraditional employment and family roles of a person the respondent considered influential in her life. Several of their measures are similar to measures available from the POS, that is, locus of control, preference for work, and work-related parental concerns. In comparing those who ever received AFDC with those who did not, Plotnick, Klawitter, and Edwards found substantial differences only in attitudes toward school. After controlling for family background, state of residence, and time period, they found that only work commitments and attitudes toward school were statistically significant predictors of who had received welfare. If they are right, then the attitudinal predictors from the POS might

[4]In FTP, for example, according to Florida administrative records, more than one-third of both the control and program groups had neither cash assistance nor earnings just three years after random assignment (Bloom et al., 1999).

not help understand the outcomes for the control group. Nevertheless, they may still help explain the impacts of welfare-to-work programs.

A. The Subgroups

1. Preference for Work. Two questions from the POS were used to measure sample members' preference for work over welfare. The first asked respondents whether they would take a full-time job today if it paid less than welfare; the second asked whether they would take a full-time job today if it paid more than welfare but they would not like the work. Individuals who said that they would take a job in either case were classified as having a "high" preference for work. Those who said that they would take a job in neither case were classified as having a "low" preference for work, and those who said they would take a job in one situation but not the other were classified as having a "moderate" preference for work.

Prior research has found a weak link between these questions and employment. A number of studies analyzed employment and mothers' responses to questions about attitudes toward working using National Longitudinal Surveys of Labor Market Experience (NLS) data (Macke, Hudis, and Larrick, 1978; Statham and Rhoton, 1983; Spitze and Waite, 1980; and Michalopoulos, 1994). In three of the four studies, women who expressed more favorable attitudes toward work did work modestly more than other women. The one study that found no relationship between attitudes and employment used a sample of younger women, many of whom were still in school. While these four studies used somewhat different attitudinal questions to gauge attitudes about work, they suggest that preferences do help understand who and who does not work.

Although preference for work may help predict who works, responses to these POS questions may confound preference for work with earnings ability. It is likely, for example, that someone who expects to earn more than her welfare grant would hold out for a higher-paying job. Thus, the relationship between preference for work and program impacts may look different if examined among people with similar earnings potential, perhaps using education or prior earnings as a proxy.

2. Work-Related Parental Concerns. Four POS items were used to construct a scale of work-related parental concerns, that is, the degree to which respondents preferred to stay with their family rather than work or go to school. The scale includes two items that relate to fears about day care. One asked respondents to evaluate the statement "I cannot go to a school or training program right now because I am afraid to leave my children in day care or with a baby sitter." The second asked them to evaluate the statement "If I got a job, I could find someone I trust to take care of my children." Of the two other items that make up this scale one relates directly to emotional work-related parental concerns ("I do not want a job because I would miss my children too much") and the other is a simple statement of preference ("Right now I'd prefer not to work so I can take care of my family full time").

To the extent that lower work-related parental concerns reflect less commitment to staying home to care for children or less fear of leaving children in day care, they would tend to be associated with higher earnings and lower AFDC payments. In the studies of the NLS data cited earlier, some of the attitudinal questions concerned the appropriateness of work for mothers. Thus, the NLS studies provide some indication that these concerns are related to employment outcomes. On the other hand, the more recent work by Plotnick, Klawitter, and Edwards found little link between welfare receipt and attitudes about family roles and work.

3. Mastery. Four POS items were used to construct a modified version of the Pearlin Mastery Scale (Pearlin et al., 1981), which is designed to measure a person's sense of mastery over external events. The scale's designers suggested that this scale may capture problems with individuals' conceptions of their control over their personal destinies as a result of "hardships that are an enduring testimony to one's lack of success or to the inadequacy of one's efforts to avoid problems" (Pearlin et al., 1981, p. 345). A low sense of mastery may dampen one's initiative to enter the job market or may lead to depression. Therefore, we expect that individuals who have a low sense of mastery will have lower earnings and higher AFDC payments than individuals with a high sense of mastery.

The four POS items asked sample members whether they strongly agreed, agreed, disagreed, or strongly disagreed with four statements: "I have little control over the things that happen to me," "I often feel angry that people like me never get a fair chance to succeed," "Sometimes I feel that I'm being pushed around in life," and "There is little that I can do to change many of the important things in my life." Sample members were divided into "high" and "low" mastery subgroups based on their responses.

4. Health or Emotional Problem. This construct measures whether sample members had health or emotional problems or family members with health or emotional problems that prevented them from working or going to school at the time of random assignment.

5. Child Care Problem. Individuals were placed in one of two groups depending on whether they indicated that they could not participate in work-related activities because they could not afford or arrange child care. Mothers who lack affordable or reliable child care are at a disadvantage in the labor market. In the pooled sample, 60 percent anticipated that they could not work or go to school because they could not either afford or arrange for child care at the time of random assignment.

6. Transportation Problem. Individuals were placed in one of two groups depending on whether they indicated that they could not participate in work-related activities because they had no way to get to the activity every day. Lacking a car or driver's license or living in an area where there is inadequate public transportation may impede a welfare recipient's ability to get and maintain a job.

B. Outcomes

1. Earnings. Table 4.2 shows average annual earnings for members of the control group in these subgroups for two periods of time: the entire three-year period after random assignment and the third year alone. The obvious hypothesis is that families with more barriers to work will earn less than families with fewer barriers. A number of characteristics were related to substantial differences in earnings, and all concurred with this hypothesis. Parents with many work-related concerns earned slightly more than half as much as parents with fewer concerns ($1,973 compared with $3,524 per year over the three-year period); parents with health, emotional, or family problems likewise earned much less than parents who did not indicate these problems ($2,097 compared with $3,525); and families with

National Evaluation of Welfare-to-Work Strategies

Table 4.2

Earnings and AFDC Payments Pooled Across Welfare-toWork Programs Administering a POS, for Control Group Members by Selected Characteristics at the Time of Random Assignment

Subgroup	Sample Size	Average Total Earnings per Year ($) Years 1-3	Year 3	Average Total AFDC Payments per Year ($) Years 1-3	Year 3	Average Total Income per Year ($) Years 1-3	Year 3
Preference for work							
Low	8,552	3,123	4,162	3,471	2,567	8,445	8,236
Moderate	8,541	3,098	3,958	3,307	2,484	8,290	7,995
High	2,782	3,444	4,243	2,955	2,176	8,285	7,929
Work-Related Parental Concerns Scale							
High	4,786	1,973	2,761	4,012	3,124	8,067	7,644
Low	15,796	3,524	4,486	3,099	2,270	8,424	8,214
Mastery Scale							
Low	7,680	2,554	3,316	3,480	2,656	8,016	7,636
High	12,911	3,503	4,500	3,196	2,347	8,495	8,296
Risk of Depression Scale[a]							
High	2,507	3,071	4,039	3,308	2,352	8,281	7,897
Moderate	4,157	3,138	3,984	3,381	2,582	8,510	8,236
Low	10,588	3,049	3,972	3,496	2,634	8,529	8,276
Health or emotional problem[b]							
Yes	5,507	2,097	2,733	3,518	2,655	7,619	7,058
No	15,181	3,525	4,529	3,220	2,389	8,563	8,395
Child care problem							
Yes	12,478	2,827	3,686	3,591	2,732	8,397	8,070
No	7,832	3,666	4,679	2,863	2,043	8,222	8,061

(continued)

Table 4.2 (continued)

Subgroup	Sample Size	Average Total Earnings per Year ($) Years 1-3	Year 3	Average Total AFDC Payments per Year ($) Years 1-3	Year 3	Average Total Income per Year($) Years 1-3	Year 3
Transportation problem							
Yes	7,212	2,026	2,636	3,736	2,893	7,753	7,194
No	13,252	3,742	4,806	3,078	2,244	8,616	8,502

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, Food Stamp records, and Private Opinion Survey (POS) data.

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

[a]Subgroups include sample members from four NEWWS sites only: Atlanta, Grand Rapids, Riverside, and Portland.

[b]Sample members in the "yes" category could have had a health or emotional problem themselves that they reported as a barrier to work or participation at random assignment or one of their family members could have had such a problem.

transportation problems earned slightly more than half as much as families without transportation problems ($2,026 compared with $3,742).

Affordability of child care was also related to earnings, but not as closely linked as other barriers. This makes sense in light of two factors. All programs that administered a POS also provided some sort of child care assistance to the control group, so that child care might not have been a major barrier to work. Nevertheless, many eligible welfare recipients did not receive child care payments, and the percentage of sample members who received transitional child care was generally small.[5]

Programs begun under the Family Support Act were likewise required to provide transportation assistance to control group members who initiated employment-related activities. In the NEWWS sites, however, very few control group members actually received transportation payments, which may be responsible for the large difference in control group earnings between sample members with transportation problems and others.

2. AFDC Payments. Patterns of welfare payment amounts for control group members are similar to patterns of earnings, but the differences across subgroups are much smaller. Parents with many work-related concerns received about $1,000 more per year from AFDC than did parents with fewer concerns. Families with transportation and child care problems received about $700 less per year than other families. For all other subgroups, annual AFDC payments were roughly between $3,000 and $3,500 per year over the three-year follow-up period.

3. Combined Income from Earnings, AFDC, and Food Stamps. There were very slight differences in combined income across the subgroups. All subgroups had average income for the three-year period roughly between $7,600 and $8,600 per year. This again reflects the tradeoff between earnings and welfare under the old AFDC program.

III. Summary of Control Group Outcomes

In general, control group members with fewer barriers to work earned more than others and received smaller AFDC payments. High school graduates earned more than nongraduates, individuals with substantial recent work experience earned more than those who had not worked recently, and new applicants and short-term recipients all earned substantially more than long-term recipients. There were likewise large differences between subgroups defined using POS data. Individuals who cited child care, transportation, or health as barriers to work earned less than others and received higher welfare payments. Individuals who did not feel in control of their life earned substantially less than those who did.

[5]Information on transitional child care receipt rates is available for the GAIN and NEWWS evaluations. Receipt rates for the program group range from zero in Riverside's GAIN program to 23 percent in Portland's JOBS program, which was the only program to have the impact on receipt rates over 10 percent. For individuals in the NEWWS Evaluation who were administered a survey after two years of follow-up, it was estimated that only about 11 percent of the control group that was eligible for transitional child care benefits (because their AFDC stopped owing to increased earnings or acquisition of a new job) across all sites used transitional child care (Freedman et al., 2000, pp. 163-168).

As a consequence, a number of subgroups have a long way to go to achieve self-sufficiency. Individuals who had not worked in the year prior to random assignment earned less than $1,800 per year in the follow-up period. High school nongraduates also earned less than $2,000 per year, while long-term recipients earned about $2,500 per year. Individuals with transportation, child care, or health barriers to work earned between $2,000 and $3,000 per year. Moreover, in each case welfare payments were larger than earnings over the three-year follow-up period. Unless welfare-to-work programs increase earnings by a substantial amount, most members of these more disadvantaged groups will earn little and remain dependent on welfare.

Chapter 5

Impacts for Subgroups Defined Using Demographic Characteristics and Welfare and Work History

Chapter 4 described outcomes for a variety of subgroups and found that, in general, people with more skills and more work experience had higher earnings and lower welfare payments than others. This chapter relates those traits to impacts of the welfare-to-work programs. It examines impacts for the subgroups based on administrative records and Background Information Forms (BIFs) — namely, recent work history, welfare history, high school credential, age of youngest child, number of children, race and ethnicity, and sex. Table 5.1 summarizes the findings, pooled and by program, for these subgroups. Four general findings provide a summary of the chapter.

- **Impacts on earnings for the less disadvantaged groups were sometimes larger and sometimes smaller than impacts for the more disadvantaged groups.**

Combining impacts across the 20 programs produced earnings gains that were larger for individuals with no recent work experience than for those with a modest amount of recent work experience and larger for parents with three children or more than for parents with one child. However, earnings gains were twice as large for new welfare applicants as for long-term recipients and larger for high school graduates than for nongraduates.

- **Impacts on welfare payments for the less disadvantaged groups were typically larger than impacts for the more disadvantaged groups.**

Although earnings gains were sometimes larger for the less disadvantaged groups, reductions in welfare benefits were as large for the more disadvantaged groups as for the less disadvantaged groups, or sometimes larger. For example, the programs increased earnings more for high school graduates than for nongraduates, but reduced welfare payments equally for the two groups. Likewise, earnings gains were significantly larger for new applicants than for long-term recipients, but welfare savings were significantly larger for long-term recipients than for new applicants.

- **Income neither increased nor decreased significantly for most subgroups**.

Despite significant earnings gains for many groups, on average, people traded their increased earnings for lower welfare benefits and ended up with neither higher nor lower total income from earnings, AFDC, and Food Stamps. There were two notable exceptions to this finding. Over a three-year period, the programs increased combined income by nearly $800 per year for new applicants and about $200 per year for families with children under age 6.

- **For the more disadvantaged groups, outcomes were fairly consistent across the 20 programs.**

In many ways, outcomes were fairly consistent across the 20 programs. Most of them significantly increased earnings and reduced welfare benefits for long-term recipients, for those with no recent work history, for families with children under age 3 (among those that enrolled families

National Evaluation of Welfare-to-Work Strategies

Table 5.1

Summary of Impacts by Subgroup Using Baseline Characteristics

Subgroup	Pooled Impacts	Impacts by Program
Recent work history	• Earnings increased and welfare payments decreased by similar amounts across subgroups. • Income did not increase for any subgroup.	• Most programs significantly increased earnings for those with no recent work history. • Income gains were similar for all groups.
Welfare history	• Earnings increased for all subgroups, but the most for new applicants. • Welfare payments decreased for all groups, but the most for long-term recipients. • Income remained unchanged for all subgroups.	• Most programs significantly increased earnings and decreased welfare benefits for long-term recipients. Results for other groups were less clear. • Few programs increased income for any subgroups, and there were few differences across groups.
High school credential	• Earnings increased and welfare payments decreased significantly for high school graduates and nongraduates. • Income increased more for high school graduates than nongraduates.	• Half of the programs increased earnings for graduates, and half increased earnings for non-graduates. • Few significant differences were found in income gains.
Number of children	• Earnings increased and welfare payments decreased less for smaller families than for larger families. • Small differences were found in effects on income across subgroups.	• Programs were more likely to produce significant earnings gains for larger families than for smaller families. • Few differences were found in effects on income.
Age of youngest child	• Earnings increased and welfare payments decreased for families with and without preschool children. • Income increased for families with young children.	• All programs with very young children significantly increased earnings for families with children under age 3. • Effects on income were rare for families with young children.
Race and ethnicity	• No comparisons were made because sample composition differed substantially across programs.	• Several programs generated significantly larger earnings impacts for nonwhite subgroups. • Few significant differences were found in income gains.
Sex	• Impacts were similar for families headed by men and by women.	• No comparisons were made because there were too few men in most programs to obtain reliable estimates.

with very young children), for parents with three children or more, and for African-American and Hispanic sample members (among those that enrolled a diverse mix of racial and ethnic groups).

Although most programs neither increased nor decreased combined income from earnings, AFDC, and Food Stamps for most subgroups, several programs did. MFIP, which is one of the two programs with financial incentives to encourage work, increased combined income across a wide range of subgroups. The GAIN programs in Riverside and Butte also increased combined income for a number of subgroups, though less consistently than MFIP. On the other hand, the HCD program in Riverside reduced combined income for a number of the more disadvantaged groups.

I. Hypotheses and Findings from Other Welfare-to-Work Studies

A. Hypotheses

For the outcomes presented in Chapter 4, the expectations were clear. Groups with less work experience and with less education were expected to work and earn less and, therefore, receive more in welfare. Because long-term welfare recipients probably had little work experience, this prediction extended to them as well and extended to characteristics associated with long-term receipt: having many children, being depressed, or having emotional or physical problems.

Expectations regarding impacts are far less clear. According to one line of thinking, impacts may be smallest for the more disadvantaged groups because they might have significant problems that keep them from either working or fulfilling participation mandates. Moreover, sanctioning might substantially reduce both their welfare benefit amounts and their income. At the same time, participation requirements may provide a disincentive to remain on welfare that the most employable will respond to by returning to work and leaving welfare.

On the other hand, the Family Support Act was explicitly designed to help individuals who are the most likely to become long-term recipients. Because the less disadvantaged groups do relatively well without welfare-to-work programs, they may benefit less from these programs. Thus, impacts might be larger for the more disadvantaged groups.

B. Prior Evidence

A body of prior research exists on subgroup impacts for programs that both preceded and followed the Family Support Act, including some of the same programs included in this report. These prior results provide little additional insight, however.

1. Findings from the GAIN Evaluation. The three-year analysis of GAIN examined impacts for a number of subgroups, including subgroups by educational attainment and need for basic education, welfare history, recent work history, race and ethnicity, number of children, and composite measures of disadvantage. The GAIN programs substantially increased earnings for those with prior work experience. Notable was Riverside's program, which had uniformly high impacts on earnings across subgroups.

Impacts on AFDC were also more common for the less disadvantaged groups than for the more disadvantaged groups, although four of the six programs did produce AFDC impacts for long-term recipients.[1]

2. Findings from the NEWWS Evaluation. An interim report in the NEWWS Evaluation included impacts for subgroups defined by educational attainment. Employment-focused programs were more likely than education-focused programs to produce impacts on employment and earnings at the two-year mark for sample members who did not have a diploma or GED at study entry. The Portland program had the most consistent and substantial earnings impacts, generating impacts for all but one of the subgroups analyzed.[2]

3. Findings from a Subgroup Analysis. In *Subgroup Impacts and Performance Indicators for Selected Welfare Employment Programs* (1988), Daniel Friedlander used data from five evaluations of state welfare-to-work programs that used random assignment (those in San Diego, Baltimore, Virginia, Arkansas, and Cook County, Illinois) to examine subgroups across programs on the basis of prior employment, welfare history, and other demographic characteristics. Friedlander found that the most job-ready and least welfare-dependent groups had below-average program impacts that were generally not statistically significant. Earnings impacts were most consistently found for moderately disadvantaged subgroups — for example, applicant returnees with earnings of less than $3,000 in the prior year. Earnings impacts were not consistently found for subgroups in the most dependent tier, including recipients with no prior earnings and those with two years or more on AFDC. The distribution of significant earnings impacts, therefore, formed an inverted U-shape, with the largest found impacts occurring among moderately disadvantaged sample members and fewer and small impacts occurring for both the more and less disadvantaged sample members.

II. Pooled Impacts for BIF Subgroups

This section presents results from the analysis of impacts on earnings, AFDC payments, and total combined income from earnings, AFDC, and Food Stamps, pooled across all 20 programs. As described in Chapter 4, subgroups defined using administrative records and BIFs included those based on recent work history, welfare history, high school credential, number of children, age of youngest child, race and ethnicity, and sex.

A. Impacts on Earnings

Table 5.2 shows the pooled impacts on annual earnings averaged across the three-year follow-up period and for the third year alone.[3] For all subgroups in both time periods, impacts on earnings were significantly positive. Beyond that, the results do not follow a clear pattern. In some cases, impacts on earnings were significantly larger for the less disadvantaged groups. For example, the programs increased earn

[1]For a complete discussion of three-year impacts in the GAIN evaluation, see Riccio, Friedlander, and Freedman (1994), Chapter 4.

[2]For a more detailed discussion of two-year subgroup impact findings from the NEWWS Evaluation, see Freedman et al. (2000), Chapter 11.

[3]For information on outcomes, see Chapter 4.

National Evaluation of Welfare-to-Work Strategies

Table 5.2

Impacts on Earnings and AFDC Payments Pooled Across Welfare-to-Work Programs, for the Full Sample by Selected Characteristics at the Time of Random Assignment

		Impacts on Average Total Earnings per Year ($)		Impacts on Average Total AFDC Payments per Year ($)		Impacts on Average Total Income per Year ($)	
Sample or Subgroup	Sample Size	Years 1-3	Year 3	Years 1-3	Year 3	Years 1-3	Year 3
Full sample	71,932	530 ***	606 ***	-385 ***	-382 ***	38	103 *
Total earnings in past 12 months					††		
No earnings	41,434	571 ***	685 ***	-416 ***	-440 ***	41	116 *
Less than $5,000	20,554	399 ***	431 ***	-359 ***	-352 ***	-58	-48
$5,000 or more	9,944	548 ***	550 **	-305 ***	-191 **	143	290
Welfare history[a]		†	†††	†	††	††	†††
Long-term recipient	43,339	544 ***	627 ***	-433 ***	-439 ***	4	59
Short-term recipient	21,333	534 ***	627 ***	-337 ***	-333 ***	94	191
New applicant	6,853	1,106 ***	1,694 ***	-218 **	-62	773 ***	1,572 ***
High school credential		††				††	
No high school diploma/GED	31,139	430 ***	546 ***	-395 ***	-401 ***	-66	19
High school diploma/GED	40,793	627 ***	683 ***	-389 ***	-389 ***	123 *	172 **
Number of children		†††	†††	†	†††		
3 or more	18,179	682 ***	818 ***	-458 ***	-498 ***	93	161
2	22,950	663 ***	726 ***	-408 ***	-430 ***	128 *	143
1	30,562	328 ***	386 ***	-326 ***	-273 ***	-65	49
Age of youngest child		††	†††			††	†††
Under 6	30,153	665 ***	900 ***	-352 ***	-401 ***	209 *	381 ***
6 or over	41,095	424 ***	455 ***	-397 ***	-365 ***	-94	-41

(continued)

Table 5.2 (continued)

Sample or Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($)		Impacts on Average Total AFDC Payments per Year ($)		Impacts on Average Total Income per Year ($)	
		Years 1-3	Year 3	Years 1-3	Year 3	Years 1-3	Year 3
Sex							
Female	65,563	530 ***	598 ***	-395 ***	-400 ***	28	77
Male	5,919	583 ***	781 ***	-321 ***	-236 **	180	479 *

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Background Information Forms (BIFs).

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

[a]Sample members were classified as new applicants if they responded on the BIF that they had never received welfare in the past; as short-term recipients if they had received welfare before on their own case or their spouse's case for a total of less than two years; and as long-term recipients if they had received welfare for two years or more prior to random assignment.

ings for new applicants by more than $1,000 per year, but increased earnings for long-term recipients by about $500 to $600 per year. Likewise, impacts were larger for high school graduates than for nongraduates. In other cases, earnings impacts were larger for groups that would be expected to have a harder time going to work. Families with three children or more increased their earnings by nearly $700 per year, but families with only one child increased their earnings by $300 to $400 per year. Families with children under age 6 increased their earnings by $665 over the three-year period compared with $424 for families with no pre-school children. Finally, in some cases, the more disadvantaged and less disadvantaged groups had similar earnings impacts. In particular, impacts were similar for individuals who had not worked in the year prior to random assignment as for those who had worked and earned $5,000 or more.

B. Impacts on AFDC Payments

Table 5.2 also shows the pooled impacts on AFDC payments. For all groups except new applicants in year 3, the programs collectively reduced AFDC payments by a significant amount. Differences across subgroups tended to be smaller than differences in earnings impacts. For example, impacts were similar for high school graduates and nongraduates, even though earnings impacts were larger for high school graduates. Impacts were about the same for families with preschool children and families with no preschool children, even though earnings impacts were larger for families with preschool children.

Where differences were significant, impacts were larger for the more disadvantaged groups. The programs reduced AFDC payments for families with two children or more by more than $400 per year, for example, but reduced payments for families with one child by only about $300 per year. This is not surprising in light of two facts: larger families qualify for larger welfare payments (and therefore have more room to have their payments reduced), and larger families had larger earnings impacts than smaller families.

More surprising were the impacts by welfare history. The programs reduced AFDC payments by more than $400 per year for long-term recipients, but reduced payments by only about $200 per year over the three-year follow-up period for new applicants, for example. At first, this does not sound surprising since long-term recipients in the control group received substantially more welfare than new applicants (see Chapter 4). However, earnings impacts were much larger for new applicants than for long-term recipients, which might imply that most new applicants who increased their earnings because of the welfare-to-work programs would have left welfare quickly anyway. It could also imply that a few applicants were responsible for both sets of impacts; that is, that they increased their earnings by substantial amounts and therefore became ineligible for welfare.

C. Impacts on Combined Income from Earnings, AFDC, and Food Stamps

Finally, Table 5.2 shows the collective program impacts on combined income from earnings, AFDC, and Food Stamps. In general, income neither increased nor decreased for most subgroups. This demonstrates a basic tradeoff faced by most welfare recipients: if they increased their earnings under the old AFDC system, they typically lost welfare benefits at almost the same rate as they increased their earnings, and their total income remained unchanged.

There were some exceptions to this general rule, though it is not clear why these groups differed from the others. For new applicants, the programs increased annual combined income by $773 over the three-year period and more than $1,500 in year 3 alone. For families with children under age 6, income increased by about $200 per year over the entire follow-up period and nearly $400 in year 3. Impacts on income were also significantly positive over the entire follow-up period for high school graduates and families with two children, though the impacts for these two groups were fairly small.

III. Impacts by Individual Program for BIF Subgroups

The section above used impacts pooled across the 20 programs to establish some broad patterns. This section examines the same impacts for the programs individually. The first question is whether individual programs exhibited the same patterns found in the pooled analysis. If they do not, the logical next question is whether there are some characteristics of the programs or sites that help to explain differences in impacts across subgroups. Three subgroups are examined: by recent work history, by high school credential, and by welfare history. In addition, Appendix Tables B.1-B.3 present results by program for number of children, age of youngest child, and race and ethnicity.

A. Earnings in Prior Year

In the results pooled across programs, impacts on earnings were positive and impacts on welfare payments were negative for the three subgroups based on earnings in the year prior to random assignment. Moreover, impacts were similar across the three groups. Table 5.3 indicates that these results are consistent across programs only in some respects.

Over the three-year follow-up period, 14 of the 20 programs significantly increased earnings for individuals with no earnings in the year prior to random assignment and 14 programs significantly reduced welfare payments for this group. In other words, the programs consistently increased earnings and reduced welfare payments for the most disadvantaged groups.

Most of the programs, in contrast, did not increase earnings or reduce welfare payments significantly for those who had earned $5,000 or more in the year prior to random assignment. In many of the programs few people earned $5,000 or more in the prior year, which is partly responsible for the lack of significant impacts. However, among programs with 400 or more individuals in this group, three had significant impacts on earnings and five did not.

Although these differences across subgroups are interesting, they may be due as much to chance as to true effects of the programs. The pooled results showed small differences across subgroups in either earnings gains or welfare reductions. The individual results confirm the lack of significant differences. In only three programs were yearly earnings gains significantly different across the three groups. In only one program (Portland) were three-year AFDC impacts significantly different across the three groups. The lack of significant differences is again largely due to the small number of individuals with substantial prior earnings in these programs.

Several programs had particularly large impacts on earnings for some subgroups. SWIM and the GAIN programs in Butte, San Diego, and Riverside increased earnings by more than $1,000 per year for individuals earning $5,000 or more in the year prior to random assignment. Note that the programs evaluated in NEWWS were less likely to have significant earnings impacts for this group both because they

National Evaluation of Welfare-to-Work Strategies

Table 5.3

Impacts on Earnings and AFDC Payments Across 20 Welfare-to-Work Programs, by Total Earnings in Year Prior to Random Assignment (in 1997 Dollars)

		Impacts on Average Total Earnings per Year ($)		Impacts on Average Total AFDC Payments per Year ($)	
Program and Subgroup	Sample Size	Years 1-3	Year 3	Years 1-3	Year 3
SWIM					
No earnings	1,948	633 ***	640 **	-843 ***	-845 ***
Less than $5,000	765	704 *	1,142 **	-634 **	-576 *
$5,000 or more	497	1,218 *	626	-422	-230
GAIN evaluation programs					
Alameda					
No earnings	915	659 **	1,065 **	-155	-305
Less than $5,000	226	347	171	-241	-307
$5,000 or more	n/a	n/a	n/a	n/a	n/a
Butte		††	†††		
No earnings	652	-209	-516	77	83
Less than $5,000	377	1,162 **	1,629 **	-473	-197
$5,000 or more	200	3,670 **	5,027 ***	199	-317
Los Angeles					
No earnings	3,435	198	263	-384 ***	-314 **
Less than $5,000	716	178	531	-389	-275
$5,000 or more	245	-1,003	-1,722	-72	7
Riverside					
No earnings	3,331	1,262 ***	1,146 ***	-679 ***	-580 ***
Less than $5,000	1,412	1,172 ***	1,088 ***	-703 ***	-586 **
$5,000 or more	765	1,917 **	2,032 **	-856 **	-742 **
San Diego					
No earnings	4,615	478 **	763 ***	-460 ***	-417 **
Less than $5,000	2,097	394	133	-282	-281
$5,000 or more	1,507	1,471 **	1,673 **	-408	-259
Tulare					
No earnings	1,294	215	565 *	-90	-266
Less than $5,000	610	-188	-102	3	-163
$5,000 or more	330	8	1,660	520	387
NEWWS programs					
Atlanta LFA					
No earnings	2,353	604 ***	714 ***	-237 ***	-226 ***
Less than $5,000	1,114	397	444	-141 *	-96
$5,000 or more	366	-3	-419	-184	-115

(continued)

Table 5.3 (continued)

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($) Years 1-3	Year 3	Impacts on Average Total AFDC Payments per Year ($) Years 1-3	Year 3
Atlanta HCD					
No earnings	2,398	529 ***	759 ***	-183 ***	-184 ***
Less than $5,000	1,110	15	26	-169 **	-160 *
$5,000 or more	373	-232	40	-101	41
Grand Rapids LFA					††
No earnings	1,527	522 ***	692 **	-791 ***	-793 ***
Less than $5,000	1,116	437 *	242	-649 ***	-393 ***
$5,000 or more	369	133	-434	-543 ***	-150
Grand Rapids HCD					†
No earnings	1,489	239	208	-496 ***	-575 ***
Less than $5,000	1,118	360	369	-540 ***	-485 ***
$5,000 or more	390	970	1,165	-252	-58
Riverside LFA					†
No earnings	4,010	782 ***	646 ***	-702 ***	-684 ***
Less than $5,000	1,596	468 *	334	-683 ***	-675 ***
$5,000 or more	1,120	17	-430	-503 ***	-195
Riverside HCD					
No earnings	2,065	300 **	472 **	-573 ***	-680 ***
Less than $5,000	686	95	476	-463 *	-577 **
$5,000 or more	384	104	-203	-851 ***	-766 **
Columbus Integrated					
No earnings	2,143	483 ***	680 ***	-351 ***	-424 ***
Less than $5,000	1,561	133	212	-395 ***	-365 ***
$5,000 or more	968	256	237	-303 ***	-285 ***
Columbus Traditional					
No earnings	2,160	585 ***	635 **	-365 ***	-400 ***
Less than $5,000	1,591	165	159	-248 ***	-231 **
$5,000 or more	978	-320	-461	-163 *	-146
Detroit					
No earnings	2,978	176	313	-83	-179 *
Less than $5,000	1,197	339	568	-323 ***	-432 ***
$5,000 or more	284	1,260	1,924	-286	-339
Oklahoma City					†
No earnings	2,581	34	129	-34	51
Less than $5,000	2,348	57	-122	-129 **	-159 **
$5,000 or more	932	-234	-392	-107	-106

(continued)

Table 5.3 (continued)

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($) Years 1-3	Year 3	Impacts on Average Total AFDC Payments per Year ($) Years 1-3	Year 3
Portland		††	††	†††	†††
No earnings	3,214	1,476 ***	1,849 ***	-876 ***	-1,062 ***
Less than $5,000	1,660	634 **	751 *	-436 ***	-492 ***
$5,000 or more	673	485	619	-540 ***	-322 *
FTP					
No earnings	1,499	352 *	496 *	-199 ***	-334 ***
Less than $5,000	922	775 ***	1,151 ***	-26	-258 ***
$5,000 or more	394	695	1,125	-247 ***	-266 ***
MFIP		†			
No earnings	797	1,074 ***	924 *	349	173
Less than $5,000	472	52	109	770 **	769 *
$5,000 or more	n/a	n/a	n/a	n/a	n/a

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Background Information Forms (BIFs).

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

N/a = not applicable because sample sizes were too small to reliably calculate impacts.

ings for new applicants by more than $1,000 per year, but increased earnings for long-term recipients by about $500 to $600 per year. Likewise, impacts were larger for HCD program came close to these levels of impacts, but neither program's impacts were statistically significant.

For the group with no prior earnings, three programs (Riverside GAIN, Portland, and MFIP) produced earnings gains of more than $1,000 over the three-year period. All three programs were employment-focused, though Riverside GAIN and Portland used a mix of first activities within their employment focus. This provides some evidence that employment-focused programs may have increased earnings more for the more disadvantaged groups, but that programs with a mix of first activities may have increased earnings more consistently across a wide range of subgroups.

In addition to increasing earnings substantially for individuals with no prior earnings, MFIP is notable for a second reason; it was the only program that significantly increased welfare payments for any subgroup. For individuals who had earned less than $5,000 in the year prior to random assignment, MFIP increased welfare payments by about $700 per year even though it did not affect earnings for that group. For those with no prior earnings, program group members received higher welfare payments than control group members, though not significantly so, even though the program significantly increased earnings for that group.[4] This unusual effect on welfare payments is probably due to MFIP's combination of relatively high welfare benefit levels and an enhanced earnings disregard that allowed welfare recipients in the program group to keep more of their benefits when they went to work.

Although no other program increased welfare payments, in several other cases reductions in welfare payments were far smaller than increases in earnings. For individuals with no prior-year earnings, this was most noticeable for the GAIN program in Riverside and the two JOBS programs in Atlanta. As a result, the programs increased combined income for this subgroup, as it did for both subgroups in MFIP (results for income not shown in Table 5.3).

B. <u>High School Credential</u>

The combined results discussed in the previous section indicated that earnings impacts were significantly positive for both high school graduates and nongraduates, but smaller for nongraduates (although not significantly so in year 3). Table 5.4 shows similar results across the 20 programs. Ten of the programs significantly increased earnings for high school graduates, and 12 significantly increased earnings for nongraduates. While differences between the two groups were statistically significant in only six programs, in four of those programs (FTP and the GAIN programs in Alameda, Riverside, and San Diego) high school graduates had larger earnings impacts than nongraduates; in the Grand Rapids LFA program and the Columbus Integrated program, impacts were larger for nongraduates than for graduates. Most of the programs also reduced welfare payments for both groups, but the differences were rarely statistically significant.

Among individual programs, Butte and Riverside GAIN significantly increased earnings for high school nongraduates by more than $1,000 per year over the three-year period. Portland and the GAIN

[4]Recall that welfare payments in MFIP also include Food Stamps and General Assistance. For all other programs, welfare includes only AFDC.

National Evaluation of Welfare-to-Work Strategies

Table 5.4

Impacts on Earnings and AFDC Payments Across 20 Welfare-to-Work Programs, by High School Diploma/GED Status

		Impacts on Average Total Earnings per Year ($)		Impacts on Average Total AFDC Payments per Year ($)	
Program and Subgroup	Sample Size	Years 1-3	Year 3	Years 1-3	Year 3
SWIM					
No high school diploma/GED	1,408	516 **	539 *	-693 ***	-655 ***
High school diploma/GED	1,802	747 **	771 *	-688 ***	-660 ***
GAIN evaluation programs					
Alameda		†††	†††		
No high school diploma/GED	444	-542	-540	-381	-371
High school diploma/GED	761	1,203 ***	1,681 ***	-113	-299
Butte				†	
No high school diploma/GED	517	1,257 ***	1,474 ***	-645	-569
High school diploma/GED	712	659	848	310	289
Los Angeles					†
No high school diploma/GED	2,873	-37	10	-244 *	-97
High school diploma/GED	1,523	318	431	-549 ***	-602 ***
Riverside		††			
No high school diploma/GED	2,613	1,029 ***	1,035 ***	-736 ***	-556 **
High school diploma/GED	2,895	1,780 ***	1,646 ***	-735 ***	-688 ***
San Diego		††	†		
No high school diploma/GED	3,520	170	302	-434 **	-425 *
High school diploma/GED	4,699	1,030 ***	1,176 ***	-418 ***	-341 **
Tulare					
No high school diploma/GED	1,224	97	412 *	-64	-165
High school diploma/GED	1,010	47	699	143	-112
NEWWS programs					
Atlanta LFA					
No high school diploma/GED	1,454	457 **	465 *	-183 ***	-197 **
High school diploma/GED	2,379	514 **	579 **	-222 ***	-171 ***
Atlanta HCD					
No high school diploma/GED	1,488	225	336	-166 **	-134
High school diploma/GED	2,393	382 *	597 **	-181 ***	-176 ***
Grand Rapids LFA		††	††		
No high school diploma/GED	1,246	838 ***	1,009 ***	-824 ***	-701 ***
High school diploma/GED	1,766	139	-66	-620 ***	-468 ***

(continued)

Table 5.4 (continued)

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($) Years 1-3		Year 3		Impacts on Average Total AFDC Payments per Year ($) Years 1-3		Year 3	
Grand Rapids HCD									
No high school diploma/GED	1,204	400	*	725	**	-590	***	-640	***
High school diploma/GED	1,793	390		192		-413	***	-368	***
Riverside LFA									
No high school diploma/GED	2,398	636	***	519	**	-768	***	-717	***
High school diploma/GED	4,328	547	***	323		-609	***	-538	***
Riverside HCD									
No high school diploma/GED	2,423	336	**	497	**	-673	***	-772	***
High school diploma/GED	712	300		472		-462	*	-463	
Columbus Integrated			††		†		††		††
No high school diploma/GED	1,951	808	***	904	***	-486	***	-535	***
High school diploma/GED	2,721	41		167		-272	***	-270	***
Columbus Traditional									
No high school diploma/GED	1,967	292		271		-324	***	-349	***
High school diploma/GED	2,762	294		293		-262	***	-254	***
Detroit									
No high school diploma/GED	1,897	368	**	652	**	-78		-208	
High school diploma/GED	2,562	394	*	572	*	-265	***	-344	***
Oklahoma City					†				
No high school diploma/GED	2,569	149		204		-34		-22	
High school diploma/GED	3,292	-60		-195		-129	**	-92	
Portland							††		
No high school diploma/GED	1,839	767	***	1,125	***	-480	***	-617	***
High school diploma/GED	3,708	1,202	***	1,417	***	-792	***	-871	***
FTP			†		††				
No high school diploma/GED	1,076	246		359		-232	***	-359	***
High school diploma/GED	1,739	833	***	1,202	***	-125	**	-280	***
MFIP									
No high school diploma/GED	441	749	**	856		690	*	516	
High school diploma/GED	922	546		360		509	**	460	*

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Background Information Forms (BIFs).

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

programs in Alameda, Riverside, San Diego significantly increased earnings for graduates by more than $1,000 per year. All of these programs used a mix of first activities, enrolling in basic education many of those in need of it, but requiring others to look for work. This provides a second piece of evidence that programs with a mix of first activities helped a broad range of groups.

It is interesting that programs with a mix of first activities did better than education-focused programs for high school nongraduates though both emphasized basic education for this group. Likewise, it is interesting that programs with a mix of first activities generally did better than job search programs for high school graduates though both emphasized job search for the less disadvantaged people. The broad success of the mixed programs may indicate that determining whether individuals need basic education is more difficult than determining whether they have graduated from high school or worked recently. Among high school nongraduates, some individuals will be ready to work. Programs with a mix of first activities may have been more likely than programs with an education focus to require these people to look for a job. In most programs with a mix of first activities, basic literacy and language skills were also used to assess job-readiness. In groups of high school graduates, likewise, there will be some individuals whose lack of language skills, for example, make it difficult for them to find work. Programs with a mix of first activities were more likely than programs that emphasize job search to enroll these people directly in basic education designed to improve their skills. In other words, programs with a mix of first activities may have been more effective at increasing earnings because they used more complex methods to determine who would benefit from job search and who would benefit from basic education.

One program — MFIP — increased welfare payments to both high school graduates and nongraduates. As discussed above, these increases are due to Minnesota's relatively high welfare benefit levels and the program's enhanced earnings disregard. Although all other programs reduced welfare payments or left them unchanged, reductions in welfare payments were far smaller than earnings gains in a number of programs, including FTP and the GAIN programs in Alameda, Butte, Riverside, and San Diego. As a result, combined income from earnings, AFDC, and Food Stamps increased for high school graduates in each of these programs (results for income not shown in Table 5.4). Only MFIP increased combined income for high school nongraduates, however.

C. Welfare History

As discussed in Chapter 4, prior welfare receipt was the best predictor of future welfare receipt. Families who had been on welfare a substantial amount of time were the most likely to continue receiving welfare after random assignment. Thus, it may be crucial for welfare-to-work programs to assist these families if they are to avoid the TANF time limits.

The results in Table 5.5 are generally positive. Fourteen of the 20 programs produced significant increased earnings for long-term recipients (those who had ever been on welfare for two years or more prior to random assignment), the group that the Family Support Act was designed to help. By comparison, of the 17 programs that enrolled short-term recipients, only six significantly increased earnings for this group.

Despite the sometimes modest effects on earnings, long-term recipients often lost substantial amounts of welfare benefits. In 17 of the 20 programs, AFDC benefits were reduced for long-term recipients. MFIP was the exception in this regard. For long-term recipients, welfare payments increased by more than $700 per year over the three-year follow-up period. In addition, for long-term

National Evaluation of Welfare-to-Work Strategies

Table 5.5

Impacts on Earnings and AFDC Payments Across 20 Welfare-to-Work Programs, by AFDC Status

		Impacts on Average Total Earnings per Year ($)		Impacts on Average Total AFDC Payments per Year ($)	
Program and Subgroup	Sample Size	Years 1-3	Year 3	Years 1-3	Year 3
SWIM[a]					
Long-term recipient	2,202	574 ***	531 *	-770 ***	-706 ***
Short-term recipient	648	607	677	-666 **	-710 **
New applicant	360	1,081	1,428	-103	-88
GAIN evaluation programs[a]					
Alameda					
Long-term recipient	1,205	545 *	839 **	-202	-316
Short-term recipient	n/a				
New applicant	n/a				
Butte					
Long-term recipient	558	1,445 ***	1,699 ***	70	262
Short-term recipient	285	885	1,060	-50	-202
New applicant	386	179	369	-344	-451
Los Angeles					
Long-term recipient	4,396	96	165	-354 ***	-277 **
Short-term recipient	n/a				
New applicant	n/a				
Riverside					
Long-term recipient	2,661	1,296 ***	1,151 ***	-780 ***	-590 ***
Short-term recipient	1,979	1,409 ***	1,235 ***	-816 ***	-747 ***
New applicant	868	1,606 ***	2,000 ***	-544 *	-541 *
San Diego					
Long-term recipient	3,948	356	381	-494 ***	-393 *
Short-term recipient	3,079	1,022 ***	1,222 ***	-313	-292
New applicant	1,192	715	961	-685 **	-683 **
Tulare					
Long-term recipient	1,397	157	678 *	20	-121
Short-term recipient	691	269	722	291	103
New applicant	n/a				

(continued)

Table 5.5 (continued)

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($): Years 1-3	Impacts on Average Total Earnings per Year ($): Year 3	Impacts on Average Total AFDC Payments per Year ($): Years 1-3	Impacts on Average Total AFDC Payments per Year ($): Year 3
NEWWS programs					
Atlanta LFA				†	†
Long-term recipient	2,495	586 ***	596 ***	-245 ***	-240 ***
Short-term recipient	1,288	233	338	-97	-39
New applicant	n/a				
Atlanta HCD					
Long-term recipient	2,543	367 **	499 **	-194 ***	-190 ***
Short-term recipient	1,275	180	402	-120 *	-85
New applicant	n/a				
Grand Rapids LFA					††
Long-term recipient	1,791	402 **	341	-769 ***	-697 ***
Short-term recipient	1,219	408	347	-576 ***	-333 **
New applicant	n/a				
Grand Rapids HCD				†	
Long-term recipient	1,775	284	266	-588 ***	-591 ***
Short-term recipient	1,215	572 *	625	-326 ***	-309 **
New applicant	n/a				
Riverside LFA			†	†	††
Long-term recipient	3,510	742 ***	679 ***	-809 ***	-800 ***
Short-term recipient	3,101	428 **	111	-531 ***	-422 ***
New applicant	n/a				
Riverside HCD					
Long-term recipient	1,841	481 ***	700 ***	-583 ***	-641 ***
Short-term recipient	1,238	91	198	-678 ***	-824 ***
New applicant	n/a				
Columbus Integrated					
Long-term recipient	3,392	513 ***	582 **	-382 ***	-427 ***
Short-term recipient	806	-404	-185	-319 ***	-290 ***
New applicant	448	541	748	-272 **	-186
Columbus Traditional					
Long-term recipient	3,415	350 **	361	-273 ***	-308 ***
Short-term recipient	793	67	277	-247 **	-187 *
New applicant	497	105	-308	-336 ***	-252 **

(continued)

Table 5.5 (continued)

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($)		Impacts on Average Total AFDC Payments per Year ($)	
		Years 1-3	Year 3	Years 1-3	Year 3
Detroit					
Long-term recipient	3,313	329 **	464 *	-237 ***	-343 ***
Short-term recipient	1,015	554	1,067 **	-34	-141
New applicant	n/a				
Oklahoma City				†	††
Long-term recipient	1,419	177	246	-122	-145
Short-term recipient	1,858	5	-146	-181 **	-159 **
New applicant	2,530	-53	-111	12	78
Portland					††
Long-term recipient	3,423	1,222 ***	1,619 ***	-766 ***	-941 ***
Short-term recipient	1,999	1,012 ***	1,096 ***	-634 ***	-596 ***
New applicant	n/a				
FTP					††
Long-term recipient	1,444	420 **	727 ***	-146 **	-376 ***
Short-term recipient	956	840 ***	1,198 ***	-103 *	-169 ***
New applicant	334	-414	-645	-87	-158 **
MFIP					
Long-term recipient	1,242	365	111	704 ***	647 ***
Short-term recipient	n/a				
New applicant	n/a				

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Background Information Forms (BIFs).

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

N/a = not applicable because sample sizes were too small to reliably calculate impacts.

[a]For SWIM and GAIN, short-term recipients are those who said they had received welfare for two years or less rather than less than two years, and long-term recipients are those who said they had received welfare for more than two years rather than two years or more.

recipients, reductions in welfare payments were substantially less than increases in earnings in three other programs — the GAIN programs in Butte and Riverside and the LFA program in Atlanta.

Because most programs increased earnings and reduced welfare payments, long-term recipients fared no worse economically overall than did the other groups; 13 of the 20 programs produced neither significantly higher nor significantly lower income amounts for long-term recipients, four programs produced significantly higher income amounts for this subgroup, and two programs produced significantly lower amounts (results not shown in Table 5.5).

Among individual programs, three increased annual earnings for long-term recipients by more than $1,000 over the three-year period: Portland and the GAIN programs in Butte and Riverside. Three programs likewise significantly increased annual earnings by more than $1,000 per year over the three-year period for short-term recipients: Portland and the GAIN programs in San Diego and Riverside. All of these programs used a mix of first activities, allowing many of those who needed to improve their skills to enroll in basic education.

IV. Summary and Implications

Several interesting patterns emerged in this chapter.

- **Welfare-to-work programs increased earnings for the both more and the less disadvantaged groups.**

In results pooled across the 20 programs, impacts on earnings were significantly positive for all subgroups. In addition, it was not clear whether the programs increased earnings more for the more disadvantaged groups or for the less disadvantaged groups. Earnings impacts were larger for new applicants than for long-term recipients and were larger for high school graduates than for nongraduates. They were larger for families with three children or more than for families with one child and as large for individuals with no recent work experience as for those with substantial recent work experience. Among the 20 programs individually, half or more increased earnings for high school graduates and high school nongraduates; most increased earnings for long-term recipients; and most increased earnings for individuals who had not worked in the year prior to random assignment.

- **Welfare-to-work programs neither increased nor decreased income for most groups.**

As discussed earlier, many welfare recipients in most of these programs had their welfare benefits reduced dollar for dollar with earnings. As a result, the programs did not significantly increase combined income from earnings, AFDC, and Food Stamps for many subgroups. Only for new applicants did income increase significantly. MFIP was the most consistent exception to this finding. While welfare payments decreased in the other programs, they generally increased because of MFIP's enhanced earnings disregard. However, combined income also increased for a number of subgroups in the GAIN programs in Butte and Riverside.

- **Programs with a mix of first activities produced the largest earnings gains across a range of subgroups. Portland and Riverside GAIN, both employment-**

focused programs in this category, were particularly effective for many subgroups.

A handful of programs increased earnings by more than $1,000 per year for at least one subgroup. In all cases except two, the programs used a mix of first activities in which individuals who were considered job-ready were required to look for work but individuals in need of basic education were allowed to build skills, either prior to or instead of being required to look for work. Among these programs, the most consistently effective were Portland and Riverside GAIN, two programs that used a mix of first activities but tried to send a message to participants that they were expected to find a job. Riverside GAIN produced some of the largest earnings gains for both high school graduates and nongraduates, long-term and short-term recipients, and individuals with no recent work experience as well as those with substantial recent work experience.

Chapter 6

Impacts for Subgroups Defined by More Than One Characteristic

According to the results presented in Chapter 5, the 20 welfare-to-work programs studied in this report generally increased the earnings of the more disadvantaged groups. More than half of the programs increased earnings for long-term recipients. About half of them increased earnings for high school nongraduates, and about half of them increased earnings for those with no recent work experience.

These results may seem inconsistent with results from prior studies. In particular, Friedlander (1988), in studying a group of very different programs from the early 1980s, found that those programs had little effect on the more disadvantaged groups. Such a finding is sometimes called an “inverted U” of program impacts, in which impacts on earnings are larger for moderately disadvantaged subgroups than for less disadvantaged or more disadvantaged groups. In particular, Friedlander instead found the largest effects on earnings of welfare re-applicants, a group similar to the short-term recipients examined in the current report. For people already receiving welfare at the time of random assignment (welfare recipients) and for people applying for welfare for the first time at random assignment (new welfare applicants), Friedlander found that the earlier programs had little effect on earnings. More generally, he found the largest earnings impacts for moderately disadvantaged sample members, but small or no impacts for either the least disadvantaged or the most disadvantaged. At the same time, the welfare-to-work programs that Friedlander studied typically showed the largest welfare savings for the most disadvantaged group, such as welfare recipients. Together, these results implied that the earlier programs were having little positive effect on the work of the most disadvantaged groups, but may have been decreasing their economic well-being, perhaps through sanctions that might have accompanied noncompliance with the programs’ requirements.

This chapter investigates the presence of an inverted U for the newer programs by defining subgroups similar to Friedlander’s. Because earnings levels appear to depend so much on education, work experience, and welfare experience, the groups are defined by three characteristics: whether they had a high school diploma or GED, whether they had received welfare for two years or more prior to random assignment (that is, were long-term recipients), and whether they had worked in the year prior to random assignment. Three groups were defined by the number of barriers to work that they had. The “most disadvantaged” group consists of long-term recipients who had not graduated from high school or worked in the year prior to random assignment. The “least disadvantaged” group consists of people who had none of these barriers. The “moderately disadvantaged” group consists of people who had one or two barriers. The results indicate that the welfare-to-work programs were more often successful for the moderately disadvantaged than for the most disadvantaged. Fewer than half of the programs increased earnings significantly for the most disadvantaged group, while most programs increased earnings for the moderately disadvantaged. The results are consistent with Friendlander’s inverted U; since half the programs increased earnings for the most disadvantaged group, however, it appears that the inverted U is somewhat flatter than Friedlander’s.

Kinds of barriers may matter more than numbers of barriers in determining the effectiveness of welfare-to-work programs. Since JOBS was explicitly designed to benefit potential and actual long-term welfare recipients, for example, the programs may be as effective for long-term recipients as for others. As part of this effort, a number of programs allowed people with few skills to enroll in basic education; lacking a high school diploma may therefore also not be as serious a barrier to success under these welfare-to-work programs as some other barriers. Section II of the chapter attempts to answer the number-versus-type question by dividing the moderately disadvantaged group into six smaller subgroups: three groups that have only one of the three barriers and three groups that have two of the three barriers. The results imply that *which* barrier may be more important than *how many* barriers. Earnings gains were larger for people who lacked recent work experience than for those who had worked recently, but they were smaller for high school nongraduates than for graduates, and similar for long-term and short-term recipients.

Various circumstances and characteristics contribute to the likelihood of staying on welfare. Long-term recipients are more likely to have physical or emotional limitations than other sample members, for example, or to have many children. It is therefore not clear whether comparisons of long-term recipients and applicants reveal the impact of having been on welfare for a substantial length of time or the effects of the many other differences between the two groups. Section III investigates "conditional impacts" of welfare-to-work programs on subgroups while holding constant a variety of other demographic characteristics. The impacts presented in Chapter 5 indicated that welfare-to-work programs increase earnings by twice as much for new applicants as for long-term recipients. According to the conditional impacts, however, welfare history was not a key factor in the success of welfare-to-work programs, but merely an indicator of other important characteristics. After adjusting for differences in demographics and work experience between groups based on welfare history, the programs increased earnings slightly more for long-term recipients than for short-term recipients and increased income more for long- and short-term recipients than for new applicants across the three years of follow-up. In most other respects, however, the conditional impacts concurred with the unconditional impacts of Chapter 5. Both sets of results indicate that earnings impacts were larger for individuals with moderate earnings than for individuals with high earnings, larger for high school graduates than for nongraduates, and larger for large families than for small families.

I. Impacts for Groups Defined by Number of Barriers

A. Pooled Impacts

Table 6.1 presents results, both pooled and by program, for three subgroups based on number of barriers to work. The most disadvantaged group had all three barriers described above: they had not graduated from high school when they were randomly assigned, they had been on welfare for two years or more at some point prior to random assignment, and they did not work in the year prior to random assignment. The least disadvantaged group had none of the barriers. The moderately disadvantaged had one or two of the barriers.

The first panel of Table 6.1, which shows results pooled across the programs, is broadly consistent with Friedlander's inverted U. Over the three-year follow-up period, the programs increased earnings by about $600 per year for the moderately disadvantaged, but only about $400

National Evaluation of Welfare-to-Work Strategies

Table 6.1

Impacts on Earnings and AFDC Payments Pooled and by Program Across 20 Welfare-to-Work Programs, by Composite Subgroup

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($): Years 1-3	Impacts on Average Total Earnings per Year ($): Year 3	Impacts on Average Total AFDC Payments per Year ($): Years 1-3	Impacts on Average Total AFDC Payments per Year ($): Year 3
Full sample		††			††
Most disadvantaged	14,393	404 ***	521 ***	-411 ***	-400 ***
Moderately disadvantaged	47,113	599 ***	704 ***	-414 ***	-427 ***
Least disadvantaged	10,019	421 ***	448 **	-282 ***	-215 ***
SWIM					
Most disadvantaged	745	469 **	448	-535 *	-497
Moderately disadvantaged	2,113	548 **	571 *	-817 ***	-798 ***
Least disadvantaged	352	1,504 *	1,600	-182	-84
GAIN evaluation programs					
Alameda					
Most disadvantaged	366	45	238	-503	-614
Moderately disadvantaged	839	789 *	1,134 **	-85	-202
Least disadvantaged	n/a				
Butte					
Most disadvantaged	n/a				
Moderately disadvantaged	807	883 *	974	-188	-273
Least disadvantaged	243	1,593	2,302 *	5	74
Los Angeles					
Most disadvantaged	2,322	84	119	-297 *	-165
Moderately disadvantaged	2,074	159	281	-429 ***	-408 **
Least disadvantaged	n/a				
Riverside					
Most disadvantaged	974	1,026 ***	1,014 ***	-354	-239
Moderately disadvantaged	3,739	1,419 ***	1,316 ***	-908 ***	-776 ***
Least disadvantaged	795	1,976 ***	2,095 **	-432	-474
San Diego		††			
Most disadvantaged	1,331	84	84	-493	-445
Moderately disadvantaged	5,405	606 **	841 ***	-450 ***	-396 **
Least disadvantaged	1,483	1,549 ***	1,461 **	-381	-313
Tulare					
Most disadvantaged	554	108	425 *	-167	-377
Moderately disadvantaged	1,423	140	563	-15	-105
Least disadvantaged	257	-120	968	515	23

(continued)

Table 6.1 (continued)

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($)		Impacts on Average Total AFDC Payments per Year ($)	
		Years 1-3	Year 3	Years 1-3	Year 3
NEWWS programs					
Atlanta LFA		†			
Most disadvantaged	828	269	357	-233 ***	-260 **
Moderately disadvantaged	2,401	705 ***	751 ***	-204 ***	-179 ***
Least disadvantaged	564	-323	-394	-109	-24
Atlanta HCD					
Most disadvantaged	860	40	80	-135	-114
Moderately disadvantaged	2,408	491 ***	711 ***	-204 ***	-193 ***
Least disadvantaged	562	-139	-36	-62	-38
Grand Rapids LFA		††	†††		††
Most disadvantaged	456	1,035 ***	1,634 ***	-966 ***	-964 ***
Moderately disadvantaged	2,124	293	95	-702 ***	-568 ***
Least disadvantaged	432	451	413	-444 ***	-110
Grand Rapids HCD				†	††
Most disadvantaged	450	546 ***	840 ***	-775 ***	-786 ***
Moderately disadvantaged	2,077	291	243	-479 ***	-498 ***
Least disadvantaged	466	603	583	-170	-31
Riverside LFA		†	†		
Most disadvantaged	1,084	668 ***	542 **	-1,059 ***	-951 ***
Moderately disadvantaged	4,374	794 ***	673 ***	-605 ***	-585 ***
Least disadvantaged	1,221	-203	-658	-566 ***	-400 **
Riverside HCD					
Most disadvantaged	1,094	438 ***	674 ***	-768 ***	-799 ***
Moderately disadvantaged	1,865	314	441 *	-526 ***	-642 ***
Least disadvantaged	n/a				
Columbus Integrated		††	†		
Most disadvantaged	899	400 **	498	-449 ***	-524 ***
Moderately disadvantaged	3,134	525 ***	655 **	-332 ***	-352 ***
Least disadvantaged	613	-997 *	-1,035	-279 ***	-221 *
Columbus Traditional					
Most disadvantaged	888	159	79	-347 ***	-373 ***
Moderately disadvantaged	3,222	355 *	423 *	-233 ***	-261 ***
Least disadvantaged	595	-242	-527	-381 ***	-250 **
Detroit					
Most disadvantaged	1,097	266	574 *	-127	-261
Moderately disadvantaged	3,033	436 **	605 **	-197 **	-301 ***
Least disadvantaged	324	30	332	-308	-237

(continued)

Table 6.1 (continued)

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($) Years 1-3	Year 3	Impacts on Average Total AFDC Payments per Year ($) Years 1-3	Year 3
Oklahoma City			†		
Most disadvantaged	291	308	495	-101	12
Moderately disadvantaged	4,057	27	84	-57	-41
Least disadvantaged	1,459	-84	-514 *	-136 *	-89
Portland		†††	†††	†††	†††
Most disadvantaged	880	701 ***	974 ***	-501 ***	-527 **
Moderately disadvantaged	3,800	1,390 ***	1,779 ***	-817 ***	-999 ***
Least disadvantaged	805	-55	-298	-285 **	-36
FTP					††
Most disadvantaged	436	295	399	-312 **	-465 ***
Moderately disadvantaged	1,783	466 **	731 ***	-142 **	-302 ***
Least disadvantaged	515	698	1,043	-36	-129 **
MFIP					
Most disadvantaged	263	1,115 ***	1,303 **	444	233
Moderately disadvantaged	1,051	442	162	652 ***	595 **
Least disadvantaged	n/a				

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Background Information Forms (BIFs).

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

N/a = not applicable because sample sizes were too small to reliably calculate impacts.

National Evaluation of Welfare-to-Work Strategies

Table 8.1

Impacts on Earnings in the LFA and HCD Programs for the Full Sample

Program	Sample Size	Impacts on Average Total Earnings ($) Year 1	Year 2	Year 3
Atlanta				
LFA	3,833	403 ***	535 ***	532 ***
HCD	3,881	65	369 **	472 **
Difference		338 **	166	60
Grand Rapids				
LFA	3,012	451 ***	425 **	350
HCD	2,997	200	622 ***	444 *
Difference		251 *	-197	-94
Riverside[a]				
LFA	3,125	780 ***	506 ***	488 **
HCD	3,135	272 **	213	485 **
Difference		508 ***	294 *	3
Pooled				
LFA	9,970	536 ***	493 ***	463 ***
HCD	10,013	170 **	396 ***	468 ***
Difference		365 ***	97	-4

SOURCE: MDRC calculations from unemployment insurance (UI) earnings records.

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

[a]In order to make the LFA and HCD groups comparable, impact estimates were compared only for those sample members lacking a high school diploma or basic skills at random assignment.

the first three years after random assignment. Results for the first year concur with the conclusion of Hamilton et al. (1997). In Atlanta, the LFA program increased earnings in year 1 by $403 per sample member, while the HCD program did not significantly increase earnings. The difference that the LFA approach made over the HCD approach is given by the difference between the impacts of the two programs: $338.[3] In Riverside and Grand Rapids, the LFA program also had significantly larger earnings impacts than the HCD program in the first year, by $251 per person in Grand Rapids and $508 in Riverside. Averaging across the three sites, the LFA programs increased earnings by $536 per person in year 1 compared with $170 for the HCD programs, a difference of $366.

The result in year 1 is not surprising. Almost by definition, employment-focused programs should increase earnings faster than education-focused programs, since few people would both go to school *and* earn as much as they could. By year 2, however, the difference in earnings impacts of the two approaches had narrowed in all three sites and was no longer statistically significant overall. The narrowing occurred not because the LFA programs became less effective — their overall earnings impact declined only from $536 in year 1 to $493 in year 2 — but because the earnings impact of the HCD programs grew, from $170 in year 1 to $396 in year 2. Even in year 2, however, in Riverside LFA program impacts were $294 higher than HCD program impacts.

By year 3, the difference between the two approaches narrowed even further; not only was the difference in impacts between the two approaches overall not statistically significant, but in each site the difference was less than $100 per person. As in year 2, the narrowing occurred primarily because the impact on earnings of the HCD program grew, particularly in Grand Rapids, and because the impact of the LFA program stayed about the same.

C. Differences in Impacts on Earnings by Subgroup

In the 1997 report, Hamilton et al. also concluded, “HCD *earnings* impacts for most subgroups had not caught up with those of the LFA programs by the end of the two-year follow-up period, but HCD *employment* impacts for some subgroups had surpassed LFA impacts as of this point.” For a number of key subgroups, Table 8.2 compares pooled impacts on earnings in year 1, year 3, and the average of all three years for the two self-sufficiency approaches. Subgroups examined in the table are defined by high school credential, level of disadvantage, risk of depression, and earnings in the past 12 months.

Overall, Tables 8.1 and 8.2 tell a similar story. In the first year after random assignment, the LFA programs increased earnings significantly more than the HCD program for all subgroups examined except one (those who earned $5,000 or more in the year prior to random assignment). The most striking difference is for those at high risk of depression, a group for whom earnings did not significantly increase overall according to the results of Chapter 3. While the HCD programs had a negative effect on their earnings, the LFA programs increased them by $675 per person in the first year; thus, the LFA programs increased earnings by nearly $900 more than the HCD programs.

[3]These impacts will differ from impacts shown in prior NEWWS reports because the regression-adjustment used in this report includes fewer covariates.

National Evaluation of Welfare-to-Work Strategies

Table 8.2

Impacts on Earnings in the LFA and HCD Programs for Selected Subgroups

	Year 1 ($)			Year 3 ($)			Years 1-3 ($)		
Program and Subgroup	LFA	HCD	Difference	LFA	HCD	Difference	LFA	HCD	Difference
By high school credential ($)									
No high school diploma or GED	658	160	498 ***	625	504	121	636	319	317 ***
High school diploma or GED	415	183	232 *	314	431	-116	366	373	-7
By level of disadvantage ($)									
Most disadvantaged	500	120	380 ***	690	493	198	602	316	286 **
Moderately disadvantaged	624	161	463 ***	449	478	-29	540	374	167
Least disadvantaged	n/a	n/a		n/a	n/a		n/a	n/a	
By risk of depression ($)									
High risk	675	-193	869 ***	175	-369	544	417	-201	618 **
Moderate risk	476	110	366 *	553	667	-114	462	363	99
Low risk	547	271	276 **	499	730	-231	540	536	4
By earnings in past 12 months ($)									
No earnings	522	158	364 ***	566	522	44	535	377	157 *
Less than $5,000	540	36	504 ***	445	264	182	468	166	302 *
$5,000 or more	525	272	253	-135	341	-477	284	289	-6

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, Private Opinion Survey (POS) data, and Background Information Forms (BIFs).

NOTES: A two-tailed t-test was applied to differences between outcomes for the two programs. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

To make the LFA and HCD groups comparable, only sample members lacking a high school diploma or basic skills at random assignment were included for Riverside.

N/a = not applicable because sample sizes were too small to reliably calculate impacts.

By year 3, there were no statistically significant differences in earnings impacts for the two approaches. The largest difference in impacts was again for those at high risk of depression, but that difference was not statistically significant. Again, this result is not very surprising since three years is enough time for most sample members to have completed their basic education courses.

Perhaps the most interesting results, however, are in the columns that combine earnings impacts for all three years. For the less disadvantaged sample members, the significant differences in year 1 were diluted by the smaller (and in some cases negative) impacts in later years, so that the difference between the two approaches was not statistically significant for any of the less disadvantaged groups. For people with a high school diploma, the LFA programs increased earnings by $366 per year compared with $373 per year for the HCD programs. For those at low risk of depression, the average impacts were $540 for the LFA programs and $536 for the HCD programs. For those who earned $5,000 or more in the year prior to random assignment, the earnings impacts were $284 and $289.

For the more disadvantaged sample members, in contrast, the LFA programs outperformed the HCD programs in every case. The largest difference remains for people at high risk of depression. The LFA programs increased earnings by $417 while the HCD programs reduced earnings by $201, a difference of $618 per person. For high school nongraduates and the most disadvantaged group, however, the differences are substantial. In both cases, the earnings impacts of the LFA programs were about double the impacts of the HCD programs. This is an especially important finding in light of the notion that education-focused programs are explicitly designed to improve the earnings ability and job prospects of people who lack basic skills, most of whom are likely to be in the more disadvantaged groups.

II. Relating Subgroup Impacts to Program Characteristics

Although the LFA-HCD comparison provides the most rigorous means of comparing program models, it has two crucial shortcomings. First, only six programs can be compared, and it is difficult to find precise patterns with so few programs. Second, while the comparison provided a rigorous means of comparing program models, it did not indicate whether other characteristics of sites might help explain variation in subgroup impacts. This section addresses both issues somewhat, looking at all 20 programs and incorporating information on local economic conditions and the generosity of the welfare system.

The final report of the GAIN evaluation (Riccio, Friedlander, and Freedman, 1994) provides one of the most comprehensive attempts to explain variation in impacts across programs. The final chapter investigated more than a dozen characteristics of the sites and their programs in an attempt to explain the substantial variation from site to site, particularly the large and consistent impacts in the Riverside program. In a thorough investigation of local economic conditions (local unemployment rate, growth in number of jobs), program participation (job search versus education and training, both overall and by subgroups related to need for basic education), characteristics of the caseload (through conditional subgroup impacts), and sanctioning, the report found no consistent relationship among these factors and program impacts. The authors of the report concluded that the large impacts might have been due in large part to the staff's near-unanimous emphasis on quick job entry, an emphasis not found in the other five GAIN sites.

A. The Factors Examined

This section relates the impacts on earnings by subgroup to four measures: local unemployment rate, AFDC grant levels, impact on participation in job search, and impact on participation in education and training.

1. Local Unemployment Rate. While the local unemployment rate serves as an indicator of general economic conditions at the time of a program's inception, it is not clear how program impacts would be affected by these conditions. Weak economic conditions imply that few people will be able to find work and that jobs will pay little. At the same time, a weak economy will result in a less disadvantaged caseload if it brings people onto the rolls who will leave welfare quickly. Both factors are true for both the control and program groups, however; impacts may therefore be either higher or lower when the economy is in bad shape.

It is also not clear how economic conditions would affect the impacts for particular subgroups. In a strong economy, the most job-ready welfare recipients might be able to find a job on their own, implying that impacts on employment or earnings would be small for the least disadvantaged, for high school graduates, and for people with considerable recent work experience. On the other hand, employment levels could be improved even for the less disadvantaged groups, and welfare-to-work programs could be a means of making that improvement. In a weak economy, the least skilled are typically the first to lose their job and have the hardest time finding a new job. Thus, we would expect the most disadvantaged members of the control group to struggle when the unemployment rate is high. For this subgroup, extra help with the job application process or with building skills through education and training might give the program group just the needed edge to get a job, implying that program impacts for the most disadvantaged might be high when the unemployment rate is high. On the other hand, welfare-to-work programs might not be able to help the more disadvantaged groups if the economy has few jobs to offer, implying that impacts would be low for them. In general, therefore, one cannot predict how impacts would vary with economic conditions.

2. AFDC Grant Levels. How do AFDC grant levels relate to program impacts? Consider the two states at opposite extremes among these programs: California and Georgia. In Riverside, the AFDC guarantee level was close to $700 per month for a single mother with two children, among the highest levels in the country. In Atlanta, in contrast, the AFDC guarantee was less than $300 per month at the beginning of the evaluation of the two Atlanta programs, implying that a person receiving welfare in Atlanta must have few other prospects for economic support. It is likely that sample members in Atlanta are more disadvantaged than sample members in Riverside because of the larger proportion of the most disadvantaged in Atlanta (19.7 percent) than in Riverside (12.3 percent) and that sample members in Atlanta will be more disadvantaged in ways we cannot observe. In other words, among the most disadvantaged, Atlanta sample members are probably even more disadvantaged in other ways than Riverside sample members. Thus, the relationship found between level of disadvantage and program impacts might be even stronger in Atlanta than in Riverside.

On the other hand, because welfare benefit levels are so low, sample members in low-grant states like Georgia are likely to be on welfare for only a short period of time. In low-grant states, almost any job will pay enough to make a person ineligible for welfare benefits; in a high-grant state, it is easier to combine work and welfare. This suggests that programs will have a harder time reducing welfare use and, presumably, increasing employment and earnings in low-grant states than in high-grant states.

The AFDC grant might also be a proxy for other factors. States with high grants have revealed a greater willingness to fund the welfare system. These states might also have put more resources into their welfare-to-work programs. In this case, AFDC grant levels would be related to program impacts not directly, but as an indicator of the intensity or value added of employment and training services. AFDC grant levels might proxy the cost of living and, by extension, the prevailing wage rates in the statewide labor market. Thus, AFDC grant levels would be positively related to earnings not because of the welfare system, but because they represent variation in wage levels from place to place.

3. Impact on Participation in Job Search. A welfare-to-work program can produce results in a number of ways. It can get more people to participate in services that will lead to more employment and earnings. It can provide more effective services that increase the ability to find a job or to find a better job. It can provide motivation for people already capable of finding a job to leave welfare or to report employment and earnings to the welfare office.

This analysis uses impact on participation in job search to measure the effectiveness of programs in encouraging or helping participants engage more in job search. The impact is measured as the difference between the proportion of the program and control groups engaging in job search activities. Although the measure differs somewhat from study to study, it is typically taken from surveys in which the clients report their activities since random assignment, usually within about one to three years of random assignment.

For 19 studies for which this measure is available (it is not available for Butte), the impact on job search ranged from about 7 percentage points in Detroit and Oklahoma City to about 53 points in SWIM. Programs with less impact on job search tended to emphasize education. In addition to Detroit and Oklahoma City, three programs had impacts on job search of 10 percentage points or less: the two Columbus programs and FTP. Likewise, programs with more impact on job search tended to be employment-focused. In addition to SWIM, the sites with impacts on job search exceeding 25 percentage points include the three LFA programs, Riverside GAIN, Portland, MFIP, and Alameda GAIN.

As a measure of the focus of a program, impact on job search might not be ideal. For one thing, it is measured in different ways for different programs. In GAIN, for example, estimates of program participation come from the five-year period after random assignment. In MFIP, they come from a 12-month follow-up survey. In JOBS, information from a two-year client survey was supplemented with reviews of participants' case files.

Nevertheless, the impact on job search may be a better measure than some others. One alternative that is available for all programs is the proportion of the program group participating in a service. This measure shares the problems of the net participation measure, however. In addition, the difference between the program and control groups would seem more important in explaining program impacts than the level of participation of the program group. For most of the 20 programs being studied, the control group was not required to participate in any employment and training service. Thus, participation by the program group might have been an excellent indicator of the impact on activities. In FTP, however, the control group was, at least in theory, required to participate in Project Independence, Florida's JOBS program. In MFIP, voluntary services were available to the control group through STRIDE, Minnesota's JOBS program.

Another measure that is available for all programs indicates whether the program emphasized job search initially for most participants, encouraged long-term education, or had some mix of the two emphases. The problem with this measure, however, is that deciding where programs fit is somewhat arbitrary, but can affect results substantially. Portland and Riverside GAIN are the two most successful programs at increasing earnings. Riverside GAIN was, on paper, a program with a mix of first activities. For participants in need of basic education, the first assignment was supposed to be education, but participants could choose job search. Staff in Riverside emphasized getting a job more than staff in the other GAIN sites. As a result, because many in-need program participants engaged in job search, the program had a large impact on job search and gained a reputation as employment-focused. Yet it would seem a mistake to group Riverside GAIN, which did increase participation in education and training considerably for people in need of basic education, with the LFA programs in which most participants engaged initially in job search. Portland presents a similar problem. Like Riverside, Portland allowed participants in need of basic education to enroll in education and training services. While staff made it clear that the objective was to get a job, they stressed even to participants in job search that they should wait for a good job rather than take the first job available. Portland's program is classified as employment-focused, but it was not as extreme in its focus as the LFA programs.

Even in the education-focused programs, many program group members eventually searched for a job. Perhaps more important than participation over several years is initial participation in various activities. This measure was examined briefly, but was not available for a number of the programs being studied, and so was not looked at extensively.

If one program has higher participation in job search than another, it might indicate that it was more effective at increasing job search, placed more emphasis on job search, or had more job-ready participants. As a result, it might be important in looking at participation to distinguish participation by job-ready sample members from participation by other sample members. Once again, this measure was investigated briefly, but was not available for a number of sites, and therefore not examined extensively.

4. Impact on Participation in Education and Training. The programs' impacts on participation in education and training may likewise reflect the ability of programs to ultimately increase employment and earnings. For the 19 studies for which this measure is available (it is not available for Butte), the impact on education and training ranges from less than 0 (the control group was more likely to engage in education and training than the program group) to about 42 percentage points. Programs with less impact on education and training tend to be employment-focused. The five programs with impacts on education and training below 10 percentage points include the three LFA programs, MFIP, and Detroit. Likewise, programs with more impact on education and training tend to be more education-focused or have caseloads that require more basic education. The six programs with impacts on education and training exceeding 25 percentage points include the HCD programs in Atlanta and Riverside; the GAIN programs in Alameda and Los Angeles, both of which enrolled only long-term recipients; the GAIN program in Tulare, which had a relatively disadvantaged caseload; and FTP, which set a low standard for judging people in need of basic education.

Using program participation as the measure of the effectiveness of a program has an important drawback. On the one hand, it seems apparent that a program will have little effect if nobody is using its services. On the other hand, people are required to participate in the welfare-to-work programs only so long as they remain on welfare. In a site or program in which welfare recipients leave the roles relatively quickly, participation might be low because many sample members have left the welfare program before being required to participate. Such programs are likely to have small impacts on employment and earnings because sample members are well enough off that they can leave welfare quickly. It would therefore be a mistake to attribute the impacts of the program, or the lack of impacts, to an impact on program participation or its lack.

Likewise, a program that generates a large increase in program participation may be serving long-term welfare recipients. This may be especially true of an education-focused program that generates a relatively large increase in job search — for example, the Riverside HCD program. Recall that all sample members in the Riverside HCD program were considered in need of basic education, a very disadvantaged sample. In this case, a relatively large impact on participation might be associated with relatively low earnings gains because the sample is disproportionately composed of long-term or potentially long-term welfare recipients.

A similar argument can be made against the use of sanctioning as a measure of the mandatoriness of a program. A program that is completely mandatory might generate no sanctioning if participants willingly comply with the program's requirements. A different program, perhaps less mandatory, might generate a larger degree of sanctioning if it is serving a less compliant caseload. In these examples, the outcome is not an independent measure of a characteristic of a program, but an amalgamation of program implementation and kinds of participants served.

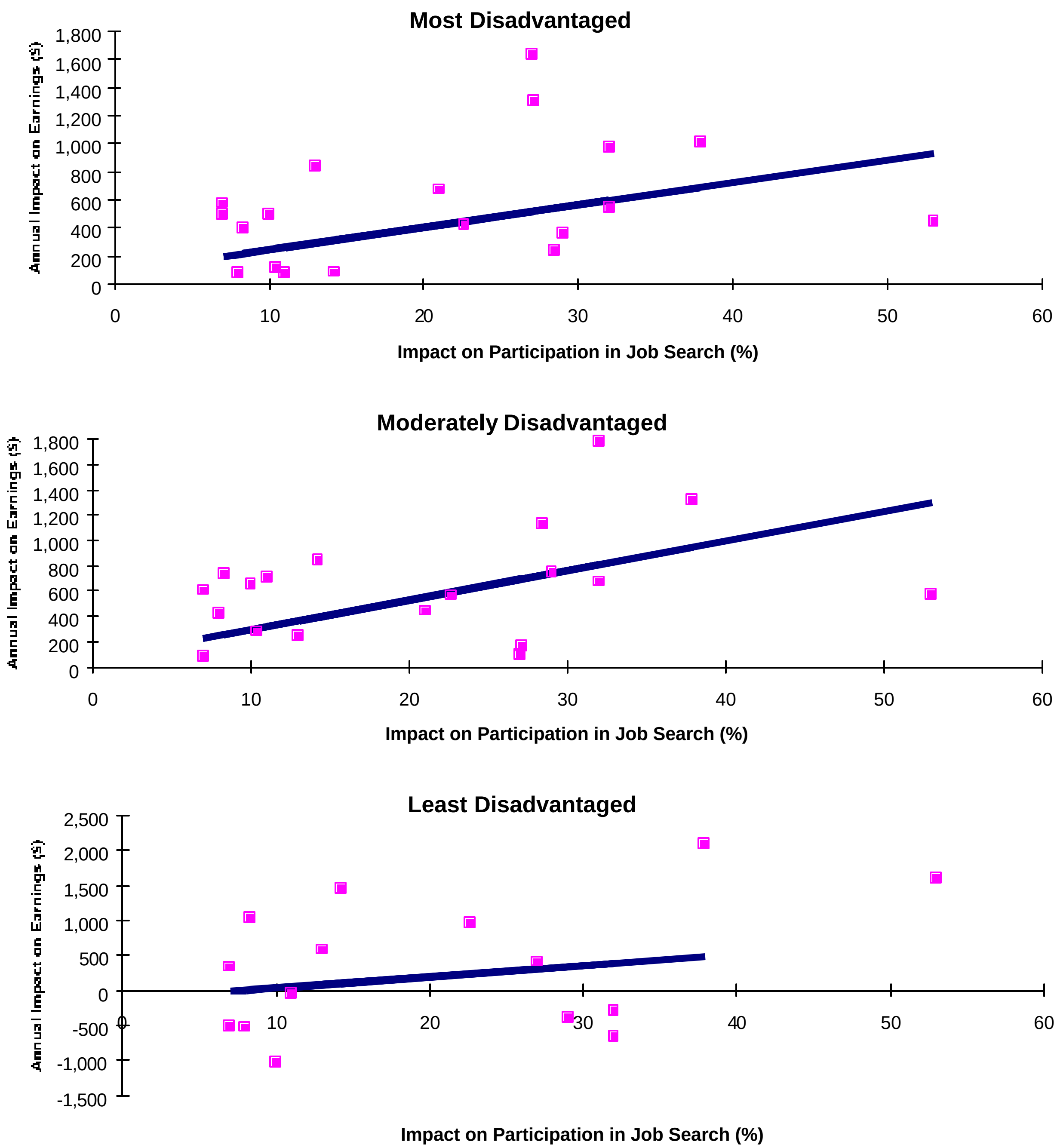
B. Impacts by Level of Disadvantage

1. Impacts on Earnings Compared with Impacts on Participation in Job Search. To provide a quick glimpse of the link, if any, between impacts on participation in services and impacts on earnings, Figure 8.1 compares program impacts on earnings over the three-year follow-up period with their impacts on participation in job search. In each panel, the horizontal axis represents the impact on participation in job search, typically in the two years after random assignment. Programs in which program group members searched for jobs at about the same rate as control group members are to the left of the panel. These tend to be programs with more of an education focus. For example, the sites with impacts on job search below 20 percentage points are, from lowest to highest, Oklahoma City, Detroit, Columbus Traditional, FTP, Columbus Integrated, Los Angeles GAIN, Atlanta HCD, Grand Rapids HCD, and San Diego GAIN. Programs to the right are likely to be more employment-focused. From right to left, the eight programs with the greatest increase in job search are Riverside GAIN, Riverside LFA, Portland, Atlanta LFA, Alameda GAIN, Grand Rapids LFA, Tulare GAIN, and Riverside HCD. Of these programs, only the Riverside HCD program was education-focused, while the other programs either strongly emphasized job search or provided a mix of first activities that required job search for the job-ready.

National Evaluation of Welfare-to-Work Strategies

Figure 8.1

Relationship Between Impact on Earnings in Years 1-3 and Participation in Job Search, by Level of Disadvantage



SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Background Information Forms (BIFs).

According to Figure 8.1, the impact on participation in job search does not appear, by itself, to provide a strong explanation for the impacts of these programs for any of the three subgroups. For each subgroup, the points are quite scattered, even within programs that increased job search by about the same amount. For example, the three programs with the greatest impact for the most disadvantaged — Portland, Riverside GAIN, and Grand Rapids LFA — are near the middle of the figure (upper panel), with increases in job search of about 30 percentage points. Yet three other programs with similar increases in job search had earnings impacts of less than $300 (Atlanta LFA, Grand Rapids LFA, and Alameda GAIN).

Nevertheless, increasing participation in job search does appear to be related to larger impacts on earnings for the most disadvantaged, seen from the solid line in the panel, which slopes somewhat upward. This line, which was chosen because it best fits the data points, indicates that earnings impacts for the most disadvantaged group increased by about $15 per year with each percentage point gain in impact on job search.[4] A typical program that increased job search by 10 percentage points would be projected to increase earnings by about $150 per year, and a program that increased job search by about 50 percentage points is projected by the line to increase earnings by about $750 per year.

A look at the corresponding panels for the moderately disadvantaged and the least disadvantaged reveals that job search helps those groups by about the same amount. For both panels, earnings impacts are also quite scattered. Nevertheless, earnings impacts increased by about $25 per year on average with each percentage point gain in the program's impact on job search for the moderately disadvantaged group and about $15 per year for the least disadvantaged group.

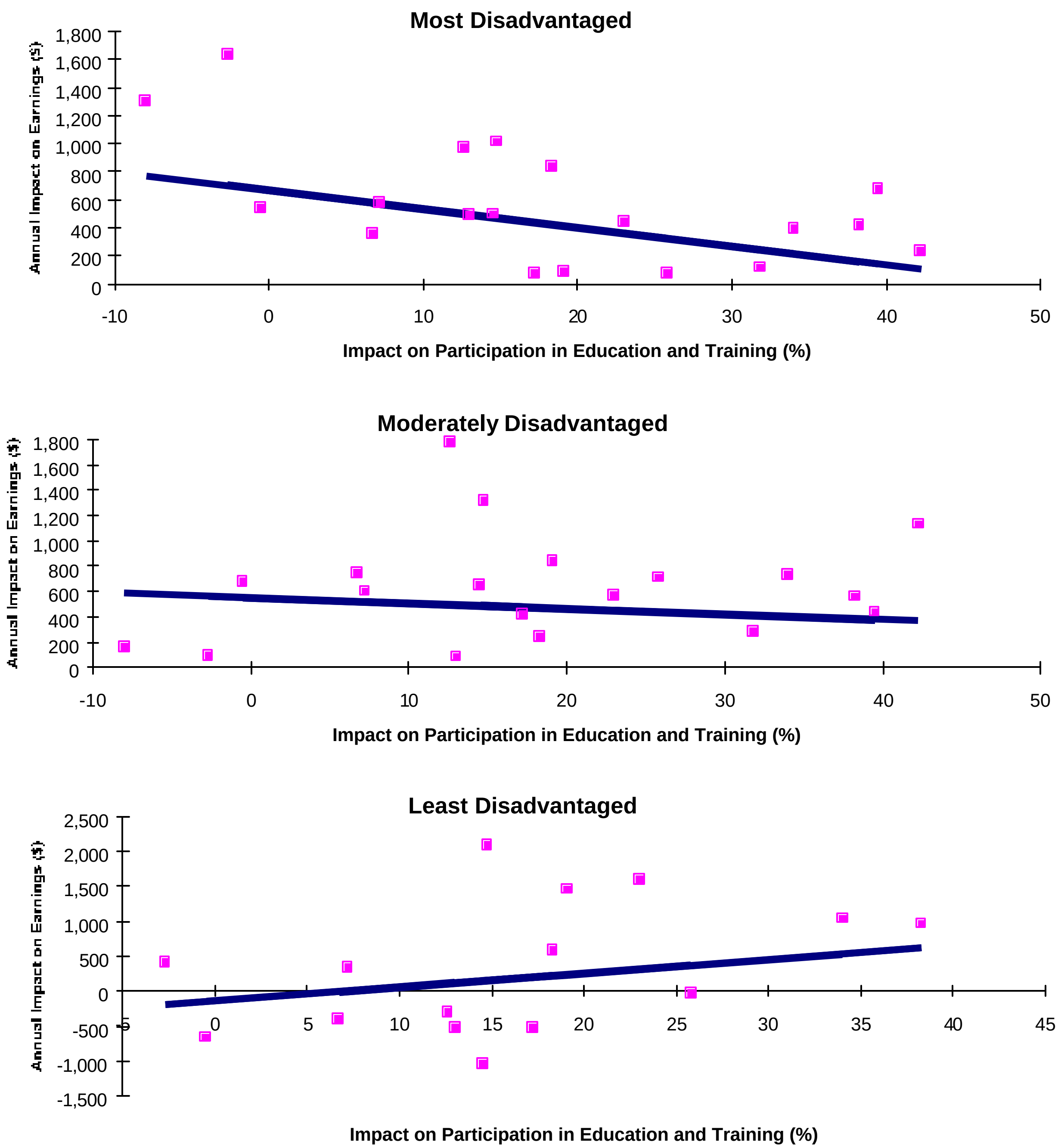
2. Impacts on Earnings Compared with Impacts on Participation in Education. Figure 8.2 compares impacts on earnings over the three-year follow-up period with impacts on participation in education and training programs. In each panel, the horizontal axis represents the impact on participation in education and training, so that programs that did not increase use of education and training very much are to the left of the panel. These tend to be programs with more of an employment focus. For example, the five sites with the impacts on education and training below 10 percentage points are, from lowest to highest, MFIP, Grand Rapids LFA, Riverside LFA, Atlanta LFA, and Detroit. (Recall that Detroit, an education-focused program, had low enforcement of its participation mandate.) Programs to the right are likely to be more education-focused. From right to left, the five programs with the greatest increase in participation in education programs are Alameda GAIN, Riverside HCD, Tulare GAIN, FTP, and Los Angeles GAIN. It is interesting that only one of these five programs was considered education-focused. The other four programs used a mix of first activities. However, Alameda and Los Angeles enrolled only long-term recipients, many of whom were in need of basic education. Likewise, in FTP most sample members were considered in need of basic education because of the criteria used to judge job-readiness; people were considered in need of basic education if they lacked a

[4]To generate the solid line in the figure, impacts by program were related to net participation in job search using weighted least squares in which each site was weighted by the inverse of the standard error of the impact on earnings so that sites with imprecisely measured impacts (typically because of small samples) would influence the line of best fit less than other sites.

National Evaluation of Welfare-to-Work Strategies

Figure 8.2

Relationship Between Impact on Earnings in Years 1-3 and Participation in Education and Training, by Level of Disadvantage



SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Background Information Forms (BIFs).

high school diploma or GED *or* they had not worked in the year prior to random assignment *or* they had failed one of the tests of basic skills.

Figure 8.2 seems to show that the link between earnings gains and participation in education are weaker than the link between earnings gains and participation in job search. For the most disadvantaged, programs with greater impacts on education and training had smaller impacts on earnings than other sites, and this relationship is about the same for impacts in the third year after random assignment (not shown). In fact, the link between participation in education and earnings impacts is about as strong as the link between participation in job search and earnings impacts. However, more job search is associated with larger earnings gains, while more education is associated with smaller earnings gains.

Somewhat paradoxically, it is the least disadvantaged who appear to gain the most from education and training. In fact, the link between earnings impacts and impacts on education are about the same as the link between earnings impacts and impacts on job search. Since the least disadvantaged, particularly because they already have a high school diploma, were not expected to benefit much from basic education, this result may imply that programs that more strongly enforce the mandate to participate generate larger impacts for job-ready people by giving them a reason to leave welfare.

C. <u>Multivariate Regressions: Relating Impacts to Several Factors at Once</u>

Although the figures presented above indicate that participation in job search and, to a lesser extent, education and training appear related to program impacts for the three subgroups by level of disadvantage, a great deal was left unexplained by either of these two factors alone. This section explores the importance of the local unemployment rate and AFDC guarantee levels in combination with impacts on program participation in explaining differences in program impacts from program to program. Least-squares regression techniques are used to explore the importance of the four factors simultaneously.

Table 8.3 presents the estimated coefficients from the multivariate regressions relating program impacts to local unemployment rate, AFDC grant level, and impacts on participation in job search and education and training. The subgroups shown in this table are by level of disadvantage, welfare history, high school credential, and recent work history. For each subgroup, the table indicates the value of the estimated coefficient and whether the factor significantly increased or decreased program impacts on earnings in the third year after random assignment.

Table 8.3 reveals several interesting patterns. The first pattern was evident in Figure 8.1: programs with greater impacts on participation in job search tend to be those with larger impacts on earnings. For the most disadvantaged subgroup, for example, an increase of 1 percentage point in the impact on job search is associated with an increased impact on earnings of nearly $11 per year. Thus, if one program increased job search by 35 percentage points and a second program did not increase job search, the regression result implies that the first program would be expected to have increased earnings by nearly $400 more (385=35*11) than the second program. Of course, the results in Table 8.3 do not show causal relationships, but merely reflect patterns in the 20 programs being studied. In other words, the $11 per year increase in earnings impacts might not be actually the *result* of increased participation in job search, but might reflect other factors that make some

National Evaluation of Welfare-to-Work Strategies

Table 8.3

Relationship Between Impact on Earnings in Years 1-3 and Local Unemployment Rate, AFDC Grant Level, and Impacts on Program Participation for Selected Subgroups

	Estimated Coefficient on			
Program	Local Unemployment Rate	AFDC Grant Level for Family of 3	Impact on Participation in Job Search	Impact on Participation in Education and Training
Level of disadvantage				
Most disadvantaged	3.24	0.356	10.82 *	-11.96 **
Moderately disadvantaged	-9.07	0.025	23.88 **	0.89
Least disadvantaged	-177.26 **	3.527 ***	7.02	15.83
Welfare history				
Long-term recipients	11.11	-0.352	18.13 ***	-1.62
Short-term recipients	-92.70 **	1.618 *	13.91	3.23
New applicants	-75.18	2.099	22.93	-24.34
Total earnings in past 12 months				
No earnings	-24.78	-0.146	24.52 ***	-3.65
Less than $5,000	-37.31	0.222	16.37 **	2.49
$5,000 or more	-131.06 *	2.595 *	13.06	6.18
High school credential				
No high school diploma/GED	4.48	-0.149	13.72 **	-10.59 **
High school diploma/GED	-42.02	0.769	21.89 **	9.25

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Food Stamp records.

NOTE: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

programs have higher participation rates than others. In particular, the estimate does not imply that a program administrator who expended resources to increase participation in job search would see an increase in the program's impacts on earnings of $11 per year.

The relationship between impacts on job search and impacts on earnings appears to be robust across a range of subgroups. For example, extra job search appears related to greater effects for both the most disadvantaged and the moderately disadvantaged, for both people with no earnings in the year prior to random assignment and those with some earnings, and for both high school graduates and nongraduates.

A second pattern appears to show little positive relationship between impacts on participation in education and training and impacts on earnings. For the two subgroups where there was a significant relationship — the most disadvantaged group and high school nongraduates — the relationship was negative. In other words, more participation in education was related to smaller impacts on earnings. It is possible that this decline stems from an initial period when people are in an education program rather than in a job. However, the relationship between impacts on participation in education and impacts on earnings is about the same in the third year after random assignment, after most people would have finished their time in education.

A third pattern is in the relationship between local economic conditions and welfare guarantee levels and impacts on earnings. For the more disadvantaged subgroups neither lower unemployment rates nor higher welfare guarantee levels appear related to the programs' impacts. This implies that impacts are equally high for the more disadvantaged sample members whether the economy is performing poorly or well at the time of random assignment.

For some less disadvantaged groups — the least disadvantaged, short-term recipients, and people who earned $5,000 or more in the year prior to random assignment — higher unemployment rates were associated with lower impacts on earnings. For example, an increase in the unemployment rate of 1 percentage point is associated with a decrease in impacts of $177 per year for the least disadvantaged group. This makes some sense since it may imply that people who are better able to find a job have an easier time responding to the help and incentives of the welfare-to-work programs when the economy is faring relatively well.

For the same groups, a higher state welfare guarantee level is associated with a higher impact on earnings, which is consistent with the discussion of welfare grant levels earlier in this section. If two welfare recipients have similar demographic characteristics and similar work and welfare histories, then the recipient in a high-grant state is likely to be less disadvantaged than the recipient in a low-grant state. Likewise, grants may be higher in some states such as California because the cost of living is higher and consequently wages are higher. Thus, the positive link between welfare guarantee levels and impacts on earnings may reflect these factors.

III. Summary

This chapter examined some of the factors that might be related to the effectiveness of welfare-to-work programs in increasing earnings for their participants. Two analyses were presented, one rigorous, because it was based on purely experimental comparisons of six programs, and one speculative, because it was based on nonexperimental correlations among the 20 programs. Although the two analyses were quite different, they yielded one consistent finding: job search increased earnings more than education for the more disadvantaged recipients.

In an experimental comparison of three labor force attachment (LFA) programs that required most participants to look for work and three human capital development (HCD) programs that allowed most to participate in basic education, both the LFA and HCD programs significantly increased earnings for most subgroups. However, the LFA programs increased earnings significantly more than the HCD programs for four groups of the more disadvantaged sample members. For the less disadvantaged sample members, in contrast, the impacts of the two programs were indistinguishable. Since programs that emphasize job search typically cost less than programs that emphasize education, however, the LFA programs might be preferred for the less disadvantaged as well as the more disadvantaged sample members.

In a nonexperimental analysis using regression techniques, programs that had greater effects on job search were also more likely to generate large earnings gains for the more disadvantaged groups, such as long-term recipients. In contrast, programs with greater impacts on participation in education tended to be programs with *smaller* impacts for the more disadvantaged groups. Especially paradoxical was a finding that programs with higher impacts on participation in education had lower earnings impacts for high school nongraduates, a group they were explicitly intended to benefit.

Both sets of results should be viewed with caution. Even though the first analysis was experimental, it involved only six programs in three sites. Because the second analysis was nonexperimental, it is possible that other factors are responsible for the patterns that were found. Finally, neither approach explicitly looked at the effects of employment-focused programs that used a mix of first activities. Therefore, these results do not imply that welfare-to-work programs should rely exclusively on job search. Indeed, the two programs that increased earnings substantially for the widest range of subgroups — Portland and Riverside GAIN — were employment-focused, but used a mix of first activities that allowed some participants to enroll in short-term basic education before looking for work.

Appendix A

Private Opinion Survey Scales and Subgroups

This appendix describes the items used in the creation of scales and subgroups based on the Private Opinion Survey (POS) conducted by MDRC as part of the NEWWS, FTP, and MFIP evaluations. These surveys are available from MDRC. The FTP and MFIP POS instruments are identical. The NEWWS POS instrument is slightly different from the others: it was designed primarily to measure barriers to participation in job search or education and training. The FTP and MFIP instruments were designed primarily to measure barriers to part-time and full-time work. Because of this difference in focus, in order to construct measures across projects it was necessary to assume that individuals with barriers to participation in job search or education and training would also have the same barriers to work.[1]

Three of the seven constructs presented below are based on scales: the Work-Related Parental Concerns Scale, the Mastery Scale, and the Risk of Depression Scale. Cronbach's alpha calculations were conducted to assess the reliability of these scales. Scale scores were created by summing the value of the responses to items in each scale. To facilitate readability, each scale score was divided by the number of items summed to approximate the original metric of the items used to construct the scale. Next, subgroups were created from the scale variables using the methods discussed below.

The four other constructs presented below were not combined into scales. Subgroups were created from these variables as discussed below.

Item responses from the POS were most often based on a 4-point scale, although some items were dichotomous. The original metric and range for each item below are presented in parentheses.

1. Preference for Work Subgroups

This measure was created from the following items:

NEWWS POS items:

17. "Would you take a full-time job today if it paid a little less than welfare?" (Dichotomous variable: "Probably yes" or "Probably no")

21. "Would you take a full-time job today if the job paid a little more than welfare, but you would not like the work?" (Dichotomous variable: "Probably yes" or "Probably no")

FTP/MFIP items:

33. "I would take a full-time job today even if the job paid less than welfare." (4-point scale: "Agree a lot" to "Disagree a lot")

---

[1]We were able to test this assumption by comparing answers on two items from the NEWWS POS that were identical except for their reference to participation in education and training or work. The correlation coefficient between these two variables was r = .59 and was statistically significant.

27. “Suppose you were offered a job that could support your family a little better than welfare. Would you take the job if you didn’t like the work?” (4-point scale: “Agree a lot” to “Disagree a lot”)

To make the NEWWS and FTP/MFIP POS items comparable, the FTP/MFIP 4-point scale items were recoded into dichotomous variables to match the NEWWS POS instrument’s “probably yes” and “probably no” categories. Then, NEWWS POS item 17 was merged with FTP/MFIP item 33 and NEWWS question 21 was merged with FTP/MFIP item 27.

From these two merged items three subgroups were created. First, sample members who would not take a full-time job that paid less than welfare and would not take a full-time job that paid more than welfare if they did not like the work were classified as having a low preference for work. Second, sample members who would take a job under one but not the other condition were classified as having a moderate preference for work. Third, sample members who would take a job that paid less than welfare and would take a job that paid more than welfare even if they did not like the work were classified as having a high preference for work.

If one or both of the values for the two items was missing then sample members was not coded into subgroups.

2. Work-Related Parental Concerns Scale and Subgroups

Cronbach’s alpha = .70

The scale used to create the Work-Related Parental Concerns subgroups is intended to measure sample members’ fears or concerns about leaving their children to go to work.[2]

The items on this scale were identical on all three POS instruments and are listed below. The NEWWS item numbers appear on the left of the slash and the FTP/MFIP question numbers appear on the right.

NEWWS and FTP/MFIP POS items:

25/20. “Right now I’d prefer not to work so I can take care of my family full time.” (4-point scale: “Agree a lot” to “Disagree a lot”; direction of scale reversed)

28/46. “If I got a job, I could find someone I trust to take care of my children.” (4-point scale: “Agree a lot” to “Disagree a lot”)

[2]The Work-Related Parental Concerns Scale is similar to the identically named scale reported in Bos et al., (2001). However, this scale included NEWWS item 1 “If you had a choice, which would you prefer, going to school to study basic reading and math or staying home to take care of your family?” and did not include NEWWS item 28 “If I got a job, I could find someone I trust to take care of my children.” Item 1 was not included in the scale because it decreased the scale’s overall reliability. Item 28 was added because it was conceptually related to the other questions in the scale.

40/26. "I do not want a job because I would miss my children too much." (4-point scale: "Agree a lot" to "Disagree a lot"; direction of scale reversed)

47/35. "I cannot go to a school or job training program right now because I am afraid to leave my children in day care or with a baby sitter." (4-point scale: "Agree a lot" to "Disagree a lot"; direction of scale reversed)

Items for which respondents indicated "don't know" or "refused" were recoded to a missing value. If three out of the four values in the scale were valid responses, the missing value was replaced with the mean of the nonmissing values on the scale for that respondent. If fewer than three out of the four values were valid responses, the respondent was given a missing value on the scale and was not coded into subgroups.

Sample members were divided into "high" and "low" barriers subgroups on the Work-Related Parental Concerns Scale using the following method. As indicated above, the direction of items 25/20, 40/26, and 47/35 were reversed so that 1 equals "disagree a lot" and 4 equals "agree a lot." Therefore, individuals with higher values on the scale have more parental concerns. Sample members whose average response across all measures in the scale was 2.5 or greater were classified as high on the scale. Sample members whose average response was less than 2.5 were classified as low on the scale. This effectively divides all sample members between those who, on average, said they disagreed or disagreed a lot on the items and those who agreed or agreed a lot.

3. Mastery Scale and Subgroups

Cronbach's alpha = .67

The items on this scale were identical on all three POS instruments and are listed below. The NEWWS item numbers appear on the left of the slash and the FTP/MFIP numbers appear on the right.

The Mastery Scale used here is a modified version of the Pearlin Mastery Scale (Pearlin et al., 1981), which is intended to measure the degree to which individuals feel in control of the direction of their life, as opposed to feeling that external factors have a dominant influence. It contains three highly intercorrelated items taken directly from the Pearlin Mastery Scale (items 27/31, 34/34, 42/23) and one other item that improved the scale's overall reliability (item 29/48). Item 29/48 coheres with the other items in this scale because it addresses respondents' feelings about lacking control over their life.

NEWWS and FTP/MFIP POS items:

27/31. "I have little control over the things that happen to me." (4-point scale: "Agree a lot" to "Disagree a lot")

29/48. "I often feel angry that people like me never have a chance to succeed." (4-point scale: "Agree a lot" to "Disagree a lot")

34/34. “Sometimes I feel that I’m being pushed around in life.” (4-point scale: “Agree a lot” to “Disagree a lot”)

42/23. “There is little that I can do to change many of the important things in my life.” (4-point scale: “Agree a lot” to “Disagree a lot”)

Items for which respondents indicated “don’t know” or “refused” were recoded to a missing value. If three out of the four values in the scale were valid responses, the missing value was replaced with the mean of the nonmissing values on the scale for that respondent. If fewer than three out of the four values were valid responses, the respondent was given a missing value on the scale and was not coded into subgroups.

Sample members were divided into “high” and “low” subgroups on the Mastery Scale using a method similar to that used on the Work-Related Parental Concerns Scale. On this scale, lower values indicate that respondents agreed to more of the questions or that they felt less in control over the events in their life. Sample members whose average response across all measures on the scale was 2.5 or less were classified as low on the scale, or as having a low sense of mastery over personal events. Sample members whose average response was greater than 2.5 were classified as high on the scale, or as having a high sense of mastery over personal events. Again, this effectively divides all sample members between those who, on average, said they disagreed or disagreed a lot on the items on the scale and those who agreed or agreed a lot.

4. Risk of Depression Scale and Subgroups

Cronbach’s alpha = 0.90

This scale was created to indicate the degree to which respondents were at risk of depression at random assignment. Data for this scale were available only for four NEWWS sites: Atlanta, Grand Rapids, Riverside, and Portland.[3]

The scale was created from the following NEWWS POS items:

“During the past week . . .”

50. “I felt sad.” (4-point scale: “Rarely” to “Most or all days”)

51. “I felt depressed.” (4-point scale: “Rarely” to “Most or all days”)

52. “I felt that I could not shake off the blues, even with the help of family and friends.” (4-point scale: “Rarely” to “Most or all days”)

53. “I felt lonely.” (4-point scale: “Rarely” to “Most or all days”)

[3]The Risk of Depression Scale and subgroups were defined following Bos et al. (2001).

To create the Risk of Depression Scale the four items above were summed and divided by the number of valid items. Items with missing values were not imputed. Instead, the scale value is simply the average of the nonmissing values of the four items.

This scale was used to define subgroups as follows. Respondents whose average score across these four items was 2 or less were classified as being at low risk of depression. Respondents whose average score was greater than 2 but less than or equal to 3 were classified as being at moderate risk of depression. Those whose average score was greater than 3 were coded as being at high risk of depression on the scale.

5. Health or Emotional Problem Subgroups

This measure was created from the following items:

NEWWS items:

45. "I cannot go to a school or job training program right now because I have a health or emotional problem." (4-point scale: "Agree a lot" to "Disagree a lot")

46. "I cannot go to a school or job training program right now because I have a child or family member with a health or emotional problem." (4-point scale: "Agree a lot" to "Disagree a lot")

FTP/MFIP item:

38. "I cannot work at a full-time job for 40 hours a week right now because I have, or a family member has, a health or emotional problem." (4-point scale: "Agree a lot" to "Disagree a lot")

Note that the FTP/MFIP POS combines the two NEWWS items into one item about the respondent *or* the respondent's child or family member. To make the NEWWS and FTP/MFIP questions comparable, the following method was used. Sample members who responded "agree" or "agree a lot" to either NEWWS item 45 or 46 were coded as having a barrier to work or participation because of a health or emotional problem or because a child or family member had such a problem. Therefore, individuals who reported that they could not attend school or work because of a health or emotional problem that they had or that a family member had were coded as "yes" to having a health or emotional problem as a barrier to work or participation. Individuals who responded that neither they nor a family member had such a barrier were coded as "no" on this measure.

6. Child Care Problem Subgroups

This measure was created from the following items:

NEWWS item:

48. "I cannot go to a school or job training program right now because I cannot afford child care." (4-point scale: "Agree a lot" to "Disagree a lot")

FTP/MFIP item:

39. “I cannot work at a full-time job for 40 hours a week right now because I cannot arrange for child care.” (4-point scale: “Agree a lot” to “Disagree a lot”)

Although the FTP/MFIP item is perhaps somewhat broader than the NEWWS item, because both pertain to logistical child care barriers that were addressed by all the programs in this report these two items were deemed comparable.

Respondents who agreed or agreed a lot to these items were coded as “yes” on this measure. In other words, these individuals stated that they were unable to attend school or work because of child care affordability or accessibility problems. Respondents who disagreed or disagreed a lot on these items were coded as “no” on this measure.

7. Transportation Problem Subgroups

This measure was created from the following items:

NEWWS item:

44. “I cannot go to a school or job training program right now because I have no way to get there every day. (4-point scale: “Agree a lot” to “Disagree a lot”)

FTP/MFIP item:

14. “I cannot work at a part-time job for 10 hours a week right now because I have no way to get there every day.” (4-point scale: “Agree a lot” to “Disagree a lot”)

Respondents who agreed or agreed a lot on these items were coded as “yes” on this measure. In other words, these individuals stated that they were unable to attend school or work because of transportation problems. Respondents who disagreed or disagreed a lot on these items were coded as “no” on this measure.

Appendix B

Supplementary Tables to Chapters 5 and 7

National Evaluation of Welfare-to-Work Strategies

Table B.1

Impacts on Earnings and AFDC Payments Across 20 Welfare-to-Work Programs, by Race and Ethnicity

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($)		Impacts on Average Total AFDC Payments per Year ($)	
		Years 1-3	Year 3	Years 1-3	Year 3
SWIM					
White	877	592	115	-864 ***	-633 **
African-American	1,361	446	512	-710 ***	-665 ***
Hispanic	814	692 *	1,054 **	-473 *	-663 **
Other	n/a				
GAIN evaluation programs					
Alameda					
White	216	57	253	37	-170
African-American	844	736 *	941 *	-270	-232
Hispanic	n/a				
Other	n/a				
Butte					
White	1,061	1,036 **	1,229 **	58	126
African-American	n/a				
Hispanic	n/a				
Other	n/a				
Los Angeles					
White	512	329	287	-413	-314
African-American	1,987	-135	-105	-511 ***	-436 **
Hispanic	1,408	194	311	-136	-108
Other	489	404 **	567 **	-239	-48
Riverside					
White	2,847	1,710 ***	1,619 ***	-777 ***	-855 ***
African-American	862	1,288 **	1,003	-783 **	-499
Hispanic	1,510	968 ***	1,126 ***	-739 ***	-507 *
Other	n/a				
San Diego		†			
White	3,478	1,148 ***	1,323 ***	-350 **	-388 **
African-American	1,865	557	609	-445 *	-477 *
Hispanic	2,094	-280	-20	-380	-205
Other	782	807 *	656	-808 *	-459
Tulare					
White	1,165	-43	360	126	-8
African-American	n/a				
Hispanic	871	231	760 *	-140	-345
Other	n/a				

(continued)

Table B.1 (continued)

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($) Years 1-3	Year 3	Impacts on Average Total AFDC Payments per Year ($) Years 1-3	Year 3
NEWWS programs					
Atlanta LFA					
White	n/a				
African-American	3,624	464 ***	483 **	-193 ***	-172 ***
Hispanic	n/a				
Other	n/a				
Atlanta HCD					
White	n/a				
African-American	3,669	299 *	437 **	-174 ***	-157 ***
Hispanic	n/a				
Other	n/a				
Grand Rapids LFA		††	††		
White	1,470	42	-222	-603 ***	-432 ***
African-American	1,214	594 ***	717 **	-761 ***	-631 ***
Hispanic	244	1,346 ***	1,462 **	-1,097 ***	-1,041 ***
Other	n/a				
Grand Rapids HCD		†	†	††	†
White	1,515	228	174	-332 ***	-373 ***
African-American	1,158	400	326	-537 ***	-417 ***
Hispanic	249	1,505 ***	2,059 ***	-1,066 ***	-1,170 ***
Other	n/a				
Riverside LFA					
White	3,464	419 **	128	-663 ***	-560 ***
African-American	1,121	433	259	-528 ***	-273
Hispanic	1,858	941 ***	994 ***	-770 ***	-839 ***
Other	255	422	-223	-35	-433
Riverside HCD		†	†		
White	1,208	218	343	-618 ***	-531 **
African-American	510	-242	-71	-279	-497
Hispanic	1,240	730 ***	1,002 ***	-796 ***	-960 ***
Other	164	-284	-694	-685	-1,085 *
Columbus Integrated					
White	2,161	309	303	-437 ***	-416 ***
African-American	2,414	430 *	647 **	-284 ***	-341 ***
Hispanic	n/a				
Other	n/a				
Columbus Traditional					
White	2,204	124	-16	-257 ***	-214 ***
African-American	2,431	420 *	522 *	-284 ***	-330 ***
Hispanic	n/a				
Other	n/a				

(continued)

Table B.1 (continued)

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($)		Impacts on Average AFDC Payments per Year ($)	
		Years 1-3	Year 3	Years 1-3	Year 3
Detroit					
White	481	783 *	795	-330 *	-418 *
African-American	3,836	344 **	626 ***	-197 ***	-294 ***
Hispanic	n/a				
Other	n/a				
Oklahoma City					††
White	3,473	5	-65	2	73
African-American	1,714	140	140	-174 **	-165 *
Hispanic	247	-234	-112	-24	-303
Other	403	-100	-291	-227	-302 *
Portland					
White	3,795	1,238 ***	1,628 ***	-753 ***	-850 ***
African-American	1,099	542 *	647	-439 ***	-585 ***
Hispanic	226	1,630 **	2,073 **	-855 **	-807 *
Other	335	1,257 *	861	-1,038 ***	-955 ***
FTP					
White	1,234	759 ***	1,127 ***	-125 **	-246 ***
African-American	1,410	393 *	630 **	-178 **	-336 ***
Hispanic	n/a				
Other	n/a				
MFIP					
White	591	437	133	770 ***	584 *
African-American	558	818 *	698	510	670 *
Hispanic	n/a				
Other	n/a				

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Background Information Forms (BIFs).

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

N/a = not applicable because sample sizes were too small to reliably calculate impacts.

National Evaluation of Welfare-to-Work Strategies

Table B.2

Impacts on Earnings and AFDC Payments Across 20 Welfare-to-Work Programs, by Age of Youngest Child at Random Assignment

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($) Years 1-3	Year 3	Impacts on Average Total AFDC Payments per Year ($) Years 1-3	Year 3
SWIM[a]					
Under 6	321	1,490 **	1,849 *	-701	-598
6 or over	2,889	563 **	556 **	-684 ***	-661 ***
GAIN evaluation programs[a]					
Alameda					
Under 6	367	762	1,027	-309	-375
6 or over	831	466	792	-201	-362
Butte					
Under 6	n/a				
6 or over	1,070	965 **	1,291 ***	82	158
Los Angeles					
Under 6	456	-372	-746	87	172
6 or over	3,875	150	254	-410 ***	-333 **
Riverside					
Under 6	890	1,913 ***	2,287 ***	-1,100 ***	-1,292 ***
6 or over	4,553	1,299 ***	1,156 ***	-682 ***	-499 ***
San Diego					
Under 6	1,056	681	1,949 ***	115	3
6 or over	7,077	593 ***	549 **	-472 ***	-394 ***
Tulare					
Under 6	329	572	1,341 *	-161	-245
6 or over	1,895	35	482	91	-87
NEWWS programs					
Atlanta LFA					
Under 3	n/a				
3 to 5	1,630	444 *	310	-240 ***	-224 ***
6 or over	2,184	551 ***	748 ***	-185 ***	-153 **

(continued)

Table B.2 (continued)

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($) Years 1-3	Year 3	Impacts on Average Total AFDC Payments per Year ($) Years 1-3	Year 3
Atlanta HCD					
Under 3	n/a				
3 to 5	1,624	252	360	-142 **	-108
6 or over	2,240	377 *	626 **	-189 ***	-188 ***
Grand Rapids LFA					
Under 3	1,396	737 ***	820 **	-748 ***	-586 ***
3 to 5	659	274	91	-458 ***	-393 **
6 or over	954	30	-164	-826 ***	-661 ***
Grand Rapids HCD					
Under 3	1,378	637 ***	773 **	-592 ***	-538 ***
3 to 5	640	565	449	-161	-237
6 or over	971	38	-16	-562 ***	-572 ***
Riverside LFA					
Under 3	388	161	504	-248	-456
3 to 5	3,328	782 ***	586 ***	-713 ***	-640 ***
6 or over	2,920	371 *	141	-678 ***	-587 ***
Riverside HCD					
Under 3	228	49	226	-916 **	-1,154 **
3 to 5	1,529	534 **	836 ***	-545 ***	-608 ***
6 or over	1,328	137	146	-622 ***	-676 ***
Columbus Integrated					
Under 3	n/a				
3 to 5	2,122	590 ***	768 ***	-349 ***	-367 ***
6 or over	n/a				
Columbus Traditional					
Under 3	n/a				
3 to 5	2,106	510 **	635 **	-294 ***	-336 ***
6 or over	n/a				
Detroit					
Under 3	1,748	245	493	-169	-341 **
3 to 5	1,110	223	426	-18	-96
6 or over	1,566	582 **	771 *	-248 **	-273 **
Oklahoma City					
Under 3	2,365	64	13	16	4
3 to 5	1,394	220	288	-114	-126
6 or over	2,056	-183	-317	-167 **	-72

(continued)

Table B.2 (continued)

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($)		Impacts on AverageTotal AFDC Payments per Year ($)	
		Years 1-3	Year 3	Years 1-3	Year 3
Portland					
Under 3	2,219	1,299 ***	1,600 ***	-794 ***	-978 ***
3 to 5	1,514	1,381 ***	1,715 ***	-741 ***	-768 ***
6 or over	1,732	739 ***	987 **	-575 ***	-621 ***
FTP					
Under 3	1,139	576 **	708 **	-223 ***	-322 ***
3 to 5	707	137	412	-79	-319 ***
6 or over	836	683 *	1,063 **	-134 *	-266 ***
MFIP					
Under 3	472	298	342	784 **	731 *
3 to 5	402	1,027 *	696	659 *	628
6 or over	456	516	385	514 *	449

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Background Information Forms (BIFs).

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent ** = 5 percent and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

N/a = not applicable because sample sizes were too small to reliably calculate impacts.

[a]Because of data restrictions, for mothers in the GAIN and SWIM programs whose youngest child was under age 6, it cannot be determined whether the child was between the ages of 3 and 5 or under age 3 at the time of program entry.

National Evaluation of Welfare-to-Work Strategies

Table B.3

Impacts on Earnings and AFDC Payments Across 20 Welfare-to-Work Programs, by Number of Children

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($) Years 1-3	Year 3	Impacts on Average Total AFDC Payments per Year ($) Years 1-3	Year 3
SWIM					
3 children or more	647	528	1,199 **	-837 **	-1,031 **
2 children	953	531	314	-643 ***	-357
0-1 child	1,609	813 ***	714 *	-749 ***	-767 ***
GAIN evaluation programs					
Alameda					
3 children or more	302	847 *	1,112	-659	-1,220 **
2 children	379	414	1,157 *	-35	-276
0-1 child	517	480	519	54	243
Butte					
3 children or more	215	1,648 **	1,471	-966	-928
2 children	406	-4	-111	352	122
0-1 child	601	1,023 **	1,587 ***	-29	106
Los Angeles					
3 children or more	1,512	-113	-111	-72	-112
2 children	1,363	755 ***	790 **	-834 ***	-724 ***
0-1 child	1,456	-287	-191	-247	-61
Riverside					
3 children or more	1,386	1,853 ***	1,703 ***	-886 ***	-761 **
2 children	1,789	1,880 ***	2,094 ***	-806 ***	-688 ***
0-1 child	2,268	754 ***	549	-688 ***	-555 ***
San Diego					
3 children or more	1,714	613	955 *	-452	-252
2 children	2,583	1,249 ***	1,581 ***	-447 **	-562 **
0-1 child	3,836	157	50	-360 **	-253
Tulare					
3 children or more	636	-588	-108	914 **	530
2 children	732	794 *	1,125 **	-192	-413
0-1 child	856	327	966 *	58	166
NEWWS programs					
Atlanta LFA					
3 children or more	1,146	647 **	746 **	-218 **	-212 *
2 children	1,257	474 *	539	-237 ***	-183 **
0-1 child	1,430	367	352	-146 **	-130 *

(continued)

Table B.3 (continued)

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($) Years 1-3	Year 3	Impacts on Average Total AFDC Payments per Year ($) Years 1-3	Year 3
Atlanta HCD					
3 children or more	1,140	290	445	-156 *	-148
2 children	1,304	223	394	-106	-66
0-1 child	1,437	357	539	-198 ***	-202 ***
Grand Rapids LFA					
3 children or more	533	683 **	521	-984 ***	-894 ***
2 children	1,056	548 *	326	-898 ***	-854 ***
0-1 child	1,423	197	299	-410 ***	-183
Grand Rapids HCD					
3 children or more	543	725 *	670	-800 ***	-750 ***
2 children	1,064	155	-4	-576 ***	-664 ***
0-1 child	1,390	510 *	700 *	-314 ***	-254 **
Riverside LFA					
3 children or more	1,914	970 ***	891 ***	-951 ***	-983 ***
2 children	2,124	258	-7	-425 ***	-305 **
0-1 child	2,688	537 ***	332	-631 ***	-545 ***
Riverside HCD					
3 children or more	1,019	597 **	785 ***	-1,177 ***	-1,318 ***
2 children	970	3	95	-354 *	-414 *
0-1 child	1,146	330	522	-257	-296
Columbus Integrated					
3 children or more	1,253	1,150 ***	1,230 ***	-497 ***	-465 ***
2 children	1,564	613 **	687 **	-477 ***	-540 ***
0-1 child	1,855	-368	-200	-193 ***	-205 ***
Columbus Traditional					
3 children or more	1,238	636 **	617 *	-478 ***	-477 ***
2 children	1,558	755 ***	763 **	-293 ***	-320 ***
0-1 child	1,933	-318	-332	-127 **	-129 *
Detroit					
3 children or more	1,185	264	619	-50	-289
2 children	1,317	242	113	-198 *	-337 **
0-1 child	1,957	518 **	879 **	-218 ***	-212 **
Oklahoma City					
3 children or more	1,095	-71	-67	-54	-54
2 children	1,794	203	122	-104	-137 *
0-1 child	2,972	-43	-100	-93 **	-19

(continued)

Table B.3 (continued)

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($) Years 1-3	Year 3	Impacts on Average Total AFDC Payments per Year ($) Years 1-3	Year 3
Portland					
3 children or more	1,457	1,353 ***	1,529 ***	-725 ***	-870 ***
2 children	1,834	970 ***	1,403 ***	-670 ***	-804 ***
0-1 child	2,256	989 ***	1,197 ***	-623 ***	-686 ***
FTP					
3 children or more	764	496	909 **	-193 *	-393 ***
2 children	814	730 **	827 *	-109	-246 ***
0-1 child	1,237	446 *	735 **	-113 **	-253 ***
MFIP					
3 children or more	417	1,005 **	1,231 **	753 **	481
2 children	443	708	338	361	351
0-1 child	503	106	-12	385	451

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Background Information Forms (BIFs).

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

National Evaluation of Welfare-to-Work Strategies

Table B.4

Impacts on Earnings and AFDC Payments Across Nine Welfare-to-Work Programs, by Preference for Work at Random Assignment

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($): Years 1-3	Year 3	Impacts on Average Total AFDC Payments per Year ($): Years 1-3	Year 3
NEWWS programs					
Atlanta LFA					
Low preference	1,100	570 *	740 *	-199 **	-168 *
Moderate preference	1,360	833 ***	800 **	-240 ***	-250 ***
High preference	485	-197	-130	-146	-39
Atlanta HCD					
Low preference	1,129	368	671 *	-114	-115
Moderate preference	1,344	720 ***	915 ***	-254 ***	-230 ***
High preference	502	-143	-110	-116	-12
Grand Rapids LFA				†	
Low preference	854	315	476	-698 ***	-595 ***
Moderate preference	837	425	156	-683 ***	-509 ***
High preference	202	232	-138	-1,335 ***	-1,056 ***
Grand Rapids HCD					
Low preference	857	541	690	-583 ***	-607 ***
Moderate preference	825	647 **	787 *	-358 ***	-340 **
High preference	190	-210	-533	-754 ***	-587
Riverside LFA					
Low preference	1,692	704 ***	613 *	-663 ***	-576 ***
Moderate preference	1,528	651 **	263	-772 ***	-650 ***
High preference	471	896 **	470	-975 ***	-930 ***
Riverside HCD					
Low preference	701	526 *	687 *	-236	-291
Moderate preference	730	542 *	761 *	-853 ***	-863 ***
High preference	235	-455	-337	-298	-330
Portland					
Low preference	2,541	1,075 ***	1,289 ***	-857 ***	-953 ***
Moderate preference	1,889	1,192 ***	1,478 ***	-610 ***	-733 ***
High preference	357	991	1,493 *	-1,037 ***	-937 ***
FTP					
Low preference	547	775 **	652	-248 **	-279 **
Moderate preference	1,102	469 *	1,010 ***	-164 **	-341 ***
High preference	761	179	322	9	-264 ***

(continued)

Table B.4 (continued)

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($)		Impacts on Average Total AFDC Payments per Year ($)	
		Years 1-3	Year 3	Years 1-3	Year 3
MFIP					
Low preference	452	936 *	661	617 *	578
Moderate preference	363	768	799	122	183
High preference	n/a				

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Private Opinion Survey (POS) data.

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

N/a = not applicable because sample sizes were too small to reliably calculate impacts.

National Evaluation of Welfare-to-Work Strategies

Table B.5

Impacts on Earnings and AFDC Payments Across Nine Welfare-to-Work Programs, by Work-Related Parental Concerns Scale Score

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($): Years 1-3	Impacts on Average Total Earnings per Year ($): Year 3	Impacts on Average Total AFDC Payments per Year ($): Years 1-3	Impacts on Average Total AFDC Payments per Year ($): Year 3
NEWWS programs					
Atlanta LFA					
High	440	513	263	-302 **	-292 *
Low	2,678	494 **	621 **	-168 ***	-134 **
Atlanta HCD					
High	427	165	350	-150	-211
Low	2,730	345 *	582 **	-148 ***	-113 *
Grand Rapids LFA					††
High	529	208	335	-996 ***	-1,016 ***
Low	1,426	404	218	-698 ***	-481 ***
Grand Rapids HCD					
High	581	516	700	-627 ***	-759 ***
Low	1,358	460	500	-447 ***	-375 ***
Riverside LFA					
High	1,033	763 ***	724 **	-720 ***	-713 ***
Low	2,783	727 ***	396	-759 ***	-645 ***
Riverside HCD					
High	557	369	444	-879 ***	-1,025 ***
Low	1,183	549 **	828 **	-420 **	-470 **
Portland					
High	1,423	1,450 ***	1,676 ***	-879 ***	-960 ***
Low	3,486	1,021 ***	1,346 ***	-687 ***	-787 ***
FTP					
High	395	771 **	1,003 **	-319 **	-392 ***
Low	2,058	431 **	717 ***	-92 *	-282 ***
MFIP			††		
High	189	1,768 **	2,890 **	85	102
Low	710	647 *	57	463 *	515

(continued)

Table B.5 (continued)

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Private Opinion Survey (POS) data.

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

National Evaluation of Welfare-to-Work Strategies

Table B.6

Impacts on Earnings and AFDC Payments Across Nine Welfare-to-Work Programs, by Self-Reported Health or Emotional Barriers to Work or Participation

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($) Years 1- 3	Year 3	Impacts on Average Total AFDC Payments per Year ($) Years 1- 3	Year 3
NEWWS programs					
Atlanta LFA					
Has barrier	815	465 *	671 *	-261 ***	-293 **
Does not have barrier	2,306	569 ***	605 **	-164 ***	-119 *
Atlanta HCD					
Has barrier	802	-47	248	-93	-65
Does not have barrier	2,367	405 *	599 **	-167 ***	-150 **
Grand Rapids LFA					
Has barrier	554	32	28	-850 ***	-749 ***
Does not have barrier	1,401	579 **	444	-771 ***	-611 ***
Grand Rapids HCD					
Has barrier	572	217	587	-609 ***	-648 ***
Does not have barrier	1,362	689 **	648 *	-483 ***	-454 ***
Riverside LFA					
Has barrier	1,000	478 **	277	-616 ***	-499 **
Does not have barrier	2,796	814 ***	538 **	-784 ***	-712 ***
Riverside HCD					
Has barrier	563	212	531	-256	-306
Does not have barrier	1,175	593 **	767 **	-639 ***	-764 ***
Portland					
Has barrier	1,385	1,338 ***	1,565 ***	-574 ***	-562 ***
Does not have barrier	3,517	1,030 ***	1,337 ***	-813 ***	-932 ***
FTP					
Has barrier	584	398	713	-183 *	-245 **
Does not have barrier	1,997	504 **	777 ***	-106 *	-302 ***
MFIP					
Has barrier	203	1,313 **	1,206	731	931 *
Does not have barrier	692	880 **	613	225	231

(continued)

Table B.6 (continued)

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Private Opinion Survey (POS) data.

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

Sample members in the "has barrier" category on this measure could have had a health or emotional problem themselves, which they reported as a barrier to work or participation at random assignment, or a family member could have had such a problem.

National Evaluation of Welfare-to-Work Strategies

Table B.7

Impacts on Earnings and AFDC Payments Across Nine Welfare-to-Work Programs, by Self-Reported Child Care Barriers to Work or Participation

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($) Years 1-3	Year 3	Impacts on Average Total AFDC Payments per Year ($) Years 1-3	Year 3
NEWWS programs					
Atlanta LFA					
Has barrier	1,759	508 **	744 **	-143 **	-131 *
Does not have barrier	1,287	550 *	411	-276 ***	-235 ***
Atlanta HCD					
Has barrier	1,763	158	379	-102 *	-67
Does not have barrier	1,316	620 **	824 **	-248 ***	-254 ***
Grand Rapids LFA					
Has barrier	1,306	177	45	-664 ***	-523 ***
Does not have barrier	631	804 **	830	-1,005 ***	-855 ***
Grand Rapids HCD					
Has barrier	1,300	413	517	-479 ***	-477 ***
Does not have barrier	612	625	647	-554 ***	-536 ***
Riverside LFA					
Has barrier	2,521	822 ***	791 ***	-783 ***	-742 ***
Does not have barrier	1,239	557 *	-137	-680 ***	-508 ***
Riverside HCD					
Has barrier	1,186	766 ***	995 ***	-686 ***	-748 ***
Does not have barrier	527	-97	117	-307	-478
Portland					
Has barrier	3,371	1,026 ***	1,303 ***	-782 ***	-899 ***
Does not have barrier	1,475	1,322 ***	1,599 ***	-672 ***	-698 ***
FTP					
Has barrier	941	805 ***	1,097 ***	-241 ***	-361 ***
Does not have barrier	1,519	269	537 *	-73	-272 ***
MFIP					
Has barrier	438	1,265 ***	1,508 **	371	353
Does not have barrier	457	703	78	306	381

(continued)

Table B.7 (Continued)

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Private Opinion Survey (POS) data.

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

National Evaluation of Welfare-to-Work Strategies

Table B.8

Impacts on Earnings and AFDC Payments Across Nine Welfare-to-Work Programs, by Self-Reported Transportation Barriers to Work or Participation

Program and Subgroup	Sample Size	Impacts on Average Total Earnings per Year ($)		Impacts on AverageTotal AFDC Payments per Year ($)	
		Years 1-3	Year 3	Years 1-3	Year 3
NEWWS programs					
Atlanta LFA					
Has barrier	1,042	840 ***	863 ***	-152 *	-118
Does not have barrier	2,057	383	493	-207 ***	-177 **
Atlanta HCD					
Has barrier	1,081	563 ***	871 ***	-159 **	-122
Does not have barrier	2,051	225	389	-160 ***	-142 **
Grand Rapids LFA					
Has barrier	670	291	138	-769 ***	-671 ***
Does not have barrier	1,272	616 **	584	-862 ***	-691 ***
Grand Rapids HCD					
Has barrier	648	302	86	-700 ***	-488 **
Does not have barrier	1,277	721 **	949 **	-448 ***	-532 ***
Riverside LFA					
Has barrier	1,395	1,040 ***	1,166 ***	-860 ***	-848 ***
Does not have barrier	2,388	640 ***	197	-711 ***	-608 ***
Riverside HCD					
Has barrier	788	459 **	517 *	-807 ***	-1,024 ***
Does not have barrier	933	508 *	875 **	-351	-359
Portland					
Has barrier	1,428	908 ***	1,419 ***	-459 ***	-539 ***
Does not have barrier	3,447	1,248 ***	1,456 ***	-863 ***	-962 ***
FTP					
Has barrier	951	759 ***	1,092 ***	-304 ***	-420 ***
Does not have barrier	1,503	336	605 *	-42	-237 ***
MFIP					
Has barrier	421	1,197 **	955	13	56
Does not have barrier	497	468	190	690 **	666 *

(continued)

Table B.8 (continued)

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Private Opinion Survey (POS) data.

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

Appendix C

Discussion of Issues Related to Pooling

In the body of the report, results are sometimes presented *pooled* across a number of welfare-to-work programs and separately for each program. Pooling provides larger samples for each subgroup which increases the precision of the estimated impacts. Increased precision allows more reliable statements to be made about subgroups that include only a few people in any one program, and makes it more likely that a difference of any given magnitude across subgroups is statistically significant.

At the same time, pooling has its drawbacks because the sample composition differed quite a bit from place to place. For example, most Hispanic sample members were in California, a state with high welfare benefit levels. In contrast, many non-Hispanic sample members came from low-grant states such as Georgia and Florida. As a result, average benefit levels for the pooled Hispanic sample will reflect the high benefits in California, whereas benefit levels for the pooled non-Hispanic sample will reflect both California's benefits and the lower benefit levels in other states in the study. Benefits for the pooled Hispanic sample will consequently be higher than for the pooled non-Hispanic sample, even if Hispanic and non-Hispanic sample members in California received similar benefits.

Table C.1 presents another example. Only eight of the 20 programs randomly assigned people who were applying for welfare (in addition to those already receiving it), and that may affect pooled comparisons of welfare applicants to short-term and long-term welfare recipients. Table C.1 shows program group and control group outcomes and impacts for annual earnings over the three-year follow-up period. Results are shown, both pooled ("full sample") and by program, for the eight programs with at least 200 new applicants (which will be called "applicants" in the remainder of this appendix).

Earnings were higher for program group applicants than for program group short-term recipients in five of the eight programs, and lower in three. However, differences between applicants and short-term recipients were much larger in sites where applicants earned more than short-term recipients. In Butte, Riverside, San Diego, and Columbus Integrated, program group applicants earned more than $500 more per year than short-term recipients. In comparison, in SWIM, Columbus Traditional, and Oklahoma City, the difference between earnings of program group applicants and program group short-term recipients was no more than $252 per year.

Since applicants earned much more than short-term recipients in sites such as San Diego and Riverside, but only slightly less in programs such as SWIM, pooled earnings for applicants might be expected to be greater than pooled earnings for short-term recipients. This is not the case. On average, program group applicants earned $236 less than short-term recipients.

The reason for the apparent discrepancy is the difference in composition across the sites. Oklahoma City, a site with generally low earnings levels, had many more applicants than any other site. Among the 6,615 new applicants in these eight sites, 2,530 (about 38 percent) were randomly assigned in Oklahoma City. Applicants in Oklahoma City earned only $2,217 on average, by far the least among the eight programs. As a result, the average program group applicant earned only $3,484, even though applicants in Columbus and FTP earned more than $5,000 on average, applicants in San Diego's GAIN program earned nearly $4,500 on average, and applicants in five of the programs earned more than this average. In comparison, only about one in every six short-term recipients was in Oklahoma City, and only about one in every 14 long-term recipients was in Oklahoma City. In fact, a

National Evaluation of Welfare-to-Work Strategies

Table C.1

Impacts on Earnings per Year
Across Selected Welfare-to-Work Programs, by AFDC Status

Program and Subgroup	Sample Size	Earnings per Year for Years 1-3 ($) Program Group	Control Group	Program Impact
Full Sample				
Long-term recipient	19,039	3,415	2,852	563
Short-term recipient	10,404	4,132	3,448	684
New applicant	6,615	3,896	3,484	412
SWIM [a]				
Long-term recipient	2,202	3,021	2,447	574 ***
Short-term recipient	648	5,044	4,437	607
New applicant	360	4,833	3,752	1,081
Butte GAIN				
Long-term recipient	558	3,049	1,604	1,445 ***
Short-term recipient	285	3,537	2,652	885
New applicant	386	4,032	3,853	179
Riverside GAIN				
Long-term recipient	2,661	3,322	2,026	1,296 ***
Short-term recipient	1,979	3,943	2,534	1,409 ***
New applicant	868	4,552	2,946	1,606 ***
San Diego GAIN				
Long-term recipient	3,948	3,082	2,726	356
Short-term recipient	3,079	4,335	3,313	1,022 ***
New applicant	1,192	5,193	4,478	715
Columbus Integrated				
Long-term recipient	3,392	4,183	3,669	513 ***
Short-term recipient	806	5,422	5,826	-404
New applicant	448	6,066	5,525	541
Columbus Traditional				
Long-term recipient	3,415	4,019	3,669	350 **
Short-term recipient	793	5,904	5,838	67
New applicant	497	5,652	5,547	105
Oklahoma City				
Long-term recipient	1,419	2,151	1,974	177
Short-term recipient	1,858	2,291	2,286	5
New applicant	2,530	2,074	2,127	-53
FTP				
Long-term recipient	1,444	3,247	2,827	420 **
Short-term recipient	956	4,452	3,612	840 ***
New applicant	334	4,675	5,088	-414

(continued)

Table C.1 (continued)

SOURCES: MDRC calculations from unemployment insurance (UI) earnings records, AFDC records, and Baseline Information Forms (BIFs).

NOTES: A two-tailed t-test was applied to differences between outcomes for the program and control groups. Statistical significance levels are indicated as * = 10 percent, ** = 5 percent, and *** = 1 percent.

An F-test was applied to differences among subgroups for each characteristic. Statistical significance levels are indicated as † = 10 percent, †† = 5 percent, and ††† = 1 percent.

Significance tests were not performed for the pooled impacts.

N/a = not applicable because sample sizes were too small to reliably calculate impacts.

[a]For SWIM and GAIN, short-term recipients are those who said they had received welfare for two years or less rather than less than two years, and long-term recipients are those who said they had received welfare for more than two years rather than two years or more.

straight average of the eight earnings levels for applicants is $4,635, or more than $700 more than $4,635, or more than $700 more than the weighted pooled average.

Just as sample composition affects the pooled program group and control group outcomes, it also affects the pooled impact. For these eight programs, pooled impacts were about $400 per year per applicant, and nearly $700 per year per short-term recipient. In this case, the pooled impact was not much different from what would be expected. While SWIM and Riverside had much larger impacts for applicants than the pooled result, FTP had much smaller impacts. Moreover, a straight average of impacts for applicants across the eight sites is $470, which is not much larger than the weighted pooled impact.

The pooled impact was closer to the average impact across the eight programs because there were much smaller differences across the sites. The difference between the largest impact for applicants — $1,606 in Riverside — and the smallest impact for applicants — a $414 reduction in FTP — was $2,020 per year. In comparison, the difference in average earnings for program group new applicants between Oklahoma City and Columbus Integrated was nearly $4,000 per year, the difference between Oklahoma City and Columbus Traditional was about $3,500 per year, and the difference between Oklahoma City and San Diego was more than $3,000 per year. Indeed, the site with the closest earnings levels to Oklahoma City was Butte, and its applicants earned nearly $2,000 more than in Oklahoma City ($4,032 compared with $2,074). If Oklahoma City had either extremely large or extremely small impacts for applicants, the pooled impacts would also have looked much different than the impacts from most of the individual programs.

References

Aneshensel, Carol S., Ralph R. Frerichs, and George J. Huba. 1984. "Depression and Physical Illness: A Multiwave, Nonrecursive Causal Model." *Journal of Health and Social Behavior* 25: 350-371.

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

Bloom, Dan, James J. Kemple, and Robin Rogers-Dillon. 1997. *The Family Transition Program: Implementation and Early Impacts of Florida's Time-Limited Welfare Program.* New York: Manpower Demonstration Research Corporation.

Bloom, Dan, Mary Farrell, James J. Kemple, and Nandita Verma. 1998. *The Family Transition Program: Implementation and Interim Impacts of Florida's Initial Time-Limited Welfare Program.* New York: Manpower Demonstration Research Corporation.

Bloom, Dan, Mary Farrell, James J. Kemple, and Nandita Verma. 1999. *The Family Transition Program: Implementation and Three-Year Impacts of Florida's Initial Time-Limited Welfare Program.* New York: Manpower Demonstration Research Corporation.

Bos, Johannes M., Susan Scrivener, Jason Snipes, and Gayle Hamilton. 2001. *Improving Basic Skills: Adult Education in Eleven Welfare-to-Work Programs*. Washington, DC: U.S. Department of Education, Office of the Under Secretary and Office of Vocational and Adult Education; and U.S. Department of Health and Human Services, Administration for Children and Families and Office of the Assistant Secretary for Planning and Evaluation.

Bruce, Martha Livinston, David T. Takeuchi, and Philip J. Leaf. 1991. "Poverty and Psychiatric Status: Longitudinal Evidence from the New Haven Epidemiologic Catchment Area Study." *Archives of General Psychiatry* 48: 470-474.

Danziger, Sandra, Mary Corcoran, Sheldon Danziger, Colleen Heflin, Ariel Kalil, Judith Levine, Daniel Rosen, Kristin Seefeldt, Kristine Siefert, and Richard Tolman. 1998. "Barriers to the Employment of Welfare Recipients." Ann Arbor, MI: Poverty Research and Training Center. University of Michigan.

Freedman, Stephen, Daniel Friedlander, Gayle Hamilton, JoAnn Rock, Marisa Mitchell, Jodi Nudelman, Amanda Schweder, and Laura Storto. 2000. *Evaluating Alternative Welfare-to-Work Approaches: Two-Year Impacts for Eleven Programs*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families and Office of the Assistant Secretary for Planning and Evaluation; and U.S. Department of Education, Office of the Under Secretary and Office of Vocational and Adult Education.

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

Friedlander, Daniel, and Gary Burtless. 1995. *Five Years After: The Long-Term Effects of Welfare-to-Work Programs*. New York: Russell Sage Foundation.

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

Hamilton, Gayle, and Thomas Brock. 1994. *The JOBS Evaluation: Early Lessons from Seven Sites*. Washington, DC: U.S. Department of Health and Human Services and U.S. Department of Education.

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

Hamilton, Gayle, Thomas Brock, Mary Farrell, Daniel Friedlander, and Kristen Harkenett. 1997. *Evaluating Two Welfare-to-Work Program Approaches: Two-Year Findings on the Labor Force Attachment and Human Capital Development Programs in Three Sites*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families and Office of the Assistant Secretary for Planning and Evaluation; and U.S. Department of Education, Office of the Under Secretary and Office of Vocational and Adult Education.

Hock, Ellen, and Debra K. DeMeis. 1990. "Depression in Mothers of Infants: The Role of Maternal Employment." *Developmental Psychology* 26(2): 285-291.

Macke, Anne Statham, Paula M. Hudis, and Don Larrick. 1978. "Sex-Role Attitudes and Employment Among Women: Dynamic Models of Continuity and Change." *Women's Changing Roles at Home and on the Job*. Special Report No. 206. Washington, DC: U.S. Department of Labor, Employment and Training Administration.

Michalopoulos, Charles. 1994. "Preferences, Attitudes, and Women's Employment." Ph.D. dissertation. University of Wisconsin-Madison.

Michalopoulos, Charles, Philip K. Robins, and David Card. 1999. *When Financial Incentives Pay for Themselves: Early Findings from the Self-Sufficiency Project's Applicant Study*. Ottawa: Social Research and Demonstration Corporation.

Miller, Cynthia, Virginia Knox, Patricia Auspos, Jo Anna Hunter-Manns, and Alan Orenstein. 1997. *Making Welfare Work and Work Pay: Implementation and 18-Month Impacts of the Minnesota Family Investment Program*. New York: Manpower Demonstration Research Corporation.

Pearlin, L. L., E. G. Monaghan, M. A. Lieberman, and J. T. Mullan. 1981. "The Stress Process." *Journal of Health and Social Behavior* 22: 337-356.

Plotnick, Robert, Marieka Klawitter, and Mark Edwards. 1998. "Do Attitudes and Personality Characteristics Affect Socioeconomic Outcomes? The Case of Welfare Use by Young Women." Institute for Research on Poverty Discussion Paper No. 1161-98. University of Wisconsin-Madison.

Quint, Janet, Kathryn Edin, Maria Buck, Barbara Fink, Yolanda Padilla, Olis Simmons-Hewitt, and Mary Valmont. 1999. *Big Cities and Welfare Reform: Early Implementation and Ethnographic Findings from the Project on Devolution and Urban Change.* 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.

Rohde, Paul, Peter M. Lewinsohn, and John R. Seeley. 1990. "Are People Changed by the Experience of Having an Episode of Depression? A Further Test of the Scar Hypothesis." *Journal of Abnormal Psychology* 99(3): 264-271.

Spitze, Glenna D., and Linda J. Waite. 1980. "Labor Force and Work Attitudes: Young Women's Early Experiences." *Sociology of Work and Occupations* 7(1): 3-32.

Statham, Anne, and Patricia Rhoton. 1983. "Attitudes Toward Women Working: Changes over Time and Implications for the Labor-Force Behaviors of Husbands and Wives." In Lois Banfill Shaw (ed.), *Unplanned Careers: The Working Lives of Middle-Aged Women*. Lexington, MA: Lexington Books.

Wells, Kenneth B., Jacqueline M. Golding, and M. Audrey Burnam. 1988. "Psychiatric Disorder and Limitations in Physical Functioning in a Sample of the Los Angeles General Population." *American Journal of Psychiatry* 145(6): 712-717.

Winefield, Anthony H., and Marika Tiggemann. 1990. "Employment Status and Psychological Well-Being: A Longitudinal Study." *Journal of Applied Psychology* 75(4): 455-459.